

# Parts in object concepts: Experiments with artificial categories

GREGORY L. MURPHY

*Brown University, Providence, Rhode Island*

Previous research has demonstrated that basic-level categories are associated with the parts of objects. Five experiments were conducted to investigate whether this correlation is caused by a psychological principle that requires basic concepts to have common parts. This hypothesis was investigated in the experiments by separating part and nonpart information in artificial categories. Basic-level structure was measured in two ways: as the level with the highest ratio of within-category similarity to between-category similarity, and as the level with the fastest categorization time in an object-identification task. The results revealed that basic-level structure could be found in categories that did not have parts in common. Furthermore, nonpart information, such as size, color, and texture, could both enhance the basic level and eliminate the basic-level advantage, depending on whether the nonpart information was consistent with or contradicted the part information. These results suggest that, psychologically, parts are neither necessary nor sufficient to form a basic level.

Although an individual object can belong to a number of categories, ranging from the very general (e.g., *physical object*) to the very specific (e.g., *long-haired Siamese cat*), there is one level of categorization that people naturally prefer (e.g., *cat*). This "basic" level is preferred in both language use and performance on identification tasks (Rosch, 1978; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). It is also apparent in language and concept learning in children (Anglin, 1977; Brown, 1958; Horton & Markman, 1980; Mervis & Crisafi, 1982).

The phenomenon of basic levels of categorization has been well documented in a variety of domains (Adelson, 1983; Berlin, Breedlove, & Raven, 1973; Rifkin, 1985; Tversky & Hemenway, 1983), but the theory behind basic-level structure is not completely settled. That is, the question of exactly what makes basic levels so natural, useful, and preferred in categorization has not yet received a complete explanation. The current article investigates one possible determinant of the basic level, the fact that object parts are typically associated with basic categories (Tversky & Hemenway, 1984).

In seeking to explain basic-level phenomena, most researchers have focused on differences in the structure of categories at different levels. One such account suggests two primary reasons for basic-level structure: a concept's *informativeness* and *distinctiveness*. A concept is

informative if it is associated with a large amount of concrete information. In general, the more specific a concept is, the more informative it is. A concept is distinctive to the degree that it is dissimilar from its contrast categories. Presumably, when two concepts are very similar, it is difficult to identify exemplars as being members of one concept and not the other. In general, the more specific a concept is, the *less* distinctive it is from its contrast categories. These principles are structural in the sense that they refer to abstract conceptual relations rather than to the category's specific content. The basic level is the level that is high on both of these factors simultaneously. For example, a concept such as *car* is fairly informative because it contains information about the shape, functions, materials, parts, and potential locations of its members. More general concepts, called *superordinates*, carry less information. So, the objects in the concept *vehicle* do not have common shapes, materials, parts, and locations. In contrast, more specific concepts, called *subordinates*, are slightly more informative than basic categories. However, they are not nearly as distinctive. *Two-door sedans* are quite similar to *four-door sedans* and even to *station wagons*, and so on. In short, the basic level is both informative and distinctive. This account of basic-level structure is consistent with experiments that contrast levels of categorization (e.g., Jolicoeur, Gluck, & Kosslyn, 1984; Mervis & Crisafi, 1982; Murphy & Smith, 1982; Rosch et al., 1976), and it is discussed in detail by Murphy and Brownell (1985).

In addition to such structural factors, however, there are also content differences between basic-level categories and categories at other levels. For example, Rosch et al. (1976, p. 392) noted that when subjects list attributes of superordinate categories, they tend to list a higher proportion of functions than for lower level categories. Artifact

---

This work was supported by Grant BNS 83-15145 from the National Science Foundation and Grant MH 41704 from the National Institute of Mental Health. The author would like to thank Edward Smith for inspiring these experiments. I am also grateful to Terry K. Au, Hiram H. Brownell, the reviewers, and the editor for helpful comments on the manuscript. Correspondence regarding this research may be addressed to Gregory L. Murphy, Department of Psychology, University of Illinois, 603 East Daniel St., Champaign, IL 61820.

superordinates such as *vehicle, furniture, tool*, and *toy* may be defined almost completely by their functions. Research has suggested that superordinates may also be associated with *collections*, or groups of objects, rather than individual objects, especially for young children (see Callanan & Markman, 1982; Markman, 1985; Markman, Horton, & McLanahan, 1980; Murphy & Wisniewski, 1989; Wisniewski & Murphy, 1989).

Tversky and Hemenway (1984) argued for another content difference between category levels. They found that the features associated with basic-level categories are primarily parts—the perceptually distinct segments that make up objects. For example, most chairs include a seat, a back, legs, and arms. Such parts make up the bulk of basic category feature lists. In contrast, superordinate categories have very few parts in common. Subordinate categories have most of the attributes of their basic categories, but these attributes, including the parts, are shared among contrasting subordinates. So, dining-room chairs have most of the same parts that other kinds of chairs do. Tversky and Hemenway found that the number of parts listed for a category increased dramatically from the superordinate to the basic level but showed little increase from the basic to the subordinate level. In contrast, nonpart attributes did increase from the basic to the subordinate level. They also found that members of the same basic category tend to share the same parts but tend to differ in other respects. Different basic categories did not share the same parts. Thus, parts can be used to distinguish basic-level categories but not higher and lower level categories.

Tversky and Hemenway (1984) discuss in some detail the potential utility of part information. They point out that the function of an object is often highly dependent on its parts; that is, the parts of an object determine how we can interact with the object and what it can do. An object with a seat can usually be sat on, an object with a handle can usually be grasped, an object with a door can usually be entered, and so on. Thus, function may be determined by the part configuration of an object, and for many concepts (especially artifacts), the function is an important component. Parts are also useful in that objects with common parts tend to have a common shape. Although it is possible to connect the same parts in different ways, given real-world constraints on biological structures and on the use of artifacts, most objects that have the same parts will tend to have the same shape.<sup>1</sup>

One possible argument against the importance of parts to all human concepts is based on nonobject categories. Although a basic level of categorization has been found in domains such as personality types (Canton & Mischel, 1979), events (Morris & Murphy, 1990; Rifkin, 1985), and computer programs (Adelson, 1983), such domains do not have perceptual shapes in the same way that objects do. However, as described in the General Discussion, these things may still have parts, which could be very important in their categorization (see also Morris & Murphy, 1990). Thus, the existence of a basic level in

nonobject domains does not by itself provide evidence against the importance of parts in categorization.

### INTERPRETATIONS OF THE ASSOCIATION OF PARTS AND THE BASIC LEVEL

Research on differences among category levels, then, provides strong evidence that there are both structural and content differences between basic and other category levels. Tversky and Hemenway's (1984) discussion of the importance of parts to conceptual structure provides a unified explanation of many of the phenomena associated with the basic level. However, there are different interpretations of their results that their data do not differentiate. In brief, their results could be interpreted as being a fact about the environment or as being a principle of the human conceptual system. The goal of this article will be to investigate the second hypothesis—whether people's conceptual system is based on part commonality.

What would it mean for the conceptual system to be organized around part information? One possibility is that there is a bias for people to attend to information about parts, to encode such information, and to use it preferentially in making category decisions. The predominant view of object identification is that it is primarily based on the object's shape (e.g., Marr, 1982; Ullman, 1989). One influential proposal by Biederman (1987) argues that recognition is accomplished by identifying the object's perceptually defined parts and their relations. Nonpart characteristics such as color and texture "will typically have only secondary roles" in identification (p. 118; see also Biederman & Ju, 1988; Ullman, 1989, p. 197). Since the goal of this system is to account for the initial categorization of objects, it suggests that people should form concepts around part commonality wherever possible.

However, Tversky and Hemenway's (1984) discovery that parts collect at the basic level might not reflect people's biases or processing strategies but could instead reflect the structure of the object categories themselves. A number of authors have argued that conceptual structure represents the structure of the environment or its interaction with the human perceptual apparatus. For example, Tversky and Hemenway (1984) argue that "the *natural breaks* among basic-level categories are between clusters of parts" (p. 186, emphasis added) and that the intercorrelation of parts, shape, and functions "seems to be a fact about the objects and organisms in the world" (p. 186). Rosch et al. (1976) argue more explicitly that categories reflect clusters of correlated attributes in the environment, that "the world does contain 'intrinsically separate things'" (p. 383). One reasonable view, then (though not necessarily one that these authors would endorse), is that the correlation of parts and basic categories is a result of the actual correlation of properties among objects in the world. However, this view does not entail that people have a bias towards detecting and using parts or that their concepts are generally organized around part information (though this could also be true). This view

only requires that people be sensitive to the correlated information in the environment—whatever it might be.

The current experiments investigate whether the conceptual system has a bias toward part information in learning and categorization—in particular, whether having common parts is a necessary condition for basic-level structure. If a basic level can be determined for categories that do not have the part structures found by Tversky & Hemenway (1984), then this would show that the conceptual system does not require part commonality in order to form basic-level structures.

Since virtually all natural categories seem to collect part information at the basic level, the present experiments used artificial stimuli whose part structure can be controlled. Previous experiments using artificial categories have obtained basic-level effects very similar to those using natural objects (Mervis & Crisafi, 1982; Murphy & Smith, 1982). The categories in Experiments 1–3 were formed so that a possible basic level was created through the clustering of attributes at one level of specificity, producing the pattern of informativeness and distinctiveness described above. However, these attributes were not parts. The question, then, is whether such categories would induce a basic category structure. Clearly, the conclusions from the experiments must be tempered by the fact that artificial stimuli were used. Furthermore, removing as many parts from the stimuli as possible might produce items quite different from most objects. However, determining the validity of a set of stimuli depends in part on the results they produce, so a complete discussion of this issue will be delayed until the General Discussion.

