Similarity- versus rule-based categorization

EDWARD E. SMITH

University of Michigan, Ann Arbor, Michigan

and

STEVEN A. SLOMAN

Brown University, Providence, Rhode Island

An influential study by Rips (1989) provides the strongest evidence available that categorization cannot be reduced to similarity. In Rips's study, subjects were presented a *sparse* description of an object that mentioned only a value on a single dimension (e.g., "an object 3 inches in diameter"), followed by two categories (e.g., PIZZA and QUARTER), where one allowed more variability on the relevant dimension than did the other (the diameter of pizzas is more variable than that of quarters). Subjects judged the described object to be more likely to be a member of the variable category (PIZZA), but more similar to the nonvariable category (QUARTER). This dissociation between categorization and similarity strongly implies that categorization was not based on similarity. In our first experiment, we used sparse descriptions like Rips's, as well as *rich* descriptions that contained features characteristic of the nonvariable category. We found that categorization tracked similarity with both kinds of descriptions. In a second experiment, we modified our procedure to be more like that of Rips's by requiring subjects to think aloud while making their decisions. Now, like Rips, we found a dissociation between similarity and categorization with sparse items; with rich descriptions, categorization again tracked similarity. These findings are consistent with the hypothesis that categorization can be done in two ways, by *similarity* and by *rule*. An exclusive reliance on rule-based categorization seems to occur only when the description of the to-be-categorized object does not contain any features characteristic of any relevant category.

Similarity is at the heart of a variety of categorization procedures. Some theories posit that we decide that a test object belongs in a category if it is sufficiently similar to an abstract prototype of the category, or to known instances of the categories. Other theories claim that similarity is determined by the match between the object's features and those of the connection strengths that encode the categories. All of these theories are consistent with the extensive evidence for "similarity-based categorization" found in studies of artificial categories, involving schematic faces and complex geometric forms (e.g., Estes, 1986; Nosofsky, 1986), and in studies that use natural categories (e.g., Smith & Medin, 1981).

Recently, numerous researchers have challenged the prevalence and importance of similarity-based categorization (e.g., Carey, 1985; Keil, 1989; Medin, 1989; Murphy & Medin, 1985; Rips, 1989). These authors have argued that often a different kind of procedure is used for

categorization, one that is driven by the categorizer's lay theory of the relevant domain. Such a procedure is particularly likely to be operative when the test object, or description of it, includes causal information about the object's origin or function that is highly diagnostic. To illustrate, consider an object described as an animal that has a white stripe down its back and that gives off a pungent odor, and that was born of raccoon parents (Keil, 1986). "Being born of raccoon parents" is both a causal and a necessary condition for being a RACCOON (capitals indicate categories). It is a necessary feature of RACCOON in that our lay theory of animals demands that they must be of the same species as their parents. Presumably, adult subjects recognize the special status of this feature, focus on it to a greater extent than on other features, and therefore categorize the test object as a RACCOON. Thus the processing involves selecting out features of an input that are necessary for some category and then intentionally applying a rule of the general form, "If necessary feature *Y*, then category *X*" (Smith, Langston, & Nisbett, 1992). We refer to this alternative as *rule-based categorization*, which we contrast to *similarity-based categorization*. In rule-based categorization, then, one decides whether a test object belongs in a category by selecting out certain special features and determining whether the object satisfies a rule suggested by these features.

A comment is in order about our characterizing the processing of interest as rule-based. Calling it "rule-based" captures the ideas that the processing is more an-

This research was supported by Air Force Contract No. AFOSR-92-0265 to E.E.S., and by a Natural Sciences and Engineering Research Council of Canada postdoctoral fellowship to S.A.S. We thank Linda Covington for her assistance in performing the experiments, and members of our lab group and reviewers for this journal for helpful comments on an earlier version of the manuscript. We are particularly indebted to Lance Rips for his advice and insights throughout the course of this research. Address correspondence to E. E. Smith, Department of Psychology, University of Michigan, 330 Packard Rd., Ann Arbor, MI 48104.

alytic and reflective than similarity-based categorization. But the notion of a rule may have some additional meaning that is not relevant here. Talk of a “born-of-raccoon-parents” rule may suggest this is the only means we have of decisively categorizing an object as a RACCOON, or that we keep this means separate from others that we possess. Clearly, though, there are several causal features that are equally necessary for RACCOON—having raccoon children, having a certain genetic structure, having a certain configuration of internal organs, and so forth—and all of these features are integrated in one’s lay theory of RACCOON. Yet although a lay theory is involved here, the application of each of its component beliefs seems rule-like. Hence, our choice of “rule-based” over “theory-based.” Other possible characterizations of the process of interest, such as “explanation-based,” are too general to be of much use here.¹

The clearest evidence for rule-based categorization, and for its difference from similarity-based categorization, comes from some studies of Rips (1989). In one experiment, on each trial a subject was presented a description of a test object that mentioned only a value on a single dimension (say, the object’s diameter). Then the subject decided which of two categories the object belonged to (it had been previously established that the object was between that subject’s extreme values for the two categories). For example, one description might be “a circular object with a 3-in. diameter,” with the associated categories being QUARTER and PIZZA. The description was constructed to be smaller than the smallest PIZZA in that subject’s experience, but larger than the largest QUARTER. Since quarters are severely restricted in size but pizzas are not, this suggests a rule that implies that the object is an instance of PIZZA. The rule might state, “If an object is more than 1 inch in diameter, it cannot be a quarter”; by default, then, the object must be a PIZZA. All of the items had this structure: one category—the “variable” category—always allowed more variability on the relevant dimension than did the other category—the “fixed” one. The important findings were that subjects were more likely to categorize the test objects as members of the variable categories (PIZZA in our example), but that they rated the objects as more similar to the fixed categories (QUARTER).

