COGNITIVE PSYCHOLOGY 7, 573–605 (1975)

Family Resemblances: Studies in the Internal Structure of Categories

ELEANOR ROSCH AND CAROLYN B. MERVIS

University of California, Berkeley

Six experiments explored the hypothesis that the members of categories which are considered most prototypical are those with most attributes in common with other members of the category and least attributes in common with other categories. In probabilistic terms, the hypothesis is that prototypicality is a function of the total cue validity of the attributes of items. In Experiments 1 and 3, subjects listed attributes for members of semantic categories which had been previously rated for degree of prototypicality. High positive correlations were obtained between those ratings and the extent of distribution of an item's attributes among the other items of the category. In Experiments 2 and 4, subjects listed superordinates of category members and listed attributes of members of contrasting categories. Negative correlations were obtained between prototypicality and superordinates other than the category in question and between prototypicality and an item's possession of attributes possessed by members of contrasting categories. Experiments 5 and 6 used artificial categories and showed that family resemblance within categories and lack of overlap of elements with contrasting categories were correlated with ease of learning, reaction time in identifying an item after learning, and rating of prototypicality of an item. It is argued that family resemblance offers an alternative to criterial features in defining categories.

As speakers of our language and members of our culture, we know that a chair is a more reasonable exemplar of the category *furniture* than a radio, and that some chairs fit our idea or image of a chair better than others. However, when describing categories analytically, most traditions of thought have treated category membership as a digital, all-or-none phenomenon. That is, much work in philosophy, psychology, linguistics, and anthropology assumes that categories are logical bounded entities, membership in which is defined by an item's posses-

This research was supported by grants to the first author (under her former name Eleanor Rosch Heider) by the National Science Foundation (GB-38245X), by the Grant Foundation, and by the National Institutes of Mental Health (1 R01 MH24316-01). We wish to thank David Johnson, Joseph Romeo, Ross Quigley, R. Scott Miller, Steve Frank, Alina Furnow, and Louise Jones for help with testing and analysis of the data. We also wish to thank Ed Smith, Ed Shoben, and Lance Rips for permission to refer to the multidimensional scaling study of superordinate categories which was performed jointly with them. Carolyn Mervis is now at Cornell Universtiy. She was supported by an NSF Predoctoral Fellowship during the research. Requests for reprints should be sent to Eleanor Rosch, Department of Psychology, University of California at Berkeley, Berkeley, CA 94720.

sion of a simple set of criterial features, in which all instances possessing the criterial attributes have a full and equal degree of membership.

In contrast to such a view, it has been recently argued (see Lakoff, 1972; Rosch, 1973; Zadeh, 1965) that some natural categories are analog and must be represented logically in a manner which reflects their analog structure. Rosch (1973, 1975b) has further characterized some natural analog categories as internally structured into a prototype (clearest cases, best examples of the category) and nonprototype members, with nonprototype members tending toward an order from better to poorer examples. While the domain for which such a claim has been demonstrated most unequivocally is that of color (Berlin & Kay, 1969; Heider, 1971, 1972; Mervis, Catlin, & Rosch, 1975; Rosch, 1974, in press-c, in press-d), there is also considerable evidence that natural superordinate semantic categories have a prototype structure. Subjects can reliably rate the extent to which a member of a category fits their idea or image of the meaning of the category name (Rosch, 1973, 1975a), and such ratings predict performance in a number of tasks (Rips, Shoben & Smith, 1973; Rosch, 1973, 1975a, in press-c, 1975b; Smith, Rips, & Shoben, 1974; Smith, Shoben, & Rips, 1974).

However, there has, as yet, been little attention given to the problem of how internal structure arises. That is, what principles govern the formation of category prototypes and gradients of category membership? For some categories which probably have a physiological basis, such as colors, forms, and facial expressions of basic human emotions, prototypes may be stimuli which are salient prior to formation of the category, whose salience, at the outset, determines the categorical structuring of those domains (Ekman, 1971; McDaniel, Note 1; Rosch, 1974, 1975b). For the artificial categories which have been used in prototype research—such as families of dot patterns (Posner, 1973) and artificial faces (Reed, 1972)—the categories have been intentionally structured and/or the prototypes have been defined so that the prototypes were central tendencies of the categories. For most domains, however, prototypes do not appear to precede the category (Rosch, in press-a) and must be formed through principles of learning and information processing from the items given in the category. The present research was not intended to provide a processing model of the learning of categories or formation of prototypes; rather, our intention was to examine the stimulus relations which underlie such learning. That is, the purpose of the present research was to explore one of the major *structural* principles which, we believe, may govern the formation of the prototype structure of semantic categories.

This principle was first suggested in philosophy; Wittgenstein (1953) argued that the referents of a word need not have common elements in order for the word to be understood and used in the normal functioning of language. He suggested that, rather, a family resemblance might be what linked the various referents of a word. A family resemblance relationship consists of a set of items of the form AB, BC, CD, DE. That is, each item has at least one, and probably several, elements in common with one or more other items, but no, or few, elements are common to all items. The existence of such relationships in actual natural language categories has not previously been investigated empirically.

In the present research, we viewed natural semantic categories as networks of overlapping attributes; the basic hypothesis was that members of a category come to be viewed as prototypical of the category as a whole in proportion to the extent to which they bear a family resemblance to (have attributes which overlap those of) other members of the category. Conversely, items viewed as most prototypical of one category will be those with least family resemblance to or membership in other categories. In natural categories of concrete objects, the two aspects of family resemblance should coincide rather than conflict since it is reasonable that categories tend to become organized in such a way that they reflect the correlational structure of the environment in a manner which renders them maximally discriminable from each other (Rosch, in press-a; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, in press).

The present structural hypothesis is closely related to a *cue validity* processing model of classification in which the validity of a cue is defined in terms of its total frequency within a category and its proportional frequency in that category relative to contrasting categories. Mathematically, cue validity has been defined as a conditional probability—specifically, the frequency of a cue being associated with the category in question divided by the total frequency of that cue over all relevant categories (Beach, 1964; Reed, 1972). Unfortunately, cue validity has been treated as a model in conflict with a prototype model of category processing where prototypes are operationally defined solely as attribute means (Reed, 1972). If prototypes are defined more broadly—for example, as the abstract representation of a category, or as those category members to which subjects compare items when judging category membership, or as the internal structure of the category defined by subjects' judgments of the degree to which members fit their "idea or image" of the category—then prototypes should coincide rather than conflict with cue validity. That is, if natural categories of concrete objects tend to become organized so as to render the categories maximally discriminable from each other, it follows that the maximum possible cue validity of items within each category will be attained (Rosch *et al.*, in press). The principle of family resemblance relationships can be restated in terms of cue validity since the attributes most distributed among members of a category and least distributed among members of contrasting categories are, by definition, the most valid cues to membership in the category in question. We use the term *family resemblance* rather than *cue validity* primarily to emphasize that we are dealing with a description of structural principles and not with a processing model. We believe that the principle of family resemblance relationships is a very general one and is applicable to categories regardless of whether or not they have features common to members of the category or formal criteria for category membership.

In all of the studies of the present research, family resemblances were defined in terms of discrete attributes such as *has legs, you drive it,* or *the letter B is a member.* These are the kinds of features of natural semantic categories which can be most readily reported and the features normally used in definitions of categories by means of lists of formal criteria. Insofar as the context in which an attribute occurs as part of a stimulus may always affect perception and understanding of the attribute, discrete attributes of this type may be an analytic myth. However, in one sense, the purpose of the present research was to show that it is not necessary to invoke attribute interactions or higher order gestalt properties of stimuli (such as those used by Posner, 1973; Reed, 1972; Rosch, Simpson, & Miller, Note 2) in order to analyze the prototype structure of categories. That is, even at the level of analysis of the type of discrete attributes normally used in definitions of categories by means of criterial features, we believe there is a principle of the structure of stimulus sets, family resemblances, which can be shown to underlie category prototype structure.

The present paper reports studies using three different types of category; superordinate semantic categories such as *furniture* and *vehicle,* basic level semantic categories such as *chair* and *car,* and artificial categories formed from sets of letter strings. For each type of stimulus, both aspects of the family resemblance hypothesis (that the most prototypical members of categories are those with most attributes in common with other members of that category and are those with least attributes in common with other categories) were tested.