Because the studies used artificial stimuli, a number of the measures of basic-level structure used by Rosch et al. (1976) were not appropriate to use in these studies. In particular, action lists, shape overlap, free naming, order of name acquisition, and lexical representation were inapplicable. Most of the present experiments used a reaction-time (RT) methodology in which subjects verify the category of each object at different levels. This has probably been the most popular measure of basic-level structure in the field (see Experiment 3), and it has generally agreed with other measures. However, the use of categorization RT does not imply that it is a superior measure to the others, but simply that it is an appropriate one for these stimuli. Experiment 2 used a different measure of category structure, which replicated the results of the RT task. Experiments 4 and 5 used different categories, in which part and nonpart information could be contrasted.

### EXPERIMENT 1

The first three experiments used an artificial hierarchy with three levels of categorization. The purpose of the first experiment was to verify the properties of the stimuli to be used in the following experiments. Experiment 1A measured the feature distribution across category levels, and Experiment 1B measured the distribution of parts.

The categories were constructed so that they would have very few parts and so that the parts would not be distinctively associated with any category level. The stimuli were all basically rectangular and so did not differ on the basis of overall shape. Other than their borders, they were not divided into segments, and no part of a stimulus had any identifiable function. Thus, the stimuli were not composed of parts in the same way that natural objects are. Instead of parts, the stimuli could be distinguished by size, color, type of border, and internal texture elements (as described in the Method section). Figure 1 presents examples of the stimuli. However, it should be emphasized that the stimuli were all brightly colored and that the size differences of the originals may not be apparent in this reproduction. Figure 2 shows the hierarchy that describes the entire stimulus set, with brief descriptions of each category's properties.

One might question whether the stimuli differed in parts in spite of the attempt to remove this source of information. In a degenerate sense, every aspect of a stimulus can be considered a "part" of the stimulus. Thus, one can stretch the language to say that part of an apple is its red color or its size. However, such terminology would then make vacuous the claim that part information is associated with a particular conceptual level, as every attribute would be a part. Thus, to investigate the collection of part information at the basic level, one must use the narrower interpretation of what parts are—perceptually distinct segments of an object, often associated with a function, that could in principle be separated from the other parts of the object. Given this definition, the borders of the stimuli appear to be a part, but the size, color, and

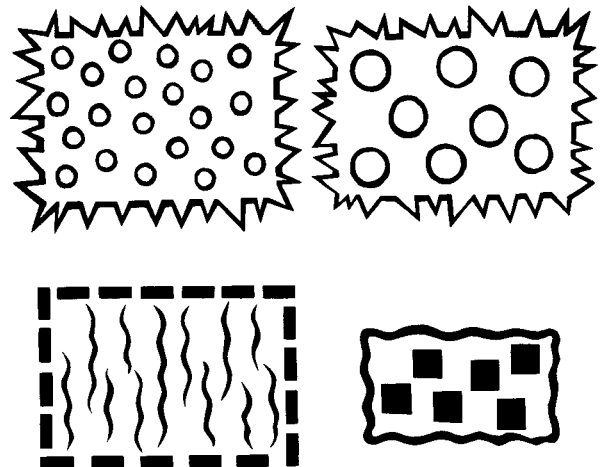


Figure 1. Examples of stimuli used in Experiments 1A, 1B, 2, and 3. The top pair of pictures represent two subordinates of the same middle-level category (a VAD on the left and a FAC on the right—see Figure 2). The difference between them is in the size of the internal circles. The bottom pictures represent different middle-level categories (a TIS and a LAR) from the same high-level category. They differ in edge type and texture elements. The original stimuli were brightly colored, and size differences appeared more dramatic in the larger originals.

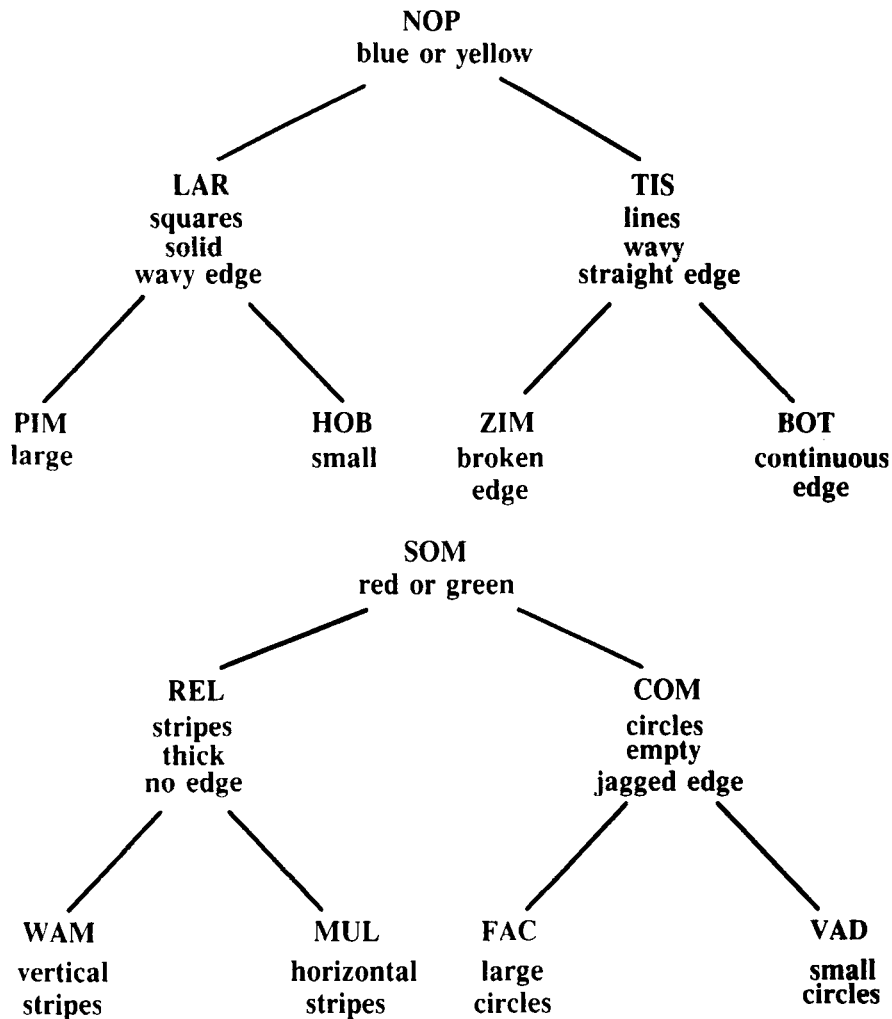


Figure 2. The names and hierarchical structure of the categories used in Experiment 1. Distinctive features of each category are listed. Each low-level category contained two exemplars differing in color and placement of the texture elements.

border features (straight, jagged, etc.) do not appear to be parts. The texture elements (small squares, circles, or lines inside the figures) are somewhat closer to parts. However, care was taken so that the different exemplars of each category did not have corresponding texture elements in the same place, so that no element could be considered a consistent part of the category. Because there was no functional or structural role played by these elements, they are more like polka dots or stripes on the surface of an object than like true parts. These claims are tested empirically in Experiment 1B.

The categories were constructed so that the middle level of the hierarchy was both informative and distinctive. The highest level categories were defined disjunctively, by having one of two colors (see Figure 2). This is similar to natural superordinates, which are often disjunctive. For example, different exemplars of *vehicle* have sails or wings or pedals but typically not all three. (However, real superordinates often contain functional features, which was not possible with these abstract stimuli.) The middle

level had many more features associated with it—the particular size, border type, color, and texture elements—and these features were distinctive to each category. Thus, the middle level had rich clusters of features, like natural basic categories (Rosch et al., 1976). The lowest level categories had a modification of one of these main dimensions. For example, one subordinate could have horizontal stripes, and its contrasting subordinate could have vertical stripes; or the texture elements could be small versus large circles. The differences between the lowest level categories, then, were similar to the differences between real subordinates, such as cuffed pants and cuffless pants or a crosscut saw and a backsaw. That is, the lowest level categories were informative but not distinctive, like natural subordinates.

Experiment 1A verified the proposed structure of the hierarchy by obtaining feature listings of the categories, similar to Rosch et al.'s (1976) Experiment 1. If the middle level is indeed a basic level, then there should be many more attributes associated with it than with the higher

level. The low level should only have a few additional attributes. Experiment 1B asked subjects to identify the parts in the categories. If the proposed analysis of the stimuli is correct, then subjects should identify very few parts in the stimuli and they should not be particularly associated with the middle level. Later experiments will then test for processing advantages at the middle level.

### Method

**Subjects.** Eight subjects, Brown University undergraduates, gave attribute listings in Experiment 1A. The four judges in Experiment 1B were Brown graduate students who were unfamiliar with this general research and naive to the specific predictions.

