COGNITIVE PSYCHOLOGY 19, 242–279 (1987)

Family Resemblance, Conceptual Cohesiveness, and Category Construction

DOUGLAS L. MEDIN AND WILLIAM D. WATTENMAKER

University of Illinois

AND

SARAH E. HAMPSON

Birbeck College, University of London

Many people have argued that natural categories are organized in terms of a family resemblance principle. Members of family resemblance categories tend to share properties with each other but have no properties that are singly necessary and jointly sufficient (defining) for category membership. This paper reports seven experiments using a sorting task to evaluate the conditions under which people prefer to construct categories according to a family resemblance principle. The first set of studies followed the typical practice of defining family resemblance in terms of independent sets of matching and mismatching values. Across a variety of stimulus materials, instructions, procedures, and category structures, family resemblance sorting was almost never observed. Despite procedures designed to prevent it, participants persisted in sorting on the basis of a single dimension. The second set of studies explored the idea that interproperty relationships rather than independent features serve to organize categories. We found that people will abandon unidimensional sorting in favor of sorting by correlated properties, especially when they can be causally connected. In addition, when conceptual knowledge is added which makes interproperty relationships salient, family resemblance sorting becomes fairly common. Implications of the results for the development of family resemblance categories and the practice of treating properties or features as additive and independent are discussed.

INTRODUCTION

It is natural to categorize. Both our language and our experience lead us to treat sets of nonidentical stimuli as in some way equivalent. The

This research was supported in part by National Institute of Mental Health Grant MH32370 and NSF Grant BNS 84-19756 to the first author, and National Institute of Mental Health Grant MH39077 to the third author. Lawrence Barsalou, Dedre Gentner, Greg Murphy, Brian Ross, and Edward Shoben provide valuable comments on earlier versions of this manuscript. Address correspondence, including requests for reprints, to Douglas L. Medin, Department of Psychology, University of Illinois, 603 E. Daniel, Champaign, IL 61820.

0010-0285/87 $7.50

categories which people normally create and use represent a tiny subset of the many possible ways in which entities and experiences could be partitioned. Therefore, a central question is what basic principles underlie category construction. That is, why do we have the categories we have and not others?

The present paper reports a series of studies where people are given a set of stimuli and asked to partition them into categories. We began this work with the idea of examining sorting based on the principle of family resemblance (e.g., Rosch & Mervis, 1975), and the first half of this paper can be thought of as a series of failed attempts to do so. Family resemblance categories are fuzzy categories where the members are generally similar to each other, but where there is no set of defining properties that any and all examples have. Family resemblance traditionally has been defined in terms of matching and mismatching properties or attributes, where the individual properties are treated as independent of and unrelated to each other. The second half of the paper examines category construction in situations where the component properties are not independent of one another but rather are related in terms of a theme or theory. Here we find what appears, on the surface, to be family resemblance sorting, but sorting which has an underlying conceptual basis in terms of the theme or theory. It is possible that there is both a similarity-based and a conceptually based family resemblance sorting principle, but our studies so far support only the latter principle. Before describing our studies in detail, however, we need to provide a more general introduction to ideas concerning structural principles in categorization.

Structural Principles and Categorization Theories

It is reasonable to expect that the most useful or natural category partitionings might be associated with lexical concepts. Theories about the structure of such concepts can be used to suggest principles of category construction.

One view of category structure is that natural language concepts are characterized by simple sets of defining features that are singly necessary and jointly sufficient to determine category membership (e.g., Katz & Postal, 1964). A candidate exemplar either does or does not possess these defining features and thereby is or is not a member of the category. The major problem with the classical view is that research suggests that the majority of natural concepts are not organized around defining features but rather are structured in terms of sets of typical or characteristic features (see Mervis & Rosch, 1981; Medin & Smith, 1984; E. E. Smith & Medin, 1981, for recent reviews).

The rejection of the classical view of categories has been associated with the ascendance of the probabilistic view of category structure. The current consensus has it that categories are "fuzzy" or ill-defined, and that they are organized around a set of properties or clusters of correlated attributes that are only characteristic of category membership. Membership in a category can thus be graded rather than all-or-none, where the better members have more characteristic properties than the poorer ones. In an attempt to be specific about the structural basis of categories organized around clusters of correlated attributes, Rosch and Mervis (1975) had subjects list properties of exemplars for a variety of concepts such as *bird, fruit,* and *tool.* They found that the listed properties for some exemplars occurred frequently in the concept, while others had properties that occurred less frequently and, most importantly, the more frequent an exemplar's properties were, the higher its rating for typicality in that category. Rosch and Mervis developed a measure called *family resemblance,* which for an exemplar increases with the frequency of the properties it shares with members of its own category and decreases with the frequency of properties it shares with members of contrasting categories (cf. Tversky, 1977, pp. 347–349). Less formally, family resemblance increases with within-category similarity and decreases with between-category similarity. Family resemblance is highly correlated with the speed with which an exemplar can be categorized as well as with other typicality effects (see Rosch & Mervis, 1975).

The present experiments are concerned with category construction and particular attention is focused on sorting on the basis of a family resemblance principle. According to this principle, sorting should be organized around exemplars that are prototypical of potential categories. In Rosch's words the idea is "that potential prototypes will tend to become centers of categories in free sorting" (Rosch, 1975b, p. 196). That is, if we construct artificial categories by selecting prototypes and generating examples to create a family resemblance structure, then these same categories should be reproduced when people are allowed to construct their own categories from these examples. Before we define the notion of family resemblance sorting more precisely, it will prove useful to describe a more formal definition of similarity.

Structural Principles and Similarity

One of the most salient properties of categories is that their members appear to be more similar to each other than to members of contrasting categories. This observation leads naturally to the suggestion that people should sort examples into categories so as to maximize within-category similarity relative to between-category similarity. Of course one cannot simultaneously maximize within-category similarity and minimize between-category similarity—maximizing within-category similarity would lead one to sort n objects into n categories, whereas minimizing between-category similarity will always lead one to sort *n* objects into exactly two categories (Medin, 1983). Therefore, some trade-off between within- and between-category similarity is inevitable. In the present studies we use average within-category similarity minus average between-category similarity as our measure of the trade-off between the two forms of similarity, but a variety of other forms of trade-off would yield essentially the same predictions.

The Rosch and Mervis (1975) measure of family resemblance treats all features or properties as equally salient (weighted only by frequency), and matches and mismatches are equally weighted. Figure 1 presents an abstract description of a set of 10 entities and two alternative means by which they might be sorted into two equal-sized categories. The dimensions correspond to types of components or features and 1 and 0 correspond to values on these dimensions. For example, D_1 might be color, and 1 might correspond to a red stimulus and 0 to a green stimulus. The abstraction notation 1111 might correspond to one large red triangle and the notation 0000 to a stimulus consisting of two small green circles.

The sort on the bottom left side of Fig. 1 is labeled as family resemblance and the topmost example in each category represents the prototype or best example of the category. Each of the other examples would match the best example on three of its four values. An alternative sorting strategy is to partition the examples on the basis of values on a single dimension (in the example in Fig. 1, the first dimension, or color). If all components are equally weighted, it is easy to see that the family resemblance sort maximizes the average within-category similarity minus the average between-category similarity. For the family resemblance partitioning there is an average of 9.6 within-category matches (the first example has 12 matches to other category members and the other four examples have 9 matches to other members; self-matches are not counted) and an average of 6.4 between-category matches, yielding a difference of 3.2 matches (mismatches will show a mirror image pattern and can be ignored in this example). For the one-dimensional partitioning there is an average of 9.2 within-category matches and an average of 7.2 between-category matches, yielding a difference of 2.0 matches. Comparing the two sorting strategies it is clear that the family resemblance partitioning produces greater within- relative to between-category similarity for the situation where the constituent dimensions are equally weighted.

A different picture would emerge if one dimension dominated all others. If the first dimension receives all the weight (no other dimensions are considered), then the one-dimensional sort produces an average difference of four matches whereas the family resemblance sort produces an average difference of two matches. Therefore, predictions based on the Rosch and Mervis family resemblance measure depart from similarity-

Example	DIMENSION D_1	D_2	D_3	D_4
1	1	1	1	1
2	1	1	1	0
3	1	0	0	0
4	0	1	0	0
5	0	1	1	1
6	0	0	0	0
7	0	0	1	0
8	1	0	1	1
9	0	0	0	1
10	1	1	0	1

Family Resemblance Sort								One-Dimensional Sort							
Category A				Category B				Category A				Category B			
Dimension				Dimension				Dimension				Dimension			
D_1	D_2	D_3	D_4	D_1	D_2	D_3	D_4	D_1	D_2	D_3	D_4	D_1	D_2	D_3	D_4
1	1	1	1	0	0	0	0	1	0	0	0	0	1	0	0
1	1	1	0	0	0	0	1	1	1	1	1	0	1	1	1
1	1	0	1	0	0	1	0	1	1	1	0	0	0	0	0
1	0	1	1	0	1	0	0	1	1	0	1	0	0	1	0
0	1	1	1	1	0	0	0	1	0	1	1	0	0	0	1

FIG. 1. The abstract notation for the 10 stimulus items presented in Experiments 1, 2b, 2c, 2d, 3, and 6. The partitioning on the bottom left represents a sort consistent with family resemblance principles and the partitioning on the bottom right represents a sort consistent with the use of a single dimension.

based predictions to the extent that one dimension is much more salient than the others. Tversky (1977) has demonstrated convincingly that the relative weighting of a feature (as well as the relative importance of common and distinctive features (see Gati & Tversky, 1984) varies with the context and experimental task. This flexibility raises the possibility that participants might, as a strategy, focus on a single dimension and make it functionally the most salient. This strategy could give rise to one-dimensional sortings even when the dimensions were perceptually equally salient. On the other hand, if family resemblance sorting is natural, then one might expect a strategic balancing of weights that would work to eliminate any effects of differences in perceptual salience.