Superordinate semantic categories are of particular interest because they are sufficiently abstract that they have few, if any, attributes common to all members (Rosch *et al.,* in press). Thus, such categories may consist almost entirely of items related to each other by means of family resemblances of overlapping attributes. In addition, superordinate categories have the advantage that their membership consists of a finite number of names of basic level categories which can be adequately sampled. Superordinate categories have the disadvantage that they do not have contrasting categories (operationally defined below); thus, the second half of the family resemblance hypothesis (that prototypical members of categories have least resemblance to other categories) had to be tested indirectly by measuring membership in, rather than attributes in common with, other superordinate categories.

Basic level semantic categories are of great interest because they are the level of abstraction at which the basic category cuts in the world may be made (Rosch, in press-a; Rosch *et al.*, in press). However, basic level categories present a sampling problem since their membership consists of an infinite number of objects. On the positive side, basic level categories do form contrast sets, thus, making possible a direct test of the second part of the family resemblance hypothesis.

Artificial categories were needed because they made possible the study of prototype formation with adequate controls. In natural language domains of any type, categories have long since evolved in culture and been learned by subjects. Both prototypes and the attribute structure of categories are independent variables; we can only measure their correlations. Artificial categories are of use because attribute structures can be varied in a controlled manner and the development of prototypes studied as a dependent variable.

PART I: SUPERORDINATE SEMANTIC CATEGORIES

Experiment 1

Although it is always possible for an ingenious philosopher or psychologist to invent criterial attributes defining a category, earlier research has shown that actual subjects rate superordinate semantic categories as having few, if any, attributes common to all members (Rosch *et al.*, in press). Thus, if the "categorical" nature of these categories is to be explained, it appeared most likely to reside in family resemblances between members. Part of the purpose of the present experiment was to obtain portraits of the distribution of attributes of members of a number of superordinate natural language categories. Part of the hypothesis was that category members would prove to bear a family resemblance relationsip to each other. The major purpose of the experiment, however, was to observe the relation between degree of relatedness between members of the category and the rated prototypicality of those members. The specific hypothesis was that a measure of the degree to which an item bore a family resemblance to other members of the category would prove significantly correlated with previously obtained prototypicality ratings of the members of the category.

Method

Subjects. Subjects were 400 students in introductory psychology classes who received this 10 min task as part of their classroom work.

Stimuli. The categories used were the six most common categories of concrete nouns in English, determined by a measure of word frequency

(Kucera & Francis, 1967). All of the categories were ones for which norms for the prototypicality of items had already been obtained for 50–60 category members (Rosch, 1975a). These norms were derived from subjects' ratings of the extent to which each item fit their "idea or image" of the meaning of the category name. (The rating task and instructions were very similar to those used in Experiment 3 of the present research. A complete account of the methods for deriving the six superordinate categories and complete norms for all items of the six categories are provided in Rosch, in press-d.) The 20 items from each category used in the present experiment were chosen to represent the full range of goodness-of-example ranks. These items are listed, in their goodness-of-example order, in Table 1.

Procedure. Each of the 120 items shown in Table 1 was printed at the top of a page, and the pages assembled into packets consisting of six items, one from each superordinate category. Items were chosen randomly within a category such that each subject who received an item received it with different items from the other five categories and received the items representing each category in a different order. Each item was rated by 20 subjects. Each subject rated six items, one from each category.

Subjects were asked to list the attributes possessed by each item. Instructions were:

> This is a very simple experiment to find out the characteristics and attributes that people feel are common to and characteristic of different kinds of ordinary everyday objects. For example, for *bicycles* you might think of things they have in common like two wheels, pedals, handlebars, you ride on them, they don't use fuel, etc. For *dogs* you might think of things they have in common like having four legs, barking, having fur, etc.
>
> There are six pages following this one. At the top of each is listed the name of one common object. For each page, you'll have a minute and a half to write down all of the attributes of that object that you can think of. But try not to *just* free associate—for example, if bicycles just happen to remind you of your father, *don't* write down *father*.
>
> Okay—you'll have a minute and a half for each page. When I say turn to the next page, read the name of the object and write down the attributes or characteristics you think are characteristic of that object as fast as you can until you're told to turn the page again.

Measurement of family resemblance. To derive the basic measure of family resemblance, for each category, all attributes mentioned by subjects were listed and each item, for which an attribute had been listed, was credited with that attribute. Two judges reviewed the resulting table and indicated cases in which an attribute was clearly and obviously false. These attributes were deleted from the tabulation. The judges also indicated any attribute which had been listed for one or more items, but was clearly and obviously true of another item in the category for which

TABLE 1
SUPERORDINATE CATEGORIES AND ITEMS USED IN EXPERIMENTS 1 AND 2

	Category					
Item	Furniture	Vehicle	Fruit	Weapon	Vegetable	Clothing
1	Chair	Car	Orange	Gun	Peas	Pants
2	Sofa	Truck	Apple	Knife	Carrots	Shirt
3	Table	Bus	Banana	Sword	String beans	Dress
4	Dresser	Motorcycle	Peach	Bomb	Spinach	Skirt
5	Desk	Train	Pear	Hand grenade	Broccoli	Jacket
6	Bed	Trolley car	Apricot	Spear	Asparagus	Coat
7	Bookcase	Bicycle	Plum	Cannon	Corn	Sweater
8	Footstool	Airplane	Grapes	Bow and arrow	Cauliflower	Underpants
9	Lamp	Boat	Strawberry	Club	Brussel sprouts	Socks
10	Piano	Tractor	Grapefruit	Tank	Lettuce	Pajamas
11	Cushion	Cart	Pineapple	Teargas	Beets	Bathing suit
12	Mirror	Wheelchair	Blueberry	Whip	Tomato	Shoes
13	Rug	Tank	Lemon	Icepick	Lima beans	Vest
14	Radio	Raft	Watermelon	Fists	Eggplant	Tie
15	Stove	Sled	Honeydew	Rocket	Onion	Mittens
16	Clock	Horse	Pomegranate	Poison	Potato	Hat
17	Picture	Blimp	Date	Scissors	Yam	Apron
18	Closet	Skates	Coconut	Words	Mushroom	Purse
19	Vase	Wheelbarrow	Tomato	Foot	Pumpkin	Wristwatch
20	Telephone	Elevator	Olive	Screwdriver	Rice	Necklace

it had not happened to be listed by any of the 20 subjects. These items were also credited with the relevant attribute. Judges were not permitted to list new attributes, and no item was credited with an attribute about which judges disagreed or about which either judge was uncertain. The total changes made by the judges were infrequent.

Each attribute received a score, ranging from 1–20, representing the number of items in the category which had been credited with that attribute. By this means, each attribute was weighted in accordance with the number of items in the category possessing it. The basic measure of degree of family resemblance for an item was the sum of the weighted scores of each of the attributes that had been listed for that item.

This basic measure of family resemblance possessed a source of potential distortion, however. In the measure, each additional item with which an attribute was credited added an equal increment of family resemblance. Thus, the measure depended upon the assumption that the numerical frequency of an attribute within a category was an interval measure of the underlying psychological weight of that attribute (e.g., the difference between an attribute which belonged to two items versus one item was equal to the difference between an attribute which belonged to 19 versus 18 items). Such an assumption is not necessarily reasonable; therefore, a second measure of family resemblance was also computed. To derive this measure, each attribute was weighted with the natural logarithm of the raw score representing the number of items in the category which had been credited with that attribute; the second measure, thus, consisted of the sum of the natural logarithms of the scores of each of the attributes that had been listed for an item.

Results and Discussion

The purpose of the study was both to provide a portrait of the structure of the categories and to test the correlation between family resemblance and prototypicality of items. In terms of structure, Fig. 1 shows the mean frequency distribution for the number of attributes applied to each number (1–20) of items/category. As had been previously found when subjects listed attributes for superordinate category names (Rosch *et al.*, in press), in the present study, few attributes were given which were true of all 20 members of the category—for four of the categories, there was only one such item; for two of the categories, none. Furthermore, the single attribute which did apply to all members, in three cases was true of many other items besides those within that superordinate (for example, "you eat it" for fruit). Thus, the salient attribute structure of these categories tended to reside, not in criterial features common to all members of the category which distinguished those members from all others, but in a large number of attributes true of some, but not all, category members.

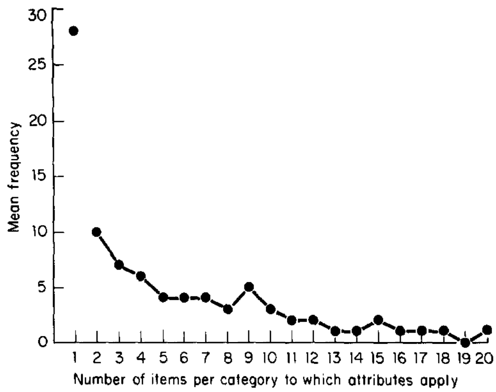

FIG. 1. Frequency distribution for number of attributes applied to each number of items/category.

