PERCEPTS, CONCEPTS AND CATEGORIES

The Representation and Processing of Information

Edited by

Barbara BURNS
Department of Psychology
University of Louisville
Louisville, KY, U.S.A.

ADVANCES IN PSYCHOLOGY

93

Editors:

G. E. STELMACH

P. A. VROON

1992

NORTH-HOLLAND
AMSTERDAM • LONDON • NEW YORK • TOKYO

Percepts, Concepts and Categories
B. Burns (Editor)

9

Basic Levels in Artificial and Natural Categories: Are All Basic Levels Created Equal?

MARY E. LASSALINE
EDWARD J. WISNIEWSKI
DOUGLAS L. MEDIN
University of Michigan

I. INTRODUCTION

This chapter addresses the nature of the basic level of conceptual structure in both naturally occurring and artificially constructed categories. We will look closely at the relationship between studies of the basic level that involve natural stimuli and those that involve artificial stimuli, and examine the extent to which structural properties of natural categories have been incorporated into artificial categories. More specifically, we evaluate studies of the basic level involving artificial categories with respect to how well conclusions from such studies extend to natural categories. In this context, we will address an important question: how confident can one be that a phenomenon called the "basic level advantage" *emerges for the same reasons* in artificial category studies as it does in natural category studies? One can view the basic level advantage as the output of a process operating on a representation. The fact that an artificial category study and a natural category study result in the same output by itself does not guarantee that they reflect the same representation and/or the same process. Before pursuing this point, however, we need to provide a bit more by way of background.

The single most important issue in the study of concepts has been the nature of conceptual structure. In the last 20 years or so, the investigation of this issue has led to a number of perspectives on concepts that are quite different from those of earlier views. As one example, the idea that all members of a category have fundamental characteristics in common that define category membership (i.e., the Classical view) has been replaced by the idea that category members *tend* to share properties that are neither necessary nor sufficient for category membership (see Smith & Medin, 1981, for a review).

Another notion that has been developed and refined over the years is that natural categories have both a vertical and horizontal component (Rosch, 1978). Drawing upon arguments presented in philosophy and logic (Wittgenstein, 1953; Zadeh, 1965), Eleanor Rosch proposed that natural categories can be viewed in terms of a gradient of typicality or goodness of category membership. Some members are the clearest cases of the category or the most prototypical members. Other members vary in terms of how good an example they are of the category (Rosch, 1975; Rosch & Mervis, 1975). Rosch further claimed that this horizontal dimension of conceptual structure is governed by the principle of *family resemblance* (see Wittgenstein, 1953). Under the family resemblance principle, category members considered the most prototypical are those with the most attributes in common with other members and the fewest attributes in common with members of contrasting categories. In a series of studies using both natural

and artificial categories, Rosch and Mervis (1975) found a positive correlation between family resemblance and ratings of how good a member was of a category.

There is also a vertical component to category structure, given that objects can be categorized at a number of levels of generality. An object being driven on a highway, with four wheels, and a top that folds back, can be called a *convertible, a car*, or *a vehicle*. The category *car* is more general than *convertible* because it includes other objects (e.g., station wagons, hard top sedans) as well as the members of *convertible*. The category *vehicle* is more general than *convertible* and *car* because it contains other objects (e.g., trucks, trains) as well as the members of these categories. Such categories form a taxonomy, or a class inclusion hierarchy. In a seminal paper, Rosch, Mervis, Gray, Johnson, and Boyes-Braem (1976) singled out one level in such a hierarchy, which they called the basic level, as playing a central role in many cognitive processes associated with categorization. For example, the category level represented by *chair, hammer, car,* and *dog* is typically considered the basic level. These concepts can be contrasted with more general "superordinate" concepts, *furniture, tool, vehicle,* and *animal*, as well as the more specific "subordinate" concepts, *recliner, ball-peen hammer, convertible,* and *labrador retriever.*

Studies using a variety of cognitive tasks show that basic level concepts have advantages or privileges over other concepts. Pictures of isolated objects are categorized faster at the basic level than at other levels (Jolicoeur, Gluck, & Kosslyn, 1984; Murphy & Brownell, 1985; Murphy & Wisniewski, 1989; Rosch et al., 1976; Smith, Balzano, & Walker, 1978). People almost exclusively use basic category names in free naming tasks (Rosch, et al., 1976). Children learn basic concepts sooner than other types of concepts (Anglin, 1977; Brown, 1958; Horton & Markman, 1980; Mervis & Crisafi, 1982; Rosch et al., 1976). Finally, different cultures tend to employ the same basic level categories, at least for living kinds (Rosch, 1974).

Researchers have examined a number of explanations for the basic level advantage. Some explanations have focused on differences in the structure of categories at different levels (Gluck & Corter, 1985; Jones, 1983; Murphy & Brownell, 1985; Rosch et al., 1976). For example, people have argued that, relative to other concepts, basic concepts are the most differentiated (Murphy & Brownell, 1985; Rosch et al., 1976). That is, their members have many common features which are also distinctive. Other explanations have suggested that differences in the *content* of categories are responsible for the advantage (Tversky & Hemenway, 1984; Murphy & Wisniewski, 1989; Rosch, et al., 1976). For example, Tversky and Hemenway (1984) argued that the basic category advantage is due to the fact that these concepts primarily represent parts of objects.

One prominent experimental strategy for analyzing the basic level is to employ artificially constructed categories. The general rationale is to identify some variable or structural property of interest, incorporate that property into artificially constructed categories, and then run experiments to evaluate the role and importance of that property in categorization. This strategy allows one to control for a variety of correlated variables which might affect performance with natural categories. For example, basic level category names tend to be both more frequent and shorter than the labels for subordinate or superordinate categories (e.g., *dog* versus *labrador retriever* or *animal*). Does the basic level advantage hinge on frequency and length of category name rather than structural properties? Using artificially constructed categories, Murphy and Smith (1982) controlled for these factors and still observed that basic level categories were easiest to learn and were associated with the fastest categorization times. Therefore, neither word length nor frequency is necessary for one to observe a basic level advantage.

This paper is concerned with the basic level in both artificial and natural categories. We will pay particular attention to relationships between artificial and natural categories. Two critical issues concern the extent to which relevant structural properties of natural categories have been successfully incorporated into artificial categories and the related issue of just what conclusions concerning basic level categories are licensed by studies involving artificial categories. To anticipate, we will argue that the link between artificial and natural categories in analyzing basic levels has not been developed carefully enough to support strong inferences from one to the other. That's the bad news. The good news is that work using artificial categories is interesting in its own right because it can be successfully linked to theories and other metrics aimed at specifying structural properties which determine the relative difficulty of different levels of categorization.

This chapter is organized in the following way. We first take a closer look at the history and rationale for research with artificially constructed categories. We then describe some initial work on basic levels and theories and metrics relevant to them. Next, we critically examine the relation between artificial and natural basic level categories and describe some recent research in our laboratory on artificial basic levels. Finally, we try to summarize the current state of affairs both with respect to theories and metrics for the basic level and with respect to mapping between the natural and the artificial.

II. ARTIFICIAL CATEGORIES

At first thought the idea of condensing the years of experience associated with the learning of natural categories into an hour or so of training in a laboratory seems very, very tenuous. Probably the strongest argument for believing that important properties of categorization of natural objects can be captured in laboratory studies with artificial categories is the remarkable parallels between the two domains. For example, in their initial pioneering studies of the family resemblance principle, Rosch and Mervis (1975) included two experiments using artificial categories. Specifically, the stimuli were strings of letters conforming to a family resemblance structure. Rosch and Mervis found that degree of family resemblance affected ease of learning, identification response latency, and typicality rating, just as it does in natural categories.

These parallels have continued. For example, one striking phenomenon is that the prototype or best example of a category may be classified more accurately in transfer tests than previously seen examples that were used during original category learning (e.g., Homa & Vosburgh, 1976; Medin & Schaffer, 1978). Furthermore, people are sensitive to correlated attributes within artificial categories (Medin, Altom, Edelson & Freko, 1982) and within natural categories (Malt & Smith, 1983). And most relevant to our purposes, it appears that a basic level advantage can be obtained with artificial materials (e.g,. Murphy & Smith, 1982).

These successful parallels may have undermined natural conservatism about mapping between natural and artificial categories. Indeed, 15 years ago the position of devil's advocate would have been that findings from artificial and natural categories are closely linked. Now, the opposite stance is required for that role. Our present strategy will be to apply fairly strict standards in evaluating parallels.

The relationship between artificial and natural category studies has important consequences for models that have addressed basic level phenomena. These models typically have simulated the results of artificial category studies. Therefore, it is important to evaluate how well such studies reflect the "true" basic level phenomena involving natural stimuli. If the basic level emerges for the same reasons for artificial and for natural categories, then a theory that explains the basic level in one domain necessarily explains the basic level in the other.

A second, independent goal of the chapter is to evaluate concept learning models with respect to how well they account for basic level phenomena. We will examine a number of classes of models, whether or not they have explicitly addressed the basic level. Given the significance of the basic level, any theory should be held accountable for the cognitive phenomena associated with it. We will evaluate four major classes of models: probability-based models that capture feature distributions within and between categories; prototype models that integrate information across category instances to form a summary representation; exemplar-based models that preserve the feature structure of individual category members; and theories based on the use of explicit rules to determine category membership. In this chapter, the classes of models mentioned above are held accountable for the vertical organization of conceptual structure, that which characterizes categorization between taxonomic levels. Before describing these models in detail, however, we take a closer look at the basic level.

III. THE BASIC LEVEL IN NATURAL CATEGORIES

Psychological research on the basic level primarily began with the work of Eleanor Rosch and her colleagues (Mervis & Rosch, 1981; Rosch, 1975; Rosch, 1978; Rosch et al., 1976). In a series of very influential studies, Rosch et al. (1976) showed that the basic level is the most general level that is *informative* and the most general level that shows processing advantages in a variety of cognitive tasks (e.g., acquisition of concepts, identification of an object from an average shape, and identity judgment). Virtually all researchers have used one or more of these tasks to study the basic level, in either natural or artificial categories. Therefore, we will describe the Rosch et. al findings in some detail.

The basic level is the most general level that is informative in the sense that category members at this level share a large quantity of information. Rosch et al. (1976) provided evidence for this claim, using a number of converging measures. In one task, designed to examine the co-occurrence of attributes in common taxonomies of natural objects, subjects listed the attributes of superordinate, basic, and subordinate categories. The nine taxonomies used as stimuli are presented in Figure 9.1. For the nonbiological taxonomies, members of basic categories shared significant numbers of attributes. Categories at a more general level (i.e., superordinates) had few attributes in common, whereas categories more specific than the basic level (i.e., subordinates) shared only a few additional attributes. In a second task, subjects described the kinds of motor movements that they made in interacting with category members (e.g., bending one's knees in a particular way to sit on a chair). Basic level categories were the most general categories that shared a significant number of motor movements that could be made toward their members. Superordinates had few common movements whereas subordinates had about the same number of common movements as basic categories. In a third task, Rosch and her colleagues computed the ratio of the area of overlap to nonoverlap of the normalized shapes of category members as an index of within category similarity. They showed that in going from the superordinate level to the basic level, there was a large increase in the similarity of basic category members. In going from the basic level to the subordinate level, there was a significant but significantly smaller increase in similarity of subordinate category members.

Superordinate	Basic Level	Subordinates	
	Nonbiological taxonomies		
Musical instrument	Guitar	Folk guitar	Classical guitar
	Piano	Grand piano	Upright piano
	Drum	Kettle drum	Bass drum
Fruit	Apple	Delicious apple	Mackintosh apple
	Peach	Freestone peach	Cling peach
	Grapes	Concord grapes	Green seedless grapes
Tool	Hammer	Ball-peen hammer	Claw hammer
	Saw	Hack hand saw	Cross-cutting hand saw
	Screwdriver	Phillips screwdriver	Regular screwdriver
Clothing	Pants	Levis	Double knit pants
	Socks	Knee socks	Ankle socks
	Shirt	Dress shirt	Knit shirt
Furniture	Table	Kitchen table	Dining room table
	Lamp	Floor lamp	Desk lamp
	Chair	Kitchen chair	Living room chair
Vehicle	Car	Sports car	Four door sedan car
	Bus	City bus	Cross country bus
	Truck	Pick up truck	Tractor-trailer truck
	Biological taxonomies		
Tree	Maple	Silver maple	Sugar maple
	Birch	River birch	White birch
	Oak	White oak	Red oak
Fish	Bass	Sea bass	Striped bass
	Trout	Rainbow trout	Steelhead trout
	Salmon	Blue back salmon	Chinook salmon
Bird	Cardinal	Easter cardinal	Grey tailed cardinal
	Eagle	Bald eagle	Golden eagle
	Sparrow	Song sparrow	Field sparrow

Figure 9.1. Taxonomies used in Rosch et al. (1976) Experiment 1.