The present studies used a variety of stimulus materials whose abstract structure was either identical or similar to that shown in the top half of Fig. 1. The general procedure involved asking the subject to examine the stimuli carefully and then place them into two equal-sized groups in "a way that seems natural or sensible." Our general interest was in the basis for sorting, and our particular interest was in family resemblance sorting.

GENERAL METHODS

Subjects

The subjects were male and female undergraduates attending the University of Illinois, who participated in partial fulfillment of course requirements in introductory psychology. Participants were run in groups that varied in size and in some cases were run individually after they had participated in an unrelated experiment.

Stimuli

The exact stimulus materials employed varied from study to study. The abstract description of the stimulus structure either was identical with or similar to that shown in Fig. 1. Whenever the structure is not identical to that shown in Fig. 1, it is described in appropriate detail. Generally, however, materials were constructed by selecting two prototypes that differed from each other along all the component dimensions (e.g., 1111 versus 0000) and then generating additional examples by performing minimal distortions of a given prototype (i.e., 1111 would yield 1110, 1101, 1011, and 0111, as examples). These new examples always shared more values with their original prototype than with the contrasting prototype. Since the prototypes also appeared as examples, they can be also designated as the best examples of the (potential) family resemblance categories.

Procedure

Each participant received a set of examples which were in a randomly scrambled order. Subjects were asked to lay out the examples, to look them over carefully, and to place the stimuli into two equal-sized groups in a way that seemed natural or sensible. The experimenter pointed out that there were many ways to classify the stimuli, that there was no one "correct" answer, and that we were interested in what the participants thought was a natural partitioning. In several of the studies additional instructions aimed at inducing family resemblance sorting were given, and these additional instructions are described at the appropriate point. Finally, after the subjects completed the sorting task they were asked to write down the basis for their sorting. This additional requirement is important in that a given partitioning could be described in a variety of different ways.

EXPERIMENTS 1, 2, 3, AND 4

The first study was conducted with the expectation that family resemblance sorting would be a very common strategy. To our surprise, no one created a family resemblance partitioning. The next several experiments employed various modifications of stimulus materials and instructions that were designed to induce family resemblance sorting.

Experiment 1

Stimuli and method. The stimuli for the first study were outline drawings of cartoonlike animals. Each animal was created by selecting values on each of the following four dimensions: head shape (angular versus rounded), number of legs (4 or 8), body markings (spots versus stripes), and tail length (short or long). The individual examples differed from each other only in the combinations of properties they possessed.

The abstract structure of the examples conformed exactly to that shown in Fig. 1. The assignment of abstract notation to particular values varied from subject to subject. For example, 1111 might correspond to round head, four legs, stripes, and long tail for one subject but angular head, four legs, spots, and a long tail for another subject. An example of one realization of the abstract structure of Fig. 1 is shown in Fig. 2, where the animals have been grouped to show the family resemblance structure. The procedures were in exact accordance with the general procedures outlined earlier.

Results and discussion. The results are easy to describe. No subject sorted by family resemblance; rather, each participant created a one-dimensional sorting conforming to the pattern in the lower right-hand side of Fig. 1. That is, each of the members of the constructed categories had one value shared by all its members. For example all four-legged animals would be in one category and all eight-legged animals in the other category for a subject who sorted on the basis of number of legs. Before drawing any conclusions from these results, we decided to examine the generality of our findings.

Experiment 2

The second experiment used a variety of stimulus materials to see if our failure to find family resemblance sorting was peculiar to the particular realization employed in the first study. Experiments 2a and 2b used drawings of cartoonlike animals, whereas Experiments 2c and 2d used descriptions of hypothetical people.

Experiment 2a. The stimuli for Experiment 2a were again cartoonlike animals, but they varied in value along six dimensions. The animals were constructed by adding the dimension of Expression (smile versus none) and Number of Antennae (1 or 2) to the animals described in Fig. 2. The best examples or prototypes became, respectively, 111111 and 000000, and six additional examples were created from each prototype by varying one of the six dimensions (e.g., to create 111110, 111101, 111011, 110111, 101111, and 011111 from 111111). Altogether there were 14 animals consisting of the two prototypes and the six additional examples constructed from each of them. The rationale for using six dimensions was twofold: first, it would increase the difference in within- minus between-category similarity associated with family resemblance sorting relative to one-di-

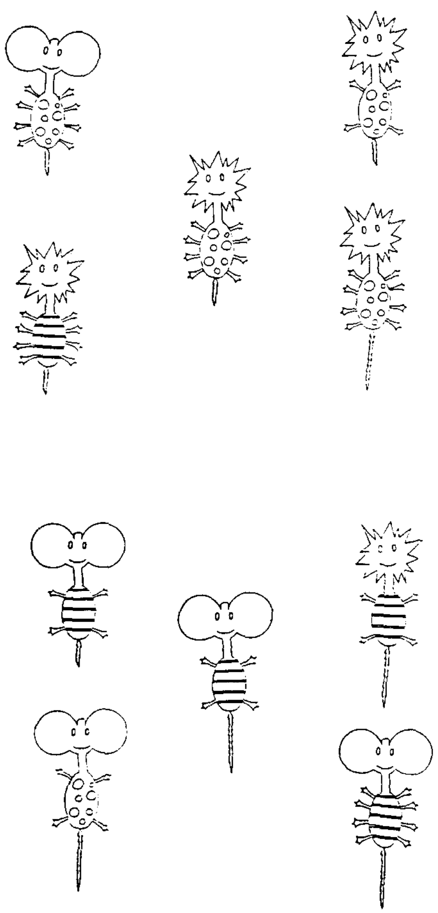

FIG. 2. The stimulus materials used in Experiments 1 and 3 grouped by a family resemblance structure.

mensional sorting, and second, there is evidence suggesting that as stimulus complexity increases, subjects are more likely to adopt a nonanalytic strategy and respond to overall similarity (e.g., Kemler, 1983; L. B. Smith, 1981). Experiment 2a employed 20 subjects.

Experiment 2b. One of the many ways in which the stimuli in the first experiment differ from more natural materials is that most natural stimuli are not simply factorial combinations of some set of properties. While robins and eagles both have wings, the wings of robins differ from the wings of eagles in a variety of ways. In general, as Medin, Dewey, and Murphy (1983) note, there are many properties that are unique to

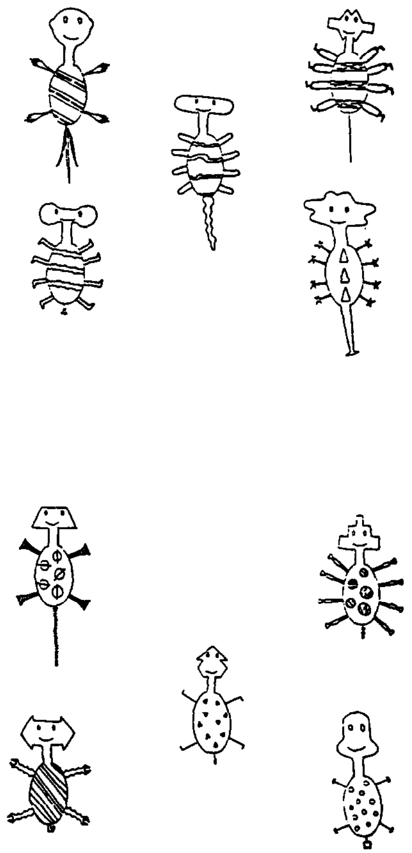

FIG. 3. The stimulus materials used in Experiments 2b and 3 grouped by a family resemblance structure.

members of categories and allow them to be identified individually. The stimulus materials in Experiment 2b were cartoonlike animals whose abstract structure corresponded to that shown in Fig. 1, but the drawing contained numerous exemplar-specific properties. These were introduced by means of variable realizations of abstract properties. For instance, although half of the animals had long tails, each one of the long tails was different from each of the other long tails. A sample set of these materials is shown in Fig. 3. We thought it possible that the addition of unique properties would encourage nonanalytical strategies and family resemblance sorting. A second major change is that the instruction specifically asked subjects to use all the properties of the stimuli rather than a single one in setting up their categories. Experiment 2b employed 16 subjects. Another set of 20 subjects was run on the stimulus materials with no unique properties. Except for the change in instructions, this condition represents a replication of Experiment 1.

Experiment 2c. The Rosch and Mervis family resemblance scores are based on subjects' listed properties. In Experiment 2c the stimuli were also property lists in the form of verbal descriptions of personality traits of people. By analogy to the visual stimuli in Experiment 1, the four traits associated with each description each varied with respect to a different personality dimension.

Investigations into the structure of the language of personality description have repeatedly identified between three and five dimensions of meaning, which encompass the entire range of personality characteristics captured by personality traits (Goldberg, 1982). The five dimensions were labeled by Norman (1963) as extroversion, agreeableness, conscientiousness, emotional stability, and culture.