**Materials.** The 16 category exemplars were created by cutting shapes from colored paper and pasting them on white paper. As described above, each stimulus was essentially rectangular. The stimuli differed in their color, size, internal texture elements, and type of edge. As Figure 2 shows, the stimuli were organized into two separate hierarchies, creating three category levels. (The names in the figure were not used in this experiment.) The highest level separated patterns that were blue or yellow from those that were red or green. The middle level separated patterns based on their edge type and texture elements. For example, the LARs (see Figure 2) consisted of serrated-edged figures with solid squares inside them; their contrast category, the TISes, consisted of straight-edged figures with wavy lines inside. The lowest category level separated the patterns according to size or specific details of the internal texture elements or the edge. For example, ZIMs and BOTs were identical, except that one had a solid border and one had a broken border. Each lower level category was represented by two stimuli that differed in color. As mentioned above, the position of the texture elements was varied across the exemplars, so that they would not be perceived as consistent parts. The figures were photographed onto color slides.

**Procedure, Experiment 1A.** In this experiment, the subjects viewed one example of each lower level category (eight pictures in all) and listed the attributes of each. They were given standard feature-listing instructions that asked them to write down all the attributes of each picture, using brief phrases and avoiding idiosyncratic reactions. They were given 1 min to write the features of each picture.

The raw listings were collected and combined by a research assistant naive to the purposes of the study. She identified common features across the listings for each picture and noted how many subjects listed each feature. Only features listed by two or more subjects were included. Then, following Rosch et al. (1976), the lists were amended so that features listed for one category were applied to its contrast category when applicable. For example, the subjects listed "continuous border" for HOB, but not for PIM, which has the identical kind of border. In such cases, the feature was applied to both categories. Finally, when a feature was found in all the subcategories of some category, it was "moved up" and also listed at the higher level. The result was a list of features of each category level, increasing in size as one goes down the hierarchy.

**Procedure, Experiment 1B.** The subjects in this experiment decided what parts were associated with each category. It was found that defining a "part" was not as simple as asking subjects to list "features," and so a group of four judges was convened to decide upon the parts for each category. (The judges also performed this function for the stimuli in Experiment 5.) The instructions told the judges that the parts of an object (1) are the segments or portions that make it up, (2) can usually be separated from the rest of the object "with a little surgery," (3) can usually be described in a sentence like "X has a Y," and (4) are not global properties of the object as a whole. These criteria (except for the second) were based on the criteria described by Tversky and Hemenway (1984). The subjects were shown all the slides in a category and then wrote

down all the parts common to the objects (two slide projectors were used and pictures were re-presented as requested so that memory for the pictures was not a problem). Then the judges discussed their proposed parts and attempted to come to a consensus. In most cases, this was possible, but in a few cases, the judges were evenly split on whether some aspect of a category was a part and no consensus was forthcoming. In these cases, the part was given credit as half a part, representing the even split in judgments. (Either eliminating these features or giving them full credit does not change the pattern of results.) The judges completed this procedure for all the categories shown in Figure 2.

### Results

Rosch et al. (1976) found that natural superordinates had few features associated with them but that this number jumped at the basic level. The number of features increased only slightly from the basic to the subordinate level. Tversky and Hemenway (1984) found a similar pattern in part listings. If the design of the current stimuli is successful, the general feature listing will follow this pattern (indicating the middle level as being basic), but the parts listing will not.

**Experiment 1A.** The mean numbers of features listed were 1.0 for the top-level categories, 6.75 for the middle level, and 7.62 for the lowest level. This pattern is not very different from the mean number of features Rosch et al. (1976) found for their nonbiological categories, 1.67, 8.72, and 10.28, respectively. (Not surprisingly, real objects have more attributes.) Thus, knowing that something is in the highest level category is not informative, but knowing that it is in the middle-level category is. The lowest level category adds relatively little information (less than one feature on average). So, all other things being equal, one would expect this stimulus hierarchy to show the usual basic-level advantage at the middle level.

In general, the features listed were quite similar to those shown in Figure 2. In addition to such features, subjects also listed more detailed aspects of the stimuli, such as that texture elements were randomly spaced, that the squares were "imperfect," that texture lines were vertically oriented, and so on. Such features may not be salient aspects of an individual stimulus, but they were characteristics that the stimuli differed on. In addition to the features counted above, the feature "rectangular" was included as an attribute of *all* the categories. (It was not counted since it does not distinguish any of the categories; including it would increase each mean by 1 feature, maintaining the apparent basic-level structure.) This particular feature is important, because it confirms the claim that the stimuli did not differ significantly in their shapes.

**Experiment 1B.** As expected, the judges found relatively few parts in the stimuli, and the parts were not concentrated on the middle level. The mean numbers of parts were .5 at the top level, 1.25 at the middle level, and 1.875 at the lowest level: Each level adds less than one part on average. Tversky and Hemenway's (1984) judge-amended tallies of natural object categories found about .4, 5.8, and 6.0 parts at the superordinate, basic, and subordinate levels (these means are taken from the graph in their Fig-

ure 1, and so are approximate). Natural categories gain over 5 parts from the superordinate to the basic level, whereas the experimental stimuli gain only .75 parts. So, based on the features in general, the middle level would be expected to be basic, but based on the parts listing, the middle level is not exceptional.

The border of each pattern was listed as a part whenever it occurred. In addition, judges apparently believed that the rest of each stimulus must be a "part" of the stimulus, but they did not identify each texture element as a part, because the elements were not constant across exemplars, as described above. Instead, the judges typically identified the entire internal pattern as one whole part. Although this choice is more liberal than was expected, it did not result in clustering of parts at the middle level, because some of these internal patterns were judged to be distinctive at the lowest level. For example, the patterns of the FAC and VAD shown in Figure 1 were associated with the lowest level categories, since judges decided that they were not the same part. If the internal pattern features are discarded, the mean number of parts would be .5, .75, and .75 for the high, middle, and low levels—again, not clustering at the middle level.

These stimuli can now be used to test the predictions described in the introduction. If parts are a psychological necessity to form a basic level, then there should not be a distinct processing advantage for the middle level in the following experiments. If parts are not a necessary ingredient of basic-level processing, then the middle level should reveal a performance advantage, since it has the same kind of feature structure as the natural basic categories investigated by Rosch et al. (1976).

## EXPERIMENT 2

Most of the experiments in this article use the timed categorization task to identify basic-level structures. Although this is only one of the many tasks used by Rosch and subsequent investigators, previous research has found that different tasks usually agree on which level is basic (e.g., Mervis & Crisafi, 1982; Rosch et al., 1976; Tversky & Hemenway, 1983). Also, as mentioned earlier, some of the traditional tasks, such as free naming and order of acquisition, are not fully appropriate to these artificial stimuli. However, one method of ascertaining the basic level is easily extended to these stimuli, namely, the comparison of within-category similarity to between-category similarity of Mervis and Crisafi (1982).

The logic of this comparison is based on the reasoning that basic-level categories are relatively informative while at the same time being distinctive (Murphy & Brownell, 1985; Rosch et al., 1986, p. 385). In contrast, superordinate categories are not very informative and subordinates are not distinctive. Therefore, when comparing two objects within a category, one would expect high similarity at the basic and subordinate levels but not at the superordinate level. When comparing two objects from contrasting categories, one would expect them to be quite

different at the superordinate and basic levels but not at the subordinate level. Mervis and Crisafi (1982) showed that the measure of within- minus between-category similarity, which they called *category differentiation*, reveals the basic level (see also Morris & Murphy, 1990).

The current experiment used the stimuli introduced earlier to discover whether the middle level was the most highly differentiated according to this measure. Subjects were shown pairs of objects from the same or different categories and judged their similarity. The results were analyzed in terms of Mervis and Crisafi's (1982) differentiation score. If a basic-level structure is possible in spite of the lack of part collection at any level, then the middle level should have the highest differentiation scores.

## Method

The subjects were 11 Brown University students who were paid for their participation. The stimuli were the pictures described in the last experiment, forming the hierarchies shown in Figure 2. For every low-level category shown in Figure 2, the following trials were constructed: same low-, same middle-, and same high-level categories, and unrelated pairs. In the same low trials, the two pictures in a low-level category were paired. In the same middle trials, two pictures from the same middle-level category but *not* in the same low-level category were paired. In the same high trials, two pictures from the same high-level categories but *not* the same middle-level categories were paired. The unrelated trials contained two pictures that were not in any category together. Thus, there were 4 trials for every subordinate, resulting in 32 trials overall. Each subject rated every pair.

Before beginning the ratings, subjects viewed about 10 of the picture pairs (including a picture from each category), so that they could calibrate the similarity scale. During the experiment, the pictures in a pair were displayed through the slide projector sequentially for 5 sec apiece. Then the subjects had as much time as necessary to rate the pair. They used a 1-9 scale, in which 1 meant the pictures were very dissimilar or had nothing in common and 9 meant the pictures were almost identical. This is the same scale used by Mervis and Crisafi (1982).

## Results and Discussion

The mean similarity scores were 1.7 for unrelated, 2.2 for same high, 6.5 for same middle, and 7.9 for same low categories. As these scores indicate, there is a jump in similarity from the high to the middle categories, but a smaller increase to the low-level categories. Mervis and Crisafi's (1982) differentiation score involves subtracting the between-category similarity of a level from its within-category similarity. (The between-category similarity of one level is the within-category similarity of the next highest level.) So, the differentiation of the low-level categories is  $7.9 - 6.5 = 1.4$ . Because most of the similarity of this level is also found at the next highest level, it is not very differentiated. The differentiation scores were 1.4, 4.3, and .5 for the low, middle, and high levels, respectively. Clearly, the middle level is much higher than the other two. In Mervis and Crisafi's study, natural categories yielded differentiation scores of .79, 3.65, and 2.09 for subordinate-, basic-, and superordinate-level categories. If anything, the current stimuli show a greater differentiation advantage for the middle (proposed basic)

level than for their natural stimuli. This result occurred even though the categories Mervis and Crisafi used, kinds of furniture and clothing, share parts at the basic level (as shown by Tversky & Hemenway, 1984) and the artificial stimuli do not concentrate parts at the middle level. Apparently, subjects do not need to perceive common parts in order for categories to have a basic level, as measured by their similarity structure.