Because categorization decisions favored one category but similarity decisions favored the other, the categorization decisions could not have been based on similarity. Thus Rips (1989) managed to show a dissociation between categorization and similarity. This is the best documented dissociation of this sort, and hence it is of particular importance. What likely lies behind this dissociation is that variability information has more effect on categorization than on similarity judgments. Furthermore, the restricted variability of the fixed category on the relevant dimension amounts to a necessary feature, and conformity to the rule based on this feature was presumably the major basis of categorization.

Thus, people seem capable of categorizing common objects on the basis of either similarity or a rule. Given

this, the question arises as to what specific aspects of a situation trigger similarity- versus rule-based processing. In the absence of an answer to this question, investigators will be unsure of what *kind* of categorization they are studying, and they will run the risk of mixing results from different kinds of categorizations and of obscuring any true generalizations that can be drawn. Applied to Rips (1989) study, the question of interest becomes what aspects of Rips’s task led to categorization that was rule based rather than similarity based.

One possibility is that whenever a necessary feature is available, people will use it as the sole basis for their final categorization. This claim is quite explicit in some treatments of theory-driven categorization (e.g., Keil, 1989), and we refer to it as the *necessary feature* hypothesis. An alternative hypothesis stems from an observation about Rips’s items. Though the description of each test object was diagnostic of the variable category (because it violated a necessary feature of the fixed category), it was *characteristic* of neither category (a feature is characteristic of a category to the extent that it is associated with that category). For example, although “3 in. in diameter” is diagnostic of the object’s being a PIZZA, it is not characteristic of either PIZZA or QUARTER. Perhaps people rely solely on rule-based categorization when they have no information that is characteristic of the relevant categories. Or to put it the other way around: once information characteristic of a relevant category is available, similarity-based processing may be triggered, and this categorization procedure, perhaps along with a rule-based one, will figure in the final decision. We refer to this possibility as the *characteristic feature* hypothesis. Note that, according to this hypothesis, an exclusive reliance on rules might be relatively rare—not only does it require that a necessary feature be present, but also that no characteristic feature be available. Hence Rips’s (1989) task may be a rarefied one.

The purpose of the present studies was to contrast the necessary feature and characteristic feature hypotheses and thereby to shed light on the determinants of similarity- and rule-based categorization. To make this contrast, we altered Rips’s (1989) items by adding to his object descriptions a feature characteristic of the fixed categories. For example, the description accompanying QUARTER and PIZZA might be “a circular object with a 3-in. diameter that is silver colored.” The description again contains the necessary feature that favors the variable category PIZZA (3-in. diameter), but also contains a characteristic feature (silver colored) that is highly associated with the fixed category QUARTER. According to the necessary feature hypothesis, subjects should favor the variable category in categorization as much with these altered descriptions as they do with the original items (because they are still relying exclusively on the same rule). According to the characteristic feature hypothesis, subjects should favor the variable category less with the altered descriptions than with the original ones (because they are relying partially on similarity).

EXPERIMENT 1

Following Rips's (1989) procedure, one group of subjects made similarity judgments, while a second group made categorization judgments. In each group, in half the items the description contained only a necessary feature ("sparse" descriptions); in the other half of the items, the description contained both a necessary and a characteristic feature ("rich" descriptions). For the sparse descriptions, we expected to replicate Rips's findings that categorization decisions favor the variable categories but similarity decisions do not. For the rich descriptions, we expected that categorization decisions would be more like similarity decisions. In both conditions, all information was presented by a computer that also recorded responses.

Method

Materials. Consider first the items that had sparse descriptions. A total of 30 items were used. Of these, 25 were selected from the items that Rips (1989) used, while the remaining 5 were newly constructed to have the same general structure as Rips's items (all items are contained in the Appendix). To illustrate, 2 of the items had the following form:

- (1) Circular object with an X-inch diameter
PIZZA QUARTER
- (2) Object X inches long
FOOT RULER PAPER CLIP

The *X* indicates that the critical value was instantiated anew for each subject. That is, prior to seeing an item, a subject was presented with the two relevant categories and he/she then estimated numerical values for them; thus, for (1), the subject estimated the diameter of the smallest pizza and that of the largest quarter, whereas for (2), the subject estimated the length of the shortest foot ruler and that of the longest paper clip. (This is the same procedure that Rips used.) The computer then determined the average of these two estimates, and that number was used to instantiate the item. In this manner, we ensured that, for each subject, the value in each description fell midway between the values associated with the appropriate end points of the fixed and variable categories (QUARTER and PIZZA in [1]).

From each sparse item, we created a rich item by adding to the original description a feature characteristic of the fixed category. We can illustrate with variants of the examples above:

- (1') Circular object with an X-inch diameter that is silver colored
PIZZA QUARTER
- (2') Object X inches long made of wood
FOOT RULER PAPER CLIP

By our judgment, the added feature is never as necessary—never as "essential"—of the fixed category as the original feature is of the variable category. Thus, while an object 3 in. in diameter must be a pizza rather than a quarter (because it is too big to be a quarter), an object that is silver colored is only more likely to be a quarter than a pizza, for nothing about the nature of pizzas precludes them from being silver in appearance. Likewise, an object 7 in. long must be a paper clip rather than a foot ruler (because it is too short to be a foot ruler), while an object made of wood is only more likely to be a foot ruler than a paper clip (because wooden paper clips are certainly possible). But while the added feature is less necessary than the original one, it is clearly more characteristic of one of the categories (the fixed one) than is the original feature.