The personality descriptions were structured around the four dimensions conscientiousness, emotional stability, culture, and agreeableness. The traits were selected from a factor analysis of ratings of 455 traits kindly made available by Lewis R. Goldberg. This analysis provided the factor loadings for each trait on all five dimensions. Traits were selected as representative of a particular dimension if they loaded above .40 and did not load higher than .40 on any of the other four dimensions.

Two sets of materials were constructed which we refer to as factorial versus related. This difference is roughly analogous to the difference between visual stimuli having or not having unique or exemplar-specific properties. For the factorial set the exemplars were composed of factorial combinations of two values on each of the four personality dimensions as is shown in Table 1. That is, each stimulus consisted of a woman's name and four descriptors. As the abstract notation in Table 1 indicates, the stimuli were organized around the two prototypes or best examples, Irene and Suzie. Irene was described with the desirable traits *relaxed* and *affectionate* on the emotional stability and agreeableness dimensions, and with the undesirable traits *careless* and *gullible* on the conscientiousness and culture dimensions. Suzie was described with desirable traits on the conscientious and culture dimensions and with undesirable traits on the emotional stability and agreeableness dimensions. This balancing was designed to circumvent sorting on the basis of an evaluative dimension.

In the related condition the exact trait descriptors used were different for each person, but they were selected on the basis of factor loadings to have values similar to those for the factorial stimuli. For example, instead of describing the five women as careless, one would be described as careless, another as sloppy, another as negligent, and so on. This is analogous to the exemplar-specific properties in Experiment 2b. The exact de-

TABLE 1
The Factorial Four-Dimensional Set Used in Experiments 2c and 2d

Irene (1111) Careless Relaxed Gullible Affectionate	Suzie (0000) Efficient Envious Knowledgeable Hard
Sandra (1110) Careless Relaxed Gullible Hard	Karen (0001) Efficient Envious Knowledgeable Affectionate
June (1101) Careless Relaxed Knowledgeable Affectionate	Amanda (0010) Efficient Envious Gullible Hard
Clare (1011) Careless Envious Gullible Affectionate	Joan (0100) Efficient Relaxed Knowledgeable Hard
Moira (0111) Efficient Relaxed Gullible Affectionate	Anne (1000) Careless Envious Knowledgeable Hard

Note. The best examples of each potential family resemblance category were Irene and Suzie, respectively. The values in parentheses show the correspondence between the trait terms and the abstract structure of Fig. 1.

scriptors used in the related set are shown in Table 2. In both the factorial and related conditions the order of the traits within descriptions and the assignment of particular names to sets of descriptions were counterbalanced. Eight subjects served in each of the two main conditions. Prior to receiving the standard sorting instructions, they were asked to read through each description and form an impression of the person.

Experiment 2d. Experiment 2d paralleled 2c except that the sortings were done from memory. Subjects were given the sets of names and traits

TABLE 2
The Related Four-Dimensional Set Used in Experiments 2c and 2d

Irene	Suzie
(1111)	(0000)
Careless	Efficient
Relaxed	Envious
Gullible	Knowledgeable
Affectionate	Hard
Moira	Anne
(1110)	(0001)
Haphazard	Conscientious
Assured	Emotional
Naive	Imaginative
Cold	Sympathetic
Sandra	Karen
(1101)	(0010)
Unpredictable	Exacting
Easygoing	Fretful
Clever	Unsophisticated
Generous	Insensitive
June	Amanda
(1011)	(0100)
Sloppy	Organized
Anxious	Unselfconscious
Ignorant	Inventive
Compassionate	Crafty
Clare	Joan
(0111)	(1000)
Dependable	Negligent
Confident	Jealous
Shallow	Artistic
Considerate	Stingy

Note. The best examples of the potential family resemblance categories were Irene and Suzie, respectively. The values in parentheses show the correspondence between the trait terms and the abstract structure of Fig. 1.

and asked to form an impression of the person described. An approximately 2-min study period was followed by a 3-min recall period where subjects were asked to write down the names of the women and the traits that went with them. Then two additional study and recall periods followed. On the second of these the experimenter supplied the names of all 10 women, and subjects were asked to list the traits associated with each name. After a short unrelated task (about 3 min long), subjects were supplied with the names from the descriptions (but not the traits) and were asked to sort the names into two equal groups in a way that seemed sensible. The idea was that judgments made from memory for properties might be biased toward overall similarity and thereby toward family resemblance. In this Experiment 16 participants were run in the factorial condition and 8 in the related condition.

Results. The results are summarized in Table 3. Sorting strategies were classified as family resemblance, one-dimensional or other (a catchall category for all sorts that were neither family resemblance nor one-dimensional). Only two subjects out of the entire set associated with the different stimulus conditions and procedures sorted by family resemblance.

For the six-dimensional stimuli of Experiment 2a, all but one of the subjects sorted on the basis of a single dimension. A breakdown of the dimensions used in sorting shows 10 people sorted by head shape, 3 by body markings, 3 by smile, 2 by number of antennae, and 1 by tail length. Apparently head shape was a fairly salient dimension.

Adding exemplar-specific information and instructing subjects to use all the information in their decisions had little discernible effect on sorting. All of the subjects produced a one-dimensional sorting. A breakdown of these sortings shows 12 people sorted by number of legs, 2 by head, 1 by spots versus stripes, and 1 by tail length. For the factorial stimuli, 8 people sorted by head shape, 5 by spots versus stripes, and 3 by number of legs.

The same absence of family resemblance sorting was found when verbal descriptions of people were the stimuli (Experiments 2c, 2d). One subject in the factorial condition of Experiment 2c produced a family resemblance sorting. Incidentally, this subject reported basing the sort on the stereotype of the "loving housewife" versus "working woman," especially "bosses." This observation is followed up in a later experiment.

TABLE 3
Frequencies of Family Resemblance (FR), One-dimensional (1-D), and Other (O) Sortings for Experiment 2

Experiment	FR	1-D	O
2a (6-Dimension animals)	1	19	0
2b (4-Dimension animals)			
Exemplar specific	0	16	0
Factorial	0	16	0
2c (Traits			
Factorial	1	4	3
Related	0	3	5
2d (Memory)			
Factorial	0	10	6
Related	0	0	8

The factorial stimulus materials produced 4 one-dimensional sorts, and in each case the subjects wrote that they had sorted on the basis of one dimension. For three subjects, the one dimension was also the first trait listed in each description (*efficient* versus *careless*, *gullible* versus *knowledgeable*, and *affectionate* versus *hard*). The fourth subject sorted on the basis of the second trait (*knowledgeable* versus *gullible*). Of the three subjects who used some other basis for their sort, one reported going by liked versus disliked names, and the other two appeared to have attempted a two-dimensional sort ("A more knowledgeable, hard group" versus "A more relaxed, affectionate group"; "More knowledgeable and efficient sounding names" versus "just efficient").

In the related condition there were 3 one-dimensional sorts. One subject reported sorting on the basis of "More intelligent" versus "Less intelligent" and used the first trait as the cue which was either positively or negatively related to the culture dimension. The second subject also used the first trait, which in this case was either positively or negatively related to agreeableness, and this subject reported sorting on the basis of "What I personally like in women" versus "What I don't like." The third subject explicitly stated sorting on the basis of the third trait, which described being "careless and unpredictable" versus being "organized and ordered."

From the subjects' descriptions of the bases for their sorts, four of the five subjects who used some other basis appeared to have used social desirability as their major cue, distinguishing between "desirable" and "good women" versus "not so desirable," "unpleasant," "evil and selfish," and "not likable women." The fifth subject's sorting structure and comments offered no clues as to what strategy was used.

In the memory condition, 10 subjects produced one-dimensional sorts with the factorial materials. Of these 10 sorts, 5 were based on *knowledgeable* versus *gullible*, and in 4 cases this was the first dimension, whereas in the fifth case this was the last dimension. Three subjects sorted on the basis of *affectionate* versus *hard*, which in 2 cases was the first dimension and in 1 case was the last dimension. One subject used *careless* versus *efficient*, which was the first dimension, and 1 subject used *envious* versus *relaxed*, which was also the first dimension. The fact that all the subjects used either the first or last dimension is in line with typical primacy and recency effects in memory.

It is difficult to classify the six subjects who sorted on some other basis. They could be described as one-dimensional sorts with one misplacement (perhaps because of a memory lapse), but there may be other interpretations of their sorting.

The sorts from memory using the related materials yielded no family resemblance or one-dimensional sorts. All eight subjects sorted on some other basis. Due to an error in recording, only data from seven subjects were available for further examination. Of these seven subjects, four may be characterized as being one-dimensional sorts with one misplacement on the first dimension, one with misplacement on the last dimension, and one on the third dimension. Culture, emotional stability, and conscientiousness were each used by two subjects, and agreeableness by one. Again, it may not be appropriate to characterize these sortings as one-dimensional.

Discussion. The persistent absence of family resemblance sorting was unexpected. The addition of neither exemplar-specific information nor two new dimensions of variation changed the basic pattern of one-dimensional sortings with the animal drawings. Furthermore, instructions to use all of the properties in developing categories had no discernible effect. The use of verbal descriptions and judgments from memory reduced the frequency of one-dimensional sortings but there was no concomitant increase of family resemblance sortings. Instead subjects appeared to assess the trait descriptions either in terms of an evaluative dimension or in terms of certain stereotypes (e.g., "loving housewife" vs "working woman").

Experiment 3

The third experiment explored variations in instructions. Informal discussion with colleagues led to the idea that the naturalness of family resemblance sorting might be tied to the idea of genetic variation and evolutionary adaptations. The rationale is that people's understanding of genetic relationships would provide an underlying basis or cause for family-resemblance-based perceptual similarity and lead to family resemblance sorting.