Basic categories also were the most general categories that had processing advantages in a number of cognitive tasks. In a fourth task, subjects were given shapes that were averages of the shapes of superordinate, basic, or subordinate category members. Averages of superordinate members could not be identified significantly better than chance. On the other hand, basic categories were the most general categories for which averages of their members could be readily identified. Averages of subordinate members were no more identifiable than the averages of basic category members. In a fifth task, the most general names that aided the detection of objects presented in noise were those of basic categories. In a sixth task, basic category names were the most general names which facilitated judgments about whether two objects were physically identical.

In the cognitive tasks described above, performance at the subordinate level actually was as good or (sometimes) better than performance at the basic level. Rosch et al. (1976) also showed that for some tasks, the basic level had advantages over both the superordinate and subordinate levels. In a categorization task, subjects more quickly identified objects as members of basic categories than as members of superordinate or subordinate categories. In a naming task, subjects overwhelmingly labeled objects with the names of basic level categories. This result occurred even though in one stimulus set, a superordinate name was sufficient to distinguish one object from another, and in another set, a subordinate name was necessary to distinguish an object from other objects.

Finally, Rosch et al. (1976) examined developmental differences in the acquisition of concepts at different levels of generality. Two studies showed that children can sort objects into basic level categories earlier than they can sort objects into superordinate categories. In a task involving oddity problems, nursery and elementary school children and adults were shown sets of three color photographs. The sets varied on whether two of the three pictures came from the same basic level category or from the same superordinate category (the "odd" picture out always belonged to a different superordinate). Subjects were instructed to "put together the two that go together, the two that are alike." At all ages, sorting into basic level categories was almost perfect, but sorting into superordinate categories showed a significant improvement with age. In the other study, elementary school children and adults were given a set of 16 pictures, either from four basic level or from four superordinate categories, and asked to sort the entire set into "the ones that go together, the ones that are the same kind of thing." Results of this second, more traditional, sorting study were similar to the first. In a third developmental study, Rosch and her colleagues examined protocols of spontaneous speech during a child's initial language acquisition. They found that virtually all of the child's first utterances of concrete nouns were at the basic level.

Rosch explained the importance of the basic level in terms of a general principle called *cognitive economy*. This principle reflects a compromise between two somewhat conflicting reasons for having concepts. On the one hand, it would seem advantageous for a concept to be as informative as possible. As a result, we could predict many properties of an object by knowing its category. Carried to the extreme, however, this tendency would favor the formation of many specific concepts--one for *each* object encountered in the world. This extreme view conflicts with another reason for having concepts: they allow us to ignore differences among objects so that we can treat them equivalently as members of the same category. Carried to the extreme (i.e., ignoring all differences among objects), this tendency would favor the formation of a *single*, very general concept--one corresponding to *all* objects encountered in the world. Rosch suggested that the basic level results from the combination of these tendencies: basic level categories are both informative *and* general.

More specifically, Rosch implied that concepts at the basic level were more *differentiated* than those at other levels. In particular, she suggested that the basic level was the most general level with categories whose members had many attributes in *common* which were also *distinct* from the attributes of other categories at that level. In contrast, the members of the more general, superordinate categories have few features in common. Furthermore, although members of more specific, subordinate categories have slightly more features in common than those of basic level categories, many of these features are not distinctive. That is, they are shared by the members of other subordinate categories (e.g., members of *kitchen chair* share a number of features with the members of other categories of *chair*).

In addition to this *structural* explanation of the basic level, Tversky and Hemenway (1984) suggested that differences in the *content* of categories were responsible for the importance of the basic level. They argued that the features of basic categories which are both common and distinctive are primarily parts of objects. Reanalyzing the attribute listing data of Rosch et al. (1976), they found that a majority of the attributes that were shared by the members of basic categories were parts (e.g., "keys," "legs," "footpedal," and "strings" for the members of *piano*). In contrast, the proportion of attributes that were parts was less for subordinate categories (and was a minority for biological categories). Subjects listed only a few parts for superordinate categories. Furthermore, in two other studies, they showed that parts were the distinctive attributes of basic categories. Specifically, different basic categories tended not to share the same parts but instead, tended to share features that were not parts (e.g., "makes sound" for musical instruments and "sweet" for fruit). In contrast, subordinate categories tended to share the same parts but to differ on features that were not parts (e.g., subordinates of the basic category "chair" have legs, a seat, a back and typically arms, but "living room chair" is distinguished from "kitchen chair" by being large and soft).

Tversky and Hemenway (1984) suggested that parts and their configurations underlie many of the cognitive tasks that converge at the basic level. Parts and the way that they are configured primarily determine the shape of an object. Shape in turn, would seem to play an important role in a number of tasks described above. Recall that in going from the superordinate level to the basic level, there was a large increase in the overlap of the shape of basic category members. Basic categories also were the most general categories for which an average shape of their members could be readily identified. These results are consistent with the finding that basic level category members have many parts in common. Furthermore, because basic level categories have distinctive parts, they may also have distinctive shapes. This factor may account for why categorization of pictures was superior at the basic level. Finally, basic level categories were the most general categories that shared a significant number of motor movements that could be made toward their members. Intuitively, it seems that many such movements involve parts (e.g., we sit on the seat of a chair, we grasp the handle of a hammer, and so on).

The findings of Rosch et al. (1976) and Tversky and Hemenway (1984) applied to object taxonomies of biological and nonbiological categories. Researchers have found evidence consistent with a basic level in other domains, including environmental categories (Tversky & Hemenway, 1983), computer programs (Adelson, 1988), personality types (Cantor & Mischel, 1979) and events (Morris & Murphy, 1990; Rifkin, 1985). We describe two of these other domains in detail. Tversky and Hemenway (1983) suggested that there was a basic level for categories of scenes. Two of the three tasks that they used were originally employed by Rosch et al. In an attribute listing task, the number of features that subjects listed for scenes described at an intermediate (basic) level of generality (e.g., home, school, beach, and mountain scenes) was substantially greater than the number they listed for scenes described at a more general level (e.g., indoors and outdoor scenes). There was only a small increase in the number of attributes that subjects listed for more specific scenes (e.g., apartment, high school, lake beach, and Rocky mountains scenes). In a naming task, subjects primarily labeled photographs of scenes with their basic level names. In a third task, subjects

primarily completed sentences describing activities performed in scenes with basic level names. The results of the last two tasks occurred even though in some cases, superordinate names were sufficient to distinguish scenes from others and in other cases, subordinate names were necessary to distinguish scenes from others. These findings for scenes parallel those of Rosch et al. (1976) involving objects.

Morris and Murphy (1990) suggested that there was a basic level for events (see also Rifkin, 1985). Three of the four tasks that they used paralleled those employed by Rosch et al. (1976). In one task, subjects listed the actions of events. The number of actions listed for events described at an intermediate (basic) level of generality (e.g., breakfast, dinner, classes, and tests) was substantially greater than the number listed for events described at a more general level (e.g., meals and school activities). There was only a small increase in the number of actions that subjects listed for more specific events (e.g., quick breakfast, family dinner, English class, multiple choice exam). In another task, subjects verified whether an action expressed as a verb phrase was part of either a subordinate, basic, or superordinate event category. So, for example, subjects verified whether the action "scream during the scary parts" was true of "horror movie" (a subordinate event), "movie" (a basic event) or "entertainment" (a superordinate event). Subjects were fastest to verify that actions were parts of basic events, slightly less faster to verify that they were parts of subordinates, and substantially slower to verify that they were parts of superordinates. In a third task, subjects were asked to choose appropriate names for events presented in story contexts. In most cases, they preferred to label the events with basic level names. One notable exception involved stories about events that could be classified in the same basic category (a horror movie and a comedy). Here, subordinate terms were necessary to distinguish these events. Unlike Rosch et al.'s findings with objects and Tversky and Hemenway's findings with scenes, subjects preferred subordinate terms slightly more than basic level terms.

A fourth task used by Morris and Murphy is interesting because it presents an alternative way of measuring the differentiation of a category. Recall that Rosch et al. viewed differentiation in terms of the *number* of attributes that were common and distinctive for a category. Such a measure does not take into account the possibility that some features are more important than others. An alternative way to measure differentiation, first suggested by Mervis and Crisafi (1982), is to subtract between-category similarity from within-category similarity. Morris and Murphy had subjects rate the similarity of category members at various levels of generality. Interestingly, they found that superordinate events were more differentiated than basic category events. They explained this apparent paradox by suggesting that functional features were more central to event concepts than to object concepts, since events are structured by goals rather than by perceptual shape. These features apparently carried a lot of weight in judging the similarity of superordinate members.

IV. THE BASIC LEVEL IN ARTIFICIAL CATEGORIES

Given the set of correlated measures that converge on the basic level, it is far from easy to identify which factors are central and which are derivative. Therefore, it is not surprising that investigators have turned to artificial category learning situations to disentangle these factors.

Murphy and Smith (1982) examined the basic level advantage for picture categorization found by Rosch et al. (1976). Looking at that study, Murphy and Smith noted that a variety of factors other than number of distinctive features may have accounted for why pictures were categorized faster at the basic level than at either the subordinate or superordinate level. In comparing subordinate categories to basic categories, subordinate names were considerably longer, some of the subordinates may have been unfamiliar (e.g., *cross-cutting hand saw*), and some of the differentiating features of subordinates may have been difficult to perceive (e.g., as in *green seedless grape*). In addition, there are some general characteristics of natural basic categories that may have accounted for their superior performance: basic categories are learned first, their names occur more frequently, their members occur more frequently than those of subordinates, and there may be a higher conjoint frequency between basic category names and objects (e.g., cars may be called *car* more frequently than they are called *vehicle* or their appropriate subordinate name).

In several studies, Murphy and Smith generally replicated the findings of the Rosch et al. picture categorization task, using artificial stimuli and removing the confounds described above. Those factors either were held constant (e.g., using CVC labels to equate category name length) or systematically varied (e.g., order of learning the different category levels was counterbalanced). In addition, Murphy and Smith explicitly designed the artificial stimuli of their first experiment so that they resembled natural subordinate, basic, and superordinate categories. Superordinates had a single attribute common to their members (namely, function), basic categories had many common features (which gave them distinctive shapes), and subordinate category members had only one attribute that could differentiate subordinates of the same basic category (e.g., there were two subordinate categories of "knife," differentiated by whether or not the blade was serrated). Figure 9.2 presents a sample of the stimuli (modeled after real-world

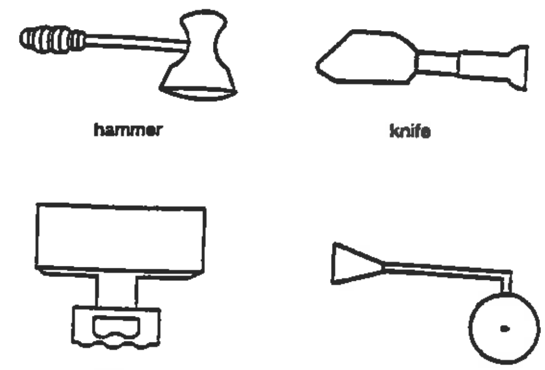

Figure 9.2. Examples of four basic tools used in Murphy and Smith (1982) Experiment 1.

tools) used in that experiment.

Although Murphy and Smith generally replicated the Rosch et al. findings, there also were important differences between their results and those of Rosch et al. In the experiment that most closely resembled that of Rosch et al., Murphy and Smith found that people categorized pictures into subordinate categories almost as quickly as they did into basic categories. In contrast, Rosch et al. found that categorization at the subordinate level was slowest. Murphy and Smith suggested that the Rosch et al. results may have resulted from some of the factors described above, including name length and perceptibility of distinguishing features.

While supporting the claim made by Rosch et al. that basic categories were associated with more distinctive attributes, Murphy and Smith argued that only perceptual features were critical to basic level superiority. The importance of perceptual features is most readily seen in comparing basic and subordinate categories to superordinate categories. Murphy and Smith (also Rosch et al., 1976) suggested that the advantage of basic and subordinate categories over superordinates arises because people typically do not have a single, perceptual representation for a superordinate. When deciding that an item belongs to a superordinate category, people must activate a *number* of representations. Such a decision involves extra capacity and more matching of features. To support this claim, Murphy and Smith conducted a study in which the members of superordinates shared a distinctive perceptual feature (size). After the experiment, subjects were divided into two groups, based on whether or not they reported attending to the size cue. For those subjects that did attend to the size cue (8 of

Table 9.1. Abstract category structure of Corter et al. (1988) Experiment 1.