### EXPERIMENT 3

Experiment 3 used the categorization reaction-time measure to identify the basic level. Many experiments have demonstrated that people are fastest at categorizing objects into basic categories, with superordinate categorization typically being the slowest (Brownell, 1978; Jolicoeur et al., 1984; Murphy & Brownell, 1985; Murphy & Smith, 1982; Murphy & Wisniewski, 1989; Rosch et al., 1976; Smith, Balzano, & Walker, 1978). Computational theories of object recognition (e.g., Biederman, 1987; Marr, 1982) have placed great emphasis on the importance of shape in identifying objects. And since the parts of an object and their relations generally determine the shape of an object, it is not at all clear that categories with few parts and very similar shapes, like those shown in Figure 1, will reveal a basic-level structure in a visual-identification task.

As the previous experiment demonstrated, the middle level of the stimulus hierarchy is distinctive and informative, thereby fulfilling the structural description of the basic level. However, parts were not concentrated in the middle level. If the conceptual system requires part collections in order to form a basic-level structure or for fast object identification, then subjects should not show the usual basic-level advantage of prior experiments. If parts are not a necessary feature of the basic level, then one would expect subjects to be able to extract the clusters of features at the middle level (see Experiment 1A) and to demonstrate a basic-level advantage. Thus, this experiment will provide a test of the two interpretations of the connection between parts and basic-level structure.

#### Method

**Subjects.** The subjects were 18 Brown University undergraduates who were paid for participating.

**Materials and Apparatus.** The 14 categories were given the nonsense syllable names shown in Figure 2, which were chosen to maximize pronounceability and to minimize "meaningfulness" (see Murphy & Smith, 1982, for details). The figures used in the previous experiments were the stimuli. The RT experiment was run by an Apple II+ microcomputer, which controlled two Kodak carousel slide projectors and two Gerbrands shutters. Subjects responded by pressing a microswitch on a button box. The computer recorded the RTs and controlled the timing of the trials.

**Procedure.** The experiment began with a learning phase, in which subjects learned the names of the 14 categories. The learning procedure closely followed that of Murphy and Smith (1982). Subjects learned each level of categories separately, since each level was composed of contrast sets (i.e., each exemplar is in only 1 category at each level). In the learning phase, subjects were given a cover sheet for each category that gave the category name and explained why all the patterns were in the same category. That is,

it mentioned the features distinctive to that category. For example, the description for the category LAR, a middle-level category, was: "There are LARs, because their edge is serrated, they have squares inside, and the squares are solid." After reading the description of a category, subjects studied the exemplars. The next category in that level was then presented. Subjects were allowed to reread the descriptions and examine the exemplars at any time before the test phase.

After subjects reported having learned the categories at a level, they were tested in a nontimed task. Subjects were shown all the pictures twice and were asked to produce the category names they had just learned for each picture. Corrective feedback was given. This testing continued until the subjects could name an entire set without error.

After all the categories had been learned, the RT task began. Subjects sat in front of two slide projectors, facing a white wall. On each trial of this task, a category name was projected onto the wall. When subjects were ready to view the exemplar, they pushed any button on the response panel. The category name immediately disappeared, and the picture appeared 200 msec later. Subjects pressed a button with their dominant hands if the picture was in the named category, and they pressed the other button if it was not. No feedback was given. Trials were separated by a 1-sec interval, during which the slide projectors were advanced. There were 192 experimental trials, divided into two blocks. The experimental trials were preceded by 24 practice trials using names and photographs of everyday objects.

**Design.** The order of learning category levels was balanced across subjects by using a Latin square to ensure that each category level was learned first, second, and third equally often. In the RT phase, there were equal numbers of true and false trials, and each pattern had 12 true and 12 false trials, evenly divided among the three category levels. The order of the 192 trials was randomized. After half of the subjects were run, the order of the trials was reversed to further reduce any possible order effects. The false trials involved some choice in stimulus selection in that there were many possible incorrect category names that could be presented. (To follow this description, the reader should refer to Figure 2.) For the high-level false trials, the only incorrect category name (either NOP or SOM) was presented. For the middle-level false trials, two of the trials had a name and picture from the same high-level category and two were from a different high-level category. For example, if a picture of a TIS was presented, the name LAR appeared in two cases and REL and COM appeared in the other two cases. For the low-level false trials, there were three levels of relatedness: The category name could be from the same middle-level category (e.g., for a picture of category HOB, the name PIM would appear), from the same high-level category (e.g., for the same stimulus, the name ZIM or BOT), or a completely unrelated name (WAM, MUL, FAC, or VAD). For each item, there was at least one of these three types of false trials, the fourth false trial for that item being chosen randomly.

#### Results

After discarding any RTs greater than 5 sec (.3% of the responses), the correct RTs were submitted to analyses of variance (ANOVAs). The resulting mean RTs for the true trials are shown in Table 1. As expected, the middle level yielded the fastest categorization decisions, with the other two levels being about equally slow. The differences between the category levels was reliable [ $F(2,34) = 3.67$ ,  $MS_e = 12,914$ ,  $p < .05$ ]. To compare the category levels in false trials, it was necessary to control for the relatedness of the picture and category name. That is, for false high-level trials, the picture and category name

**Table 1**  
**Mean Categorization Times (in Milliseconds), Experiment 3**

	Category Level		
	Low	Middle	High
True	776	688	779
False	735	728	854

were always completely unrelated, whereas for middle- and low-level trials, some of the category names and pictures were related by sharing membership in higher level categories. Therefore, the RTs from only unrelated false trials at each level were collected and entered into an analysis. The mean RTs are also shown in Table 1. In this analysis, middle-level categories yielded slightly faster decisions than low-level trials, whereas high-level trials were considerably slower [ $F(2,34) = 5.74$ ,  $MS_e = 15,737$ ,  $p < .01$ ].

Another question regarding false trials is that of the relatedness of category name and picture. As is well known in the semantic memory literature, when subjects make categorization or set-membership decisions, the more related the two items (category and exemplar), the harder it is to respond negatively (e.g., McCloskey & Glucksberg, 1979; Smith, Shoben, & Rips, 1974). For example, it would be harder to respond "no" to *A fish is a mammal* than to *A fish is a vegetable*, because *fish* and *mammal* are similar, both being kinds of animals. The same sort of effect may be found with artificial stimuli (Murphy & Smith, 1982). As just mentioned, some of the categories are related by being in the same middle- or high-level categories. If a picture from one category is shown paired with a related category name, subjects may find it harder to answer "no."

For the low-level categories, there are three levels of relatedness in false trials: The category name and the picture could come from the same middle-level category, they could come from the same high-level category, or they could be unrelated. Reaction time decreased with decreasing relatedness [ $F(2,34) = 7.28$ ,  $MS_e = 23,885$ ,  $p < .005$ ]. When the name and picture were in the same middle-level category, the mean RT was much slower (931 msec) than when they were in the same higher level category (833 msec) or were unrelated (735 msec). A separate analysis showed that when the two items were from the same middle-level category, responses were slower than when they were from the same high-level category [ $F(1,17) = 6.87$ ,  $MS_e = 12,611$ ,  $p < .025$ ]. The difference between the high-level and unrelated conditions was not reliable [ $F(1,17) = 3.03$ ,  $MS_e = 28,812$ ,  $p = .10$ ].

For the trials using middle-level categories, there are only two levels of relatedness: The exemplar and category name could be from the same high-level category, or they could be unrelated. In this case, there was no difference between the two (732 and 728 msec, respectively) [ $F(1,17) < 1$ ,  $MS_e = 4,233$ ]. This result reinforces the null result above, that being in the same high-level category did not reliably slow categorization times.

The error rate was low overall (5.5%) and was positively correlated with RT ( $r = .55$ ), ruling out a speed-accuracy tradeoff. In particular, for the true trials, the error percentages were 5.9, 5.6, and 7.5 for low-, middle-, and high-level categories, respectively. Thus, the middle level was both fastest and most accurate. The large number of cells with no errors precluded performing an analysis of variance on these data.

## Discussion

The stimuli in this experiment had few, if any, parts of the kind defined earlier. In overall shape, the stimuli were identical. They differed in size, characteristics of the texture elements, color, and edge type. Nonetheless, the results gave the pattern expected from stimuli that have a basic level of categorization at the middle level. The middle level, which was both informative and distinctive, was fastest for true responses, with the other levels being about equally slower. Furthermore, in the false responses, when the picture and category name were from the same middle-level category (e.g., the picture was of a VAD and the name was FAC—see Figure 2), responses were quite slow. In contrast, when the picture and category name shared the same high-level category, there was no significant slowing relative to unrelated items. This suggests that middle-level-category membership was more salient than membership in the higher level—that is, that the middle level was a basic category. Murphy and Smith (1982) found similar results with artificial categories that were based on shape and part commonality at the basic level.