Experiment 3a. In Experiment 3a subjects were given the drawings of animals used earlier and asked to examine each of the properties and think about how they might represent good adaptations to life on some planet. They were also told that stimuli comprised two groups of genetically related animals and that they should use all of the information in coming up with their categorization decision. A total of 30 subjects was run in this condition, 15 with the animals with factorial properties and 15 with animals having exemplar-specific properties.

Experiment 3b. Experiment 3b was like Experiment 3a in terms of instructions concerning genetic relatedness and adaptive value of properties. Subjects were also told that one group of animals (bugs) lived on the top of the water and that the other half were bottom dwellers. In addition, the prototypes for each potential family resemblance category were specifically labeled as the "best adapted top dweller" and "best adapted bottom dweller," respectively. Twenty subjects participated in Experiment 3b, 9 with the factorial set and 11 with the idiosyncratic set.

Results. The results are simple to describe. All but one of the subjects used a one-dimensional sorting strategy. In Experiment 3a with the factorial stimuli, 6 people sorted by headshape, 6 by body markings, and 3 by number of legs. With the stimuli having exemplar-specific or idiosyncratic properties, 12 people sorted by number of legs, 2 by headshape, and 1 by body markings. In Experiment 3b with the factorial stimuli, 4 people sorted by number of legs, 4 by head shape, and 1 by body markings. For the stimuli with exemplar-specific information, 8 people sorted by number of legs and 2 by head shape. One subject sorted the animals into categories by whether or not they were "rough and tough looking" based, by our guess, on how angular the drawing was as a whole.

The descriptions of the sortings in Experiment 3b suggest that the instructions did have an effect. Almost all of the descriptions mentioned the adaptive value of different properties and these descriptions typically extended beyond the dimension used in sorting. Although this indicates that the subjects took the instructions seriously, the sorting strategies were unchanged.

Discussion. The instructions apparently led subjects to think about all the dimensions in relation to adaptive values, but family resemblance sorting was not observed. Even when the best examples or prototypes were explicitly labeled and specific suggestions for habitat (top and bottom water dwellers) were given, the sortings were exclusively one-dimensional. The descriptions of the sortings suggest that participants sorted by a single dimension but then justified their partitionings in terms of the adaptive value of not only the dimension selected but also, in some cases at least, with respect to the (partially) correlated values on other dimensions that were associated with their partitioning (note in Fig. 1 that sorting on the first dimension results in the value 1 being more frequent in the first category for each of the other dimensions). In brief, again we were unsuccessful in producing family resemblance sorting.

Experiment 4

So far the most consistent result observed is the predominance of one-dimensional sortings. It is possible that the family resemblance sorting is natural, but it must compete with an even easier unidimensional focus. Although almost any set of stimuli can be dichotomized along some dimension (e.g., the presence or absence of some property), the observed preference may be peculiar to our use of the combination of binary-valued dimensions with two equal-sized groups.

Experiment 4a. Experiment 4a used trinary-valued dimensions coupled

DIMENSION

Example	D_1	D_2	D_3	D_4
1	1	1	1	1
2	1	1	1	2
3	1	1	2	1
4	2	1	2	1
5	2	2	1	1
6	1	2	1	2
7	2	2	3	3
8	2	3	2	3
9	3	2	3	2
10	3	3	3	2
11	3	3	2	3
12	3	3	3	3

FIG. 4. Abstract description of the stimulus materials used in Experiment 4.

with the requirement that exactly two categories be created. Consequently there is no straightforward way to use a one-dimensional sorting. The abstract description of the stimulus materials is given in Fig. 4 where family resemblance sorting would consist of partitioning the first six and second six examples into separate groups. The values 1, 2, and 3 designate three distinct values, and the ordering is not to be taken to imply that the values are ordered on a given dimension.

One should also note that for a family resemblance sorting, the value 1 and the value 3 are *sufficient* to determine category membership. Although no one has proposed that sufficient properties serve to organize fuzzy categories, if family resemblance sorting is natural for these materials, then the role of sufficient features would merit closer scrutiny.

It is also important to bear in mind that a family resemblance sorting does not provide unequivocal evidence that the categories were constructed by a family resemblance principle. As was mentioned earlier, a given partitioning can be created and described in a number of alternative ways. For example, a subject may begin with a single dimension (say D_1) and place Examples 1, 2, 3, and 6 in Fig. 4 into one category and Examples 9, 10, 11, and 12 into the other category. This leaves examples 4, 5, 7, and 8. They might then notice that only D_4 has exactly two different values and then place Examples 4 and 5 into the first category and 7 and 8 into the second, thereby producing a sorting that is nominally family resemblance. To distinguish this strategy from the use of a family resemblance principle one needs to look at how the sorting is described. For example, a description of the first category constructed corresponding to "Value 1 on D_1, or Value 2 on D_1 and Value 1 on D_4" would support the strategy just described rather than a true family resemblance principle.

Stimuli and procedures. The stimulus materials consisted of cartoonlike animals which could have one of three values along each of four dimensions. The dimensions were number of legs (4, 8, or 12), body markings (dots, diamond spots, or stripes), head shape (square, circle, or triangle), and tail length (short, medium, or long). The prototype corresponding to the abstract notation 1111 had a square head, 8 legs, medium tail, and diamond spots. The 3333 counterpart had a round head, 4 legs, long tail, and stripes. Note that the Abstract Values 1, 2, and 3 on the dimension of number of legs and tail length are not ordered in terms of increasing quantity or length, respectively.

The same procedure used in the earlier experiments was again followed. Participants were asked to study the animals and produce two equal-sized categories. They were then asked to describe the basis for their partitioning. Altogether 41 subjects participated in the experiment.

Results. Of the 41 sortings, 8 corresponded to a family resemblance partitioning. As will be seen, however, examination of subjects' justifications for their partitionings undermines the idea that these instances represent a true family resemblance principle. The remaining 33 sortings were scattered among 18 distinct types. The most common alternative partitioning was 1111, 1112, 1212, 2211, 3232, and 2233 versus the rest, created by six subjects, and 1121, 2121, 1212, 2211, 2323, and 3323 versus the rest, created by five subjects. It is perhaps notable that the latter partitioning results in as many between-category attribute matches as within-category attributes matches. One other partitioning was used by three subjects, 4 were used by two subjects, and the remaining 11 were unique.

The above somewhat scattered pattern of sorting becomes more clear when one examines subjects' justifications for their sortings. There were five types of sorting strategies: (1) single dimension, (2) primary dimension plus conjunction, (3) primary dimension plus disjunction, (4) coordinate, and (5) chaining.

Although the experiment was set up to prevent unidimensional sorting, it was the most common strategy (used by 14 subjects). For example, a subject might put the drawings with short tails in one category, those with long tails in another category, and then closely scrutinize the figures with medium length tails for any differences in tail length. Inevitably there were tiny differences and a subject could use these to assign two medium length tails to the "long" category and two to the "short" category. Six subjects used tail length to create a unidimensional sort. Four subjects used an analogous strategy with the dimension of number of legs and here the intermediate values (eight legs) were assigned on the basis of slight differences in how spread out the legs appeared to be. The other unidimensional descriptions were produced by 3 subjects who classified the animals of the basis of whether they seemed overall to be "simple" or "complex," and 1 subject partitioned the drawings on the basis of minute differences in how far a drawing was from the top of the card. Simplicity and complexity appeared to be based on the number of line segments, but it is perhaps notable that each of the 3 subjects reporting this strategy created a different partitioning.

The second most frequent strategy, used by 13 subjects, was to start by sorting eight of the drawings on the basis of a single dimension and then to sort the remaining drawings having the third value by means of a conjunction with other values. For example, a category might be described as "four legs, or eight legs and a square head" or as "square head, or triangle head and diamond body markings." The primary dimension was number of legs for 6 subjects, head shape for 5 subjects, and tail length for the other 2 subjects.

Five subjects used a primary dimension plus a disjunction to handle the remaining four drawings. For example, a category might be described as "triangle-shaped head or more than eight legs." The primary dimension was head shape for four subjects and body markings for the remaining subject.

Three subjects created coordinate categories. For example, each category might contain two drawings of each head type with the selection of which two depending on matching values on a second dimension (e.g., tail length). This type of sorting strategy has been reported before (e.g., Handel & Rhodes, 1980).

Two subjects reported a chaining strategy. For example, one subject started with the two most similar (round head, stripes, four legs) animals, then added two more by dropping the body marking requirement (they had round heads and four legs) and then completed the category by requiring that the drawing have stripes. One of the two subjects (the one just described) had created a family resemblance sorting and this description seems consistent with a family resemblance construction principle (as would several other strategies, such as grouping on the basis of overall similarity).

One description began with a primary dimension of body marking and then combined it with an and/or description of a second dimension. The remaining three descriptions were vague and, to the extent we could understand them, not clearly consistent with the partitioning that had been created. None of these vague descriptions was associated with a family resemblance partitioning.

When one examines the strategy reports it becomes clear that almost all of the family resemblance sortings were not based on a family resemblance construction principle. Of the eight family resemblance sortings, five were described as primary dimension plus conjunction, one as unidimensional (based on tail length), one as primary dimension plus disjunction, and one in terms of chaining. Only the last strategy seems consistent with a family resemblance construction principle.