Stimulus	Categories: Top	Middle	Bottom	Feature Dimensions: Gums	Eyes	Rash
1	queritism	burlosis	jirenza	1	1	1
2	queritism	burlosis	malenza	1	2	1
3	queritism	cretosis	gilenza	2	3	2
4	queritism	cretosis	surenza	2	4	2
5	philitism	midosis	habenza	3	1	2
6	philitism	midosis	kelenza	3	2	2
7	philitism	nitosis	tumenza	4	3	1
8	philitism	nitosis	valenza	4	4	1

12), categorization times were faster for superordinates than for intermediate level categories, which were not defined perceptually.

To test Murphy and Smith's claim that only perceptual features are critical for the basic level advantage, Corter, Gluck, and Bower (1988) conducted two experiments using artificial categories that were defined in terms of verbal or conceptual features rather than perceptual ones. The categories were diseases, characterized by three symptoms (gums, eyes and rash), each having four possible values (swollen, discolored, bleeding, or sore gums; puffy, sunken, red, or burning eyes; and blotchy, spotted, itchy, or scaly rash). Table 9.1 shows the diseases, symptoms, and abstract category structure of the experiment.

The category structure closely paralleled that used by Hoffman and Ziessler (1983), but substituted verbal or conceptual features for the perceptual features that they used. Based on the table, one can see that top-level categories are defined by a disjunction of values of the first symptom, middle-level categories are defined by a single value of the first symptom, and bottom-level categories are defined by a conjunction of values of the first and second symptoms. Given this structure, Corter et al. argued that the middle level should be the preferred level because it is the highest level for which category members share values on one or more symptoms. Like Hoffman and Ziessler (1983), they found that people categorized items faster at the middle level than at other levels, and learned the middle-level categories the fastest. These results suggest that perceptual features are not necessary for a basic level advantage.

In a developmental study, Mervis and Crisafi (1982) tested the claim made by Rosch et al. that children first learn to categorize at the basic level before learning to categorize at other levels. Like Rosch et al., Mervis and Crisafi

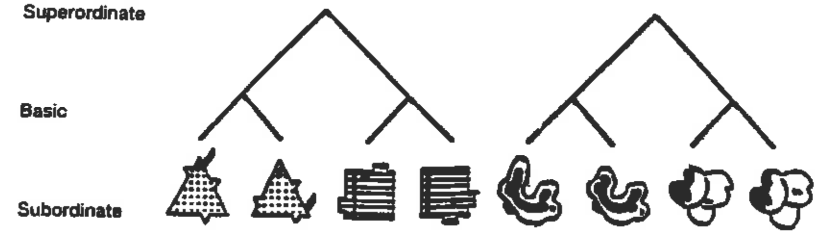

Figure 9.3. Examples of stimuli used in Mervis and Crisafi (1982) Experiment 1.

used a task involving oddity problems (see above). Figure 9.3 presents a sample of their stimuli.

Specifically, children were shown a picture of an object (the standard) and asked which of two other pictures went with the standard. The standard and one of the other pictures either came from the same subordinate, basic, or superordinate category. However, their task provided a more stringent test of the Rosch et al. claim in three ways. First, for the sets involving basic categories, children only could solve the problems by attending to cues at the basic level. In basic level sets used by Rosch et al., the "odd" item out came from a different superordinate, so it was possible to put the other two items together because they belonged to the same superordinate and not because they belonged to the same basic category. So for example, in the set *beagle, cocker spaniel,* and *jet*, one can put *beagle and cocker spaniel* together because they are both animals, rather than because they are both dogs. Mervis and Crisafi removed this possibility by using sets of pictures in which two items came from the same basic category but all of the items came from the same superordinate (e.g. *beagle, cocker spaniel and deer*). Second, Mervis and Crisafi used artificial stimuli which were never labeled, thereby ruling out some of the same alternative explanations noted by Murphy and Smith (see above). Third, this study also examined acquisition of subordinate categories (it was not considered in the Rosch et al. study).

Like Murphy and Smith, Mervis and Crisafi designed their artificial stimuli to mimic the structure of natural categories (see Figure 9.3). Basic category members had very similar overall shapes and shared a number of distinctive attributes. Superordinates shared very general abstract attributes. Subordinate category members shared more features than members of basic categories, but their features were not as distinctive.

As predicted, children were able to categorize at the basic level at a younger age than at other levels. For the basic-level sets, all children (ages 2-6, 4, and 5-6 year-olds) performed significantly more accurately than would be expected by chance. For the superordinate level sets, 4 and 5-6 year-olds performed significantly more accurately than chance. For the subordinate-level sets, only the 5-6 year-olds performed significantly more accurately than chance. These results imply a particular order of category learning. Basic level concepts are learned first, followed by superordinate concepts, and then subordinate concepts.

The stimuli used may have been problematic, though, on two counts. First, it is unclear that members of subordinate categories were discriminable from members of their contrast categories. For natural categories, subordinates may be much more distinct from one another. More generally, experimenters have the freedom to make subordinate categories arbitrarily similar to each other, and it is not clear how similar subordinates "should" be. This factor weakens the primary conclusion drawn from this work, namely that categorization at the basic level is acquired first.

Recently, Murphy (1991) has investigated whether the importance of parts in basic level categories (as detailed by Tversky and Hemenway, 1984) is caused by a psychological principle that requires basic concepts to represent such information. That is, the importance of parts could reflect an intrinsic bias of the human conceptual system. Alternatively, the importance of parts may reflect the structure of object categories in the world, rather than a special property of the conceptual system. By this account, people generally are sensitive to the common and distinctive information of categories, whether or not it is associated with the parts of objects. It just happens to be true of objects that this information is based on their parts. Murphy (1990) provides evidence that parts are neither necessary or sufficient for a basic level advantage. He found a basic level advantage for artificial categories whose members did not have parts in common but rather shared distinctive, nonpart information such as size, texture, and type of border. He also found that the reaction time advantage of basic categories (containing *distinctive parts*) could be increased even further by adding *distinctive nonpart* information to them. On the other hand, this advantage could be eliminated when the distinctive nonpart information was associated with another category.

Although none of the studies so far described are flawless, they encourage the view that one can reliably obtain an advantage for an intermediate level of categorization. Regardless of whether the parallels to the basic level with natural categories are precise, these findings do have implications for theories of categorization. As we shall see, it is nontrivial to develop a model that favors an

intermediate level of categorization. Let's take a closer look at some of these metrics and models.

V. METRICS, THEORIES, AND THE BASIC LEVEL

There have been two theoretical approaches to analyzing the basic level. One strategy has been to specify what types of information or structural properties might be maximized at the basic level. The other strategy has been to develop and evaluate categorization models in terms of their ability to account for basic level advantages. In some cases, the metrics and theories are linked in that the theory may explicitly embody the metric of interest. We first describe some metrics and then turn to associated theories.

A. Metrics

Cue validity. One possibility is that the basic level maximizes the cue validity of features. The cue validity of a feature *f* is the probability that an item belongs to a category *c*, given that it has the feature, that is, $P(c \mid f)$. In fact, Rosch et al. (1976) suggested that the most differentiated categories (i.e., basic categories) maximized a related measure, which they called *category cue validity*. This measure is the sum of the cue validities of features for the category. Carried to the extreme, maximizing either measure would favor the formation of a single, very general concept of all objects encountered in the world. Furthermore, a number of researchers (e.g., Medin, 1983; Murphy, 1982) have shown that for nested categories, category cue validity increases with the generality of a category, or stays the same. A simple example illustrates this point. Consider the nested categories *sparrow, bird,* and *animal*, and the feature "flies." Let v be the probability that an item with the feature "flies" is a *sparrow*. The probability w that an item with the feature "has wings" is a *bird* must be greater than v because there are more birds than sparrows that fly. Furthermore, the probability x that an item with the feature "flies" is an animal must be greater than w because there are more animals than birds that fly (e.g., bats, some insects, perhaps some fish, etc).

Category validity. Another possibility is that the basic level maximizes the category validity of features. Category validity is the probability that an item has a particular feature *f*, given that it belongs in some category *c*, that is, $P(f \mid c)$. This measure emphasizes how well unknown features of an object can be predicted from knowledge of its category membership. Carried to its extreme, maximizing category validity would favor the formation of a concept for each object encountered in the world. Therefore, it always will pick out the subordinate or least inclusive level as the most advantageous (Medin, 1983), a problem directly opposite from that faced by cue validity (favoring the most inclusive level). Consider again the nested categories *sparrow, bird, and animal*, and the feature "flies." Let v be the probability that given that an item is a sparrow, it has the feature "flies" and let w be the probability that given that an item is a bird, it has the feature "flies." Because there are more birds than sparrows that do not fly (e.g., penguins, ostriches), w must be less than v. Furthermore, the probability x, given that an item is an animal it will have the feature "flies," must be less than w because there are more animals than birds that do not fly. Weighting category validity by a category's base rate avoids the extreme consequence of favoring the least inclusive level. Anderson (1990) uses this metric in his rational theory of categorization and has successfully modeled the basic level findings of Murphy and Smith (1982) and Hoffman and Ziessler (1983).

The product of cue and category validity. Jones (1983) suggested that the product of cue and category validity (which he called *category-feature collocation*) might determine the preferred level of categorization. Jones suggested applying this measure in the following way to predict which level should be basic. For each feature in each category of a nested hierarchy, compute its collocation score $P(f \mid c)*P(c \mid f)$. Then, "assign" each feature to that category in the hierarchy which has the highest collocation score for that feature. Basic categories should be those categories which have the greatest number of features assigned to them. Corter and Gluck (1990) have suggested an alternative way of applying this measure. For each category, one might compute the average of the collocation scores for all features. Basic categories should be those categories with the highest average collocation. Using a number of hierarchies involving natural and artificial categories, Corter and Gluck (1990) evaluated both measures along with a third metric called category utility (described below). They found that only category utility reliably predicted which categories should be basic.

Category Utility. Gluck and Corter (1985) developed a context sensitive metric analogous to H_t (information transmitted) of information theory, which they call *category utility*. It is the increase in the expected number of features that can be correctly predicted given knowledge of category membership versus no such knowledge. Specifically, the utility of a category C_k is the difference between the sum of squared category validity probabilities of its features, $\Sigma_i P(f_i \mid C_k)^2$, and the sum of squared base rates of those features, $\Sigma_i P(f_i)^2$, weighted by category's base rate, $P(C_k)$:

(1) $$\text{Category Utility } (C_k) = P(C_k)\,[\Sigma_i P(f_i \mid k)^2 - \Sigma_i P(f_i)^2]$$

Gluck and Corter assume that category utility should be calculated over a universal set of features, including those that do not occur in one category but do occur in a contrast category (that is, over any feature that appears in any category). For example, if feature x appears only in Category B, but not in Category A, when calculating the category utility for Category A, x is included as a feature. This assumption prevents the addition of a single instance from radically changing the feature domain. In addition, though, this assumption raises the problem of determining the relevant universe of features, a problem which has not been fully addressed.

Category utility predicts that a novel item will be placed into the category whose utility increases the most when that item is considered a member of the category. Furthermore, it is assumed that the level whose categories have the highest utility is the basic level. Corter and Gluck (1990) showed that the level predicted by category utility to be basic corresponded to the empirically determined basic level in the experiments of Rosch et al. (1976), Murphy and Smith (1982), and Hoffman and Ziessler (1983).

B. Theories of Categorization

Computational models of the basic level. Fisher (1987) developed an incremental conceptual clustering program, called COBWEB, that uses category utility. Conceptual clustering is a task in which items are partitioned into categories based on some criterion and a concept is formed for each category. It differs from other concept formation tasks in which people explicitly are given feedback on whether or not an item belongs in a category. Using category utility as its clustering criterion, COBWEB adds a novel item to an existing category (i.e., one that it has previously constructed) or places the item into a new, singleton category. Specifically, COBWEB tentatively places the item into an existing category and calculates the average category utility for that partition of categories:

(2) $$\frac{\Sigma_k P(C_k)\,[\Sigma_i\Sigma_j P(f_{ij} \mid k)^2 - \Sigma_i\Sigma_j P(f_{ij})^2]}{n}$$

In this formula, $P(f_{ij} \mid k)$ is the probability of attribute i possessing value j in category k, $P(f_{ij})$ is the base rate of attribute i possessing value j, and n is the number of categories in the partition. Note that this is simply category utility summed across categories and divided by the number of categories.

This process is repeated, placing the item into a different category each time. In addition, COBWEB calculates the category utility for the partition that results when the item is placed in a new category by itself. COBWEB then classifies the item into the category that results in the partition with the highest category utility.

COBWEB incrementally constructs categories by examining novel items sequentially. Eventually, it constructs a hierarchy of categories, related to each other by class inclusion. Fisher (1988) modified the program to account for basic level phenomena. Specifically, the program augments the category hierarchy with *collocation-maximizing indices*: pointers from the root of the hierarchy to those categories that maximize the collocation of an attribute value f_{ij}. Collocation, again, is the product of cue and category validity (see Jones, 1983). An object is classified into the category with the greatest sum of collocation. Fisher (1988) has shown that his system categorizes items at the level corresponding to the empirically determined basic level in the experiments of Murphy and Smith (1982) and Hoffman and Ziessler (1983).