Hence our designation of the original feature as "necessary," and of the added feature as "characteristic."

Procedure. Consider first the procedure for the similarity condition. On each trial, subjects first estimated numerical values for the two relevant categories. Then they were presented the instantiated description and beneath it, the two categories arrayed horizontally. Subjects were instructed to "choose the category that the described object is more similar to." Subjects indicated their choice by pressing either "1" or "0" on the computer keyboard, where a "1" indicated a choice of the category on the left and a "0" indicated a choice of the category on the right. Subjects were also informed that the number that appeared in each description was the average of their two estimates for that trial. Each subject was given 30 trials, half with items containing sparse descriptions and half containing rich descriptions. A given subject never saw the sparse and rich descriptions of the same object. Sparse and rich items were intermixed, with the order of the 30 items being randomized anew for each subject.

The procedure for the categorization condition was comparable to the preceding. Again each trial started with the subject estimating values for the two categories, followed by presentation of the whole item. But this time the instantiated description appeared alone first, and once the subject indicated that he/she had finished reading it (by pressing a "Ready" button), the two categories appeared, arrayed horizontally. The subjects were instructed to "choose the category that the described object is more likely to be," and to indicate their choice by pressing either a left-hand or right-hand button. The subjects were further instructed to respond "as quickly as you can WHILE BEING AS ACCURATE AS POSSIBLE." Other aspects of the procedure—for example, the number of trials, the intermixing of sparse and rich descriptions—were the same as in the similarity condition.²

The remaining instructions for the categorization condition strongly encouraged rule-based categorization (we did this because pilot work suggested that subjects rarely made categorization decisions on the basis of rules). Specifically, subjects were told:

Please be aware that for each item there is in fact always a correct answer. Sometimes the correct answer won't be obvious. As an example, consider the following [item]: A black animal with a stripe down its back, it can sometimes give off an odor, and its parents are raccoons.

RACCOON SKUNK

Even though the description might seem at first glance like that of a skunk, the animal has to be a raccoon because its parents are raccoons. So the correct answer in this case would be to press the button on the left, indicating RACCOON.

In addition to the biasing instructions, categorization subjects were also given seven practice trials with feedback, which were intended to further encourage rule-based categorization in general, and to highlight the notion of a necessary feature in particular. The following illustrates one practice item and its associated feedback:

A piece of sheet metal with a rough edge used to pound in nails.

HAMMER SAW

It has the function of pounding in nails, it must be a hammer.

Subjects. Forty University of Michigan undergraduates were paid for their participation; 20 served in the similarity condition, and 20 in the categorization condition.

Results and Discussion

The data of interest are the percentage of trials on which subjects chose the variable category, which is the answer dictated by rule-based categorization—PIZZA in our standard example. Table 1 presents these percentages, averaged over subjects, separately for similarity

and categorization decisions about sparse and rich items. All trials in which subjects judged the smallest member of the larger category to be smaller than the largest member of the smaller category were removed before all analyses. Such reversed estimates presumably reflect a lack of relevant knowledge on the subject's part. (Reversed estimates occurred on 24% of the similarity trials and 23% of the categorization trials; data for these trials are also presented in Table 1.) Also included in Table 1 are the comparable numbers from Rips (1989). Two facts are immediately apparent. First, our results with sparse items failed to replicate the categorization-similarity dissociation. Second, variable categories were chosen less frequently with rich than with sparse items. (Both of these conclusions seem to hold up in the data that were excluded because of reversed estimates.)

Both conclusions are also supported by statistical analyses. An analysis of variance (ANOVA) with subjects as the random effect showed no effect of type of decision (similarity vs. categorization) [$F(1,38) < 1$], a significant effect of description richness [$F(1,38) = 91.93$, $p < .01$], and no interaction between these factors [$F(1,38) = 1.20$, $p > .20$]. Subsequent contrasts also failed to find any hint of a difference between the similarity and categorization conditions [for both sparse and rich items, $t(16) < 1$]. Comparable results were obtained in an ANOVA with items as the random effect: no effect of type of decision [$F(1,57) < 1$], a significant effect of description richness [$F(1,57) = 40.33$, $p < .01$], and no interaction between the two factors [$F(1,57) < 1$]. Subsequent contrasts again showed no hint of a difference between the similarity and categorization conditions [for both sparse and rich items, $t(26) < 1$]. Even if we restrict our analyses to just those 25 items that Rips (1989) used, we find no effect of type of decision.

Rather than a dissociation between categorization and similarity, our results for the sparse items suggest a correlation between the two tasks. To pursue this point, we determined the product moment correlation, across the 30 sparse items, between choices of the variable category in the categorization condition and such choices in the similarity condition. The resulting r , .61, is highly significant ($p < .001$). This apparent reliance of categorization on similarity is striking in light of our heavy-handed instructions to encourage rule-based categorization and suggests that similarity-based categorization may be hard to inhibit. (See Allen & Brooks, 1991, for comparable findings with artificial categories.)

Table 1
Average Percentage of Trials on which Subjects
Chose the Variable Category in Experiment 1,
Along With Those in Rips (1989)

Decision Type	Item Type				Rips
	Sparse		Rich		
	Avg.	Rev. Est.	Avg.	Rev. Est.	
Categorization	55	58	23	19	63
Similarity	52	55	26	15	31

Note—Rev. est. = percentages based on the data from trials with reversed estimates (see text).