Those attributes unique to a single member are not of primary interest for the present study since they do not contribute to the structure of the category per se. In actual fact, the number of unique attributes applicable to items was evenly distributed over members of the categories; for none of the six categories was the number of unique attributes significantly correlated with prototypicality. Of the attributes applicable to two or more members, Fig. 1 shows that the number of attributes decreases as the number of items to which the attribute is applicable increases. In summary: the majority of attributes listed for items in the six categories demonstrated a family resemblance relationship; that is, they were common to only some of the category members.

The major hypothesis of the experiment was that this family resemblance structure would prove significantly correlated with the prototypicality of items. Correlations were computed separately for each of the two measures of family resemblance and separately for each category. The measure of prototypicality was the mean rating on a 7-point scale of the extent to which items fit subjects' idea or image of the meaning of the category names (Rosch, 1975a). The basic measure of degree of family resemblance for an item was the sum of the weighted raw scores of each of the attributes listed for the item. The logarithmic measure of family resemblance was the sum of the natural logarithms of the scores of each of the attributes that had been listed for an item. Items in each category were ranked 1–20 on the basis of prototypicality and were ranked 1–20 on the basis of each of the measures of family resemblance. Spearman rank–order correlations between the ranks of items on family resemblance and their ranks on prototypicality were performed separately for each of the measures of family resemblance and for each of the categories. These correlations, for the basic measure of family resem-

TABLE 2
NUMBER OF ATTRIBUTES IN COMMON TO FIVE MOST AND FIVE LEAST PROTOTYPICAL MEMBERS OF SIX CATEGORIES

Category	Most typical members	Least typical members
Furniture	13	2
Vehicle	36	2
Fruit	16	0
Weapon	9	0
Vegetable	3	0
Clothing	21	0

blance, were: furniture, 0.88; vehicle, 0.92; weapon, 0.94; fruit, 0.85; vegetable, 0.84; clothing, 0.91. These correlations for the logarithmic measure of family resemblance were: furniture, 0.84; vehicle, 0.90; weapon, 0.93; fruit, 0.88; vegetable, 0.86; clothing, 0.88. All were significant ($p < .001$).

Such results strongly confirm our hypothesis that the more an item has attributes in common with other members of the category, the more it will be considered a good and representative member of the category. Furthermore, the similarity in results obtained with the basic and the logarithmic measures of family resemblance argues that this relationship is not dependent upon the properties of the particular scale used in measurement. Specifically, items in a category tended to be credited with approximately equal numbers of attributes, but the less prototypical the item, the fewer other items in the category tended to share each attribute. Thus, the ranks for the basic and logarithmic measures of family resemblance were almost identical, and the correlations between family resemblance and prototypicality were scarcely affected by the change in measure. The relationship between degree of family resemblance and prototypicality for these categories, thus, appears to be a robust one.

A corollary of this finding may account for one of the persistent illusions concerning superordinate categories. Subjects, upon receiving feedback from the experiment, and audiences, upon being told of it, generally argue that they feel positive that there are many attributes common to all members of the category even when they cannot think of any specific attributes for which there are not counterexamples. If the more prototypical members of a category are those which have most attributes common to other members of the category it is probable that they are most likely to have attributes in common with each other. To investigate this possibility, the number of attributes common to the five most and five least prototypical items in each category were compared. The number of attributes are shown in Table 2. It is clear from this count that, while category members as a whole may not have items in common, the five most typical items of each category tend to have many items in common. Thus, if subjects think of the best examples of the category when hearing the category name (Rosch, 1975a), the illusion of common elements is likely to arise and persist—an illusion which may be what makes definition of categories in terms of criterial attributes appear so reasonable.

A second corollary of the finding of a strong relationship between family resemblance and prototypicality concerns the structure of the semantic space in which items of a category are embedded. Previous studies of the nature of the semantic spaces of superordinate categories have focused on the dimensionality of the space (Henley, 1969; Rips *et al.*, 1973; Smith, Shoben, & Rips, 1974). However, there are other properties of semantic spaces which can be of interest. For example, items which are perceived as closest to all members of a group of items should fall in the center of the space defined by means of proximity scaling of those items. For purposes of the present study, we can predict that items with the greatest family resemblance should fall in the center of the semantic space defined by proximity scaling of the items in a category; such an effect can be predicted regardless of the dimensionality or lack of dimensionality of the semantic space. If, in addition, items are perceived as similar to each other in proportion to the number of attributes which they have in common, multidimensional scaling of the similarity judgments between all pairs of items in a category should result in a semantic space in which the distance of items from the origin of the space is determined by their degree of family resemblance.

A multidimensional scaling study of the categories *furniture, vehicle, weapon, fruit,* and *vegetable* was performed as part of a larger study.[1] Stimuli were the same 20 items in these categories shown in Table 1 plus the superordinate category name. All possible pairs of the 21 items in each category were printed in a booklet and were rated on a 9-point scale for degree of similarity between the items. Fifteen subjects rated the items in each category. The similarity ratings were scaled by M–D scale (Shepard, 1962; Shepard, Romney, & Nerlove, Vol. 1, 1972). Results showed that, while the dimensionality of the scaling solutions was generally difficult to interpret, in all cases the category name and the most prototypical items appeared to be the most central in the scaling solution regardless of the number of dimensions or the rotation used. To check this finding, Spearman rank–order correlations between degree of

[1] The larger study was performed in collaboration with E. E. Smith, E. J. Shoben, and L. J. Rips of Stanford University. Half of the subjects were tested at the University of California, Berkeley, half at Stanford University. The multidimensional scaling was performed entirely at Stanford.

family resemblance and distance of an item from the origin in the three-dimensional scaling solution with minimum stress were performed for the five categories. These correlations were: furniture, 0.89; vehicle, 0.94; weapon, 0.95; fruit, 0.92; and vegetable, 0.90. All were significant ($p < .001$).

In the use of proximity scaling for items in semantic categories, it is customary to rely for interpretation on dimensions which characterize the space as a whole (see Shepard, Romney, & Nerlove, Vol. II, 1971). Such a trend is similar to the tradition of treating categories only in terms of logical defining features which are common to all members of the category. The present example of the use of scaling shows that, although family resemblance was defined in terms of discrete features no one of which was common to all category members, and although the dimensionality of the categories was not obvious in the scaling solutions, the property of centrality of items in the semantic space was still interpretable; that is, degree of family resemblance was highly predictive of centrality in a semantic space defined by global similarity ratings of the items in the category.

In summary: The hypotheses of Experiment 1 were confirmed. For six superordinate categories, 20 members of the category were characterized by attributes which were common to some, but not all, members. The degree to which a given member possessed attributes in common with other members was highly correlated with the degree to which it was rated prototypical (representative) of the category name. In addition, degree of family resemblance predicted the centrality of items in the semantic space generated by multidimensional scaling of similarity ratings between items in the category.

Experiment 2

The initial hypothesis behind Experiment 2 was the direct converse of that of Experiment 1, namely that the most prototypical members of categories would not only have the greatest family resemblance to members of their category but would also be maximally distant from and, thus, have the least attributes in common with members of other categories at the same level of linguistic contrast. We found that this hypothesis could not be tested directly for superordinate categories.

The standard empirical method for deriving linguistic contrast sets from research participants is some variant of the question, "If X is not a Y, what is it (might it be)?" (Frake, 1969). We pretested both the simple form of the question and the elaborated instructions used in Experiment 4. However, for the six superordinate categories of the present research, such instructions failed to produce consistent responses from subjects; those subjects who were able to respond at all tended to produce individual creative answers which were not considered reasonable by other subjects to whom they were shown. For superordinate categories, we, therefore, turned to an indirect test of the hypothesis by means of measurement of overlap in category membership. If the best examples of superordinate categories are those with least in common with other categories they should be dominant members of few (or no) categories other than the superordinate in question. Thus, prototypicality should be correlated with a measure of the dominance of a category over its members (Loftus & Scheff, 1971). Subjects could readily list superordinates for category members. The hypothesis of Experiment 2 was, thus, that the more prototypical a member of a superordinate category, the less dominant its membership would prove to be in categories other than the superordinate in question.

Method

Subjects. Subjects were 400 students in introductory and upper-division psychology classes, none of whom had participated in Experiment 1. They participated in the experiment as part of their classroom work.

Stimuli. Stimuli were the same members of five of the six most common superordinate categories of concrete nouns that had been used in Experiment 1 (*clothing* was erroneously omitted). The items were assembled in the same manner as described for Experiment 1. The only difference in format was that under each item, three lines labeled "1, 2, and 3" were printed on the page.