Anderson (1990) has developed and extensively evaluated a clustering model which, like Fisher's COBWEB, embodies the idea that the organism's goal is to maximize the inferences that can be drawn from category membership. In the model, a novel item is placed into its most *probable* category, which either may be an existing category (i.e., one that previously has been constructed) or a new, singleton category. The probability P_k that an item belongs to a category k is $P(k \mid F)$, or the conditional probability that the item belongs to the category k, given its set of attribute values F. This probability can be expressed as the following equation:

$$(3) \qquad Pk = P(k \mid F) = \frac{P(k)P(F \mid k)}{\Sigma k P(k)P(F \mid k)}$$

The probability P(k) in turn, has the following form:

$$(4) \qquad P(k) = \frac{cn_k}{(1 - c) + cn}$$

where n_k is the number of items assigned to category k so far, n is the total number of items seen so far, and c is the *coupling probability*. This probability (a fixed value) indicates how likely two objects come from the same category (for most of Anderson's simulations, this probability was set to .30). For large n, P(k) approximates n_k/n, strongly biasing the system to place items into large categories.

The probability that an item displays a set of attribute values F, given membership in category k, is computed by the following equation:

$$(5) \quad P(F \mid k) = \Pi \, P(ij \mid k)$$

where $P(ij \mid k)$ is the probability of an item displaying value j on attribute i, given that it comes from category k. This probability in turn, has the following form:

$$(6) \quad \frac{n_{ij} + 1}{n_k + m_i}$$

where n_k is the number of items in category k that have a value j on attribute i, n_{ij} is the number of items in the category that have the same value as the item to be classified, and m_i is the number of values on the dimension to be classified. Notice that the *normative* value for $P(ij \mid k)$ is calculated by n_{ij} / n_k (and this equation approximates that value for large n). However, the term m_i is included in the equation above to deal with problems of small samples. As Anderson notes (page 105), if a system has just seen one item in a category and it was red, one would not want to assume that all items in the category were red (as would be the case if the probability was simply n_{ij} / n_k. Assuming seven colors, equally probable on prior grounds, P (red | k) equals .25 by direct substitution into equation 6.

Anderson has accounted for the basic level findings of Murphy and Smith (1982) and Hoffman and Ziessler (1983) in the following way. He presented the model with random sequences of the stimuli used in those experiments, encoded as bit patterns (with unique subpatterns corresponding to different values on the various feature dimensions). The system forms categories at various levels by varying the value of the coupling probability.

Prototype and exemplar models of categorization. In general, these models have not been held accountable for hierarchical categorization. Rather, they usually characterize such within-level categorization phenomena as typicality effects and new/old recognition. Much of this work contrasts the predictions of prototype versus exemplar-based theories of categorization. (e.g., Medin & Smith, 1984; Nosofsky, 1988). To our knowledge, however, these models previously have not been compared with respect to their predictions concerning levels of categorization. We will examine a pair of models that are examples of general classes of models: the adaptive network model (Gluck & Bower,1988), which is an example of *multiplicative-similarity prototype* models (see also Massaro's fuzzy logical model, Massaro & Freidman, 1990 and Nosofsky, 1990), and the context model (Medin & Schaffer, 1978; Nosofsky, 1986), which is an example of *multiplicative-similarity exemplar* models (see also the array model of Estes, 1986, and Hintzman's MINERVA, 1986).

The adaptive network model. Prototype models combine information about individual examples into a summary representation. In the case of the adaptive network model, this information is represented as a set of weights that connect input nodes of the system to output nodes. The system learns categories in the following manner. A category item is presented as a pattern of activity across the input nodes. The pattern of activity corresponds to a set of attribute values that describe the item. The system responds by generating a pattern of activity across its output nodes. Each output node represents a category. The response of a given output node is:

(7) $o_j = \Sigma_i w_{ij} a_i$

which is simply the sum of the activations of the input units to which the output node is connected, the a_i, multiplied by their weights, the w_{ij}. An input node has an activation value of 1 if the attribute value that it represents is present in the item. Otherwise, the node has a value of 0. The system is taught to activate the node of the correct category and to inhibit activation of the other nodes, by adjusting its weights, using the delta or least mean squares learning rule (Duda & Hart, 1973; Rescorla & Wagner, 1972):

(8) $\Delta w_{ij} = \beta a_i (o_j - \Sigma_i w_{ij} a_i)$

Here, the change in the weight between input unit i and output unit j, Δw_{ij}, is the difference between the desired output from unit j, o_j (which is 1 if the stimulus belongs to category j and is 0 otherwise), and the sum of the activations of the input nodes to which the output node is connected, the a_i, multiplied by their weights, the w_{ij}. This difference is multiplied by the activation of input node i, a_i and by the learning rate, ß, which is left as a parameter ($0<\beta<1$). This learning rule is a gradient descent procedure leading to minimum squared error over a set of learning items.

After some amount of learning, the system is typically tested on how well it has learned. An unclassified stimulus pattern is presented and the system generates a pattern of activity across its output nodes. The output is converted into a response probability for each of the possible categories, $P(R_j)$, by computing the ratio of the output for a category o_j, to the sum of the outputs for all categories i, Σo_i:

(9) $$p(R_j) = \frac{oj}{\Sigma o_i}$$

There are several problems with models of these types. First, it is well known that these relatively simple models only can learn *linearly separable* categories (e.g., Hinton, 1987; Minsky & Papert, 1969). Categories are linearly separable if their members can be correctly classified using a weighted, additive function of their features. Nosofsky (1991) also has shown that the adaptive network model is a special case of a class of multiplicative prototype models which have the same limitation. Apparently the constraint that categories be linearly separable does not apply to human category learning (e.g. Medin & Schwanenflugel, 1981). Second, the multiplicative prototype model also is not sensitive to correlated attributes within categories and in general, does not fare well in contrast with exemplar models that use multiplicative similarity (see Nosofsky, 1991 for a review).

To address some of these limitations, Gluck, Bower, and Hee (1989) slightly modified the adaptive network model by adding input nodes to the model which correspond to the *conjunctions* of single features. For example, a model with nodes for the features *a*, *b*, *c*, and *d*, would also have nodes for the conjunctions of those features (i.e., a+b, a+c, a+d, b+c, b+d, and c+d). By encoding the input in this way, the model is sensitive to correlated attributes. Furthermore, such an encoding may transform a learning problem that is not linearly separable with single features into one that is linearly separable in terms of combinations of features. Gluck, Corter, and Bower (1990) have shown that this model (which they call the configural-cue network model) reproduces the basic level advantage observed by Murphy and Smith (1982), Hoffman and Ziessler (1983), and Corter et al. (1988).

The context model. In contrast to prototype models, standard similarity-based exemplar models explicitly represent information about individual category members. Exemplar models assume that category items are stored in memory and retrieved in a probabilistic fashion. Also, an item is classified into the category to which it is most similar. For the Medin and Schaffer context model, the classification rule takes the form of Luce's choice rule (Luce, 1959). In this case, the probability of classifying an item as a member of a target category is equal to the sum of the similarities of that item to each category member, divided by the sum of the similarities of the target item to all items (both in the target category and in the contrast categories). Furthermore, the similarity of an item to a category member is a *multiplicative* function of the similarity of the item's attribute values and to those of the category member (this is not specified by Luce):

(10) P(classifying item in category A) =

$$\frac{\Sigma_a \Pi s_d}{[\Sigma_a \Pi s_d + \Sigma_b \Pi s_d]}$$

where a represents the items in category A, b represents the items in a contrast category B, d represents the attributes that make up each item, and s represents the similarity of the item's attribute values to those of a category member. If the target item and the comparison item have a matching value on attribute d, then target item and the comparison item have a matching value on attribute d, then s for that attribute equals 1; otherwise, the similarity parameter takes on the best fitting value between 0 and 1. Matches between the target item and comparison items in memory increase similarity; mismatches decrease similarity.

Exemplar models of this form, which combine information about category size with similarity information, will always derive a higher probability of classification for the category with the most exemplars, as long as the categories are nested, because within category similarity increases (or at least remains constant) with increases in category size. *Average* within-category similarity (dividing the sum of the similarities between the target item and each of the category items by the number of category items) may decrease as one moves up a nested hierarchy but *summed* similarity (simply summing the similarities between the target item and each category item across all category items, without dividing by the number of category items) can only increase.

Some provisos are in order concerning the predictions of exemplar models. First of all, exemplar models have almost exclusively been applied to transfer tests given after learning. They have not been developed as computational models of learning. It may be that some form of exemplar model that incorporates competitive learning could predict a basic level advantage. Our predictions hold only for the class of models in which learning is a monotonic function of the classification probabilities, as indexed by the equation above. One also must keep in mind that the context model permits selective attention to constituent dimensions. Therefore, a second hedge is that the equation above should only hold when one allows the similarity parameters to adjust in the direction of optimizing performance. With nested category structures, selectivity at a lower level will also benefit performance at higher levels of categorization. Therefore, the prediction that higher levels will yield better performance ought to hold even

Table 9.2. Similarity analysis of multiplicative exemplar model using average rather than summer similarity applied to Corter, Gluck and Bower (1988).

	Average Similarity		
Level	Within	Between	
Superordinate	$\frac{1+S_2+2S_1S_2S_3}{4}$	$\frac{S_1S_2S_3+S_1S_3+2S_1S_2}{4}$	
Basic	$\frac{1+S_2}{2}$	$\frac{3S_1S_2S_3+S_1S_3+2S_1S_2}{6}$	
Subordinate	1	$\frac{S_2+3S_1S_2S_3+S_1S_3+2S_1S_2}{7}$	
with $S_1 = 0$:			
Superordinate	$\frac{1+S_2}{4}$	0	
Basic	$\frac{1+S_2}{2}$	0	
Subordinate	1	$\frac{S_2}{7}$	
with $S_1 = S_2 = S_3 = .30$			Ratio
Superordinate	.338	.052	.867
Basic	.650	.058	.918
Subordinate	1.00	.093	.915

S_1= similarity parameter on first dimension (gums)
S_2= similarity parameter on second dimension (eyes)
S_3= similarity parameter on third dimension (rash)

when selective attention is allowed. The sole exception that comes to mind is the possibility that selective attention might operate more effectively at one level than others. For example, errors may drive the adjustment processes faster than correct guesses, and correct guesses should occur more frequently for the superordinate level than for the basic level. Still, it appears that the most natural application of exemplar models fails to predict a basic level advantage. At the very least, this represents a serious challenge to exemplar models that rely on summed similarity.

Alternatively, one could formulate an exemplar model based on average rather than summed similarity (with both the average and the sum taken over all members of a category). As a result, the model would favor the basic level over the superordinate in experiments such as those of Corter et al. (1988). To see this, consider the design in Table 9.1 in terms of average within-category similarity versus average between-category similarity. This is summarized at the top of Table 9.2. The special case in which $S_1=0$ is clear, as shown in the bottom half of Table 9.2. One immediately sees that the basic level will be favored over the superordinate level. Furthermore, as S_2 increases, the basic level will be favored over the subordinate (the exact point of the shift depends on whether one assumes performance is a function of differences versus ratios). Another way to see this is to let $S_1 = S_2 = S_3 = .30$ and note that the ratio of average within-category similarity to average within plus between category similarity is maximized at the basic level. This case is shown in the bottom third of Table 9.2. For at least some values of similarity, the average similarity exemplar model picks out the basic level.

We will return to the multiplicative exemplar model using average or normalized similarities a bit later. For the moment we simply note the following problem. The average similarity exemplar model implies an insensitivity to category size or frequency but the data suggest that people's categorization is, in fact, at least partially sensitive to frequency (e.g., Medin & Florian, 1991; Nosofsky, 1988). The asymptotic performance of the adaptive network model is also insensitive to category frequency. The summed similarity exemplar model handles frequency effects but does not pick out the basic level as easiest. Overall, we are in the uncomfortable situation of having no model which simultaneously gives an account of category frequency and basic levels effects. Rather than dwell on these difficulties, we introduce a few further ones. First, we take a second look at natural categories and then describe some additional results growing out of this analysis.

VI. FURTHER COMPLICATIONS

A. Natural categories reconsidered

Although both natural and artificial categories have yielded results favoring a basic level of categorization, one cannot be sure that the underlying basis for these results is the same. In fact, a closer analysis of the category structures used with artificial stimuli gives cause to fairly strong reservations. Consider again Table 9.1, summarizing the Corter et al. (1988) design. Note that for the middle level of categorization there is always a feature that only and always is associated with it. In short, a defining feature is present at the middle level. But we began this chapter by noting the strong evidence and near uniform consensus that natural concepts do not have defining features (let alone a single defining feature). The same problem arises for the Hoffman and Ziessler (1983)

Table 9.3. Abstract category structure of Murphy and Smith (1982) Experiment 1.