Returning to the data in Table 1, a specific comparison of our results for sparse items with those of Rips (1989) is illuminating. In comparison with Rips, we found more choices of the variable category in the similarity condition, but fewer choices of the variable category in the categorization condition (see Table 1). The first of these results may be a consequence of our including items with rich descriptions along with items with sparse descriptions. With a rich description there is a substantial difference in similarity between the description and the fixed category versus that between the description and the variable category; for example, a *circular object with a 3-in. diameter that is silver colored* is much more similar to QUARTER than it is to PIZZA. Hence, rich descriptions give rise to relatively large similarity differences. In contrast, sparse descriptions give rise to relatively small similarity differences; a *circular object with a 3-in. diameter*, for example, is not much more similar to QUARTER than to PIZZA. In the context of rich descriptions and large similarity differences, similarity differences with sparse items that favor the fixed category (as Rips found) would have been perceived as minimal. This contrast effect would account for our findings that, with sparse items, similarity subjects chose each category roughly 50% of the time, whereas with rich items, similarity subjects clearly favored the fixed category.

What remains to be explained about the data in Table 1 is why our subjects did not favor the variable category in categorization despite instructions encouraging them to do so. Or to put it at a theoretical level, why did our subjects *not* engage in rule-based categorization as much as Rips's subjects did? Again the difference in results may be due to a procedural difference. In Rips's study, subjects in the categorization condition were instructed to talk aloud as they made their choices. This instruction may have suggested to subjects that they base their categorization decisions on factors that could be articulated and that seemed principled; this could have fostered rule-based processing. Along similar lines, there was no mention of speed of response in Rips's instructions (as there was in ours), and a lack of concern with speed might also encourage a more reflective or rule-based approach. These possibilities were tested in our next experiment.

In sum, our failure to replicate Rips and find evidence for rule-based categorization with sparse items calls into question the necessary feature hypothesis. But this conclusion depends on a null result. Before we can assess the triggering conditions of a phenomenon, we should be able to reproduce it.³

EXPERIMENT 2

Method

This study was identical to the preceding one except that subjects in both the categorization and the similarity conditions were now instructed to talk aloud while making their decisions, and there was no mention of speed in the instructions nor any attempt to surreptitiously time subjects' responses. Subjects' verbal pro-

protocols were recorded and subsequently analyzed. Having subjects give protocols made the use of a computer awkward, and consequently the study was conducted entirely by a human experimenter (who, among other things, on each trial determined the average of a subject's estimates for the two categories). Again, 20 subjects served in the similarity condition, and 20 in the categorization condition. For sparse descriptions, we expected to replicate Rips's (1989) finding that categorization decisions favor the variable categories but similarity decisions do not, whereas for rich descriptions, we expected categorization and similarity decisions to be more alike.

Results and Discussion

Similarity and categorization decisions. Table 2 presents the average percentage of trials on which subjects chose the variable category, separately for similarity and categorization decisions about sparse and rich items. Again, we excluded from the analyses all trials in which subjects' size estimates were not in the order required by the logic of the study. (Reversed estimates occurred on 21% of the similarity trials and 17% of the categorization trials; data for these trials are also presented in Table 2.) The comparable data from Rips (1989) are again included in the table. The results for sparse items now show the kind of phenomenon that Rips found. Whereas similarity decisions were equally divided among variable and fixed categories, categorization decisions favored the variable category. Thus, like Rips, we have obtained a dissociation between similarity and categorization judgments that is indicative of rule-based categorization. Of course, unlike us, Rips (1989) found that similarity decisions actually favored the fixed categories. But as we pointed out in the discussion of our previous experiment, our failure to obtain this similarity result may have been due to a contrast effect resulting from our inclusion of rich as well as sparse descriptions. For rich descriptions, both similarity and categorization decisions favored the fixed category, and to roughly the same degree. This pattern of results is consistent with the characteristic feature hypothesis, and inconsistent with the necessary feature hypothesis.⁴

Statistical analyses support these conclusions. An ANOVA with subjects as the random effect showed main effects of type of decision [$F(1,38) = 9.22, p < .01$] and description richness [$F(1,38) = 50.37, p < .01$]. Although the interaction between these two factors failed to reach an acceptable level of significance [$F(1,38) = 1.78, .10 < p < .20$], subsequent contrasts provided some evidence that similarity and categorization decisions

were based on different factors when the items were sparse but not when they were rich. For sparse items, categorization decisions differed significantly from similarity decisions [$t(16) = 2.86, p < .05$]; for rich items, categorization and similarity decisions did not differ [$t(16) < 1$]. We observed a similar pattern of results in an ANOVA with items as the random effect. There were main effects of type of judgment [$F(1,58) = 3.47, p = .07$] and of item richness [$F(1,58) = 58.94, p < .01$]. Again, although the interaction between the two factors was not significant [$F(1,58) < 1$], subsequent contrasts showed a significant difference between similarity and categorization decisions for sparse items [$t(26) = 2.27, p < .05$], but not for rich items [$t(26) = 1.25, p > .20$].

Protocol analysis. The protocols, which tended to be rather brief, were analyzed for which features were mentioned. For sparse items, the possibilities included (1) mention of the necessary feature (diameter value in the PIZZA-QUARTER example) *plus* mention of the inferred rule concerning limited variability (e.g., "QUARTERS can't be that large"); (2) mention of just the necessary feature; or (3) no explicit mention of either. The results for sparse items are presented in Table 3, with the first two columns providing results for categorization decisions and the next two columns for similarity decisions. Within each pair of columns, the first presents the number of trials appearing in each cell, and the second presents the percentage of choices of the variable category, conditionalized on each possibility.