Procedure. Instructions were as follows:

> On each of the pages given you, you will see a noun and three lines. On each line, we want you to write a category to which the noun belongs. For example, if the noun were "collie," you might write *dog, animal,* or *pet* (etc.).
>
> Note that all of the words you see are to be interpreted as concrete nouns, not as verbs. For example, if you saw the word "dress," interpret it as the article of clothing "dress" and not the action of getting dressed.
>
> Be sure to write three categories to which the noun belongs for each noun.

Computation of category membership score. Categories listed in first, second, and third place were weighted accordingly: three for first place mention, two for second place mention, one for third place mention. Since our hypothesis concerned single versus multiple category memberships and salient category memberships in any category other than the designated superordinate, we required a measure of the degree of dominance of the designated superordinate over the other most frequently mentioned superordinates. For each item, this was the following weighted measure: (designated superordinate minus most frequently mentioned other superordinate) plus (designated superordinate minus second most frequently measured other superordinate). This produced a single measure of category dominance for each item.

Results and Discussion

Category dominance of each item was scored as described above; the items within each category were ranked in accordance with their relative degree of category dominance. A Spearman rank–order correlation was performed for each category between category dominance and prototypicality. These correlations were: fruit, 0.71; furniture, 0.83; vegetable, 0.67; vehicle, 0.82; weapon, 0.77. All were significant ($p < .001$).

Our hypothesis was that the more prototypical an item in a given category, the less it would bear a family resemblance to items in other categories, and, thus, the less likely it would be to have salient membership in those other categories. Membership in other categories was the variable which it proved possible to measure. The strong positive correlations between prototypicality and dominance of membership in the category for which prototypicality had been measured confirms this hypothesis.

PART II: BASIC LEVEL CATEGORIES

It has been previously argued (Rosch, in press-a; Rosch *et al.*, in press) that there is a basic level of abstraction at which the concrete objects of the world are most naturally divided into categories. A working assumption has been that, in the domains of both man-made and biological objects, there occur information-rich bundles of attributes that form natural discontinuities. These bundles are both perceptual and functional. It is proposed that basic cuts are made at this level. Basic objects (for example, *chair, car*) are the most inclusive level of abstraction at which categories can mirror the correlational structure (Garner, 1974) of the environment and the most inclusive level at which there can be many attributes common to all or most members of the categories. The more abstract combinations of basic level objects (e.g., categories such as *furniture* and *vehicle* used in Experiments 1 and 2) are superordinates which share only a few attributes; the common attributes are rather abstract ones. Categories below the basic level are subordinates (e.g., *kitchen chair, sports car*). Subordinates are also bundles of predictable attributes and functions, but contain little more information than the basic level object to which they are subordinate. Basic categories are, thus, the categories for which the cue validity of attributes within categories is maximized: Superordinate categories have lower cue validity than basic because they have fewer common attributes within the category; subordinate categories have lower cue validity than basic because they share attributes with contrasting subordinate categories (e.g., *kitchen chair* shares most of its attributes with *living room chair*).

In a converging series of experiments (Rosch *et al.*, in press), it was confirmed that basic objects are the most inclusive categories in which clusters of attributes occur which subjects agree are possessed by members of the category; sets of common motor movements are made when using or interacting with objects of that type; commonalities in the shape, and, thus, the overall look, of objects occur; it is possible to recognize an averaged shape of an object of that class; and it is possible to form a representation of a typical member of the class which is sufficiently concrete to aid in detection of the object in visual noise. In addition, basic objects were shown to be the first categorizations made by young children, and basic object names the level of abstraction at which objects are first named by children and usually named by adults.

The present research concerned the question of whether the family resemblances of items in basic level categories were related to prototypicality in the way in which it had proved to be in the superordinate categories studied in Experiments 1 and 2. Do subjects agree concerning which members of basic object categories are the more prototypical—do they agree, for example, about which cars more closely fit their idea or image of the meaning of *car*? And, if agreement in prototypicality ratings is obtained, does it hold, as it did in the case of superordinate categories, that the more prototypical category members are those with most resemblance to members of that category and least resemblance to other categories? In Experiment 3, the hypothesis was tested that prototypicality ratings and degree of family resemblance were positively correlated. Experiment 4 tested the converse hypothesis that prototypicality ratings were negatively correlated with the degree to which an item possessed attributes which were also possessed by members of contrasting categories.

Experiment 3

Method

Subjects. Subjects were 182 paid undergraduate volunteers who participated as a part of a fundraising for a student organization. None had participated in the superordinate category experiments. Thirty-two subjects rated the stimuli for goodness-of-example; 150 listed attributes.

Stimuli. Superordinate categories have a finite number of members designable by words, with norms available for the frequency with which the members are listed by subjects (Battig & Montague, 1969). The members of basic level categories, however, are actual objects, an essentially infinite population. Six categories were chosen for the present experiment, which had been shown to be at the basic level of abstraction by the convergent techniques used in Rosch *et al.* (in press). The categories were: car, truck, airplane, chair, table, and lamp. Each of these was a category for which pictures of many objects could be readily obtained and a category which had the property that the atrtributes of the object most listed by subjects (Rosch *et al.*, in press) could be seen in pictures of the object. Pictures to be used in the present research were selected from a large sample of pictures (described in Rosch *et al.*, in press); two judges chose 15 pictures by the following method—for each category, they first found the picture they felt most prototypical of the category, then the picture they felt was the worst example of the category (but still clearly called a *car, chair,* etc.). They then selected 13 other pictures which they agreed spanned the distance between the two extreme pictures in as equal subjective steps as possible given the available pool of pictures. The 90 pictures, 15 in each category, chosen in this manner served as stimuli in the experiment.

Procedure

1. Prototypicality ratings. Subjects were given essentially the same instructions as had been given subjects who rated the prototypicality of members of superordinate categories. Basically, subjects were asked to rate, on a 7-point scale, the extent to which an instance represented their idea or image of the meaning of the category name. Precise instructions were:

> This study has to do with what we have in mind when we use words which refer to categories. Let's take the word *red* as an example. Close your eyes and imagine a true red. Now imagine an orangish red . . . imagine a purple red. Although you might still name the orange-red or the purple-red with the term *red,* they are not as good examples of red (as clear cases of what *red* refers to) as the clear "true" red. In short, some reds are redder than others. The same is true for other kinds of categories. Think of dogs. You all have some notion of what a "real dog," a "doggy dog" is. To me a Retriever or a German Sheperd is a very doggy dog while a Pekinese is a less doggy dog. Notice that this kind of judgment has nothing to do with how well you like the thing; you can like a purple–red better than a true red but still recognize that the color you like is not a true red. You may prefer to own a Pekinese without thinking that it is the breed that best represents what people mean by dogginess.
>
> In this study you are asked to judge how good an example of a category various instances of the category are. The members of the category are pictures; you will be told the name of the category and shown 15 pictures of items in the category. On your answer sheet are six columns of 15 numbers. After each number is a blank. You are to rate how good an example of the category each picture is on a 7-point scale. A 1 means that you feel the picture is a very good example of your idea or image of what the category is; a 7 means you feel the picture fits very poorly with your idea or image of the category (or is not a member at all). A 4 means you feel the picture fits moderately well. Use the other numbers of the 7-point scale to indicate intermediate judgments.

> Don't worry about why you feel that something is or isn't a good example of the category. And don't worry about whether it's just you or people in general who feel that way. Just mark it the way you see it.

Slides of the 15 pictures in a category were shown the subjects once through rapidly in random order; then, each slide was shown the group for 30 seconds while subjects made their ratings. Means of the ratings of the 32 subjects in the experiment formed the basis for ranking the items.

2. *Attribute listing.* Subjects were given the same instructions for listing attributes as the subjects in Experiment 1, with the exception that they were told they would be seeing pictures and were asked to list the attributes of the item in each picture. Each subject listed attributes for six pictures, one from each of the basic level categories. Sets of pictures were assembled by the same principles as the sets of words had been in testing superordinate categories. Ten subjects listed attributes for each picture. Subjects were allowed 1.5 min to list attributes for each slide.