Experiment 4b. In Experiment 4a, despite the use of trinary values, some subjects created unidimensional sorts by attending to minute differences in intermediate values. To further reduce the possibility of unidimensional sorts, Experiment 4b used values that would be more difficult to combine. The abstract structure shown in Fig. 4 was also used in this experiment, but the stimuli were cartoonlike drawings of men that differed on the dimensions of color of body stripes (green, red, blue), orientation of body stripes (diagonal, horizontal, vertical), body shape (triangle, square, circle), and type of hat (fedora, top, pointed). As in the previous experiment subjects were instructed to place the drawings into two equal-size categories. Eighteen subjects participated in the experiment.

Results. These stimuli successfully eliminated the use of one-dimensional sorts, but still only 4 out of 18 participants sorted according to family resemblance, and none of these four descriptions was consistent with a family resemblance explanation. Three of these people described their partitioning by a conjunction of disjunctive values within a dimension (e.g., "red or blue color *and* triangle or square body shape") and the fourth used a single disjunction ("*green* or *red and fedora hat*"). The 14 other participants all used a primary dimension plus either a conjunction or a disjunction as was described in the introduction to Experiment 4a. Of these 14 descriptions, 10 were disjunctive and can further be broken down as follows: 2 were simple disjunctions, 4 were single disjunctions with a conjunction used with it (e.g., *cone hat* or *fedora hat and square body*), and 4 were disjunctions that began with a major classifier and then listed the exceptions in a nondiscriminating manner (e.g., [diagonal or vertical lines] versus [horizontal or diagonal lines]). The outer 4 descriptions consisted of 2 conjunctive rules involving higher order subjective dimensions (e.g., tall hat and athletic body) and 2 involving a matching or coordinate strategy (e.g., each category has two of each color and three of one shape, two of another, and one of a third) where two categories are created having equal properties (see also Imai, 1966).

Discussion. The procedure of using trinary values was at least moderately successful in preventing unidimensional sortings, but it did not lead to family resemblance sorting. Despite the use of trinary values, about a third of the subjects in Experiment 4a produced a unidimensional sorting by combining or attending to minute differences in intermediate values, or in a few cases by creating a new dimension based on overall complexity. The large majority of the remaining strategies also reflect a primary focus on a single dimension. A single dimension was commonly used to partition eight of the drawings and the remaining four were then partitioned on the basis of their value on a second dimension. Experiment 4b effectively eliminated unidimensional sorting, but it did not otherwise change the pattern of results. The predominant strategy was to use a single dimension as the major classifier and then add either a conjunction or disjunction of values on a second dimension to create two equal-sized categories.

Across both experiments, approximately 20% of the partitionings created corresponded to a family resemblance sorting. The descriptions of these categorizations, however, undermine the idea that a family resemblance principle was used to construct the categories. Only a single subject provided a description that fits the notion of family resemblance. The remaining subjects mostly focused on a single dimension and used a value on another dimension to complete the categories, and the family resemblance partitioning appears to be an incidental byproduct of this strategy.

Discussion of Experiments 1–4

From the point of view of our interest in family resemblance sorting, the results of the first four experiments are not very encouraging. Experiments 1, 2, and 3 employed a variety of category structures and stimulus structures designed to induce family resemblance partitioning, but almost all subjects preferred a unidimensional sorting strategy. The fourth experiment was designed to prevent one-dimensional sorting, but we still observed that the sortings were organized around a single primary dimension. Even though family resemblance partitionings appeared with a modest frequency, the associated descriptions undermine the idea that the categories were constructed by a family resemblance principle.

The above results do not offer any clear suggestions regarding the most promising research strategy concerning family resemblance sorting. It certainly is possible to think of numerous variations on stimulus properties, category structures, and instructions that might be tried in order to find family resemblance sorting. One obvious criticism of the preceding studies is that adequate care was not taken to ensure that the dimensions were equally salient. Therefore, one could argue that the most appropriate follow-up work would involve a preliminary similarity scaling study to select equally salient component dimensions. On the other hand, an interpretation in terms of salience is both post hoc and inconsistent with the observation that subjects used minute differences to create unidimensional sortings in Experiment 4a. In addition, it is unrealistic to think that the dimensions or properties of real world entities are equally salient. In any event, we have not pursued this line of research, in large part because of experiments conducted by Imai and Garner (1965, 1968) which we became aware of only when the present project was well under way. Imai and Garner did take the care to do appropriate similarity scaling before running free and constrained classification (sorting) conditions. Like us, however, they found that sortings were not determined by simple overall similarity and difference relations among the stimuli. That is, subjects did not maximize within-category similarity relative to between-category similarity. Consider, for example, the situation where the stimuli can be described in terms of their values on two dimensions and where the stimulus set consists of the following four stimuli: (1,2) (1,3) (2,1) (3,1). On the basis of overall similarity, subjects should prefer to create two categories consisting of the first two and second two stimuli. Instead, Imai and Garner (1968) found that people preferred three categories based on values along one of the dimensions, even though there is other evidence which suggests that people prefer to create equal-sized groups (Handel & Imai, 1972).

Although the work of Imai and Garner does not encourage the idea that when salience is properly controlled sorting by overall similarity emerges, we do not mean to suggest that testing the category structures employed in our studies with controlled salience would be a waste of time. The Imai and Garner studies did not use four or more dimensions, and it is possible that sorting by overall similarity would emerge with more complex stimuli. In addition, there are numerous other avenues concerning stimulus properties that could justifiably be explored.

A second response to the first four studies is to suggest that some alternative measures of family resemblance be developed and evaluated. Perhaps there needs to be some mixture of properties that are necessary but not sufficient, sufficient but not necessary, characteristic properties, and exemplar-specific information, in order for family resemblance sorting to be natural. We have not, however, followed this line of research either. Instead, we have been led to reevaluate our original assumption that categories can be defined simply in terms of matching and mismatching properties.

The Rosch and Mervis (1975) measure of family resemblance in terms of matching and mismatching properties is a convenient simplification and served their purposes rather nicely. This approach implicitly assumes, however, that the properties listed by subjects can be treated as primitives which can be independently added together. But as Armstrong, Gleitman, and Gleitman (1983) noted, the fact is that most concepts are not a simple sum of features whatever they might be. All the features that are characteristic of a bird do not make it a bird—unless these properties are held together in a bird structure. The properties typically listed for the concept bird—laying eggs, flying, having wings, having hollow bones, building nests in trees, having feathers, and singing—each represents a complex concept with both internal structure and an external structure based on interproperty relationships. Building nests is linked to laying eggs, and building them in trees poses logistic problems whose solution involves other properties such as wings, hollow bones, and flying. Therefore, it seems that the properties associated with many lexical concepts are far from being an independent set of nondecomposable primitives.

One idea growing out of this analysis is that the key to family resemblance sorting is not attributes but rather relational properties having attributes as arguments. Presumably it is relational properties that make a bird more than a list of bird properties. Perhaps people will find it natural to sort by family resemblance when it is defined in terms of matching and mismatching *relational* properties (such as inside, on top of, brighter than, and so on). We did some preliminary work along this line, but the results were not encouraging. First of all, it is difficult to set up a family resemblance structure on relational properties without inadvertently introducing defining features based on something other than the intended set of relational properties. Even with this difficulty eliminated, we still did not observe family resemblance sorting. People preferred to partition the categories on the basis of a single relational property. Given evidence that relational properties per se did not promote family resemblance sorting, our attention turned to the nonindependence of attributes induced by interproperty linkages.

Our main idea is that a web of interproperty relationships links properties to each other and provides conceptual cohesiveness to concepts like *bird* (see Murphy & Medin, 1985, for an amplification of this argument). If this is so, then it may be a mistake to look for family resemblance sorting in contexts where the component properties bear little or no conceptual relationship to each other.

The remainder of this paper contains two sets of experiments. The first is concerned with correlated attributes as a structural principle in categorization (see Medin, 1983; Mervis & Rosch, 1981; Rosch, 1975, for more detailed presentations). Certain attributes in nature tend to cooccur—e.g., animals with feathers are likely to have wings and beaks, whereas animals with fur are unlikely to have either wings or beaks—and people may prefer to construct categories that follow these natural clusters of correlated attributes. Experiment 5 shows that people may sort by correlated attributes but given alternative sets of correlated attributes they prefer to sort by those for which a (causal) linkage is salient or can readily

DIMENSION

Example	D_1	D_2	D_3	D_4	D_5
1	1	2	1	2	1
2	1	2	2	2	2
3	1	1	1	1	1
4	1	1	2	1	2
5	2	1	1	1	1
6	2	1	2	1	2
7	2	2	1	2	1
8	2	2	2	2	2

FIG. 5. Abstract description of the stimulus materials used in Experiment 5.

be developed. The second set of experiments returns to family resemblance sorting where attempts are made to make interproperty relationships salient.

EXPERIMENT 5

The correlated attributes principle is an attractive one, especially when one considers the alternative possibility that the attributes of the world are randomly spread across entities. One problem with the correlated attributes notion is that it is not clear how the correct ones get picked out (see Keil, 1981, for an elaboration of this point). Medin, Altom, Edelson, and Freko (1982) found in experiments with artificial categories that people are sensitive to feature correlations but, notably, during debriefing, the participants frequently offered reasons for *why* the correlation was present. The purpose of Experiment 5 was to see if the readiness with which causal linkages could be established between attributes influenced how subjects would sort examples into categories.