Starting with categorization decisions, note that mention of a necessary feature plus the corresponding rule happens infrequently, on only 46 of 248 opportunities. Assuming that mention of a necessary feature along with the corresponding rule is a signature of rule-based categorization, such categorizations seem to occur infrequently even with sparse items. When the necessary feature and rule are mentioned, the variable category is virtually always chosen (see the second column of Table 3); this finding strengthens our belief that mention of the feature plus rule is indeed an indicator of rule-based categorization. However, occasional cases of rule-based categorization may have occurred that were unaccompanied by explicit mention of limited variability. For when the protocols mention only the necessary feature or no feature at all, categorization decisions are not evenly divided between the two categories, but rather favor the variable category (see the second column of Table 3). This suggests that subjects sometimes engaged in rule-based categorization but neglected to mention the necessary feature or the corresponding rule. The percentage of trials on which this occurred, given some simple assumptions, would be on the order of 20%, which would still put the total incidence of rule-based categorization at less than 50%.⁵

The data for similarity decisions with sparse items (last two columns of Table 3) permit a more straightforward interpretation. Only rarely do subjects explicitly consider variability in their decision (16 of 228 cases), but when they do, they inevitably select the variable cat-

Table 2
Average Percentage of Trials on which Subjects
Chose the Variable Category in Experiment 2,
Along With Those in Rips (1989)

Decision Type	Item Type				Rips
	Sparse		Rich		
	Avg.	Rev. Est.	Avg.	Rev. Est.	
Categorization	67	66	32	23	63
Similarity	50	55	26	29	31

Note—Rev. est. = percentages based on the data from trials with reversed estimates (see text).

Table 3
Analysis of Verbal Protocols in Terms of Features Mentioned
During Categorization and Similarity Decisions:
Sparse Items (Experiment 2)

Features Mentioned	Decision Type			
	Categorization		Similarity	
	No. Trials	% Choices Variable	No. Trials	% Choices Variable
Necessary feature & rule	46	96	16	100
Necessary feature	122	58	130	45
No feature or rule	80	65	82	49
	$\Sigma = 248$		$\Sigma = 228$	

egory, suggesting they are engaging in a kind of rule-based processing. On the vast majority of similarity trials, subjects do not mention variability and select the two categories about equally often. Putting together the similarity and categorization results for sparse items, we are led to tentatively conclude that rule-based decisions occur more frequently in categorization than in similarity, but that they occur relatively infrequently even in categorization.

We turn to the protocol results for rich items, presented in Table 4. Now we have five possibilities to consider: (1) mention of the necessary feature plus mention of the rule (in the PIZZA-QUARTER example, diameter value plus limited variability); (2) mention of just the necessary feature (diameter value); (3) mention of just the characteristic feature (silver colored); (4) mention of both the necessary and the characteristic features; and (5) no explicit mention of any of these. The first two columns of Table 4 give the results for categorization decisions, and the latter two columns, the results for similarity decisions.

We start with the categorization results, using mention of the necessary feature plus its corresponding rule as a signature of rule-based categorization. This kind of categorization occurred even more infrequently than it did with sparse items (only 35 out of 246 opportunities), but when it did occur, subjects were likely to choose the variable category. When subjects focused instead on just the necessary feature, their choices were evenly divided between the two categories (unlike the situation with sparse items, when choices favored the variable category). On trials in which subjects mentioned the characteristic feature, or gave no features, they preferred the fixed category (roughly 70% of the time), which indicates a preference for the category that is more similar to the description. The data for similarity decisions with rich items (last two columns of Table 4) show that necessary features plus corresponding rules were virtually never mentioned (a mere 5 of 232 opportunities). With respect to other kinds of features mentioned, the similarity data are comparable to the categorization data. When necessary features are focused on, choices are roughly equally divided between the two categories, but when a characteristic feature is included in the focus or

no feature is explicitly mentioned, the fixed category is strongly preferred. Putting together the similarity and categorization results for rich items, we see that rule-based decisions occur very rarely, but more frequently in categorization than similarity; and, in cases in which the focus does not include a necessary feature, the choices in categorization match those in similarity. These results—particularly the extent to which categorization judgments mirror similarity judgments—seem completely at odds with the necessary feature hypothesis.

To support the preceding qualitative observations, we performed statistical analyses on the protocol data. Only trials in which subjects explicitly mentioned some feature were considered. First, collapsing over sparse and rich items, we found that subjects were more likely to mention a necessary feature plus corresponding rule—and, hence, presumably engage in rule-based processing—when making categorization decisions, 22%, than when making similarity-based decisions, 7% ($z = 5.47$, $p < .01$). This supports the key assumption explaining Rips's (1989) results. Second, given trials in which subjects explicitly mentioned a necessary feature plus rule, we can ask whether they inevitably chose the variable category (as assumed in Rips, 1989). The answer is clearly "yes" for sparse items, because they chose the variable category 96% of the time; for rich items, the answer is more equivocal, since the variable category was chosen only 88% of the time. Although these percentages are obviously close to 100%, the differences cannot be tested statistically because a population score of 100% has no variance. Third, given trials in which subjects focused on the necessary feature but did not mention the rule, we may ask whether categorization judgments tracked similarity judgments. For sparse items, the answer is "no." With individual subjects as the unit of analysis, the percentage choice of the variable category in categorization, 58%, exceeded that in similarity, 45% [$t(34) = 2.61$, $p = .01$]; this difference is what led us earlier to propose that subjects occasionally engage in rule-based categorization yet neglect to mention the

Table 4
Analysis of Verbal Protocols in Terms of Features Mentioned
During Categorization and Similarity Decisions:
Rich Items (Experiment 2)