Results and Discussion

Methods used for computing family resemblance and for computing the correlation between family resemblance of attributes and prototypicality ratings were the same methods as had been used in Experiment 1. As expected (Rosch *et al.*, in press), basic level categories differed from the subordinates in that many more attributes were common to all members of the basic level categories. However, there were also many attributes listed which were not common to all members. These attributes were used in the correlation between family resemblance and prototypicality. The Spearman rank–order correlations between the basic measure of family resemblance and prototypicality were: car, 0.94; truck, 0.84; airplane, 0.88; chair, 0.81; table, 0.88; and lamp, 0.69. The correlations between the logarithmic measure of family resemblance and prototypicality were: car, 0.86; truck, 0.88; airplane, 0.88; chair 0.79; table, 0.85; and lamp, 0.64. All were significant ($p < .01$). Thus, we have verified for pictures of basic level objects, as well as for names of members of superordinate categories, the more prototypical items are those which have most attributes in common with other members of the category. As in the case of superordinate categories, this relationship was not dependent on the particular scale used to measure family resemblance.

Experiment 4

The purpose of both Experiments 2 and 4 was to provide data complementary to that of Experiments 1 and 3. The basic hypothesis of both experiments was that categories tend to become organized in such a way that they are maximally discriminable from other categories at the same level of contrast; hence, the most prototypical members of a category are those with least resemblance to, or membership in, other categories. For superordinate categories, it had not been possible to obtain contrast sets and, thus, not possible to measure commonality of attributes between contrasting categories directly; instead, the hypothesis had been tested indirectly by means of an item's membership in multiple categories. For members of basic level categories, the hypothesis proved testable directly.

The basic design of the experiment was: (a) to determine which categories were seen in direct contrast to a sample of the basic level categories for which we had obtained prototypicality ratings and attribute lists in Experiment 2, (b) to obtain lists of attributes for pictures representing items in the contrasting categories, and (c) to correlate the number of attributes which items shared with contrasting categories with prototypicality ratings for the items; a negative correlation was predicted.

Method

Subjects. Subjects were 44 students in psychology classes who performed the task as part of their classroom work; 24 of the subjects served in the contrast set portion of the experiment; 20 subjects listed attributes.

Stimuli and procedure. The first part of the experimental procedure required obtaining contrast sets of the basic level categories to be used. Subjects were read the following instructions:

> Suppose that you are participating in a communication task experiment. Another person is describing "items" to you, and you have to figure out what kind of "item" he is describing. The person tells you about each item's *physical attributes* (what it looks like, what parts it has, etc.), and about its *functions* (what people do with it), and about its *actions* (what it does). Suppose, also, that you have guessed once for each item, and you have been told that your answer was not correct, but was very close to the correct one. Assume that each word I read was your first answer to one item. After I read each item, write down what your second answer would be. Remember that your first guess was very close to being correct. Think of something that has physical attributes, functions, and actions very similar to the ones your "first answer" had.

Subjects were then given the six names of basic level items used in Experiment 2 and asked to write their first guess as to what the item might be. Thirty seconds per item were allowed.

Subjects' responses were tallied. From the six basic level categories, two were selected for which the most consistent responses had been given. These two, *chair* and *car,* were used for the second part of the experiment.

Stimuli for the attribute listing consisted of pictures of two examples of each of the three most frequently given contrast items for *chair* and *car.* These were: for chair—sofa, stool, and cushion; for car—truck, bus, and motorcycle. The pictures were chosen randomly from the pool of available pictures of these items, with the restriction that all of the pictures chosen had been rated (by two judges) as good examples of their category.

Attribute lists had already been obtained for the chair and car pictures in Experiment 3. Attributes for the six contrast categories were obtained by the same procedures as used in Experiment 3; subjects were read the same instructions as in Experiment 3, were shown slides of the pictures in the contrast categories in random order, and were given 1.5 min to list attributes for each picture. Each subject saw six pictures, one of each contrast item. Each picture was seen by 10 subjects.

Results and Discussion

For each of the 15 chair and 15 car pictures, a tally was made of the number of attributes listed for that picture, which had also been listed for at least one of the pictures of one of the three contrast categories. This tally was used as the measure of amount of overlap between the attributes of a given item and the attributes of items in the closest contrasting categories. A Spearman rank–order correlation was performed between the prototypicality and attribute overlap ranks of the 15 chair and 15 car pictures. Results were: chairs $r = -.67$; cars, $r = -.86$. Both were significant ($p < .01$). In short, it was clearly confirmed for two basic level categories that the more prototypical of the category a picture had been rated, the fewer attributes it shared with categories in direct contrast with that category.

PART III. ARTIFICIAL CATEGORIES

In the four preceding experiments, it was shown for a sample of naturally occurring categories that items rated more prototypical of the category were more closely related to other members of the category and less closely related to members of other categories than were items rated less prototypical of a category. Categories designated by the words of natural languages have the advantage for study that they have evolved and occur in actual human usage; however, they have the disadvantage that the variables of interest occur in uncontrolled and, thus, unanalyzable conjunction with each other and with other extraneous factors. In the previous experiments, the object was to determine the structure of preexisting categories. In the following two experiments, artificial categories were constructed in which items differed only in the degree of family resemblance within categories or amount of overlap of attributes between categories. In these experiments, the structure was provided as an independent variable; our hypothesis was that this structure would affect rate of learning of category items; reaction time in judging category membership once the categories were learned; and ratings of prototypicality of items.

In Experiment 5, only the family resemblance of items within categories was varied; the categories which subjects learned to discriminate contained no elements in common and, thus, no overlap. In Experiment 6, categories which had previously been learned in Experiment 5, where there was no overlap, were taught in conjunction with categories whose attributes overlapped some items. Thus, changes in learning, reaction time, and judgments of prototypicality created by the difference in contrast sets could be observed.

Experiment 5

Method

Subjects. Subjects were 30 students in introductory psychology classes who received course credit for participation.

Stimuli. All of the stimuli were constructed out of strings of letters. The digits 1–9 were mixed with the letters when more symbols were needed. Vowels were used only in one stimulus type and only at the ends of strings so that pronounceability was not a factor.

Three types of family resemblance structures were used. Table 3 shows one example of each structure with the items ordered so that the nature of the structure can be seen. After each letter string, the family resemblance score of that string is shown. This score is computed as described for the basic family resemblance measure in Experiments 1 and 3. Each letter received a weight (1–5) representing the number of strings in the category in which it occurred; the weights of each letter in a string were summed to generate the family resemblance score of that string. (A letter could not receive a weight of 6 because no letter occurred in all six items in a category.)

In the control group structure, all members of the category overlapped other members (possessed letters in common with other members) to an equal degree; thus, the family resemblance scores of each item were the same. The structure of the control group categories was generated in an ad hoc manner in order to achieve equal family resemblance scores for all items. (To understand better the derivation of the family resemblance score, the reader might wish to count the number of strings in the category in which each letter occurred and then sum the weights of the letters in a given string.)

Two experimental groups were used in which members of categories possessed letters in common with other members to an unequal degree and, hence, possessed differing degrees of family resemblance. One of these groups was constructed with a symmetric structure. Each string differed systematically by one letter from the preceding string, resulting in two "central" strings which were maximally overlapped with the other strings in the category. The nature of the symmetric category

TABLE 3
ARTIFICIAL CATEGORY STRUCTURES USED IN EXPERIMENTS 5 AND 6

		Type of category structure							
		Control set			Symmetric experimental set			Asymmetric experimental set	
Use of the category	Item in category	Letter string	Family resemblance score	Overlap score	Letter string	Family resemblance score	Overlap score	Letter string	Family resemblance score
Basic category structure	1	HPNWD	12	0	JXPHM	15	0	DLT83	16
	2	HPNSJ	12	2	XPHMQ	19	1	DLT8A	15
	3	GKNTJ	12	4	PHMQB	21	2	DLTPM	14
	4	4KCTG	12	5	HMQBL	21	3	DLGKI	12
	5	4KC6D	12	3	MQBLF	19	4	D9H60	10
	6	HPC6B	12	1	QBLFS	15	5	3YH7V	7
Nonoverlap contrast category (Experiment 5)	1	R7QUM	12		CTRVG	15		SXB25	16
	2	R7QXV	12		TRVGZ	19		SXB2Q	15
	3	Z5Q2V	12		RVGZK	21		SXBRE	14
	4	L5F27	12		VGZKD	21		SXVFW	12
	5	L5F1M	12		GZKDW	19		S4Z1&	10
	6	R7F19	12		ZKDWN	15		5JZCN	7
Overlapped contrast category[a] (Experiment 6)	1	8SJKT		4[b]	GVRTC		0		
	2	8SJ3G		3	VRTCS		1		
	3	9UJCG		3	RTCSF		2		
	4	4UZC9		2	TCSFL		3		
	5	4UZRT		3	CSFLB		4		
	6	MSZR5		3	SFLBQ		5		

[a] Overlap is with the basic category structure not the nonoverlap contrast category.
[b] Contrast strings in control do not have same structure as initial category strings and were not analyzed in Experiment 6.

structures should be apparent from the items shown in Table 3. The two central, two intermediate, and two peripheral items of the symmetric categories possessed the same degrees of family resemblance. The other experimental group was constructed with an asymmetric structure. These categories were generated so that one letter only occurred in five strings, another letter in four strings, another letter in three strings, three letters occurred in two strings, and all other letters occurred only once. This structure yielded differences in degree of family resemblance between all six of the strings of the asymmetric categories. Specifically, the strings of the asymmetric categories possessed the following letter weights: string 1 = 5, 4, 3, 2, 2; string 2 = 5, 4, 3, 2, 1; string 3 = 5, 4, 3, 1, 1; string 4 = 5, 4, 1, 1, 1; string 5 = 5, 2, 1, 1, 1; string 6 = 2, 2, 1, 1, 1. The reader can verify these letter weights by counting the number of strings in which each letter occurs. For each string, the letter weights are summed to produce the family resemblance scores shown in Table 3.