Stimulus	Feature Dimensions: Handle	Shaft	Head-Blade	Size
1	2	2	0	0
2	2	2	0	1
3	2	2	1	0
4	2	2	1	1
5	0	3	4	0
6	0	3	4	1
7	1	3	4	0
8	1	3	4	1
9	3	4	2	0
10	3	4	2	1
11	3	4	3	0
12	3	4	3	1
13	4	0	5	0
14	4	0	5	1
15	4	1	5	0
16	4	1	5	1

studies as their design also maps onto that of Table 9.1. It is not entirely clear how the abstract structure of the Murphy and Smith (1982) study should be represented but one plausible scheme, taken from Corter et al. and shown in Table 9.3, reveals that numerous defining features were available at the basic level and not at the subordinate or superordinate level. It also appears that there were defining features in Mervis and Crisafi's stimuli (see Figure 9.3).

The basic level advantage for artificial categories may be related to defining features. However, we have noted that natural categories are not characterized by defining features. As a result, the explanation for the basic level advantage in natural categories may have little to do with the explanation for this advantage in artificial categories. It is possible that a measure such as category utility picks out the basic level for both artificial and natural categories but the presence of defining features undermines the parallels. It is difficult to assess the seriousness of this defining feature issue. We turn attention now to some recent studies of levels effects in which defining features were not present. Having adopted a skeptical frame of mind, however, we should note that these studies are themselves far from immune to criticism.

B. Fuzzy artificial categories and the basic level

Recently, the first author studied levels effects using artificial categories without defining features (Lassaline, 1990). The studies were also designed to evaluate concept learning models with respect to the basic level. We decided that there was no absolute need to have three levels of categorization, as the theories and metrics we have described require only two levels for a contrast. Therefore the artificial categories were organized into a two-level hierarchy with two General categories and four Specific categories, two nested in each General category. In addition, training level was varied between rather than within subjects. One reason for this choice is that it has proven surprisingly difficult for subjects to learn multiple levels of categorization simultaneously. Previous studies with artificial categories have used the procedure of training people on one level at a time and only later integrating across levels.

The first experiment was an exploratory one, designed to show that a levels effect can be obtained for category structures that do not contain defining features. The design of the first study is summarized in Table 9.4. The stimuli were sixteen objects, each consisting of a value on four different dimensions, D1, D2, D3, and D4 (two shape dimensions and two texture dimensions), modeled after the tool-like stimuli of Murphy and Smith. The General level consisted of two categories (Categories A and B). There were four categories for the Specific level (Categories C, D, E, and F). Unlike previous work, there was no single feature or combination of features at any level that was necessary and sufficient to determine category membership.

Subjects were trained to verify category membership upon presentation of a picture of one of the objects and a category name. After training on the categories, subjects were tested with a speeded version of the category verification task. Average response latencies and error rates across subjects, by level, were used to evaluate predictions for ease of category learning and verification made by three theories: category utility, the adaptive network model, and the exemplar similarity model.

Category utility, the adaptive network model, and the summed similarity exemplar model predict an advantage for the General level of categorization. The general level partition has a category utility value of .266 whereas the specific level has a value of .211. The predictions of the average similarity exemplar model depend on values of the similarity parameters--large values favor the Specific level whereas smaller ones favor the General level. The predictions of the adaptive network model were derived through simulations assuming a configural cue

Table 9.4. Lassaline (1990) Experiment 1 Stimulus Structure.

Example	D1	D2	D3	D4	Category General	Category Specific
1	1	1	1	0	A	C
2	1	0	1	0	A	C
3	1	0	0	0	A	C
4	0	1	1	0	A	C
5	1	1	0	1	A	D
6	0	1	0	0	A	D
7	0	1	0	1	A	D
8	1	0	0	1	A	D
9	2	2	2	0	B	E
10	2	0	2	0	B	E
11	2	0	0	0	B	E
12	0	2	2	0	B	E
13	2	2	0	2	B	F
14	0	2	0	0	B	F
15	0	2	0	2	B	F
16	2	0	0	2	B	F

representation and setting any negative output values to zero before computing the response probabilities. Across variations in the simulations, these predictions did not vary. Specifically, whether or not negative outputs were transformed to zero, whether or not the stimulus representation was supplemented with configural cues, and across various learning rates, the adaptive network model predicts a General level advantage.

The results of this first experiment were consistent with the General level advantage predicted by category utility, the adaptive network model, and the summed and average similarity exemplar models. As Figure 9.4 indicates, the General level proved easier to learn than the Specific level. Although this experiment does not distinguish between models (since all of them predict a general level advantage), it did demonstrate that a levels effect can come about with fuzzy artificial categories.

The second experiment was somewhat more diagnostic in that the various models did not all make the same prediction. The abstract structure of the experiment is shown in Table 9.5. Each of the twelve stimuli was constructed from three dimensions, D1, D2, and D3 (two shape dimensions and one texture dimension), each of which had three possible values. Again, no defining feature is present. In addition, no single feature is even sufficient to determine category membership at either level, although the Specific categories do have a feature with perfect category validity. That is, all members of Category C have a value

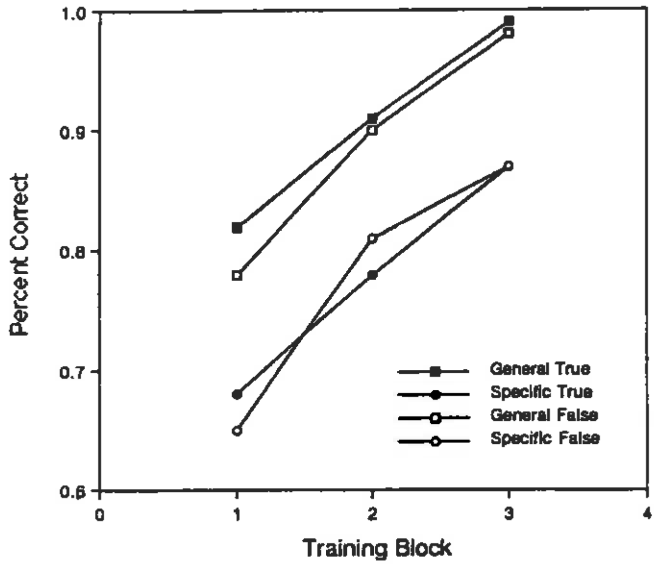

Figure 9.4. Training accuracy for General and Specific categories (Lassaline, 1990, Experiment 1).

of 1 on D2; all D's have a value of 1 on D3; all E's have a value of 2 on D2; and all F's have a value of 2 on D3.

The category utility value is .118 for the General categories and .267 for the Specific categories. Thus, category utility predicts a Specific level advantage. The average similarity exemplar model also predicts a Specific advantage. The summed similarity exemplar model predicts, of course, an advantage for the General level, as within-category similarity increases with category size. In our simulations, the adaptive network model also predicts an advantage for the General level.

The Specific level proved to be much easier to learn than the General level. These learning differences carried over to speeded categorization trials given after training, as shown in Figure 9.5. In brief, the data support category utility and the average similarity exemplar model over the summed similarity exemplar model and the adaptive network model.

The most challenging data come from a third experiment whose design is summarized in Table 9.6. Structure (One Dimension or Four Dimensions) was varied between subjects, such that each subject learned the twelve items in either

Table 9.5. Lassaline (1990) Experiment 2 Stimulus Structure.

Example	D1	D2	D3	Category General	Category Specific
1	1	1	2	A	C
2	1	1	0	A	C
3	2	1	0	A	C
4	1	2	1	A	D
5	1	0	1	A	D
6	0	0	1	A	D
7	2	2	1	B	E
8	2	2	0	B	E
9	1	2	0	B	E
10	2	1	2	B	F
11	2	0	2	B	F
12	0	0	2	B	F

the One Dimension or the Four Dimensions structure. Each of the sixteen items was constructed from four dimensions, D1, D2, D3, and D4 (two shape and two texture dimensions) each of which had four possible values. Note that we have introduced a defining feature for the Specific level of categorization, and thus a disjunction of two features is defining at the General level. Of primary interest was the effect of the distribution of the defining feature across dimensions. For the One Dimension case, the defining feature for each of the Specific categories consisted of a value on the same dimension (values 1, 2, 3, and 4 on D1 for Categories C, D, E, and F, respectively). For the Four Dimensions case, the defining feature was taken from a different dimension for each of the Specific categories (a value of 1 on D1, D2, D3, or D4 for Categories C, D, E, or F, respectively).

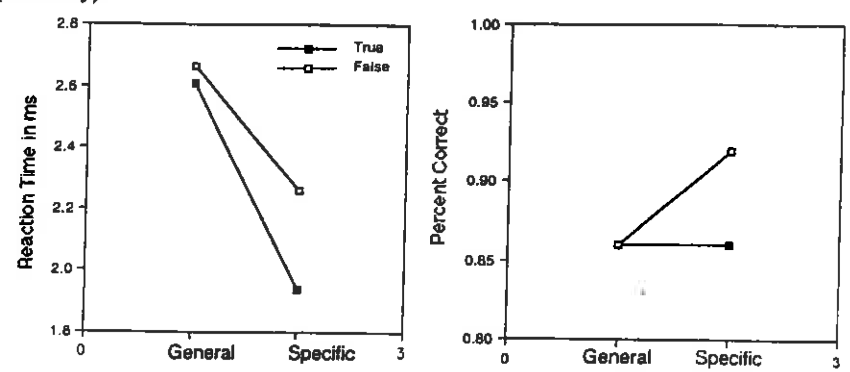

Figure 9.5. Category verification times and classification accuracy for General and Specific categories (Lassaline, 1990, Experiment 2).

Table 9.6. Lassaline (1990) Experiment 3 Stimulus Structure.

	Structure								Category	
	One Dimension				Four Dimensions					
Example	D1	D2	D3	D4	D1	D2	D3	D4	General	Specific
1	1	1	3	4	1	2	4	3	A	C
2	1	2	4	1	1	3	2	4	A	C
3	1	3	1	2	1	4	3	2	A	C
4	2	4	2	1	2	1	3	2	A	D
5	2	1	3	2	3	1	4	2	A	D
6	2	2	4	3	4	1	2	3	A	D
7	3	3	1	3	4	3	1	2	B	E
8	3	4	2	4	2	4	1	3	B	E
9	3	1	3	1	3	2	1	4	B	E
10	4	2	4	2	3	4	2	1	B	F
11	4	3	2	3	4	2	3	1	B	F
12	4	4	1	4	2	3	4	1	B	F

Neither the category utility measure nor the adaptive network model is sensitive to how the defining features are distributed, and thus predict no difference in category learning or verification between the One-dimension and the Four-dimension cases. Versions of the summed similarity exemplar model that allow selective attention can predict that the One-dimension case will be easier than the Four-dimension case. Given that the average similarity exemplar model also allows for selective attention, it can also predict that the One-dimension case will be easier than the Four-dimension case.

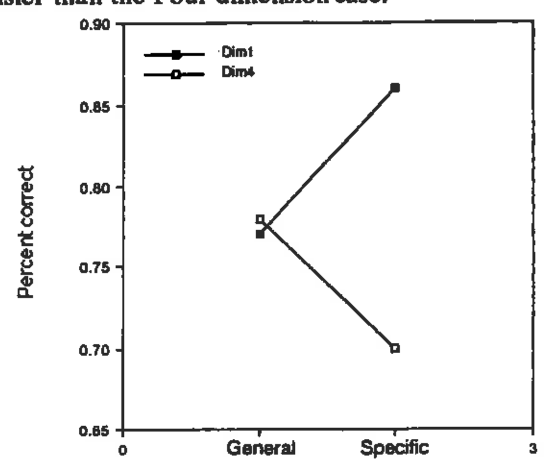

Figure 9.6. Training accuracy for General and Specific categories (Lassaline, 1990, Experiment 3).

With respect to the contrast between levels, the category utility is .167 for the General level and .250 for the Specific level, and thus category utility predicts a Specific advantage. Both the summed similarity exemplar model and the adaptive network model predict that the General level will be easier for both the One-dimension and the Four-dimensions cases. The average similarity model, like category utility, predicts that the Specific level will be easier than the General level in either case.

Somewhat surprisingly, levels interacted strongly with how the defining feature was distributed. As Figure 9.6 indicates, for the One-dimension case, the Specific level was much easier than the General level both in learning and at test, but for the Four-dimension case the General level was easier than the Specific level.

C. Summary

The first two experiments show that one can obtain clear levels effects for category structures that do not contain defining features. In these experiments, the favored level was picked out by the category utility measure. The third experiment, however, yielded an interaction of level with feature distribution that is outside the scope of category utility. Table 9.7 presents a summary of the predicted and obtained results of these experiments.