An abstract description of the stimulus materials is shown in Fig. 5. The values on the second and fourth dimensions are perfectly correlated as are the values on the third and fifth dimensions. Although the examples could readily be sorted into two equal groups on the basis of their value on the first dimension, our interest was in sorting on the basis of either of the two sets of correlated attributes.

Two main conditions were run. In Experiment 5a the stimuli were descriptions of hypothetical diseases and the values consisted of symptoms. For a given subject, one pair of correlated symptoms could be more readily linked or related (e.g., dizziness and earache) than the other pair of symptoms (e.g., sore throat and skin rash). In Experiment 5b the stimulus structure was the same but materials consisted of descriptions of animals. For a given subject one pair of correlated properties could be more readily linked by some general notion of adaptation (e.g., brightly colored and poisonous) than the other pair of correlated properties (e.g., long tailed and slow).

In short, there were three basic ways the stimulus materials could be sorted: (1) using the first (single) dimension, (2) using the correlated properties that were easily linked, and (3) using the correlated properties that were less easily linked. Any preference between the last two strategies would suggest that the ease of conceptually linking properties influences category partitioning.

Experiment 5a

Materials and procedure. Each example consisted of a set of five symptoms typed on an index card. The first symptom was always either loss of sleep or stiff muscles. The other eight possible symptoms always consisted of two pairs that could readily be related and two pairs that were less easily related. Across subjects, each symptom was in a related pair half the time and in an unrelated pair half the time. Based on our informal guesses the related pairs were sore throat, white cell count up; dizziness, earache; skin rash, itchiness; and weight gain, high blood pressure. The unrelated pairs were dizziness, weight gain; earache, high blood pressure; white cell count up, itchiness; and sore throat, skin rash. The related and unrelated pairs were always selected such that the set of eight symptoms was exhausted and conformed to the abstract notation in Fig. 5. For half the subjects the related symptoms appeared in the second and fourth position, and for half they appeared in the third and fifth positions. Altogether 38 subjects participated in the study. Subjects were told to study the examples and put them into two equal groups in a way that made sense.

Experiment 5b

Materials and procedure. This study was exactly analogous to Experiment 5a, the only difference being that the stimulus materials consisted of verbal descriptions of animals. The first dimension was always bears twin young or bears quadruplets. The related pairs were lives in a tree, long tail; slow, armored; brightly colored, poisonous; and sleeps in the day, well-hidden home. The unrelated pairings were slow, long tail; armored, lives in a tree; brightly colored, sleeps in the day; and well-hidden home, poisonous. A total of 24 subjects participated in the study.

Results

People preferred to sort by correlated attributes that could easily be conceptually linked. In Experiment 5a, out of 38 subjects, 21 sorted by the linked correlated attributes, 7 sorted by the unrelated correlated attributes, 4 sorted by the first dimension, 2 set up two subcategories, and 2 sorted in a pattern we did not understand. The difference between related and unrelated pairs was statistically reliable (χ^2 (28) = 7.0, $p < .05$). In Experiment 5b, out of 24 subjects 16 sorted by the related pair of correlated attributes, 5 by the unrelated pair, 2 used the first dimension, and 1 set up two subcategories. Again the difference between related and unrelated pair sorting is statistically reliable ($\chi^2(21) = 5.76, p < .05$).

Discussion

People showed a strong tendency to cluster on the basis of correlated attributes for which a causal or explanatory link could readily be made. Furthermore, subjects mentioned such linkages to justify their sorting. For example, a subject might say that an ear infection could produce both earaches and dizziness (by disturbing the vestibulary organ). This study shows then that interproperty relationships influence category construction. The fact that very few subjects employed a one-dimensional sorting undermines the idea that participants always choose the "easiest" or most simple sorting strategy.

EXPERIMENT 6 AND 7

If interproperty relationships help to organize categories, then many concepts with a family resemblance attribute structure may derive their coherence in part from these underlying relationships among attributes. Although this claim is difficult to establish for lexical concepts, it suggests that one ought to be able to obtain family resemblance sorting when the attribute structure maps onto some unifying theme. Experiments 6 and 7 test the idea that salient interproperty relationships are a sufficient condition to induce family resemblance sorting.

Experiment 6

Experiment 6 was designed to bring out *interproperty* relationships among trait descriptors. The trait descriptors were related either to a basic personality dimension (*introversion*) or to an occupational stereotype.

Stimuli and procedures. Experiment 6 used stimulus descriptions consisting of a first name and four trait descriptors as in Experiment 2. The abstract design corresponded exactly to that shown in Fig. 1, and again subjects were asked to sort the examples into two equal-sized groups.

In Experiment 6a, the trait descriptors were all relevant to introversion and extroversion. The traits used in these person descriptions were selected from Cantor and Mischel (1977, Table 1, p. 42) who reported the rated relationship (high, moderate, unrelated) between 48 traits and extroversion and introversion. The prototypical extrovert was described with four traits rated as highly or moderately related to extroversion, whereas the prototypical introvert was described with four traits rated as highly or moderately related to introversion. As in Experiment 2, both a factorial and a related set of descriptors were employed as listed in Tables 5 and 6.

Half of the factorial and related sets had one random assignment of names to descriptions and the other half of these sets had another, different, random assignment of names to the descriptions. In addition, the order of the four trait descriptors within each description was

TABLE 4
The Factorial Extrovert and Introvert Set of materials Used in Experiment 6

Extroverts	Introverts
Carrie	Miranda
(1111)	(0000)
Outgoing	Sad
Energetic	Self-conscious
Entertaining	Inhibited
Bold	Day dreamer
Susan	Dawn
(1110)	(0001)
Outgoing	Sad
Energetic	Self-conscious
Entertaining	Inhibited
Day dreamer	Bold
Olivia	Charlotte
(1101)	(0010)
Outgoing	Sad
Energetic	Self-conscious
Inhibited	Entertaining
Bold	Day dreamer
Felicity	Vanessa
(1011)	(0100)
Outgoing	Sad
Self-conscious	Energetic
Entertaining	Inhibited
Bold	Day dreamer
Wendy	Rosemary
(1000)	(1000)
Sad	Outgoing
Energetic	Self-conscious
Entertaining	Inhibited
Bold	Day dreamer

Note. The best examples of the potential family resemblance categories were Carrie and Miranda, respectively. The values in parentheses show the correspondence between the trait terms and true abstract structure of Fig. 1.

varied systematically such that each trait appeared in each position. This resulted in four versions of both the factorial and related descriptions.

Experiment 6b was identical to the factorial condition of Experiment 6a, but participants were given the additional information that 5 of the women were extroverts and 5 introverts. Thirteen subjects participated in this condition.

Experiment 6c used the person descriptions that had been employed in Experiment 2, where the four dimensions corresponded to four basic dimensions of personality. The participants were given additional instruction to the effect that the women were either "spies" or "truckstop waitresses." These category labels were chosen because they seemed to be two occupations which, in terms of everyday stereotypes as (inaccurately) portrayed on

TABLE 5
The Related Extroverts and Introverts Set of Materials Used in Experiment 6

Extroverts	Introverts
Carrie	Miranda
(1111)	(0000)
Outgoing	Sad
Energetic	Self-conscious
Entertaining	Inhibited
Bold	Day dreamer
Susan	Dawn
(1110)	(0001)
Impulsive	Withdrawn
Exuberant	Solemn
Boisterous	Soft-spoken
Meditative	Spirited
Olivia	Charlotte
(1101)	(0010)
Vigorous	Timid
Active	Lonesome
Courteous	Vigorous
Talkative	Oversensitive
Felicity	Vanessa
(1011)	(0100)
Friendly	Studious
Hesitant	Daring
Venturesome	Subtle
Lively	Reserved
Wendy	Rosemary
(0111)	(1000)
Bashful	Eager
Impulsive	Quiet
Ambitious	Discreet
Dominating	Melancholy

Note. The best examples of the potential family resemblance categories were Carrie and Miranda, respectively. The values in parentheses show the correspondence between the trait terms and the abstract structure of Fig. 1.

television, summarized the combination of personal dimensions represented by each of the two best examples. The related materials were used for 9 subjects and the factorial materials for an additional 19 subjects.

Results. A substantial number of family resemblance sortings were obtained. The relevant data are summarized in Table 6. When the descriptors mapped onto introversion and extroversion, 25% of the subjects in the factorial condition and over 70% of the subjects in the related condition sorted by family resemblance. Although few subjects used the terms extrovert and introvert in their explanations, they did mention superordinate concepts that were clearly very similar (e.g., "fun to be with" versus "not the sort of person to take to a party"). Providing the explicit label of introvert and extrovert (Experiment 6b) increased the proportion of family resemblance sortings associated with the factorial stimuli.

Finally, supplying the occupational stereotypes was, at best, only modestly successful in inducing family resemblance sorting. Only 4 of 19 subjects used a family resemblance sorting for the factorial stimuli compared with 1 out of 8 for the corresponding condition in Experiment 2c. For the related stimuli there were no family resemblance sortings. For the factorial stimuli 4 of the 10 one-dimensional sorts were on the basis of the first trait (for 2 of them the first trait was *efficient* versus *careless;* for the others it was *knowledgeable* versus *gullible,* and *affectionate* versus *hard*), 3 used the last trait (for 2 of them this was *careless* versus *efficient,* for 1 it was *knowledgeable versus gullible*), 2 used the second trait (both *gullible* versus *knowledgeable*), and 1 used the third trait (also *gullible* versus *knowledgeable*).