There were six items (six letter strings) within each category. Each letter string was typed horizontally in capital letters in the center of a 12.70 × 20.32 cm white card. Ten different sets of letter combinations were used for each category type, all of which possessed a structure identical to the set of that type shown in Table 3. Each subject received a different set. The letters in a string were not presented to the subject in their "structural" order as shown in Table 3, but were randomly ordered within each string. An item appeared in the same order throughout the task for a given subject.

Procedure. The letter strings were displayed in a Harvard two-field tachistoscope. Subjects were told that they were to learn to distinguish two categories, *1s* and *2s*. They pressed a telegraph key with the forefinger of the dominant hand to indicate a *1* and another key with the forefinger of the other hand to indicate a *2*. They were first shown the strings which made up each category. First the *1s* and then the *2s* were displayed in random order, 4 sec/item, with an approximate interstimulus interval of 10 sec. The 12 strings which composed the two categories were then displayed in random order. Subjects responded by pressing one of the keys to indicate their judgment of the category to which the item belonged. They received verbal feedback of *correct* and *wrong* from the experimenter until they had achieved two errorless runs. At that point, the experimenter told the subject that he had learned the categories, but would continue seeing the items and that he was to respond with the category designation as quickly as possible without making errors. The subject continued for 15 additional trials. At the conclusion, the subject was given the six cards in each category, read the instructions for rating prototypicality, and asked to rank order the six cards in terms of the degree to which each fit the idea or image which he had developed of the category.

Results and Discussion

There were three dependent variables of interest; rate of learning, reaction time, and rankings of prototypicality. Rate of learning was measured by the total number of errors subjects made in classifying an item; reaction time was the time a subject took to respond to an item in his last run; and rankings of prototypicality were the subjects' rankings of the prototypicality of the items.

The two structural types of experimental stimuli were each divided into three parts: the two stimuli with greatest family resemblance, the two with middle degree of family resemblance, and the two with least family resemblance. (For the symmetric group, the two items within each of the three parts possessed equal degree of family resemblance.) For the control set, no such division was possible, and the six items of the set were analyzed separately. Table 4 shows the means for all three experimental variables for the three types of category structures.

A one-way analysis of variance (ANOVA) for correlated scores was performed on the high, medium, and low family resemblance items for the two experimental sets and on all six of the items for the control set. For the control group, results were not significant for any of the three variables. For both experimental groups, however, the results were significant for all three variables: Symmetric categories—rate of learning, $F(2,9) = 7.09$, $p < .05$; reaction time, $F(2,9) = 6.41$, $p < .05$; prototypicality rating, $F(2,9) = 11.35$, $p < .01$; Asymmetric categories—rate of learning, $F(2,9) = 9.48$, $p < .01$; reaction time, $F(2,9) = 7.91$, $p < .05$; prototypicality rating, $F(2,9) = 14.66$, $p < .01$. Thus, the predicted results were obtained. In artificial family resemblance sets, when no single attribute is common to all members of the set, even when contrast sets have no elements in common with each other, items that have greater degree of family resemblance with the members of their own set

TABLE 4
EFFECT OF DEGREE OF FAMILY RESEMBLANCE ON RESPONSE MEASURES

Stimulus type	Number of errors			Reaction time (msec)			Prototypicality rating		
	Hi[a]	Med	Lo	Hi	Med	Lo	Hi	Med	Lo
Symmetric experimental	2.8	4.4	5.5	560	617	692	5.0	3.4	2.1
Asymmetric experimental	2.4	6.8	9.5	532	619	765	5.5	3.5	1.6
Control	6.5	6.4	6.7	670	651	644	3.7	3.4	3.4

[a] Hi, Med, and Lo refer to family resemblance scores.

are learned more rapidly, identified more rapidly even after practice, and judged as more prototypical members of the category than are items with a lesser degree of family resemblance.

Experiment 6

For the natural categories in the first four experiments, we found that the more prototypical items had less resemblance to contrasting categories than the less prototypical items. In natural categories, greater family resemblance within a category and less resemblance to contrasting categories are inseparable if categories are assumed to form in accordance with the natural contingency structure of an environment in which attributes occur in correlated clusters (Rosch *et al,* in press). In artificial categories, however, it is possible to separate those two principles. Experiment 5 showed prototype formation to be a function of greater family resemblance when there was no influence of overlapping attributes from contrasting categories. The hypothesis of Experiment 6 is that items are considered more prototypical of a category to the extent that they do not overlap with contrasting categories.

The experiment was performed for two sets of stimuli. For the control stimuli of Experiment 5, no item had greater family resemblance with any other item, and no differences in learning, reaction time, or prototypicality ratings were found. For categories with this structure, the hypothesis is that prototypicality can be induced purely as a function of extent of overlap with a contrasting category. For the experimental groups of Experiment 5, the hypothesis is that extent of overlap and family resemblance within the category will combine to produce prototypicality under conditions of overlap. Only the symmetrically structured group was used. (The asymmetrical structures proved impossible to learn in the hour of subject time available when the items with greatest family resemblance within the category overlapped a contrast category.)

Method

Subjects. Subjects were 20 students in introductory psychology who had not participated in the previous experiment. They received course credit for their participation.

Stimuli and procedure. Stimuli were constructed as described for Experiment 5. The contrast sets used were the overlapping contrast sets shown in Table 3. Scores for the amount of overlap with the contrast category were computed by the same methods as described in Experiment 4; each letter which occurred at least once in the contrast category received a weight of "1." These weights were summed to give the overlap score of a string. The overlap score, thus, represents simply the number of letters in a string which occur in the contrast category. The contrast categories for the categories which had been controls in Experiment 5 were constructed in an ad hoc manner in order to provide overlap scores 0–5 for the strings of the initial category. These contrast category strings did not possess the same structure or same overlap scores as the initial category strings. Contrast categories for the symmetrical groups were constructed with the same symmetric structure and same degree of overlap as the initial categories. The overlap scores for each string in the initial category and the overlap contrast category are shown in Table 3. (To understand better the derivation of the overlap score, the reader might wish to count the number of letters in a string which occur in at least one string of the contrast category and sum the number of such letters.)

Procedures were identical to those of Experiment 5.

Results and Discussion

The control and experimental sets of Experiment 5 had different structures. For the control set, the items had shown no difference in learning, reaction time, or prototypicality ratings when learned in conjunction with nonoverlapping contrast sets. By the method of measuring amount of overlap described under *Method* above, the contrast sets in the present experiment changed this set into one in which items ranged from zero attributes of overlap to all five attributes overlapped (only the initial category and not the contrast category itself was analyzed for this set—see Table 3).

Table 5 shows learning, reaction time, and prototypicality scores for the greatest, middle range, and least overlapped items. A one-way ANOVA for correlated scores was performed for these measures for the three dependent variables. All three were significant: rate of learning, $F(2,9) = 8.36$, $p < .01$; reaction time, $F(2,9) = 10.75$, $p < .01$; prototypicality, $F(2,9) = 11.19$, $p < .01$.

For the symmetric experimental set, overlap is in conflict with the family resemblance structure internal to the category. There are two hypotheses which we wished to test with the analysis: The first was that, with internal family resemblance held constant by the analysis, extent of overlap with the contrast category would show a significant effect on

TABLE 5
EFFECT OF DEGREE OF OVERLAP ON RESPONSE MEASURES FOR CONTROL SET

Response measure	Degree of overlap		
	Low	Medium	High
Number of errors	7.1	9.4	12.6
Reaction time (msec)	909	986	1125
Prototypicality rating	5.3	3.4	1.8

learning, reaction time, and prototypicality. The second was that internal family resemblance would still have an effect and would tend to counter the effect of degree of overlap for the three items in which these two principles were opposed.