Table 9.7. Summary of predicted and obtained level advantage from Experiments 1, 2, and 3.

	Predicted Advantage				Obtained Advantage
	Category Utility	Context Model	Adaptive Network	Rule Account	
Experiment 1					
Level	General	General	General	General	General
Experiment 2					
Level	Specific	General	General	?	Specific
Experiment 3					
Level(1D)	Specific	General	General	Specific	Specific
Level(4D)	Specific	General	General	?	General
Structure	No Diff	One Dim[a]	No Diff	One Dim	One Dim

[a] One Dimension structure.

To our knowledge, no current model can account for the full pattern of Lassaline's results. Although the multiplicative average similarity exemplar model can predict these results, these predictions are parameter specific. We have been unable to find a principled reason for varying the parameters in the manner necessary to accommodate this set of results.

D. Rule-Based Accounts

Going beyond the formal models and metrics we have presented thus far, we will next briefly discuss a class of models that use rules to define category membership. Many of the participants in these studies report using rules. It turns out that a rule-based account may well capture many of Lassaline's results, but it is difficult to evaluate such an account quantitatively, as there are no well-specified rule-based models.

Turning to a rule-based account, the ordering of levels with respect to difficulty of learning and verifying category membership was, for the most part, consistent with subjects' use of rules across the three experiments. Subjects' use of rules to learn categories was tested for each experiment by deriving the simplest logical rules that captured the category structure used. The general idea is that categories described by simple rules should be easier to learn than categories requiring complex rules. Of course, it is possible that there might be some simple rule that is difficult to discover. In Lassaline's experiments, however, the dimensions and values are all fairly salient so it is plausible that simple rules were easier to discover than complex rules.

The rule account can also be used to predict performance on individual items. For example, if a subject learning the Specific categories from Experiment 1 uses the rule "a value of 1 or 2 on D3 leads an item to be placed in Category C or Category E, respectively; a value of 1 or 2 on D4 leads on item to be placed in Category D or F, respectively". If subjects learning the Specific categories used this rule, then performance involving the one member of each Specific category that doesn't possess the feature specified in the rule (Items 3, 6, 11, and 14 from Categories C, D, E, and F, respectively) should be poorer than for the other items. The results are consistent with use of this rule. The only inconsistency between the data and the simplest-rule account came from Experiment 2 (see Table 9.5 for category structure). The simplest-rule account predicted an advantage for certain items in this experiment that was not obtained. The derivation of predictions for Experiment 2 for a rule-based account was not as straightforward as that for the first and third experiments, as explicit rules to

capture category membership were not as obvious. For the General categories, a value of 1 on the first dimension for Category A and a value of 2 on the first dimension for Category B characterizes four of the six members of each category. It is also possible to capture General category membership by using a conjunction of the necessary features that characterize the Specific categories, i.e. for Category A, conjoining the necessary feature for Category C, a 1 on the second dimension, and the necessary feature for Category D, a 1 on the third dimension. For each Specific category, in contrast, its necessary feature is not sufficient as one nonmember possesses it. There is one exception to using a rule based on the necessary feature for each Specific category (i.e. Item 10 for Category C, Item 7 for Category D, Item 4 for Category E and Item 1 for Category F); whereas for each General category, there are two exceptions to using the rule based on a conjunction of the second and third dimensions (i.e. Items 7 and 10 for Category A and Items 1 and 4 for Category B), and three exceptions to using the rule based on the first dimension (i.e. include Items 3 and 6 but exclude Item 9 for Category A, and include Items 9 and 12 but exclude Item 3 for Category B).

To determine whether or not subjects used any of the rules described above, data from Experiment 2 were analyzed by item. Verifying category membership for Item 6 for the General level and Item 4 for the Specific level was significantly more difficult than for all other items. There were no other item differences. While these two differences are both consistent with use of the rules described above, the results of the item analysis as a whole do not strongly suggest that subjects used rules as a strategy to learn categories and verify membership in Experiment 2. This inconsistency is not terribly damaging to the rule-based account, though, as it is possible that subjects were simply employing different rules than those proposed by the experimenter. Until a model based on use of logical rules is formally specified and supplemented with a mechanism for exemplar retrieval, it is impossible to quantitatively evaluate the rule-based account. Yet the notion that subjects were employing logical rules in classification may plausibly characterize performance across the three experiments.

Martin and Caramazza (1980) suggest that previous research with artificial categories which has been used to support similarity-based models may actually reflect use of a logical rule system for classification. These authors claim that even using categories with no set of defining features, subjects attempt to learn categories by a sequence of feature tests evaluating logical rules for category membership. This claim is supported by a series of experiments using artificial categories in which the pattern of reaction times for category verification and typicality judgments was consistent with performance based on sequential feature testing. Martin and Caramazza emphasized the importance of analyzing individual subjects' data, as subjects may be using the same process to perform the categorization task (in this case the process would be developing rules to capture category membership), but may differ in the specific rules developed.

Mixed support for a rule-based account comes from experiments by Nosofsky, Clark and Shin (1989) testing the extent to which exemplar-based similarity models can account for classification of rule-described categories. The rule account formulated by these authors assumes that subjects adopt the simplest possible rule for partitioning categories (Shepard, Hovland & Jenkins, 1961). In two of the experiments by Nosofsky et al. (1989), when subjects were explicitly instructed to use specific rules, classification performance was better predicted by the minimal-rule account than by the exemplar model, but could be predicted by the minimal-rule account. In contrast, when subjects were not explicitly instructed to use rules, the exemplar model provided a better account of the findings.

Experiments involving category construction do provide support for a rule-based account (Ahn and Medin, 1989). These authors proposed a two-stage categorization model that accurately predicted when family resemblance sorting will occur. In the first stage, initial categories are created based on values on a single primary dimension, a defining feature for the initial categories. Remaining exemplars not classified by values on the defining feature are classified in the second stage according to overall similarity to members of the initial categories. This model integrates use of an explicit unidimensional rule with exemplar similarity computations to produce categories. Models based on category utility or similarity alone were unable to predict when family resemblance categories would and would not be constructed.

In formulating rule-based models, psychologists have tended to assume that simplicity governs the choice of rule derived to capture category membership. Medin, Wattenmaker and Michalski (1987) argue instead that human rule induction is constrained by a category validity bias and a bias toward rules that are more specific than necessary, perhaps to maximize inferences and protect the rule system from over-generalizing. By exploring such constraints, one might develop more precise rule-based models. In addition, some mechanism governing exemplar storage and retrieval is needed to supplement the rule system. As things stand now, we cannot conclude a great deal about the viability of a rule-based account of hierarchical categorization, as current rule models are imprecise.

E. Identification Versus Classification

A limiting case of levels effects is identification learning, or learning in which each example must be placed into a separate category. The classic studies of Shepard, Hovland and Jenkins (1961) looked at the relative difficulty of different types of category partitions. The stimuli were the eight possible combinations of three binary valued dimensions (111, 110, 101, 011, 001, 010, 100, and 000). For two equal-sized categories, there were six distinct types of category partitions (one corresponding to a single defining feature, one to correlated attributes within a category, one to a linearly separable category with characteristic features, and so on). Shepard et al. found that classification learning was easier than identification learning except for the worst category structure, where no difference was found. Nosofsky (1984) demonstrated that the multiplicative summed similarity exemplar model could account for the ordering of learning difficulty, if selective attention to or weighting of dimensions is allowed. Surprisingly, however, the category utility measure is completely unable to account for the Shepard et al. results. Category utility predicts that identification learning should be easier than classification learning for four of the category types and that there should be no difference for a fifth type. Only for the defining features case does category utility predict classification to be superior to identification.

F. Summary

The initial rosy picture of parallels between artificial and natural basic level categories has turned somewhat problematic. Studies with artificial categories have often used defining features. This undermines the idea that relevant structural properties of natural categories have been incorporated into artificial categories but leaves the hope that some metric or theory will correctly predict levels effects in both domains. To our chagrin, however, none of the metrics and theories hold up well, even within the domain of artificially constructed categories. Lassaline's finding of an interaction of level with feature distribution was not predicted by any metric or theory. Before attempting to draw out the implications of this series of challenges, we take one further look at linkages between artificial and natural categorization.

VII. ARTIFICIAL VERSUS NATURAL CATEGORIES: FURTHER OBSERVATIONS

There are a number of recently observed aspects of natural categorization that typically have not been examined in artificial category studies. We see these characteristics as challenges to any complete theory of categorization. Furthermore, we are fairly sure that these aspects of natural categorization play an important role in the basic level advantage. Below, we describe these aspects and speculate on how they might be relevant to basic categories.

A. Conceptual Functions

Certainly, our concepts have uses or functions. In almost all research on concepts, including that involving the basic level, two conceptual functions have been emphasized. First, we use concepts to *classify* or identify items. Classification is a prerequisite for many cognitive and behavioral tasks. For example, building a bookcase might require nails, a hammer, a saw, and wood. It seems necessary that one be able to identify such objects before carrying out such a task. Second, we use concepts to *predict* information about an item. For example, if one has classified an object as a bird, one might reasonably conclude that it will fly. The importance of prediction is fairly obvious. If we can predict things about an object or event, we can anticipate potentially harmful or beneficial consequences, and take appropriate steps to avoid or insure them.

As we have noted, models typically have accounted for the basic level advantage by suggesting that this level maximizes some function of the cue and category validity of features. Cue validity emphasizes the classification function of concepts. It measures how distinctive a feature is with respect to other categories. A feature with maximum cue validity allows one to classify with certainly an item possessing that feature. Such a feature would be unique to members of the category and therefore very distinctive. Category validity emphasizes the prediction function of concepts. It measures how common a feature is with respect to other categories. Given that an item belongs to a category, one can predict a feature of a category member with certainty, if the feature has maximum category validity (i.e., is common among all members of the category).

Recently, some researchers have emphasized that people use concepts for a *variety* of purposes (Matheus, Rendell, Medin, & Goldstone, 1989; Medin, 1983). Concepts have other functions besides classification and prediction. Below, we briefly describe other possible functions of concepts.

Learning. We use concepts to learn other concepts. There are probably a variety of ways that concepts facilitate learning. For example, one might learn about an unfamiliar concept (e.g., *electricity*) using reasoning by analogy to a

familiar concept (e.g., *water flow*). Gentner and Gentner (1983) discuss this type of learning. Existing concepts might also provide expectations or biases that guide learning of new concepts. For example, when you first learned about computers you previously may have been told that they could be used for text processing and mathematical calculations. Learning about computers may have been a process in which you related these abstract expectations to more concrete features of the computer (particular keypress sequences, displays on the computer screen, etc). Researchers in the field of machine learning have called this type of learning *operationalization* (Dejong & Mooney, 1986; Mostow, 1983). In general, it involves translating abstract, high level expectations into more specific, low level information.

Explanation. We use concepts to explain and understand why things happen. For example, one might explain why a bottle of Coca-Cola shattered after being left in the freezer for several hours by noting that water expands when it freezes, Coca-Cola is composed of almost entirely of water, and so on. The role of explanation is related to prediction: if we can explain an event (for example), we generally know the conditions under which it will occur. Therefore, if we can identify these conditions in a new situation, we can predict the event.

Conceptual combination. We also combine existing concepts to create new ones (see Murphy, 1988; Wisniewski & Gentner, 1991 for reviews). For example, one might combine the concepts *elephant* and *box* to form the new concept *elephant box*, which might mean "a special type of box for transporting elephants." Conceptual combination is another way that we learn concepts. It is also an efficient way of adding new terms to a language (Downing, 1978).

Communication. Of course, we also use concepts to communicate with other people. One role of concepts in communication is to efficiently convey knowledge. There is a wealth of information associated with a concept (Barsalou, 1989). Furthermore, people believe that their concepts are similar to other people's (Rey, 1983; Miller & Johnson-Laird, 1976). Therefore, in uttering a phrase, a speaker may mean more than what is explicitly stated and believe that the listener will have access to the knowledge implicit in the utterance. For example, a speaker uttering the sentence, "She began eating the soup and it burned her mouth" assumes that the listener has a similar concept of *soup* and does not have to explicitly state why the soup burned the woman's mouth.

Theories of the basic level have assumed that this level best captures the functions of classification and prediction. The design of category structures involving artificial stimuli primarily has varied factors associated with these functions (e.g., commonality and distinctiveness of features). Little work has attempted to relate the basic level to other conceptual functions. Furthermore, one intriguing possibility is that concepts at different levels in a taxonomy may differ in their conceptual functions. One possible reason for the basic level advantage in so many cognitive tasks is that these tasks correspond to the functions ideally suited for the basic level, and not other levels. Further research needs to examine the functions of other concepts like superordinates and subordinates.