Five subjects in the *occupational label,* factorial condition sorted on some other basis. Three of these subjects formed two groups with one prototype in each and two subjects put the prototypes in the same group. The subjects' explanations for their sorts yielded no pattern or insight into what was going on here beyond a clear indication from most of them that they were considering most, if not all, of the traits in their sorting strategy.

In the *occupational label,* related condition there was only 1 one-dimensional sort based on the first trait, which was either positively or negatively related to agreeableness, and the subject described the groups as "shifty" versus "really caring for people." The remaining eight subjects all sorted on some other basis, and no clear indication of their guiding strategy emerged from examining their explanations.

Discussion. When the trait descriptors mapped onto the abstract concepts of introversion and extroversion, the majority of subjects produced family resemblance sortings. Our interpretation of this result is that this

TABLE 6
Frequencies of Family Resemblance (FR), One-dimensional (1D), and Other (O) Sorts in Experiment 6

Experiment	Factorial			Related		
	FR	1D	O	FR	1D	O
6a	3	7	2	10	0	4
6b	8	5	0	—	—	—
6c	4	10	5	0	1	8

theme provided interproperty linkages that led subjects to integrate the descriptors into an overall impression. When unrelated dimensions were used, but an occupational stereotype was suggested that seemed to correspond to the prototypes, there was some tendency to create family resemblance sortings, but this was true for a minority of subjects in the factorial condition and none of the subjects in the related condition. Of course, we do not know how well the subjects' stereotypes mapped onto our guesses, and any differences would produce a different sorting. One should, therefore, probably be cautious in drawing inferences from the occupational stereotype results. Despite our reservations about Experiment 6c, Experiments 6a and 6b show that making interproperty relationships salient is sufficient to lead subjects to abandon a unidimensional sorting strategy in favor of family resemblance partitionings.

Experiment 7

Experiment 7 was similar to Experiment 6 except that the stimuli were cartoonlike animal drawings. Interproperty coding was encouraged by selecting properties that could be linked both to each other and to a higher order property (flying ability).

Stimuli and procedures. The stimulus materials for Experiment 7 were cartoonlike animal drawings. The drawings varied along five dimensions: body size (large or small), foot type (paws or webbed feet), body covering (a few feathers or a few hairs), ear size (large or small), and mouth type (beak or mouth). One prototype consisted of an animal with a small body, webbed feet, a few feathers, small ears, and a beak, whereas the other prototype had opposite values on each of these dimensions. Five examples were generated from each prototype by switching the value of one of the five dimensions (e.g., from the 11111 prototype, the examples 01111, 10111, 11011, 11101, and 11110 were generated).

The basic task was to sort the 10 examples into two equal-sized groups. The pattern of correlated attributes was designed such that one prototype contained attributes that might be associated with birds and flying. Participants were told that half the animals were fliers and half were nonfliers. The hint about flying was designed to allow subjects to relate the attributes to each other. Since it was quite possible that the effect of this additional information would be that all sortings would be done on the basis of feathers versus hair, an attempt was made to reduce the salience of the body covering dimension by including only a few feathers or a few hairs in the drawings. Twenty-four subjects ran in the study.

Results. The information that half of the animals were fliers was effective. Although the majority (14) of the sortings were unidimensional, 9 subjects sorted by family resemblance. One other subject's sorting was not unidimensional, but it could not be readily classified. Even the unidimensional sortings reflected an influence of the information about flying —8 people classified by hair versus feathers and an additional four by body size (the other 2 were based on type of foot). The rationales for these sortings were typically linked to flying. For example, one subject who sorted by body size reasoned that the greater mass of the large animals would require more food energy in order for them to fly. The explanations of family resemblance sortings were also linked to flying. Subjects mentioned characteristics that would facilitate flying and in some cases mentioned explicitly that they had picked out the "best example" and then categorized the other animals on the basis of how well they matched the best example. Finally, several subjects included analogies to real world animals in explaining their sortings.

The concern that the information about flying might lead subjects to focus exclusively on body covering was only partially justified. A plurality of participants produced family resemblance sortings and their justifications of their partitioning involved integrated information across component dimensions.

Discussion of Experiments 6 and 7

For the first time we were able to observe family resemblance sorting supported by explanations that indicated that participants had not simply responded on the basis of one or two of the components. The information about flying, the mapping of trait descriptors onto the abstract notions of introversion and extroversion, and to a modest extent the suggestion of appropriate occupational stereotypes led to family resemblance sortings. Presumably this is attributable to both interproperty relationships being made salient and to the availability of a higher order linkage to a theme (e.g., flying). These results are in sharp contrast to the first four experiments, which either did not observe family resemblance sorting or were supported by justifications which were inconsistent with a family resemblance principle of category construction.

These data suggest that one route to family resemblance sorting is provided by information that brings out interproperty relationships. When these interproperty relations allow integration across component dimensions, then family resemblance sorting becomes natural.

GENERAL DISCUSSION

The general pattern of our results is easy to describe. When examples are constructed which have the potential to be clustered according to a family resemblance principle defined in terms of matching and mismatching properties, family resemblance sorting is very far from being automatic. In fact, it is only a slight exaggeration to say that it did not appear at all in the first four experiments which used a variety of stimulus materials and instructions. Instead, people found unidimensional sortings to be more natural. There was no evidence that subjects used overall similarity or some other means of integrating component information to construct categories. Although some variation on stimulus structure, category structure, or procedure could conceivably lead to the emergence of family resemblance sorting, we can state with confidence that it is considerably more difficult to observe family resemblance sorting than we initially imagined.

These initial studies led us to question the generality of the idea that attributes or features can be treated as independent primitives or that categories are little more than a collection of features. At least for the types of attributes that derive from subject property listings, it seems likely that components of concepts are linked by a variety of interproperty relationships. The fifth experiment showed that people prefer sorting by two correlated attributes to sorting on the basis of a single dimension and, more significantly perhaps, that given two sets of correlated attributes, people prefer to sort in terms of the correlated attributes for which a causal linkage is readily perceived. The final two experiments showed that when interproperty relationships suggest that integrating across components is appropriate, family resemblance sortings become fairly frequent, and such sortings are described in a way that is compatible with a family resemblance principle of category construction.

Our results provide a new perspective on fuzzy categories and family resemblance sorting: they suggest that *the apparent use of family resemblance rules may be masking the use of a deeper principle that some core factor or cause is present which probabilistically leads to surface structure (family resemblance) features*. This idea is generally consistent with approaches to concepts which draw a distinction between core and identification features. For example, the concept *man* can be defined in terms of core features like "adult, male, human," and identification features like hair length, the presence of a beard or mustache, or characteristic gait may be used to decide that some person is a man. Most importantly, we wish to argue that core and identification properties are intimately linked and that the core properties may give rise to the identification properties. For example, being male is partly a matter of hormones which directly influence physical attributes such as facial hair (see E. E. Smith, Medin, & Rips, 1984, for a more extensive development of this approach to core and identification properties). This underlying concept or core provides the basis for interproperty linkages and helps to structure what would otherwise be independent properties. Family resemblance categories may be organized not so much in terms of their surface features or properties but in terms of a deeper underlying concept that may give rise to them. For example, characteristic properties for the category *bird*, such as having hollow bones, feathers, wings, building nests in trees, and even singing, can all be seen as adaptations to allow for consequences of flying.

Relation to Other Work

Our initial failures to observe family resemblance sorting in many ways parallel work examining relationships between different types of category structures and ease of learning. There is no evidence that switching to a learning paradigm would have led to the emergence of family resemblance category structures as being natural. For example, particular attention has been directed at linearly separable versus not linearly separable categories. Linearly separable categories are categories that can be partitioned by summing the component information and both a family resemblance metric and, more generally, probabilistic view models predict that linearly separable categories should be easier to learn than categories that are not linearly separable. In a series of four experiments, however, Medin and Schwanenflugel (1981) found no evidence that linearly separable categories were easier to learn than categories that were not linearly separable. The parallel between the present results and the results of experimental work on linear separability is extended by some recent work in our laboratory that introduced interproperty relations to the linear separability paradigm. Wattenmaker, Dewey, Murphy, and Medin (1986) found that the presentation of themes that encouraged integrating information across dimensions greatly facilitated the acquisition of linearly separable categories.

For example, in one study the descriptions were properties of objects, and the categories were structured such that the typical attributes for one category would all be desirable properties if one were to search for a substitute for a hammer (e.g., flat surface, easy to grasp). In one condition subjects were given the notion of hammer substitutes and in another condition they were not. When the hint about hammer substitutes was given, linearly separable categories were easier to learn than nonlinearly separable categories. Although participants did use a summation strategy under these conditions, when themes were presented that highlighted conjunctions of features and mapped onto nonlinearly separable categories, the reverse result was observed. This interaction between category structure and knowledge was interpreted in terms of the forms of interproperty codings that were induced. In the special case where conceptual knowledge induces a summing strategy and properties are encoded relative to an integrated theme, linearly separable categories will be easy to learn. However, a multiplicity of interproperty codings is possible, and consequently, in many instances this special case will not hold. Conceptual knowledge can promote several different types of relations between features other than a summing strategy.

Exactly the same conclusion may hold for sorting studies. That is, there may be no general answer to the question of which partitioning of some abstract structure of a set of examples is most natural. Rather, naturalness may hinge on the types of interproperty relationships that are made salient when the abstract structure is realized in concrete examples. We were successful in producing family resemblance sorting only where interproperty relationships were made salient and where these relationships were compatible with or encouraged integrating information across components.