In terms of internal family resemblance, the symmetric category can be divided into two halves (items 1, 2, and 3 versus items 4, 5, and 6) which are structurally identical. In relation to the present contrast set, however, items 1, 2, and 3, had, respectively, 0, 1, and 2 overlapped attributes; while items 4, 5, and 6, had, respectively, 3, 4, and 5 overlapped attributes. Our prediction was that, while in the tasks of Experiment 5, the two halves showed no significant differences, such differences would be apparent in the present experiment. *T*-tests were performed separately for each dependent variable for the means of items in the two halves for the data from Experiment 5 and for those means in the present experiment. None of these tests was significant for the Experiment 5 data. However, all were significant for the data from the present experiment: rate of learning, $t(9) = 5.57, p < .001$; reaction time, $t(9) = 6.08$, $p < .001$; prototypicality, $t(9) = 9.38$, $p < .001$. Thus, degree of overlap with the contrasting category clearly influenced performance even in a set which had an internal family resemblance structure.

The second part of the hypothesis was that the internal family resemblance structure would also influence performance. In three of the items (items 1, 2, and 3 in Table 3), internal family resemblance structure and degree of overlap were contradictory. (For the Experiment 5 data, in which only family resemblance within the category varied, results had been significant for all the variables: learning, $F(2,9) = 4.56$, $p < .05$; reaction time, $F(2,9) = 5.16$, $p < .05$; prototypicality, $F(2,9) = 9.24$, $p < .01$.) For the data from the present experiment, neither the learning rate nor the prototypicality results were significant. The one significant variable, reaction time—$F(2,9) = 5.29$, $p < .05$—showed shorter reaction times for items with less overlap, although these items also had less family resemblance within the category. Thus, internal family resemblance, when in conflict with overlap, serves to mitigate the effect of overlap.

In sum, the results of the present experiment have demonstrated that extent of overlap with a contrast category serves to structure categories in which items did not previously differ in degree of family resemblance and to influence the structure of categories in which items did previously differ in degree of family resemblance.

GENERAL DISCUSSION

The results of the present study confirmed the hypothesis that the most prototypical members of common superordinate, basic level, and artificial categories are those which bear the greatest family resemblance to other members of their own category and have the least overlap with other categories. In probabilistic language, prototypicality was shown to be a function of the cue validity of the attributes of items. In the particular studies in this paper, we defined and measured family resemblance in terms of discrete attributes; however, previous studies indicate that the principle can be applied, to some extent, to other types of categories, such as dot patterns distorted around a prototype and categories consisting of items composed of continuous attributes which have a metric (Posner, 1973; Reed, 1972; Rosch, Simpson & Miller, Note 2). In such categories, the prototype dot pattern and the pattern with attributes at mean values have more in common with (are more like) the other items in the category than are items further from the prototype or the mean. Family resemblances (even broadly defined) are undoubtedly not the only principle of prototype formation—for example, the frequency of items and the salience of particular attributes or particular members of the categories (perceptual, social, or memorial salience) as well as the as yet undefined gestalt properties of stimuli and stimulus combinations, undoubtedly contribute to prototype formation (Rosch, 1975b)—however, the results of the present study indicate that family resemblance is a major factor.

Such a finding is important in six ways: (a) It suggests a structural basis for the formation of prototypes of categories, (b) It argues that in modeling natural categories, prototypes and cue validity are not conflicting accounts, but, rather, must be incorporated into a single model, (c) It indicates a structural rationale for the use of proximity scaling in the study of categories, even in the absence of definable category dimensionality, (d) It offers a principle by which prototype formation can be understood as part of the general processes through which categories themselves may be formed, (e) It provides a new link between adult and children's modes of categorization, and (f) It offers a concrete alternative to criterial attributes in understanding the logic of categorical structure.

Family resemblance as a structural basis for prototype formation. The origin of prototypes of categories is an issue because, as outlined in the introduction, there is now considerable evidence that the extent to which members are conceived typical of a category appears to be an important variable in the cognitive processing of categories (Rosch, 1975a, 1975b in press-a, b, c, d). From that previous work alone, it could be argued that ratings of prototypicality are only measures of the associative linkage between an item and the category name and that it is such associative strength which determines the effects of typicality on processing tasks such as those used in semantic memory. While in a processing model, associative strength may, by definition, be directly related to typicality effects, associative strength need not be conceived only as the result of the frequency of (arbitrary or accidental) pairings of the item with the category name. The present experiments have attempted to provide a structural principle for the formation of prototypes; family resemblance relationships are not in contradiction to, but, rather, themselves offer a possible structural reason behind associative strength.

The principle of family resemblance is similar but not identical to two recent accounts of prototype effects: the attribute frequency model (Neumann, 1974) and an element tag model (Reitman & Bower, 1973). Both of these models were designed to account, without recourse to an "abstraction process," for the findings of several specific previous experiments—primarily those of Bransford and Franks (1971) and Franks and Bransford (1971). Both models predict memory (particularly the mistaken memory for prototype items which were not actually presented) from the frequency with which elements appear in a learning set.

A family resemblance account of prototypes is of greater generality than these models. In the first place, it accounts for prototypes in terms of distributions of attributes rather than in terms of the simple frequency of attributes (a factor which also distinguishes family resemblances from a narrow definition of cue validity). In the second place, it includes an account of the distribution of attributes over contrasting categories rather than focusing only on the category in question. That it is distribution rather than simple frequency of attributes which is most relevant to prototypes in natural categories is argued by two facts: (a) The measure of distribution used in the present study was highly correlated with ratings of prototypicality for superordinate categories, whereas, a measure of the frequency of items (which is necessarily correlated with frequency of attributes) in the category is not correlated with prototypicality (Mervis, Rosch, & Catlin, Note 3), and (b) The overlap of attributes with contrasting categories is itself a distributional property not a property of simple frequency. (In the artificial categories of Experiment 5 of the present paper, distributional and simple frequency were equivalent; however, in the other experiments, they were not—clarification of the relations between distribution and frequency of attributes is an issue which requires further research.) That the distribution of attributes over contrasting categories is as important a principle of prototype formation as distribution of attributes within a category is argued by the results of Experiments 2, 4, and 6.

At this point, it should be reiterated that the principle of family resemblance, as defined in the present research, is a descriptive, not a processing principle. Family resemblances are related to process models in two ways: (a) Any account of the processes by which humans convert stimulus attributes into mental or behavioral prototypes (such as an attribute tag model) should be able to account for the family resemblance attribute structure of categories outlined by the present research, and (b)

Classification by computation of cue validity and classification by matching to a prototype have been treated as alternative process models which are in conflict; however, the principle of family resemblance suggests that, for natural categories, both should be aspects of the same processing model.

Family resemblance as an argument for the compatibility of cue validity and prototype models. Probability models, such as cue validity, and distance models, such as matching to a prototype, have been treated as two fundamentally different forms of categorization model whose conflicting validities must be tested by empirical research (Reed, 1972). However, the present study has shown that empirically defined prototypes of natural categories are just those items with highest cue validity. Such a structure of categories would, in fact, appear to provide the means for maximally efficient processing of categories. Computation and summation of the validities of individual cues is a laborious cognitive process. However, since cue validity appears to be the basis of categories (Rosch *et al.*, in press), it is ecologically essential that cue validities be taken into account, in some manner, in categorization. If prototypes function cognitively as representatives of the category and if prototypes are items with the highest cue validities, humans can use the efficient processing mechanism of matching to a prototype without sacrificing attention to the validity of cues. (Note that such an account is similar to the compromise model which ultimately proved the most predictive for Reed's 1972 categories of schematic faces—a prototype matching model in which the importance of each feature in the prototype was weighted in accordance with its cue validity.) In short, humans probably incorporate probabilistic analysis of cues and computation of distance from a representation of the category into the same process of categorization; future research on categorization would do well to attempt to model the ways in which that incorporation can occur rather than to treat cue validity and prototypes as conflicting models.

Family resemblance as a basis for proximity scaling. Just as it has been customary to treat categories in terms of logical defining features which were assumed to be common to all members of the category, it is also not uncommon to treat proximity scaling of items in categories only as a means of determining the general dimensions along which items of the category are seen to differ. However, the results of the multidimensional scaling of the items of the superordinate categories in Experiment 1 (performed with Smith, Shoben, and Rips) indicated that family resemblance was predictive of centrality of items in the derived similarity space regardless of interpretability of dimensions or of item clusters. It should, in general, be the case that the more that items have in common with other items in a class (the closer the items are to all other items irrespective of the basis of closeness), the more central those items will be in a space derived from proximity measures. The demonstration of the importance of family resemblances (and of prototypicality) in classification provided by the present research suggests that the dimension of centrality may itself be an important aspect of and deserve to be a focus of attention in the analysis of proximity spaces.