Features Mentioned	Decision Type			
	Categorization		Similarity	
	No. Trials	% Choices Variable	No. Trials	% Choices Variable
Necessary feature & rule	35	88	5	60
Necessary feature	33	48	52	42
Characteristic feature	82	6	61	3
Necessary & characteristic	46	33	33	36
No feature or rule	50	20	81	28
	$\Sigma = 246$		$\Sigma = 232$	

rule. For rich items, though, categorization did track similarity. We performed a two-factor ANOVA on the percentage choices, with the factors being type of decision (similarity vs. categorization) and features mentioned (necessary vs. characteristic vs. both) and found no significant effects due to decision type or to the interaction of the two factors ($F < 1$ in both cases). These findings support the characteristic feature hypothesis, which assumes that once characteristic features are present, categorization begins to track similarity to at least some extent. Fourth and finally, we found that categorization subjects were more likely to engage in rule-based processing with sparse items, 27%, than rich ones, 18% ($z = 2.18, p < .05$). This finding also supports the characteristic feature hypothesis as opposed to the necessary feature hypothesis.

GENERAL DISCUSSION

Summary

Our empirical focus has been on a dissociation between categorization and similarity. With common objects of the sort we have studied, this dissociation occurs only under special circumstances. One such circumstance, or constraint, is that the description of the to-be-categorized object has to lack features that are characteristic of potential categories. The strongest evidence for this constraint comes from the choice data in Experiments 1 and 2; we never obtained a dissociation of categorization and similarity when the items contained a characteristic feature—that is, when they were items with rich descriptions. A second possible constraint on the dissociation is that the situation encourage people to articulate and explain their reasons for categorization, rather than encourage rapid judgments. Evidence for this second constraint comes from a comparison of the choice data between Experiments 1 and 2; we obtained the dissociation of interest for sparse items only when subjects had to talk aloud while making their decisions and when there was no hint of time pressure.

These findings seem incompatible with the necessary feature hypothesis, which claims that the mere presence of a necessary feature inevitably leads to total reliance on it in categorization decisions (which should result in widespread categorization–similarity dissociations). In contrast, our constraints on the dissociation fit nicely with the characteristic feature hypothesis, which holds that categorization decisions will be based on characteristic as well as necessary features, once the former are available.

Our major theoretical claims, then, are that the necessary feature hypothesis has failed, and that this failure is due to the fact that both similarity- and rule-based categorization may operate in the domain of common objects. In what follows, we consider rejoinders to each of these claims.

An Attempt to Rescue the Necessary Feature Hypothesis

Perhaps our evidence against the necessary feature hypothesis reflects not so much the falsity of this hy-

pothesis as an inappropriate selection of necessary features on our part. To use our standard example: for QUARTER, limited size is not as salient a necessary feature as “being a coin.” And for natural-kind categories, such as CAT, there seem to be clear-cut necessary features, such as information about parentage (“parents were cats” or “gave birth to a kitten”), or development (“was once a kitten”), or biology (“has a certain genetic structure”). Furthermore, the latter features seem to be part of a person’s lay theory in a way that beliefs about coins are not. All of this makes one wonder what would have been the fate of the necessary feature hypothesis had we, for example, used only natural-kind categories and their appropriate necessary features.

Coincidentally, Rips (1989) performed a study that bears on this issue. In one condition of this study, subjects were told about an animal that initially had the characteristic features of a bird but then, because of chemical contamination, underwent changes in its surface characteristic features so that they now resembled those of an insect; eventually, the animal mated with another of its original kind and produced an offspring with the features of a bird. The fact about mating constitutes a necessary feature of BIRD. When asked to categorize the animal, subjects favored BIRD over INSECT. Thus we have a situation in which the characteristic features favor one category, INSECT, the necessary feature favors another category, BIRD, and subjects’ categorizations follow the necessary feature. According to the necessary feature hypothesis, the preference for BIRD should not increase if one were to add to the description features that are only characteristic of BIRD, but might well decrease if one were to remove the necessary feature about mating. Both of these predictions turn out to be wrong. In one of Rips’s control conditions, the description of the animal included only features characteristic of BIRD (no necessary feature of BIRD); subjects’ categorization decisions now favored BIRD more than they did in the original condition. Even with natural-kind categories, the necessary feature hypothesis failed.⁶

Still more evidence against the necessary feature hypothesis comes from recent studies by Malt (in press). She focused on the category WATER, for which H₂O presumably is the necessary feature. When subjects were asked which descriptions of liquids were “types of water,” their categorization decisions clearly reflected features in addition to H₂O content. The subjects considered the location of the liquid (domestic vs. “in the wild”), its immediate source (nature vs. human intervention), and its customary use. These three factors seem more like characteristic features, and so we have yet another demonstration that categorization considers characteristic as well as necessary features.

As best we can tell, then, there is no support at all for the idea that people restrict their categorization decisions to a few essential features. Yet this idea seems to play a key role in some recent discussions of theory-driven categorization. What is wrong with the necessary feature hypothesis is its assumption that people can ignore char-

acteristic features of a to-be-categorized object. This assumption is reminiscent of earlier models of categorization, which assumed that categorization consisted solely of determining whether a test object contained the *defining* features of the relevant concepts. Such models were undone, in part, by the evidence that categorization decisions consider features that are only characteristic of the categories (Smith, Shoben, & Rips, 1974).

Single- Versus Dual-Categorization Procedures

Given that characteristic as well as necessary features are considered in categorization, we have assumed throughout that one needs to posit both similarity-based and rule-based mechanisms (as in Allen & Brooks, 1991, and Smith et al., 1992). But there are alternative views that rely on only a single mechanism. One possibility is that all categorization is similarity based; another is that all categorizations result from inferential processes. We treat these possibilities in turn.