When the category levels were compared in false responses (for unrelated pictures and category names), the low- and middle-level responses were about equally fast, with high-level categories engendering much slower times. Why should it be that the middle level is quite a bit faster for true trials but not for false trials? The model of Murphy and Smith (1982) provides a possible answer to this seeming contradiction. In making a "true" judgment, the subject must decide that the features in the picture match the features associated with the category to some criterion. Because subordinate categories are so similar to one another, subjects must establish a higher criterion to answer "yes" (to avoid categorizing a sports car as a *sedan*, for example). Similarly, in this experiment, to correctly categorize something as a BOT, subjects had not only to identify the straight edge of the figure and its internal wavy lines, but also had to ensure that the edge was a solid one. Thus, to correctly categorize stimuli at the lowest level, subjects had to verify more features than at higher levels. For false trials, however, subjects can make a decision as soon as they find any mismatching feature. For example, if a subject expected a BOT, he/she could answer "false" immediately upon noticing a wavy edge or internal circles or the color red, and so on. In this way, more specific categories can have an advantage in false trials (holding relatedness constant), because they are associated with more features than higher level categories, giving more chances for a mismatching



feature. Similarly, the relatedness effects are easily explainable as a function of the number of mismatching features between the category representation and the picture presented.

One alternative interpretation of this result is that the RT advantage of the middle level is an artifact of some other aspect of the stimulus structure. Perhaps just being the middle level of a hierarchy or perhaps the number of categories at that level caused the results. Fortunately, Murphy and Smith (1982) were able to rule out such explanations in their studies, which used the same hierarchical structure, number of categories, and category names. Furthermore, the current Experiments 4 and 5 demonstrate modifications of the middle-level advantage using the identical category structure. Thus, it is unlikely that the results were caused by such an artifact.

In summary, this experiment provides evidence that the presence of parts is not necessary to ensure a basic-level advantage in a visual-identification task. Although the stimuli shared an overall shape and had no obvious parts, the middle level showed an advantage very similar to that of previous basic-level experiments. The middle-level categories shared a large number of features and were very distinctive relative to the other categories at their level. The lower level categories also shared many features, but they were not nearly as distinctive. The higher level categories shared no single feature, but had a disjunctive pair of features. In these respects, the category levels resembled basic, subordinate, and superordinate categories. The results suggest that this feature structure is sufficient to generate the category-level phenomena noted by Rosch et al. (1976) and others and that the presence of parts is not necessary.

#### EXPERIMENT 4

Although parts may not be absolutely necessary to create a basic-level advantage, the stimuli of previous experiments in the literature have in fact collected parts at the basic level. One possible criticism of Experiments 2 and 3 is that the nonpart information created a basic level only because there was no consistent part information available. In natural objects, which contain both part and nonpart attributes, perhaps people form their concepts and identify objects only on the basis of part information. If this were true, Experiments 2 and 3 do not address the possibility that parts have a privileged place in the conceptual system, since parts were not available. That is, although parts may not be necessary to form a basic level, perhaps when they are present, they fully determine category structure. A recent comparison of shape and surface information by Biederman and Ju (1988; discussed below) gives evidence for this hypothesis. One way to test this possibility would be to discover whether nonpart information influences categorization—specifically, the basic-level advantage—even when parts are available.

This experiment took the artificial-tool stimuli used by Murphy and Smith (1982) and added other features to the

shape features that primarily defined those categories. If the other features make no contribution to basic category structure, then their addition will not change the pattern of categorization times. In particular, it will not change the basic-level advantage.

Experiment 4 compared two stimulus conditions, a simple condition and an enhanced condition. In the simple condition, the stimuli were exactly the same as those used by Murphy and Smith (1982); they consisted of black-and-white outline drawings of artificial tools. These stimuli have already been shown to yield a basic-level advantage for the middle category level. In the enhanced condition, other distinctive features were added to the middle categories. The tools of one middle category were drawn in red and were covered with red dots; the tools of other categories were drawn in green with stripes, in yellow with circles, or in solid blue. These features were distinctive to the middle categories in that contrasting subordinate categories shared these extra features, and superordinate exemplars did not have any one of these features in common to all its members. Thus, these features might be expected to increase the basic-level advantage. If such nonpart attributes are not relevant to basic-level structure, then they should have little effect on the basic-level advantage.

How should "the basic-level advantage" be measured in this experiment? In general, it is the difference between basic categories and subordinate and superordinate categories on some performance measure. (The absolute speed of processing across the two conditions is not of interest, since the basic level is defined *relative* to the other levels. Adding new features for subjects to learn might slow them down overall.) However, the theory that basic-level structure reflects informativeness and distinctiveness makes fairly specific predictions on how the addition of these nonpart features should influence categorization. By adding features to categories at the middle level, one is making those categories both more informative and more distinctive. Lower level categories are being hurt by this manipulation, because they are now less distinctive than before. For example, the WAM and MUL (the hierarchy is the same as in Figure 2) in the simple condition shared the same handle, shaft, and head shape but differed in the head width. In the enhanced condition, they shared two more features, red color and dots, but still had only one distinctive feature. Thus, the lower level categories become less distinctive in the enhanced condition, a factor known to slow categorization (McCloskey & Glucksberg, 1979; Murphy & Brownell, 1985): When viewing a WAM, it is now harder for subjects to be sure that it is not a MUL instead. Adding features made high-level categories slightly more distinctive than before, however. By perceiving a tool as red or dotted, subjects could identify it as a SOM, whereas these features were unavailable in the simple condition. This effect is not as beneficial to the high-level categories as it is to the middle-level ones, however, because these features were not common to all the high-level exemplars. Thus, if the explanation offered earlier for basic-level structure is correct, the difference between low-

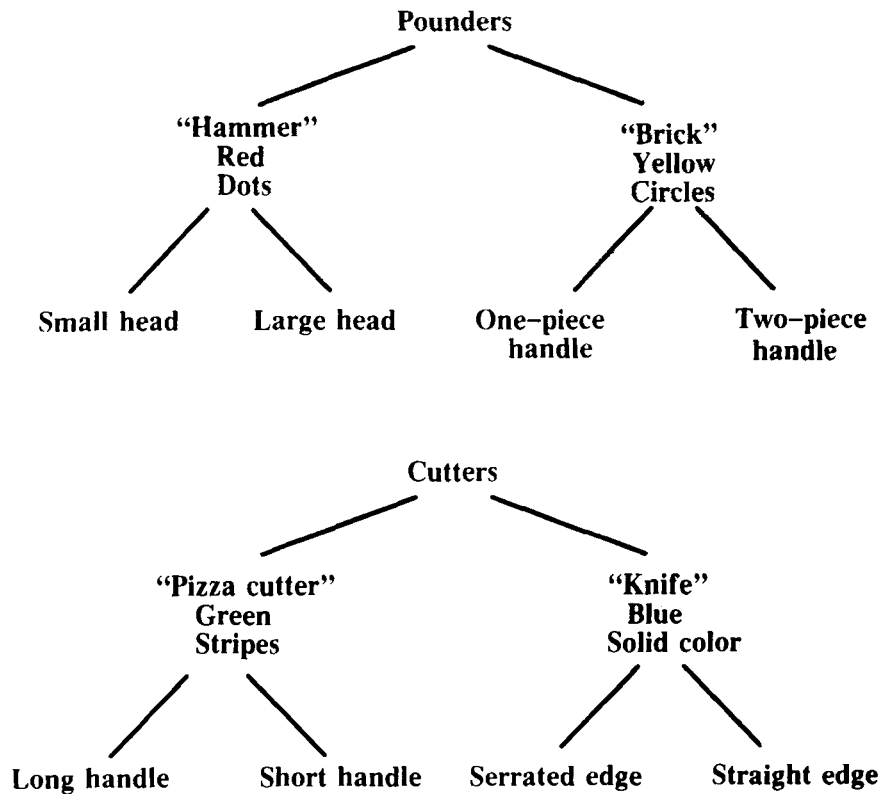


Figure 3. The hierarchical structure of the enhanced categories used in Experiment 4. The descriptions in each category summarize the features associated with that category. The names "hammer," "brick," "knife," and "pizza cutter" refer to the overall shape of the object. These names were never presented to subjects.

and middle-level categories should be greater in the enhanced condition but the difference between middle- and high-level categories should be the same or slightly less.

### Method

**Subjects.** The subjects were 24 Brown University undergraduates who were randomly assigned to the two conditions with the constraint that equal numbers of subjects be assigned to each one. Subjects were paid for participating.

**Stimuli.** Sixteen black-and-white drawings of tools, as described by Murphy and Smith (1982), were the category exemplars in the simple condition. These drawings were traced over with colored markers to create the enhanced exemplars. Also, the enhanced exemplars had internal texture added to them, as described above. The category structure of the enhanced condition is outlined in Figure 3. (The names in quotes in that figure refer to the object shapes; these names were not used in the experiment and are for expository purposes only.) As it shows, the color and texture features were common to all exemplars in a middle-level enhanced category, but these features did not help distinguish contrasting subordinate categories. The color and texture were distinctive of superordinate categories, but as only half of a superordinate's exemplars had the same color, these would not be as helpful here as at the basic level. Also, following Murphy and Smith's (1982) experiment, all subjects were told a function for superordinate categories. SOMs were said to be pounding tools and NOPs cutting tools. In terms of shape, the simple and enhanced categories were identical. Color slides were made of the stimuli to be used during both training and the RT task.