Family resemblance as a part of the general process of category formation. The concept of family resemblances is also of general use because it characterizes prototype formation as part of the general process by which categories themselves are formed. It has been argued by Rosch *et al.*, (in press) that division of the world into categories is not arbitrary. The basic category cuts in the world are those which separate the information-rich bundles of attributes which form natural discontinuities. Basic categories have, in fact, been shown to be the most inclusive categories in which all items in the category possess significant numbers of attributes in common, and, thereby, are used by means of similar sequences of motor movements and are like each other in overall appearance. Basic categories are the categories for which the cue validity of attributes within categories is maximized since superordinate categories have fewer common attributes within the category than do basic categories and subordinate categories share more attributes with contrasting categories than do basic categories. Basic categories are, thus, the categories which mirror the correlational structure of the environment.

The present study has shown that formation of prototypes of categories appears to be likewise nonarbitrary. The more prototypical a category member, the more attributes it has in common with other members of the category and the less attributes in common with contrasting categories. Thus, prototypes appear to be just those members of the category which most reflect the redundancy structure of the category as a whole. That is, categories form to maximize the information-rich clusters of attributes in the environment and, thus, the cue validity of the attributes of categories; when prototypes of categories form by means of the principle of family resemblance, they maximize such clusters and such cue validity still further within categories.

Family resemblance as a link with children's classifications. The principle of family resemblances in adult categories casts a new perspective on children's classifications. Young children have been shown to classify objects or pictures by means of *complexive classes,* that is, classes in which items are related to each other by attributes not shared by all members of the class (Bruner, Olver, & Greenfield, 1966; Vygotsky, 1962). For example, Vygotsky (1962) speaks of the child in the "phase of thinking in complexes" starting with a small yellow triangle, putting with it a red triangle, then a red circle—in each case matching the new item to one attribute of the old. Bruner *et al.* describe the young child's tendency to classify by means of "complexive structures," for example, "banana and peach are yellow, peach and potato round. . . ." Such complexive classes have been considered logically more primitive than the adult preferred method of grouping taxonomically by "what a thing is"—that is, grouping by superordinate classes and justifying groups by their superordinate names. However, the present research has shown that family resemblances, a form of complexive grouping, appears to be one of the structural principles in the composition of the superordinate classes themselves, and, thus, one of the structural principles in adult classification. Since adult taxonomic classes such as *furniture* or *chair* themselves consist of complexive groupings of attributes, it would appear appropriate to study the development of the integration of complexive into taxonomic categories rather than the replacement of the former by the latter.

Family resemblance as a logical alternative to criterial attributes. There is a tenacious tradition of thought in philosophy and psychology which assumes that items can bear a categorical relationship to each other only by means of the possession of common criterial attributes. The present study is an empirical confirmation of Wittgenstein's (1953) argument that formal criteria are neither a logical nor psychological necessity; the categorical relationship in categories which do not appear to possess criterial attributes, such as those used in the present study, can be understood in terms of the principle of family resemblance.

REFERENCES

Beach, L. R. Cue probabilism and inference behavior. *Psychological Monographs*, 1964, **78**, 1–20.

Berlin, B., & Kay, P. *Basic color terms: Their universality and evolution*. Berkeley: University of California Press, 1969.

Bransford, J. D., & Franks, J. J. Abstraction of linguistic ideas. *Cognitive Psychology*, 1971, **2**, 331–350.

Bruner, J. S., Olver, R. R., & Greenfield, P. M. *Studies in cognitive growth*. New York: Wiley, 1966.

Ekman, P. Universals and cultural differences in facial expressions of emotion. In J. K. Cole (Ed.), *Nebraska symposium on motivation*. Lincoln NE: University of Nebraska Press, 1971.

Frake, C. O. The ethnographic study of cognitive systems. In S. A. Tyler (Ed.), *Cognitive anthropology*. New York: Holt, Rinehart & Winston, 1969.

Franks, J. J., & Bransford, J. D. Abstraction of visual patterns. *Journal of Experimental Psychology*, 1971, **90**, 65–74.

Garner, W. R. *The processing of information and structure*. New York: Halsted Press, 1974.

Heider, E. R. Universals in color naming and memory. *Journal of Experimental Psychology*, 1972, **93**, 10–20.

Henley, N. M. A psychological study of the semantics of animal terms. *Journal of Verbal Learning and Verbal Behavior,* 1969, **8,** 176–184.

Kucera, H. K., & Francis, W. N. *Computational analysis of present-day American English.* Providence, RI: Brown University Press, 1967.

Lakoff, G. Hedges: A study in meaning criteria and the logic of fuzzy concepts. *Papers from the Eighth Regional Meeting, Chicago Linguistics Society.* Chicago: University of Chicago Linguistics Department, 1972.

Loftus, E. F., & Scheff, R. W. Categorization norms for fifty representative instances. *Journal of Experimental Psychology,* 1971, **91,** 355–364.

Mervis, C. B., Catlin, J., & Rosch, E. Development of the structure of color categories. *Developmental Psychology,* 1975, **11,** 54–60.

Neuman, P. G. An attribute frequency model for the abstraction of prototypes. *Memory and Cognition,* 1974, **2,** 241–248.

Posner, M. I. *Cognition: An introduction.* Glencoe, IL: Scott, Foresman, 1973.

Reed, S. K. Pattern recognition and categorization. *Cognitive Psychology,* 1972, **3,** 382–407.

Reitman, J. S., & Bower, G. H. Storage and later recognition of concepts. *Cognitive Psychology,* 1973, **4,** 194–206.

Rips, L. J., Shoben, E. J., & Smith, E. E. Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior,* 1973, **12,** 1–20.

Rosch, E. On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language.* New York: Academic Press, 1973.

Rosch, E. Linguistic relativity. In A. Silverstein (Ed.), *Human communication: Theoretical perspectives.* New York: Halsted Press, 1974.

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

Rosch, E. Universals and cultural specifics in human categorization. In R. Brislin, S. Bochner, & W. Lonner (Eds.), *Cross-cultural perspectives on learning.* New York: Halsted Press, 1975. (b)

Rosch, E. Classifications of real-world objects: Origins and representations in cognition. *Bulletin de Psychologie,* in press-a.

Rosch, E. Cognitive reference points. *Cognitive Psychology,* in press-b.

Rosch, E. Human categorization. In N. Warren (Ed.), *Advances in cross-cultural psychology.* (Vol. 1) London: Academic Press, in press-c.

Rosch, E. Cognitive representations of semantic categories. *Journal of Experimental Psychology: Human Perception and Performance,* in press-d.

Rosch, E. The nature of mental codes for color categories. *Journal of Experimental Psychology: Human Perception and Performance,* in press-e.

Rosch, E., Mervis, C. B., Gray, W., Johnson, D., & Boyes-Brian, P. Basic objects in natural categories. *Cognitive Psychology,* in press.

Shepard, R. N. The analysis of proximities: Multidimensional scaling with an unknown distance function. I and II. *Psychometrika,* 1962, **27,** 125–140, 219–246.

Shepard, R. N., Romney, A. K., & Nerlove, S. B. *Multidimensional scaling: Theory and applications in the behavioral sciences.* (Vols I. & II:). New York: Seminar Press, 1972.

Smith, E. E., Rips, L. J., & Shoben, E. J. Semantic memory and psychological semantics. In G. H. Bower (ED.), *The psychology of learning and motivation.* (Vol. 8.) New York: Academic Press, 1974.

Smith, E. E., Shoben, E. J., & Rips, L. J. Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review,* 1974, **81,** 214–241.

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

Wittgenstein, L. *Philosophical investigations.* New York: Macmillan, 1953.

Zadeh, L. A. Fuzzy sets. *Information and Control,* 1965, **8,** 338–353.

REFERENCE NOTES

1. McDaniel, C. K. *Hue perception and hue naming*. Unpublished B.A. thesis, Harvard College, 1972.
2. Rosch, E., Simpson, C., & Miller, R. S. *Structural bases of typicality effects*. Manuscript submitted for publication, 1975.
3. Mervis, C. B., Rosch, E., & Catlin, J. *Relationships among goodness-of-example, category norms, and word frequency*. Unpublished manuscript, 1975. (Available from the second author.)

(Accepted June 3, 1975)