Similarity only. According to the similarity-only view, similarity and categorization decisions are both based on a similarity computation, but some of the features that enter into the computation may be weighted differently in the two tasks. Consider how this idea might apply to the results of Experiment 2. In the similarity condition, when presented a sparse description like *3-in. in diameter*, subjects might judge this feature to be roughly equally similar to the diameter feature of PIZZA and that of QUARTER. In the categorization condition, in contrast, subjects might require a closer match for a less variable feature of a category—the diameter feature of QUARTER might require a closer match than the diameter feature of PIZZA. As a consequence, the description would be less similar to QUARTER than to PIZZA, leading to the categorization of PIZZA. Hence, for sparse descriptions, the similarity and categorization conditions could lead to different decisions. With rich descriptions, the addition of a characteristic feature that matches QUARTER so well tips the similarity balance in favor of the latter, even in the categorization condition. Hence, categorization will track similarity with rich descriptions (B. Ross, personal communication).

There is an obvious problem with the preceding account: It rests on an ad hoc assumption—that closer matches for less variable features are more important for categorization than for similarity judgments. There seems to be no rationale for this assumption other than to explain the results of interest. In contrast, our dual-mechanism account does have some independent rationale. A rule-based mechanism should be invoked in categorization decisions but not similarity ones, because categorization decisions can be correct or incorrect and in need of justification, whereas similarity judgments lack these qualities.

Another aspect of the similarity-only view is worth noting. Although a similarity computation with variable feature weights can produce the same *outcome* as our dual-mechanism proposal, the *processing* hypothesized in the two views is different. According to the similarity-

only view, the same kind of matching process underlies similarity and categorization judgments. According to the dual-mechanism view, different kinds of processes underlie the two types of judgments. That is, while similarity judgments are presumably based almost exclusively on a matching process (this is similarity-based categorization), categorization judgments also reflect a rule mechanism, which involves first selecting out critical features and then applying a rule. These hypothesized processing differences might well serve as a way of distinguishing between single- and dual-mechanism views in subsequent research.

Inference only. Another alternative to our dual-mechanism view was suggested by Rips in his remarkable (1989) paper. The idea—which concerns only categorization decisions—is that categorization is theory driven not in the sense that it is rule based but in the sense that it is a species of “inference to the best explanation.” This means that the final categorization of an object is the best explanation of all available data, where “all available data” includes characteristic as well as necessary features. We can relate this idea to the categorization results in Experiment 2 as follows. With sparse descriptions, essentially there is only one datum to be explained, the necessary feature (e.g., 3 in. in diameter), and the variable category (PIZZA) provides a better account of it than does the fixed category (QUARTER). With rich descriptions, there are two pieces of data to be explained, the necessary and characteristic features (e.g., 3 in. in diameter and silver colored), and the variable category does not provide a particularly good account of them (because pizzas are rarely silver colored). On the other hand, the fixed category offers a better account of the rich data than it does of the sparse data (because quarters are usually silver colored); hence, our finding that choice of the variable category decreases in moving from sparse to rich descriptions.⁷

Although Rips's (1989) version of inference to the best explanation is compatible with our categorization results, it has a major drawback: it has nothing to say about similarity. It offers no account of the similarity judgments, and it gives similarity *no* role in categorization; yet no one would deny that similarity plays *some* role. Thus, in our experiments, inference to the best explanation can tell us nothing about why, with rich descriptions, categorization judgments tracked similarity judgments as closely as they did (see in particular the results in Table 4). One way around this problem is to propose that (1) similarity judgments, too, are based on inference to the best explanation; (2) the weights given to the component inferences differ when one is making similarity as opposed to categorization decisions; and (3) the kind of inferencing done for similarity judgments can intrude on categorization decisions. This inference-only proposal suffers from the same problems that plague the similarity-only view. Specifically, it is based on the ad hoc assumption that the weights given to the component subprocesses differ for the categorization and similarity tasks. And it blurs the important distinction that

categorization decisions, but not similarity ones, can be correct or incorrect and in need of justification.

In sum, we hold that the most parsimonious and principled account of the current findings, as well as of results from other comparable studies (e.g., Keil, 1989; Rips, 1989), is that two distinct processes can be used to categorize common objects. One of these processes is deliberative, analytic, and capable of providing explicit justification for a categorization decision; we have treated this process as rule based, and we have noted that it is clearly relevant to theory-driven categorization. The other process is more automatic and holistic, and it cannot be used to supply convincing justifications for categorization decisions. We have treated this process as similarity, or matching, based. In our account, categorization is like language and reasoning in that it often reflects the operation of both similarity and rules (Pinker, 1991; Smith et al., 1992).