Some work has addressed the idea that superordinate and basic concepts have different functions. Based on a linguistic analysis of basic and superordinate terms, Wisniewski and Murphy (1989) argued that people primarily use superordinate terms to conceptualize related groups of objects and use basic concepts to conceptualize single objects. To test this hypothesis, Murphy and Wisniewski (1989) had people categorize objects that were isolated or in scenes. The scenes were naturalistic depictions of actual places that contained members of a superordinate category. Consistent with many studies of the basic level, when the object was isolated, people were faster at categorizing an object at the basic level than at the superordinate level. However, when the object was in a scene, superordinate and basic categorization took about the same amount of time. Furthermore, superordinate categorization was more disrupted than basic categorization when the object was in an improbable scene. These studies suggest that superordinate and basic concepts have different functions.

B. Selective Induction Versus Constructive Induction

Artificial category studies of the basic level have the virtue that, in general, the category structure created by the experimenter closely approximates the structure perceived by the subject. The features typically used are well-specified and unambiguous. As a result, the experimenter and subject probably agree on the features that are present in the stimuli. (This characteristic also is true of other artificial category studies that do not explicitly address the basic level.)

However, many researchers have suggested that the crucial problem in induction is determining the units of analysis or constituents upon which to perform induction (e.g., Medin, Wattenmaker, & Michalski, 1987). That is, induction may involve selecting features from a feature space (i.e., selective induction) but more importantly, it involves constructing that feature space (i.e., constructive induction). In artificial category studies of the basic level, the

constructive induction problem has been "solved" for the learner by the experimenter, who typically uses stimuli with well-specified, unambiguous constituents. This is as true for our studies as for any others.

Some researchers have argued that constructive induction may result from an interaction of people's prior expectations or background knowledge and information delivered by the perceptual system (e.g., Wisniewski & Medin, 1991). One criticism of artificial category studies is that they create situations in which a person's use of background knowledge is unnecessary for category learning. For example, we are sure that people do not need to use prior knowledge to learn the categories shown in Table 9.1.

Some evidence for this view of constructive induction comes from studies involving children's drawings. Wisniewski and Medin (in press) had two groups of subjects learn about the same categories, but the groups were given different, meaningful labels for the categories. Specifically, one group of subjects was told that the drawings were done by creative and noncreative children. The other group was told that the same drawings were done by farm and city children. The purpose of giving the groups different labels for the categories was to activate different, prior knowledge in the two groups during learning.

This simple experimental manipulation had important effects on the *kinds* of features that people perceived in the drawings. Furthermore, it appeared that people's prior knowledge helped determine these features. A striking finding was that different people sometimes perceived the *same* part of a drawing as a *different* feature. For example, consider the two drawings shown in Figure 9.7.

Many subjects in the Creative/Noncreative group believed that creative children were likely to draw detailed pictures. In contrast, many people in the Farm/City group believed that city children would draw people that they were likely to encounter in the city. Accordingly, one person in the Creative/Noncreative group interpreted the circled part of drawing 1 as *buttons* (evidence for detail), whereas a person in the Farm/City group interpreted the same part as a tie (evidence for a business person). In drawing 2, a person in the Creative/Noncreative group interpreted the circled part as a *pocket* (again, evidence for detail), whereas a person in the Farm/City group interpreted the same part of the drawing as a *purse* (evidence of a woman from the city).

Of course, such features were not constructed solely from people's prior expectations. For example, consider again the circled line configuration in drawing

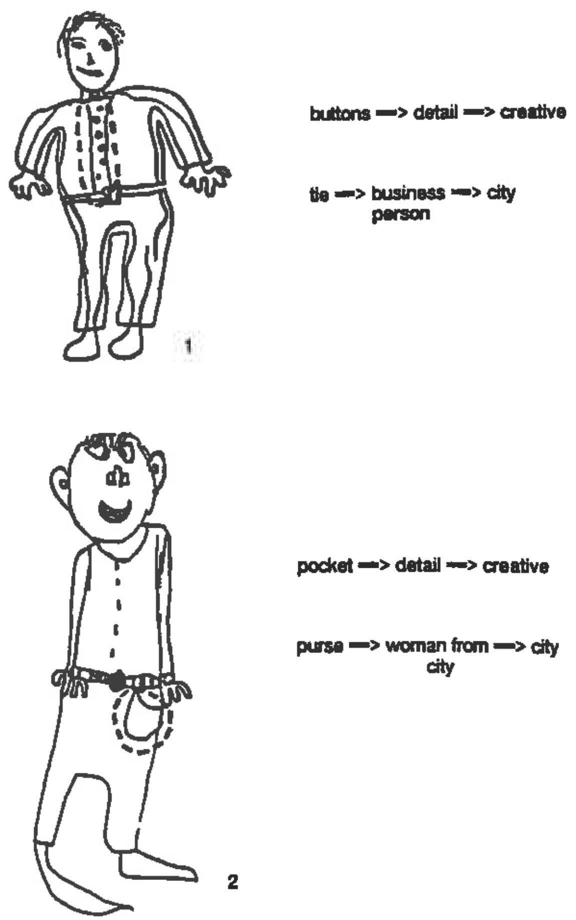

Figure 9.7. Constructive induction processes in classifying children's drawings (Wisniewski and Medin, in press).

2 of Figure 9.7. While a person interpreted the configuration as a *pocket*, therefore providing evidence of detail, the person did not interpret the configuration as *shoelaces, eyebrows, buttons*, etc., even though such features were mentioned by other subjects as evidence of detail. Most likely, the perceptual system provides important information that constrains which features are plausible, such as low-level shape descriptions and locations (e.g., Biederman, 1985; Marr & Nishihara, 1978).

In more natural learning situations, it may be more accurate to view features as *hypothetical entities* that people believe exist among the category members with varying degrees of certainty. As they experience more members and environmental feedback, they adjust these certainties. Sometimes the certainty of an hypothesized feature becomes very low and the hypothesis is abandoned. A new hypothetical feature may take its place. In fact, Wisniewski and Medin (1991) found that when people incorrectly categorized a drawing, they sometimes abandoned those feature interpretations that supported the incorrect category and reinterpreted features of the drawing in a way that was consistent with the correct category.

We should at least consider the possibility that determining the features of natural categories is a very important aspect of learning about them. Studies of the basic level using artificial stimuli have not addressed this aspect of category learning. There are at least two potential problems with ignoring constructive induction. First, constructive induction may be related in some way to the basic level advantage. Second, because of constructive induction processes, a person's perceived category structure could be different from the category structure assumed by the experimenter.

C. Features at Different Levels

A number of researchers have suggested that concepts at different levels in a taxonomy have different features, or similar features that differ in importance. (This suggestion also is consistent with the idea that conceptual function may vary with taxonomic level.) Rosch et. al (1976) suggested that the attributes of superordinate and basic concepts are qualitatively different. In particular, basic concepts contain many perceptually salient features whereas superordinate concepts contain more abstract or functional features (e.g., "plays music" for *musical instrument*). As previously discussed, Tversky and Hemenway (1984) argued that, compared to other concepts, basic concepts tend to represent the parts of objects, and that this difference accounts for the basic level advantage. Murphy and Wisniewski (1989) suggested that superordinate concepts not only contain abstract or functional features, but information about the scenes in which their members are likely to be found. For example, *furniture* might include a representation of a living room scene, linking the concepts *couch, lamp, coffee table, and chair* together, and a representation of a bedroom scene, linking the concepts *bed, dresser, and mirror* together. Furthermore, these basic concepts could be connected by relations. For example, *lamp and table* could be connected by the relation "sits on" and *chair and table* by the relation "oriented towards." On this account, superordinates are partly defined through their associated basic concepts and the relations between them.

Importantly, most artificial concept studies of the basic level have not varied the kinds of features that describe different levels. One notable exception is the Murphy and Smith (1982) studies. Basic level categories were described by parts and superordinates were described by their functions. However, as previously mentioned, basic categories in those studies could be defined by necessary and sufficient features.

D. Dimensions of Features

Artificial concept studies of the basic level typically have characterized features along the dimensions of commonality and distinctiveness. On the one hand, it seems clear that a model of concept learning should be sensitive to these dimensions. As we have suggested, they play an important role in the classification and prediction functions of concepts.

On the other hand, it seems unlikely that commonality and distinctiveness are the *only* important dimensions that characterize the features represented in concepts. At least in the case of artifacts, features can be characterized in terms of how relevant they are to the artifact's function. To take a simple example, consider the features "uses water" and "has a door that opens downward," associated with the concept *dishwasher*. It seems that "uses water" is much more relevant to the function of a dishwasher than "has a door that opens downward." Intuitively, this feature seems more important to our concept of dishwasher. This is the case even though "has a door that opens downward" is just as common among dishwashers as "uses water," and furthermore, is more distinctive of dishwashers than "uses water." Specifically, artifact concepts might include information about *which* features are related to the function and *how* they are related to the function. Among other things, such information would be crucial in understanding how to carry out the function associated with the artifact and in understanding how to repair an artifact that is not functioning properly.

Recent research supports the idea that the relevance of a feature to an artifact's function affects categorization. A study by Rips (1989) suggests that functional relevance is more important for categorization than other information, such as appearance. He gave subjects descriptions of artifacts that underwent two types of transformations. In one type, the artifact was changed so that its appearance resembled members of another category, while preserving its intended

use. In another type, the artifact was changed so that its intended use resembled those of the members of another category, while preserving its appearance. In both cases, subjects tended to classify the artifact into the category whose intended use most resembled that of the artifact. More recently, in a study involving artificial artifact concepts, Wisniewski and Medin (1992) examined the classification of items composed of features that were relevant to the function of one category and features that were irrelevant to the function of another category, but more diagnostic of membership in that category. They found that subjects tended to classify the items into the category containing the functionally relevant features, even though those features were less diagnostic.

Besides functional relevance, features can be viewed at multiple levels of abstraction. For example, consider a drawing of a person that belongs in the category "drawn by a creative child." A low level feature such as a particular *configuration of lines* might be viewed as a *pocket*, which might be an example of *detail*, which in turn, might be an example of *creativity*. One consequence of this process is that features can be treated equivalently. So, the features *shoelaces* and *pocket* are different, but when viewed as examples of *detail*, they are similar.

As evidence for this claim, Wisniewski and Medin (1992) showed that when different features could be conceptualized as examples of the same, higher-level feature, people rated them as more similar than in a case in which they could not be conceptualized in that manner. For example, people were given two descriptions of objects, one containing the feature "uses a poisonous substance" and the other containing the feature "emits microwaves." In a neutral context, both of these features are very different. When told that both objects were "used for killing bugs," people rated the two features as more similar than when told that both objects were "used for transporting people underwater." Presumably, people inferred that "emits microwaves" and "uses a poisonous substance" were both examples of "methods for killing bugs." They were unable to conceptualize these features as examples of the same higher-level feature associated with "used for transporting people underwater" and thus rated them as dissimilar.

This dimension of features could have important effects on how concepts are learned. To take one simple example, consider complex stimuli such as drawings done by emotionally disturbed children. One indicator of emotional problems is the omission of body parts, such as hands, feet, and so on (Koppitz, 1964). Consider two conceptual clustering models that are shown drawings done by emotionally disturbed children. One model groups drawings based on category utility and uses selective induction. The second model views features at multiple levels of abstraction. (The second model also has prior knowledge about emotionally disturbed children.) Both models will cluster the drawings in different ways. The first model will group drawings into categories with common, overlapping features and maximize category utility. On the other hand, the second model will group drawings on the basis of whether or not they have missing body parts, even though such features may not overlap. For example, the feature "missing hands" in one drawing does not overlap with "missing feet" in another drawing. However, because the second model views features at multiple levels of abstraction, it will consider the features "missing hands" and "missing feet" equivalent and group these drawings together.

To summarize, it is clear that features have other important dimensions besides commonality and distinctiveness. Research on the basic level using artificial categories, ours included, has almost exclusively focused on these two dimensions. Exactly how other dimensions are related to the basic level is unclear. Nevertheless, we can speculate on some of the possible relations between the basic level and the dimensions described above. For example, with artifact categories, it may not be that the basic level is the one with the greatest number of common and distinctive attributes, *per se*. Rather, it may be that it is the level with the greatest number of common and distinctive attributes relevant to the category's function. Some evidence for this claim comes from Tversky and Hemenway's (1984) studies. They showed that categories at the basic level had the greatest number of part features, which typically are relevant to a category's function. Another possibility is that the commonality and distinctiveness of features depend *crucially* on the level of abstraction at which people view features. In the example above, it might be more accurate to view the features two members of the category "drawings of emotionally disturbed children" as having a common feature ("missing body parts") even though at a lower level, they possess different features (e. g., "missing hands" and "missing feet"). On the other hand, it might be more accurate to view the feature "round" as somewhat distinct in the categories basketball and cantalope, since at a higher level of abstraction they are different. One can dribble and shoot a basketball and "round" is important in these activities. It would be very difficult to do either activity with a cantalope.