**Procedure and Design.** The procedure of the experiment was the same as that of Experiment 3. The design was also very simi-

lar, except for the addition of the stimulus condition (simple or enhanced) as a between-subjects factor. The order of learning each category level was again determined by a Latin square for each condition. And again, each stimulus order in the test phase was determined randomly and then reversed after half the subjects had been run.

### Results

The results from the true trials of both conditions are shown in Table 2. First, the simple condition will be described and analyzed to discover whether it replicated the results of Murphy and Smith (1982, Experiment 1). In general it did: The basic categorizations were the fastest, with the subordinates relatively close behind and the superordinates markedly slower. The differences among the three levels were reliable [ $F(2,22) = 7.09$ ,  $MS_e = 12,792$ ,  $p < .005$ ]. Furthermore, for the false trials

Table 2  
Mean Categorization Times (in Milliseconds), Experiment 4

	Category Level		
	Low	Middle	High
Simple Condition			
True	862	811	980
False	949	792	983
Enhanced Condition			
True	1,132	854	955
False	1,000	806	875

(with relatedness held constant), the basic level was by far the fastest [ $F(2,22) = 4.11$ ,  $MS_e = 29,077$ ,  $p < .05$ ]. Thus, the results for the simple condition give ample evidence that a basic-level advantage can be found for these stimuli.

When the stimuli were enhanced by the addition of color and texture cues, the basic-level advantage was again found. As Table 2 shows, the basic level is still the fastest, with subordinate categories now quite a bit slower. The total differences among the levels was reliable [ $F(2,22) = 9.13$ ,  $MS_e = 26,035$ ,  $p < .005$ ]. For false trials, the basic level was also the fastest, and the differences were again reliable [ $F(2,22) = 5.02$ ,  $MS_e = 23,229$ ,  $p < .025$ ].

These findings are probably not too surprising in that they essentially replicate the Murphy and Smith (1982) results. The main question of interest is whether the addition of the color and texture cues actually influenced the basic-level advantage, that is, whether there was an interaction of stimulus condition and category level. (Comparisons of a single level across groups cannot reveal the preferred level of categorization.) The true RTs were combined into a larger analysis of variance, with stimulus type as a between-subjects factor. As one would expect from the earlier analyses, the differences among the three category levels was highly significant [ $F(2,44) = 9.54$ ,  $MS_e = 19,413$ ,  $p < .001$ ]. There was no main effect of stimulus type (simple vs. enhanced) ( $F < 1$ ,  $MS_e = 282,114$ ). More important, the interaction between stimulus type and category level was significant [ $F(2,44) = 7.38$ ,  $MS_e = 19,413$ ,  $p < .005$ ], showing that the differences among category levels were different in the two conditions. In particular, the difference between the low and middle levels increased from 51 msec in the simple condition to 278 msec in the enhanced condition. In contrast, the difference between middle and high levels decreased slightly, from 170 to 101 msec. Orthogonal contrasts of the interaction effect revealed that the differences between the low level and the two higher levels accounted for a significant portion of variance [ $F(1,44) = 14.03$ ,  $p < .001$ ], but that the differences between the middle and high levels did not [ $F(1,44) < 1$ ]. That is, the interaction seems to be due primarily to the slowing of the subordinate level relative to the others.

The relatedness factor in the false items was not as much an issue in this experiment, but the data were analyzed as in Experiment 3 nonetheless. First, the low-level trials will be presented. For the simple stimuli, when the picture and category name were in the same middle-level category, subjects found it difficult to say "no" relative to the pairs that shared the same high-level category (1,065 vs. 924 msec, respectively) [ $F(1,11) = 8.44$ ,  $MS_e = 14,142$ ,  $p < .025$ ]. However, stimuli sharing the same high-level category were no slower than unrelated stimuli (949 msec) [ $F(1,11) < 1$ ,  $MS_e = 9,380$ ]. This shows that the middle-level category was an important one in judging stimulus identity, whereas the high level was not.

For the enhanced stimuli, the items that shared the same middle category were again slower than those that shared only a high category (1,318 vs. 1,100 msec) [ $F(1,11) =$

8.55,  $MS_e = 33,468$ ,  $p < .025$ ]. In this condition, however, the latter items were also slower than the completely unrelated exemplar-category pairs (874 msec) [ $F(1,11) = 5.76$ ,  $MS_e = 52,659$ ,  $p < .05$ ], contrary to the results in the simple condition. In other words, adding the color and texture to the stimuli apparently made the high-level categories somewhat more salient than they were in their simple versions. This result is reminiscent of the speed-up of the superordinates in the true trials, and it will be discussed shortly.

For middle-level false trials, the picture and category name could either share the same high-level category or be unrelated. For both conditions, the difference between these two was nonsignificant for both the simple and enhanced conditions [ $F(1,11) = 4.24$ ,  $MS_e = 13,947$ ,  $p > .05$ ;  $F(1,11) = 1.52$ ,  $MS_e = 6,243$ ,  $p > .20$ ].

The error rates for the simple and the enhanced conditions were 5.0% and 6.1%. Errors were positively correlated with RT for both conditions,  $r = .28$  for the simple stimuli and  $r = .72$  for the enhanced stimuli. Interestingly, one cell had markedly more errors than any other: In the enhanced condition, for low-level false trials in which the picture and name were in the same middle category, there were 18.3% errors. In contrast, the same cell in the simple condition had 2.5% errors. This is more evidence that the addition of nonpart features enhanced the basic category advantage (making it harder to respond negatively to stimuli from the same middle category).

## Discussion

The results paralleled the predictions to a large degree. Both conditions evinced a significant RT advantage for the middle level of categorization. More significantly, the enhanced condition showed a greater advantage than did the simple condition, as shown by the reliable interaction of condition and category level. The particular pattern of this change is of considerable interest. When the new cues were added, the difference between the middle- and low-level categories increased markedly, whereas the difference between the middle- and high-level categories actually decreased somewhat. This pattern is consistent with the explanation of basic-level structure described earlier, in that the new features were stored primarily with the middle categories, making the lower level categories less distinctive. Since adding the features made high-level categories slightly more distinctive, it is not surprising that they became relatively faster in the enhanced condition, though this was not a reliable effect. This finding provides crucial support to the hypothesis that parts are not the only features that contribute to a category level having superior performance. Because color and texture pattern cannot be plausibly defined as parts of an object (at least, when they do not correspond to component boundaries), this experiment shows that categorization into basic categories uses all kinds of perceptual information—not just part-based information.

A recent finding by Biederman and Ju (1988) may seem to contradict the results of this study. They found that subjects were just as fast in identifying black-and-white con-

tour drawings of objects as they were in identifying color photographs. The texture, color, and other surface information in the photos apparently had little effect on RT. This result does not necessarily contradict the conclusions of the present experiment, however. Biederman and Ju did not compare different category levels but only one level of naming and categorizing, the basic level. Thus, their result does not contradict the main findings here, since they did not compare the same conditions. Also, they used very familiar concrete objects whose shapes were highly associated with the objects in long-term memory. Although they agree that one *could* use surface information to help identify objects in some cases (pp. 59-62), it may be that form information is so quickly processed in these familiar objects that the additional surface information gives little help. Finally, some of the nonpart information used in the present experiments was different from that found in most everyday objects and in their experiment. If, for example, all telephones were the exact same size or had bright green stripes, then people might well use this information to identify them.

In short, Biederman and Ju's (1988) study cannot be meaningfully compared with the present one, since they did not compare different levels of categorization. The fact that they found little use of surface features in identification may be an important commentary on the real-world association of these features to object categories more than on whether people *can* use this information to categorize objects. The present experiment shows that subjects can and do use such features when they are available.

### EXPERIMENT 5

The hypothesis under consideration claims that the conceptual system prefers to use parts to organize concepts. The first two experiments have demonstrated that nonpart attributes are important in establishing a basic level. Unfortunately, it is virtually impossible to test whether part or nonpart features are more important. It would be very difficult to equate the features for salience, ease of encoding, and similar variables. Even such basic considerations as equating the number of features present in a stimulus would be difficult to carry out. Is size just one feature? How many features are in a part, like the seat of a chair? And if one could control these variables, this might remove the interesting differences between parts and nonpart attributes.

Because of such problems, this experiment does not investigate the general question of whether parts or nonpart features are more important. Instead, it investigates the narrower question of whether nonparts can weaken or eliminate a basic-level advantage created by parts. The stimuli were the tools used in Experiment 4, again augmented by color and texture cues. Unlike in Experiment 4, the two kinds of cues were pitted against one another rather than picking out the same basic level. In this experiment, the shape and parts of the tools were common at the middle level (as before), but the color, size, and texture cues

were associated with the highest level. If these cues are important in category representation and identification, then the highest level may become as fast as or faster than the middle level. This manipulation should influence categorization of both the middle and lower levels, since both are being made less distinctive. However, the middle level should still be faster than the lower level, as it still has a number of distinctive features (i.e., the parts and shape), whereas the lowest level has only one (e.g., head width, type of handle).

### Method

Subjects were 18 Brown University students who were paid for their participation. Four other subjects were eliminated because of overly high error rates (greater than 30% errors in one of the true conditions).