One clear result of the present experiments is that in the absence of interproperty relations, people tend to focus on a single dimension in constructing categories. In fact, people were quite ingenious in constructing unidimensional sortings in the fourth experiment that used trinary values and required a binary partitioning. These results suggest that in many circumstances people will prefer to use simple rules based on single dimensions over rules that preserve cue validity information from several dimensions. The tendency to use simple unidimensional rules over similarity-based family resemblance structures in category formation is consistent with the work of Martin and Caramazza (1980). They found that subjects attempted to learn categories with a family resemblance structure by constructing well-formulated hypotheses that often involved single dimensions. The focus on single properties to the exclusion of other relevant features has also been observed in language learning. For example, children are often observed to overextend their use of a word based on one salient property (e.g., the use of moon to refer to everything round).[1] On the other hand, in Experiment 5 people preferred to sort on the basis of correlated properties rather than on the basis of a single dimension. Therefore, it is not the case that subjects necessarily will classify in a unidimensional manner whenever such a strategy is available.

Nonanalytic Family Resemblance?

Based on the work of Rosch, Mervis, and others and the near-ubiquitous typicality effects obtained with other procedures, it is difficult to escape the intuition that family resemblance sorting *ought* to be natural. We believe that our idea that conceptual knowledge makes interproperty relationships salient and serves to organize family resemblance categories goes a long way toward explaining these intuitions. One remaining possibility is that family resemblance categories are usually learned in circumstances that make a family resemblance structure more natural.

Family resemblance structures may become important when the learner is not being analytic. That is, category construction based on nonanalytic learning may operate in terms of overall similarity. For example, it has often been argued that there is a developmental shift from holistic representations based on overall similarity to analytic representations (e.g., Kemler & Smith, 1978; Smith & Kemler-Nelson, 1984; Vy-

[1] Greg Murphy suggested this parallel.

gotsky, 1962). Holistic representations preserve characteristic features along several dimensions rather than focusing on a single defining feature. In support of this idea Kemler-Nelson (1984) found that 5-year-olds had more difficulty learning categories defined by a single feature than categories defined by overall similarity, but this was not true for 10-year-olds. Although responding in terms of overall similarity canot be directly equated with a family resemblance principle, these results do show that young children often do not rely on a single defining dimension.

In some related research, Keil and Batterman (1984) have found evidence for a developmental shift in the representation of word meanings from collections of characteristic features to more nearly defining features. For example, kindergartners preferred a description of an island as a place that is warm and has coconut trees, palm trees, and girls with flowers in their hair, even though the technical definition of an island was violated in the description. In contrast, fourth graders preferred a description that had none of the characteristic features of an island, but contained the crucial information that the land was surrounded by water on all sides. These results appear to parallel Kemler–Nelson's findings.

It is likely that responding in terms of overall similarity is not only characteristic of young children but also holds when adults are nonanalytic. In support of this possibility, Kemler-Nelson encouraged adult subjects to be either analytic or nonanalytic in learning categories that could be partitioned by either a defining feature or in terms of overall similarity structure. Classification of novel patterns following learning revealed that while the categorization decisions of subjects in the analytic group were based on the defining feature, the subjects in the nonanalytic conditions preferred to categorize by overall similarity. These results raise the possibility that family resemblance structures are more important in circumstances where adult subjects are learning concepts less analytically.

A second possibility with regard to standard typicality effects is that normally examples of concepts may be encountered one at a time rather than simultaneously, and category construction might be based on one exemplar reminding the learner of similar exemplars. If category formation is governed by similarity to stored exemplars, then dissimilar exemplars (those with atypical properties) may not be directly incorporated into the category representation.

The above discussion suggests that the use of family resemblance information might arise from nonanalytic processing, whereas the results of the current experiments imply that family resemblance partitionings emerge when subjects are able to apply higher order conceptual processes to illuminate interproperty relations. Although there is no developmental evidence on family resemblance sorting in tasks similar to ours, the potential commonality is intriguing in that what appears to be a psychologically primitive mechanism (nonanalytic processing and responding to overall similarity) and a higher order cognitive process (the conceptually motivated encoding of interproperty relations) can lead to sensitivity to the same type of information. In the present setting, these two different mechanisms, overall similarity and conceptually motivated explanations, could conceivably lead to the same output (family resemblance partitionings). This overlap might illustrate a general tendency for theory-laden cognitions to preserve some of the constraints associated with more basic learning and memorial mechanisms (see Medin & Wattenmaker, 1986, for an extended discussion of this speculation, and Newport, 1984, for some related ideas). Note that the use of overall similarity includes the indirect encoding of complex interproperty relations. Until the corresponding research is conducted, however, these potential parallels remain as speculations.

Summary and Conclusions

Our emphasis on conceptual knowledge reflects a need to broaden the study of concepts. Instead of focusing on features as independent entities, it is necessary to consider the types of relations, operations, and transformations that features participate in. Murphy and Medin (1985) have argued that categorization models that view concepts as a collection of attributes and categorization as based solely on attribute matching are insufficient to explain conceptual coherence. Rather, concepts should be viewed as embedded in theories and are coherent to the extent they fit people's background knowledge or naive theories about the world. Furthermore, the developmental research reviewed by Murphy and Medin (1985) supports the claim that theories help structure categories in even very young children.

In concluding, we would like to emphasize the notion that the constituent components of concepts are not irreducible primitives but rather participate in a variety of interproperty relationships. These relational properties may be partly perceptual, but they may also involve conceptual relationships associated with the theories people develop to relate the concepts in a domain and help structure the attributes internal to a concept. Natural categories may be structured not so much in terms of numbers of characteristic properties as in terms of the web of relationships in which these properties participate. That is, when people build family resemblance categories they may not only use bricks, but also mortar.

REFERENCES

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

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

Gati, I., & Tversky, A. (1984). Weighting common and distinctive features in perceptual and conceptual judgments. *Cognitive Psychology*, **16**, 341–370.

Goldberg, L. G. (1982). From Ace to Zombie: Some explorations in the language of personality. In C. D. Spieldberger & J. N. Butcher (Eds.), *Advances in personality assessment* (Vol. 1, pp. 203–234). Hillsdale, NJ: Erlbaum.

Handel, S., & Imai, S. (1972). The free classification of analyzable and unanalyzable stimuli. *Perception & Psychophysics*, **12**(1B), 108–116.

Imai, S. (1966). Classification of sets of stimuli with different stimulus characteristics and numerical properties. *Perception & Psychophysics*, **1**, 48–54.

Imai, S., & Garner, W. R. (1965). Discriminability and preference for attributes in free and constrained classification. *Journal of Experimental Psychology*, **69**, 596–608.

Imai, S., & Garner, W. R. (1968). Structure in perceptual classification. *Psychonomic Monograph Supplements*, **2**(9, Whole No. 25).

Katz, J. J., & Postal, P. M. (1964). *An integrated theory of linguistic descriptions*. Cambridge, MA: MIT Press.

Keil, F. C. (1981). Constraints on knowledge and cognitive development. *Psychological Review*, **88**, 197–227.

Keil, F. C., & Batterman, N. (1984). A characteristic-to-defining shift in the development of word meaning. *Journal of Verbal Learning and Verbal Behavior*, **23**, 221–236.

Kemler, D. G. (1983). Holistic and analytic modes in perceptual and cognitive development. In T. Tighe & B. E. Shepp (Eds.), *Perception, cognition, and development: Interactional analyses* (pp. 77–102). Hillsdale, NJ: Erlbaum.

Kemler, D. G., & Smith, L. B. (1978). Is there a developmental trend from integrality to separability in perception? *Journal of Experimental Child Psychology*, **26**, 460–507.

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

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

Medin, D. L. (1983). Structural principles of categorization. In T. Shepp & T. Tighe (Eds.), *Interaction: Perception, development and cognition* (pp. 203–230). Hillsdale, NJ: Erlbaum.

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

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

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

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

Medin, D. L., & Wattenmaker, W. D. (1986). Category cohesiveness, theories, and cognitive archeology. In U. Neisser (Ed.), *Categories reconsidered: The ecological and intellectual bases of categories*. London/New York: Cambridge Univ. Press.

Medin, D. L., Wattenmaker, W. D., Michalski, R. S. (1986). *Constraints in inductive learning: An experimental study comparing human and machine performance*. Manuscript submitted for publication.

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

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, **92**, 289–316.

Newport, E. (1984). Constraints on learning: Studies in the acquisition of American sign language. *Papers and Reports on Child Language Development*, **23**, 1–22.

Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology*, **66**, 564–583.

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

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

Rosch, E. (1975c). The nature of mental codes for color categories. *Journal of Experimental Psychology: Human Perception and Performance*, **1**, 303–322.

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

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

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

Smith, J. D., & Kemler-Nelson, D. G. (1984). Overall similarity in adults' classification: The child in all of us. *Journal of Experimental Psychology: General*, **113**, 137–159.

Smith, L. B. (1981). Importance of the overall similarity of objects for adults' and children's classifications. *Journal of Experimental Psychology: Human Perception and Performance*, **7**, 811–824.

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

Vygotsky, L. S. (1962). *Thought and language*. New York: Wiley.

Wattenmaker, W. D., Dewey, G. I., Murphy, T. D., & Medin, D. L. (1986). Linear separability and concept learning: Context, relational properties, and concept naturalness. *Cognitive Psychology*, **18**, 158–194.

(Accepted April 24, 1986)