VIII. SUMMARY AND CONCLUSIONS

We now return to two issues with which we began this chapter. First, we have been concerned with the extent to which parallels can be drawn between work using natural versus artificial categories in studying the basic level. It is clear from our analysis that most studies using artificial categories have not carefully

incorporated structural properties of natural categories. In particular, experimenters have tended to employ artificial categories that possess defining features, which is inconsistent with what is typically assumed about natural categories. In addition, artificial categories have, for the most part, ignored differences in the types of features and in conceptual function across levels of categorical structure.

Do these differences truly matter? We do not know. A potentially fruitful strategy is to be more rigorous in bringing what's known about natural categories into the laboratory with artificial categories. For the moment, anyway, we believe that caution should be exercised in drawing conclusions about hierarchical categorization in general from studies involving artificial categories. What then, can be gained from the work done thus far using artificial categories?

The primary contribution made by studies of the basic level using artificial categories has been to make possible the evaluation of various classification theories with respect to their account of hierarchical categorization. This is the second concern that this chapter was intended to address. The metrics and theories considered here are constrained by simplifying assumptions about the representation of items to be classified, and are thus difficult to evaluate using natural categories, where the representation is less clear than with artificial ones.

Although category utility does an adequate job of predicting difficulty ordering in the set of basic level studies using artificial categories discussed here, it assumes that features are independent, and is thus insensitive to any nonindependence among features. This is clearly inconsistent with what is known about human categorization. In addition, category utility fails to predict difficulty ordering between identification and classification learning in all but the simplest case, the one in which categories possess defining features.

Exemplar models, in contrast, do well on capturing categorization within horizontal structure, including cases where categories possess correlated attributes. Exemplar models that compute similarity as a sum across items, like cue validity, predict that the most inclusive level of classification will be the easiest to learn. As exemplar models are extended to learning we may get a better picture of how and whether this generalization needs to be qualified. Average similarity exemplar models, in contrast, are not sensitive to category size, but the accuracy of their predictions depends on the particular best-fitting value of the similarity parameter. In addition, although an insensitivity to category size frees the average similarity exemplar model from the prediction that the most inclusive level is basic, it also makes it difficult for such a model to predict category size effects.

The adaptive network model, like the exemplar models, is sensitive to violations of feature independence. At least for Lassaline's studies, though, it did not correctly predict difficulty ordering between levels, but instead predicted a General advantage across experiments.

Finally, accounts based on use of logical rules to capture category membership remain too underspecified to be carefully evaluated. In addition, while the notion that humans learn categories by deriving the simplest logical rule by which to classify members has intuitive appeal and some empirical support with artificial stimuli, the derivation of rules governing natural category membership is not as straightforward.

In summary, the classification metrics and theories we have considered here do not fully account for the basic level advantage found with studies using artificially constructed categories. The prospects for an account of hierarchical categorization in natural categories are perhaps even more dim, as it is unclear how to apply these theories to natural categories. Not wishing to be harbingers of grim news, nor end this chapter on a completely pessimistic tone, we nonetheless find ourselves forced to conclude that our work in understanding the basic level is not done, and further, that without understanding a phenomenon so prominent in categorization as the basic level, we have a long way to go.

ACKNOWLEDGEMENTS

We gratefully acknowledge Woo-kyoung Ahn, Barbara Burns, Jim Corter, Judy Florian, Mark Gluck, Rob Goldstone, Evan Heit, Matt Kurbat, Joshua Rubinstein, David Thau, and Phil Webster for their helpful comments and suggestions on this chapter. We also thank Bob Dylan for inspiration.

REFERENCES

Adelson, B. (1983). Constructs and phenomena common to semantically-rich domains. Cognitive Science Tech. Report No. 14, New Haven CT: Yale University.

Ahn, W.K., & Medin, D.L. (1989). A two-stage categorization model of family resemblance sorting. In Proceedings of the Eleventh Annual Conference of the Cognitive Science Society, Ann Arbor, MI.

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

Anglin, J.M. (1977). Word, object and conceptual development. New York: W.W. Norton.

Barsalou, L.W. (1989). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.), Concepts and conceptual development: Ecological

and intellectual factors in categorization. Cambridge: Cambridge University Press.

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

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

Cantor, N., & Mischel, W. (1979). Prototypes in person perception. In L. Berkowitz (Ed.), Advances in experimental social psychology. New York: Academic Press.

Corter, J.E., Gluck, M.A., & Bower, G.H. (1988). Basic levels in hierarchically structured categories. In Proceedings of the Tenth Annual Conference of the Cognitive Science Society, Montreal, Quebec, Canada.

Corter, J. E., & Gluck, M. A. (1990). Explaining basic categories: Feature predictability and information. Unpublished Manuscript.

Dejong, G., & Mooney, R.J. (1986). Explanation-based learning: An alternative view. Machine Learning, 1, 145-176.

Downing, P. (1978). On the creation and use of English nominal compounds. Language, 53, 810-842.

Duda, R. O. & Hart, P. E. (1973), Pattern classification and scene analysis, New York: Wiley.

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

Fisher, D. (1987). Knowledge acquisition via incremental conceptual clustering. Machine Learning, 2, 139-172.

Fisher, D. (1988). A computational account of basic level and typicality effects. In Proceedings of the Seventh National Conference on Artificial Intelligence, 233-238. Saint Paul, MN: Morgan Kaufman.

Gentner, D., & Gentner, D.R. (1983). Flowing waters or teeming crowds: Mental models of electricity. In D. Gentner & A.L. Stevens (Eds.), Mental models. Hillsdale, NJ: Erlbaum.

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

Gluck, M.A., Bower, G. H., & Hee, M.R. (1989). A configural-cue network model of animal and human associative learning. In Proceedings of the Eleventh Annual Conference of the Cognitive Science Society, Ann Arbor, Michigan. Hillsdale, NJ: Erlbaum.

Gluck, M.A., & Corter, J.E. (1985). Information, uncertainty and the utility of categories. In Proceedings of the Seventh Annual Conference of the Cognitive Science Society, Irvine, CA.

Gluck, M.A., Corter, J.E., & Bower, G.H. (1990). Basic levels in the learning of category hierarchies: An adaptive network model. Unpublished manuscript.

Hinton, G. E. (1987). Connectionist learning procedures. Technical Report CMU-CS-87-115. Carnegie Mellon University.

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

Hoffman, J., & Ziessler, C. (1983). Objectidentifikation in kunstlichen begriffshierarchien. Zeitscrift fur Psychologie, 16, 243-275.

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

Horton, M.S. & Markman, E. (1980). Developmental differences in the acquisition of basic and superordinate categories. Child Development, 51, 708-715.

Jolicoeur, P., Gluck, M.A., & Kosslyn, S.M. (1984). Pictures and names: Making the connection. Cognitive Psychology, 16, 243-275.

Jones, G.V. (1983). Identifying basic categories. Psychological Bulletin, 94, 423-428.

Koppitz, E. M. (1964). Evaluation of human figure drawings, New York: Grune & Stratton.

Lassaline, M.E. (1990). The basic level in hierarchical classification. Unpublished master's thesis, University of Illinois.

Luce, R.D. (1959). Individual choice behavior. New York: Wiley.

Malt, B.C., & Smith, E.E. (1983). Correlated properties in natural categories. Journal of Verbal Learning and Verbal Behavior, 28, 250-269.

Marr, D., & Nishihara, H.K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. Proceedings of the Royal Society of London, B200, 269-294.

Martin, R.C., & Caramazza, A. (1980). Classification in well-defined and ill-defined categories: Evidence for common processing strategies. Journal of Experimental Psychology: General, 109, 320-353.

Massaro, D.W., & Freidman, D. (1990). Models of integration given multiple sources of information. Psychological Review, 97, 225-252.

Matheus, C.J., Rendell, L.R., Medin, D.L., & Goldstone, R.L. (1989). Purpose and conceptual functions: A framework for concept representation and learning in humans and machines. The Seventh Conference of the Society for the Study of Artificial Intelligence and Simulation of Behavior. Sussex, England.

Medin, D.L. (1983). Structural principles in categorization. In Tighe, T.J. & Shepp, B.E. (Eds.), Perception, cognition, and development: Interactional analyses. Hillsdale, NJ: Erlbaum.

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

Medin, D.L., & Florian, J.E. (1991). Abstraction and selective coding in exemplar-based models of categorization. In A.F. Healy, S.M. Kosslyn, & R.M. Shiffrin (Eds.), From Learning Processes to Cognitive Processes: Essays in Honor of William K. Estes, Vol. 2. Hillsdale, NJ: Erlbaum.

Medin, D.L., & Schaffer, M.M. (1978). Context theory of classification learning. Psychological Review, 85, 207-238.

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

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

Medin, D.L., Wattenmaker, W.D., & Michalski, R.S. (1987). Constraints and preferences in inductive learning: An experimental study of human and machine performance. Cognitive Science, 11, 299-339.

Mervis, C.B., & Crisafi, M.A. (1982). Order of acquisition of subordinate-, basic-, and superordinate-level categories. Child Development, 53, 258-266.

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

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

Minsky, M. & Papert, S. (1969). Perceptrons. Cambridge, MA: MIT press.

Morris, M.W., & Murphy, G.L. (1990). Converging operations on a basic level in event taxonomies. Memory and Cognition, 18, 407-418.

Mostow, J. (1983). Machine transformation of advice into a heuristic search procedure. In R.S. Michalski, J.G. Carbonell, & T.M. Mitchell (Eds.), Machine Learning: An Artificial Intelligence Approach, Vol. 1. Los Altos, CA: Morgan Kaufman.

Murphy, G.L. (1982). Cue validity and levels of categorization. Psychological Bulletin, 91, 174-177.

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

Murphy, G. L. (1991). Parts in object concepts: Experiments with artificial categories, Memory and Cognition, 19, 423-438.

Murphy, G.L., & Brownell, H.H. (1985). Category differentiation in object recognition: Typicality constraints on the basic category advantage. Journal of Experimental Psychology: Learning, Memory and Cognition, 11, 70-84.

Murphy, G.L., & Smith, E.E. (1982). Basic level superiority in picture categorization. Journal of Verbal Learning and Verbal Behavior, 21, 1-20.

Murphy, G. L., & Wisniewski, E.J. (1989). Categorizing objects in isolation and in scenes: What a superordinate is good for. Journal of Experimental Psychology: Learning, Memory, and Cognition, 15, 572-586.

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

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

Nosofsky, R.M. (1991). Exemplars, prototypes and similarity rules. In A. Healy, S. Kosslyn, & R. Shiffrin (Eds.), From Learning Theory to Connectionist theory: Essays in honor of W. K. Estes, Vol. 1. Hillsdale, NJ: Erlbaum.

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

Rescorla, R.A., & Wagner, A.R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A.H. Black & W.F. Prokasy (Eds.), Classical conditioning II: Current research and theory. New York: Appleton-Century-Crofts.

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

Rifkin, A. (1985). Evidence for a basic level in event taxonomies. Memory and Cognition, 13, 538-556.

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

Rosch, E. (1974). Universals and cultural specifics in human categorization. In R. Breslin, W. Lonner, & S. Bochner (Eds.), Cross-cultural Perspective on Learning. London: Sage Press.

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

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

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

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

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

Smith, E.E., Balzano, G.J., & Walker, J. (1978). Nominal, perceptual, and semantic codes in picture categorization. In J.W. Cotton & R.L. Klatzky (Eds.), Semantic factors in cognition. Hillsdale, NJ: Erlbaum.

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

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

Tversky, B., & Hemenway, K. (1983). Categories of environmental scenes. Cognitive Psychology, 15, 121-149.

Tversky, B., & Hemenway, K. (1984). Objects, parts and categories. Journal of Experimental Psychology: General, 113, 169-913.

Wisniewski, E.J., & Gentner, D. (1991). On the combinatorial semantics of noun pairs: Minor and major adjustments to meaning. In G. B. Simpson (Ed.), Understanding word and sentence. Amsterdam: North Holland.

Wisniewski, E.J., & Murphy, G.L. (1989). Superordinate and basic category names in discourse: A textual analysis. Discourse Processes, 12, 245-261.

Wisniewski, E.J., & Medin, D.L. (1991). Harpoons and long sticks: The interaction of theory and similarity in rule induction. In D. Fisher & M. Pazzani (Eds.), Concept Formation: Knowledge and Experience in Unsupervised Learning. San Mateo, CA: Morgan Kaufman.

Wisniewski, E.J., & Medin, D.L. (1992). Is it a mornek or a plapel? Prior expectations and functionally relevant features in category learning. Unpublished Manuscript.

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

Zadeh, L.A. (1965). Fuzzy sets. Information and Control, 8, 338-353.