The stimuli were the eight simple tool exemplars used in Experiment 4, modified in the following way. The SOMs were all red and dotted, had continuous edges, and were large; the NOPs were green and striped, had broken edges, and were small. As before, each high-level category also shared a common function. The middle level was still defined primarily by overall shape and the low level by one feature difference. As mentioned in Experiment 1, four judges examined these stimuli to determine their parts. The reason for this was to ensure that the added features of color, size, edge, and dots or stripes were not considered parts. And in fact, none of the added features was listed as a part; the only parts were the components of the tools (such as the head and handle) that were also present in the black-and-white drawings. Thus, the features added at the high level did not change the part configuration of the categories. For this experiment, the computer used was an IBM XT-PC running the Digitry reaction-time system.

The procedure and design were the same as in Experiments 3 and 4. The descriptions of the categories presented during the learning phase were altered to reflect the changes in the stimuli.

### Results

The data were analyzed through the standard ANOVAs, followed by two orthogonal contrasts. One contrast examined the difference between the high-level and middle-level categories; the other contrast compared low-level categories with the mean of the other two. The assumption was that the addition of nonpart attributes to the high level would eliminate the difference between it and the middle level (found to be basic in the previous experiment), as measured by the first contrast. However, low-level categories should still be slow relative to the other levels, since they were not altered. Thus, only the second contrast should reveal significant differences.

The data from the true trials, shown in Table 3, followed the predicted pattern. The highest level categories were slightly faster than the middle-level categories, with the low-level categories being quite a bit slower. The over-

Table 3  
Mean Categorization Times (in Milliseconds), Experiment 5

	Category Level		
	Low	Middle	High
True	1,072	881	854
False	961	827	914

all differences between the three levels were reliable [ $F(2,34) = 9.65$ ,  $MS_e = 899,734$ ,  $p < .001$ ].<sup>2</sup> Furthermore, the planned contrasts revealed that there was no difference between the high and middle levels [ $F(1,34) = .25$ ], but that the lower level was slower than the other two [ $F(1,34) = 19.04$ ,  $p < .001$ ]. In short, adding nonpart features to the highest level eliminated the middle RT advantage in true trials. The low-level trials were still slower, as predicted. In fact, the difference between middle and low levels is greater here than for the simple pictures of Experiment 4.

For false trials, the first analysis looked only at items in which the category name and picture were completely unrelated. As Table 3 shows, the pattern here was somewhat different from that of the true trials. For false trials, the middle level was fastest, followed by the high and the low levels. But unlike in previous experiments, the three levels were not reliably different [ $F(2,34) = 1.82$ ,  $MS_e = 1,553,642$ ,  $p < .20$ ], so the added features apparently reduced the potency of the middle-level advantage.

The relatedness analyses of false items also provides an estimate of whether high-level categories achieved greater coherence with their added features. For middle-level trials, the name and picture could be from either the same or a different high-level category. The mean RTs were 1,024 and 827 msec, respectively [ $F(1,17) = 5.96$ ,  $MS_e = 989,166$ ,  $p < .03$ ]. Being in the same high-level category slowed the subjects' ability to say "no," suggesting that these concepts had achieved a certain amount of coherence. This effect was not found in the previous experiments with the original high-level categories.

For false trials with low-level names, the results were not so clear-cut. As before, when a picture and name were from the same middle-level category, subjects took much longer to respond "no" (1,582 msec) than when they were from the same high-level category (978 msec) or were unrelated (961 msec) [ $F(2,34) = 8.29$ ,  $MS_e = 9,231,960$ ,  $p < .002$ ]. The latter two conditions were not very different, in spite of the difference found with middle-level names.

The error rate was 5.1% overall, and errors were positively correlated with RT,  $r = .59$ , ruling out a speed-accuracy tradeoff. Importantly, the high-level true trials had fewer errors (3.5%) than the middle- (8.2%) or low-level (5.9%) true trials. Thus, the highest level was both the fastest and most accurate.

### Discussion

These results demonstrate that nonpart information can influence the utility of category levels even when it contradicts part information. In contrast to the results of Murphy and Smith (1982) and of Experiment 4, which used similar stimuli, the high-level categories were now as fast and accurate in categorizing stimuli as the middle-level was. Clearly, the subjects represented the nonpart information and used it in making category decisions. The results for false trials are a little more equivocal. The middle level was nonsignificantly faster than the other two. If false decisions are being made primarily on the basis

of mismatching features, this result may simply reflect the advantage more specific concepts have in detecting mismatches. Alternatively, the difference in shapes at the middle level may have been more salient than the differences in color or textures associated with the superordinate level. Nonetheless, the effect of relatedness in the false trials with middle-level names indicates that subjects were still sensitive to the nonpart features. In fact, this comparison was not significant in any of the previous experiments, suggesting that the high-level categories were more coherent in this experiment, as predicted.

One unpredicted result was the relatively large difference between the middle- and low-level categorizations. It may simply be that the low-level categories, which started out with low distinctiveness, simply became too similar for subjects to distinguish quickly. Because the middle-level categories could still be distinguished by shape, they were not as badly affected.

One might object to this demonstration by arguing that there should always be exactly one basic level and that all its categories must be superior to categories formed at higher and lower levels of abstractions. The results found here would be viewed as unnatural, since there was no single preferred level. Such a view is oversimplified. Studies of natural categories have shown that there is considerable variation of "basicness" across different categories. Jolicoeur et al. (1984) and Murphy and Brownell (1985) demonstrated that certain subordinate concepts have the same processing advantages as basic concepts. Furthermore, studies of expertise have shown that experts may have excellent categorization skills at levels lower than the basic level (Rosch et al., 1976; Tanaka & Taylor, 1991) or higher than the basic level (Dawson, Zeitz, & Wright, 1989; Murphy and Wright, 1984). At the same time, experts certainly have no problem categorizing at the basic level. Thus, the current results are by no means inconsistent with findings from natural categories.

### GENERAL DISCUSSION

Tversky and Hemenway's (1984) important paper on object parts and categorization pointed out a striking correlation in natural object categories. It showed that members of the same basic category but not members of superordinate categories share common parts. Furthermore, subordinate categories are differentiated by other kinds of features. Things that have the same parts tend to have the same shapes, and this is an important source of information for categorizing objects. Sharing parts could also explain why people interact with basic category members using the same behaviors (Rosch et al., 1976). However, what was not clear from Tversky and Hemenway's empirical results was whether their finding reflected a psychological principle that gives special weight to part attributes. Is having common parts a necessary component of basic-level structure? And if parts are associated with a particular category level, is that sufficient for the level to be basic? The present study used novel categories in

order to experimentally manipulate the presence of part and nonpart attributes.

Experiments 2 and 3 provided evidence that parts are not necessary aspects of basic category structure. The stimuli in those experiments shared the same overall shape and varied in color, size, and texture elements. The stimuli did not have obvious segments, and the few parts identified by judges were not concentrated at the middle level. Nonetheless, the results showed a basic-level advantage that was predictable from the number of features associated with each level. The categories in the middle level had a large number of features associated with them, and they were distinctive, just as natural basic concepts are. They showed a normal basic-level advantage in the similarity and categorization tasks. Thus, even with stimuli that did not vary in parts, a basic-level structure emerged. This structure is not simply an artifact of being the middle level of an artificial hierarchy (or of the precise number of categories at each level, etc.), as Experiments 4 and 5 modified that effect using the same three-level hierarchy.

Experiments 4 and 5 demonstrated that having part information associated with one level is not sufficient to create a basic level. First, Experiment 4 showed that nonpart information could enhance the basic level, which rules out the view that nonpart features are overwhelmed by the shape-part information when both are available. Second, Experiment 5 showed that by massing nonpart information (color, size, and texture) at a higher level, the basic-level advantage due to parts could be eliminated. That is, the informativeness of the nonpart information at one level counteracted the informativeness of the part information at another level. Thus, simply collecting part features at one level of categories does not ensure that that level will be the basic level.

Thus, parts are neither necessary nor sufficient for establishing a basic-level structure. It is important to re-emphasize that these studies do not question Tversky and Hemenway's (1984) empirical finding that parts are generally collected at the basic level. The issue is whether this is a property of conceptual structure. It might well be that people are particularly sensitive to part information of objects and that their conceptual apparatus uses that information exclusively to form useful concepts. Alternatively, it may be that the distinctive categories in the world tend to share parts, which in turn cause them to share other features (functions, shapes, materials, common locations, etc.), but that the human conceptual system has no bias in favor of parts.

The results presented here favor the second interpretation of Tversky and Hemenway's (1984) findings. If parts are neither necessary (cf. Experiments 2 and 3) nor sufficient (cf. Experiment 5) for a category to demonstrate a basic-level advantage, then it seems unlikely that part commonality is an essential constraint on forming basic concepts. The remaining explanation, which was not directly tested here, is that the best categories of objects in the

world happen to share parts. That is, this result most likely reflects environmental structure.

The experiments presented here have not attempted to explain why parts normally *are* so important. Nonetheless, the question should be addressed as to why parts seem so important in both object identification and categorization (Biederman, 1987; Tversky & Hemenway, 1984) and yet did not appear essential in the current studies. One possible explanation is that parts are crucial to some earlier component of the categorization process. That is, concept learning involves both a conceptual-memory component in which the category representation is formed and a perceptual component in which the stimulus is analyzed. (Of course, these may not be completely independent.) The extraction of parts may be performed by the perceptual system, in that parts can usually be determined by perceptual discontinuities, without requiring considerable top-down processing (Biederman, 1987). And since parts are important to determining shape, perception may automatically detect them. Thus, the reason that parts are normally so important in concept representations could be that they are the features that the perceptual system picks up and inputs to the concept-formation process. Furthermore, as Tversky and Hemenway (1984) showed, parts are particularly helpful in distinguishing natural basic concepts, and so it is not surprising that they are encoded and used in categorization.

In short, I am suggesting that the concept-learning process, which forms prototypes and links concepts into hierarchical structures, is neutral on the issue of parts. However, our perceptual apparatus may pick out parts in object perception and present them as the features to be learned in concept formation.

The current experiments used only artificial stimuli, as it was impossible to find actual category taxonomies that did not have parts collected at the basic level. Clearly, these stimuli limit the generality of the results. The main problem is probably the lack of richness in artificial concepts compared with the wealth of knowledge people have about real objects. However, in defense of the present categories, it should be noted that the stimuli of the first three studies followed the same pattern of results as Rosch et al.'s (1976) natural categories. That is, the middle category level was "basic" according to the feature lists (Experiment 1A), the category similarity measure (Experiment 2), and categorization time (Experiment 3). These results provide converging evidence that the category levels were comparable to natural categories in relevant respects. Furthermore, research on categorization has provided many examples of similarities in learning actual and artificial concepts. For example, the acquisition of prototypes seems to be similar in abstract, artificial stimuli and in natural object concepts (e.g., Rosch & Mervis, 1975). Previous work on basic concepts has shown highly consistent similarities between studies using artificial categories and those using natural categories (see Mervis & Crisafi, 1982; Murphy & Smith, 1982). Thus,

artificial concepts provide an important means to test hypotheses arising from natural concepts, even if they cannot be a complete substitute.

Nonobject categories, such as scenes, personality types, events, and computer programs (Adelson, 1983; Cantor & Mischel, 1979; Rifkin, 1985; Tversky & Hemenway, 1983) provide an interesting domain in which to further study the parts hypothesis. Consider event concepts. The parts of events are not exactly like object parts in that they do not combine to form shapes, as object parts do. That is, all the parts of *riding a train* or *assault* do not form a single perceptual entity in the way the parts of a chair or an iguana do. For example, the beginning and end of an event (and its component subevents) are often difficult to determine, whereas most objects have clear boundaries. Thus, if perceptual shape information is crucial, one would not have expected event concepts to follow Tversky and Hemenway's (1984) generalization. However, there is a basic level of events that shares parts and shows a number of the expected processing advantages (Morris & Murphy, 1990), which is further evidence that shapes per se are not the crucial element for forming a basic level. Research on the part structure of other nonobject domains would further our understanding of what is common to all concepts.

#### REFERENCES

- ADELSON, B. (1983). *Constructs and phenomena common to semantically rich domains* (Cognitive Science Technical Report No. 14). New Haven, CT: Yale University.
- ANGLIN, J. M. (1977). *Word, object and conceptual development*. New York: Norton.
- BERLIN, B., BREEDLOVE, D. E., & RAVEN, P. H. (1973). General principles of classification and nomenclature in folk biology. *American Anthropologist*, *75*, 214-242.
- BIEDERMAN, I. (1987). Recognition-by-components. A theory of image understanding. *Psychological Review*, *94*, 115-147.
- BIEDERMAN, I., & JU, G. (1988). Surface versus edge-based determinants of visual recognition. *Cognitive Psychology*, *20*, 38-64.
- BROWN, R. (1958). How shall a thing be called? *Psychological Review*, *65*, 14-21.
- BROWNELL, H. H. (1978). *Picture perception and semantic memory*. Unpublished doctoral dissertation, Johns Hopkins University, Baltimore.
- CALLANAN, M. A., & MARKMAN, E. M. (1982). Principles of organization in young children's natural language hierarchies. *Child Development*, *53*, 1093-1101.
- CANTOR, N., & MISCHEL, W. (1979). Prototypes in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 12, pp. 3-52). New York: Academic Press.
- DAWSON, V. L., ZEITZ, C. M., & WRIGHT, J. C. (1989). Expert-novice differences in person perception: Evidence of experts' sensitivities to the organization of behavior. *Social Cognition*, *7*, 1-30.
- HORTON, M. S., & MARKMAN, E. M. (1980). Developmental differences in the acquisition of basic and superordinate categories. *Child Development*, *51*, 708-719.
- JOLICOEUR, P., GLUCK, M., & KOSSLYN, S. M. (1984). Pictures and names: Making the connection. *Cognitive Psychology*, *16*, 243-275.
- MARKMAN, E. M. (1985). Why superordinate category terms can be mass nouns. *Cognition*, *19*, 31-53.
- MARKMAN, E. M. (1989). *Categorization and naming in children*. Cambridge: MIT Press.
- MARKMAN, E. M., HORTON, M. S., & McLANAHAN, A. G. (1980). Classes and collections: Principles of organization in the learning of hierarchical relations. *Cognition*, *8*, 227-241.
- MARR, D. (1982). *Vision*. New York: W. H. Freeman.
- McCLOSKEY, M., & GLUCKSBERG, S. (1979). Decision processes in verifying category membership statements: Implications for models of semantic memory. *Cognitive Psychology*, *11*, 1-37.
- MERVIS, C. B., & CRISAFI, M. A. (1982). Order of acquisition of subordinate, basic, and superordinate level categories. *Child Development*, *53*, 258-266.
- MORRIS, M. W., & MURPHY, G. L. (1990). Converging operations on a basic level in event taxonomies. *Memory & Cognition*, *18*, 407-418.
- 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, & Cognition*, *11*, 70-84.
- MURPHY, G. L., & MEDIN, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*, 289-316.
- MURPHY, G. L., & SMITH, E. E. (1982). Basic-level superiority in picture categorization. *Journal of Verbal Learning & 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, & Cognition*, *15*, 572-586.
- MURPHY, G. L., & WRIGHT, J. C. (1984). Changes in conceptual structure with expertise: Differences between real-world experts and novices. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *10*, 144-155.
- RIFKIN, A. (1985). Evidence for a basic level in event taxonomies. *Memory & Cognition*, *13*, 538-556.
- ROSCH, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27-48). 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., JOHNSON, D., & BOYES-BRAEM, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*, 382-439.
- 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* (pp. 137-168). Hillsdale, NJ: Erlbaum.
- SMITH, E. E., SHOBE, E. J., & RIPS, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, *81*, 214-241.
- TANAKA, J. W., & TAYLOR, M. E. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology*, *23*, 457-482.
- 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*, 169-193.
- ULLMAN, S. (1989). Aligning picture descriptions: An approach to object recognition. *Cognition*, *32*, 193-254.
- WISNIEWSKI, E. J., & MURPHY, G. L. (1989). Superordinate and basic category names in discourse: A textual analysis. *Discourse Processes*, *12*, 245-261.

#### NOTES

1. An anonymous reviewer pointed out that many animals share many parts, such as eyes, hearts, and mouths, and that a number of artifacts might share parts like batteries, gears, switches, and so on. Sharing such parts does not confer common shapes. This is a valid point, and it reveals two subtle issues, neither of which can be done justice here. First, the fact that two parts have the same name does not necessarily mean that they are conceptually the same. Birds and cats both have "heads,"

"tails," and "legs," but they are clearly not the same heads, tails, and legs (e.g., imagine a bird with a cat's legs). It is not clear how to tell whether two parts with the same name are psychologically the same part. Second, examination of feature lists will reveal that subjects generally do not list such internal parts as hearts, kidneys, wires, gears, and so forth. This is part of the problem that every pair of objects has a large, probably infinite number of features in common, but people only attend to a small number of them (for detailed discussions, see Markman, 1989, and Murphy & Medin, 1985). Fortunately, parts listed by

subjects do not have the overlap problem posed by the above examples, as Tversky and Hemenway's (1984) data show. The discussion in the main text may only apply to such psychologically salient parts.

2. A different statistical package was used for Experiment 5 than for the previous experiments. Although the  $F$  values do not change as a function of the statistical program, the  $MS$ s do change, as they depend on the exact computational formula used. Therefore, the  $MS$ s of this study are not comparable to those of the previous studies, but the  $MS$ s within this experiment can be meaningfully compared.

(Manuscript received October 16, 1989;  
revision accepted for publication December 12, 1990.)

## Notices and Announcements

### **21st Annual Meeting of the Society for Computers in Psychology San Francisco, California November 21, 1991**

The 21st Annual Meeting of the Society for Computers in Psychology will be held at the Hyatt Regency Hotel at Embarcadero Center in San Francisco on November 21, 1991, the day before the annual meeting of the Psychonomic Society. The meeting will include presentations, discussions, tutorials, and times for software and hardware demonstrations. All areas of psychology are featured, including research, education, clinical practice, and industrial applications.

For further information regarding the conference, contact William L. Palya, Department of Psychology, Jacksonville State University, Jacksonville, AL 36265 (BITNET address FWLP@JSUMUS, phone (205) 782-5641, FAX (205) 782-5680).

### **32nd Annual Meeting of the Psychonomic Society San Francisco, California November 22-24, 1991**

The 32nd Annual Meeting of the Psychonomic Society will be held in San Francisco, November 22-24, 1991. The meetings will begin Friday morning and continue until Sunday at noon. The headquarters hotel will be the Hyatt Regency San Francisco at Embarcadero Center.

The program and hotel reservation cards are being mailed to members and associates this month (September). A copy of the program will be published in the November issue of the *Bulletin of the Psychonomic Society*.

For further information, please contact the secretary-treasurer of the Society: Cynthia H. Null, P.O. Box 7104, San Jose, California 95150-7104 (telephone: 415-604-1260).