REFERENCES

- ALLEN, S. W., & BROOKS, L. R. (1991). Specializing the operation of an explicit rule. *Journal of Experimental Psychology: General*, **120**, 3-19.
- BRUNER, J. S., GOODNOW, J. J., & AUSTIN, G. A. (1956). *A study of thinking*. New York: Wiley.
- CAREY, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.
- ESTES, W. K. (1986). Array models for category learning. *Cognitive Psychology*, **18**, 500-549.
- KEIL, F. C. (1986). The acquisition of natural kind and artifact forms. In W. Demopoulos & A. Marras (Eds.), *Language learning and concept acquisition* (pp. 135-153). Norwood, NJ: Ablex.
- KEIL, F. C. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.
- MALT, B. C. (in press). Water is not H₂O. *Cognitive Psychology*.
- MEDIN, D. L. (1989). Concepts and conceptual structure. *American Psychologist*, **44**, 1469-1481.
- MURPHY, G. L., & MEDIN, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, **92**, 289-316.
- NOSOFSKY, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, **115**, 39-57.
- PINKER, S. (1991). Rules of language. *Science*, **253**, 530-535.
- RIPS, L. J. (1989). Similarity, typicality, and categorization. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 21-59). New York: Cambridge University Press.
- SMITH, E. E., LANGSTON, C., & NISBETT, R. E. (1992). The case for rules in reasoning. *Cognitive Science*, **16**, 1-40.
- SMITH, E. E., & MEDIN, D. L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.
- SMITH, E. E., SHOBEN, E. J., & RIPS, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, **81**, 214-241.
2. The purpose of separating the presentation of the description from that of the categories was to allow us to record the reading time for the description separately from the categorization time. This desire to collect latencies was also why we mentioned speed of response in the instructions. We had thought that categorization latencies might actually be faster for rich than for sparse descriptions, which we premised on the notion that similarity-based categorization might be faster than rule-based categorization. As we will see, though, we found no evidence for rule-based categorization, making the latency prediction beside the point.
3. The only other data from Experiment 1 are the latencies referred to in Note 2. Subjects took considerably longer to read rich than sparse descriptions [4.67 vs. 2.82 sec; $t(19) = 7.97, p < .001$], but no longer to categorize rich than sparse descriptions [2.16 vs. 2.34 sec; $t(19) < 1$].
4. The dissociation between categorization and similarity for the sparse items is not complete. There is still a correlation of .40 ($p < .05$) between percentage choices of the variable category in the categorization condition and such choices in the similarity condition. This correlation, though, is appreciably lower than that obtained in Experiment 1, where there was no indication of the dissociation of interest.
5. The simple assumptions are that (1) on some proportion of the trials in which the rule was not mentioned, subjects did in fact engage in rule-based categorization, in which case they chose the variable category with probability 1.0; (2) on the remaining proportion of these trials, subjects engaged in similarity-based categorization, in which case they chose the variable category with probability .5; and (3) rule- and similarity-based categorization are mutually exclusive.
6. Rips's (1989) results in the original condition of this study—that categorizations are more in line with the necessary feature than with the characteristic features—differ from our results for rich descriptions. This difference in results is likely due to the fact that Rips used more salient necessary features than we did; the mating activities of birds, for example, is a more salient necessary feature than the limited size of quarters.
7. This description of an inference to the best explanation process is quite abstract. For example, it does not specify exactly how one might determine that QUARTER provides a better account than does PIZZA of an object that is 3 in. in diameter and silver colored. One means for implementing such inferences would be a scenario construction. Thus, it might be easier to mentally construct a scenario in which a silver colored quarter could become 3 in. in diameter (e.g., by pounding the quarter) than to construct a scenario in which a pizza could become 3 in. in diameter and silver colored.

APPENDIX Items

For each item, the relevant pair of categories is given first (with the variable category on the left), followed by the sparse description, followed by the rich description. Only the five asterisked items were *not* taken from Rips (1989).

1. PIZZA	QUARTER
	Circular object with an X inch diameter.
	Circular object with an X inch diameter that is silver colored.
2. PEACH	NICKEL
	Circular object with an X inch diameter.
	Circular object with an X inch diameter that is silver colored.
3.* SAUCER	PENNY
	Circular object with an X inch diameter.
	Circular object with an X inch diameter that is copper colored.
4. FOUNTAIN PEN	YARDSTICK
	Object X inches long.
	Object X inches long that has marks along its length.

NOTES

1. There are other reasons why rules and theories cannot be equated. For one thing, rule-based categorization can occur even with novel, artificial categories (as in the classic studies of Bruner, Goodnow, & Austin, 1956), though presumably we have no theory of such categories. In addition, there are examples of theory-driven categorization that seem not to involve rules (see, e.g., some of Murphy & Medin's, 1985, examples). We emphasize cases of rule-based categorization that connect to lay theories because of the special importance of such cases to theoretical discussions about categorization.

5. PAPER CLIP	FOOT RULER	19. MOVIE	VETERANS DAY
Object X inches long.
Object X inches long that is made of wood.		Something that lasts X hours.
Something that lasts X hours and commemorates soldiers.	
6. DESK	LEGAL PAD	20. FINAL EXAM	VALENTINE'S DAY
Object X inches long.
Object X inches long that has parallel lines drawn across it.		Something that lasts X hours.
Something that lasts X hours and celebrates love.	
7.* FLAG POLE	VOLLEYBALL NET	21. DINNER PARTY	BASKETBALL GAME
Object that is X feet high.
Object that is X feet high and made mostly of mesh.		Something that lasts X hours.
Something that lasts X hours and involves tall people playing.	
8. CEREAL BOX	STOP SIGN	22. DICTIONARY	SILVER DOLLAR
Object that is X feet high.
Object X feet high with one word printed on it.		Object that weighs X pounds.
Object that weighs X pounds and is made by a mint.	
9. TABLE	SPEED LIMIT SIGN	23. ENCYCLOPEDIA	VIOLIN
Object X feet high.
Object X feet high with a number printed at the top.		Object that weighs X pounds.
Object that weighs X pounds and has four strings.	
10. WASHING MACHINE	GALLON JUG	24. LAMP	GOLF BALL
Object that can hold X gallons.
Object that can hold X gallons with a stopper at the top.		Object that weighs X pounds.
Object that weighs X pounds and is spherical.	
11. COFFEE MUG	FIFTH OF LIQUOR	25. CHESS CHAMP	EMERITUS PROFESSOR
Object that can hold X quarts.
Object that can hold X quarts and is a bottle.		Someone X years old.
Someone X years old who teaches.	
12. WINE GLASS	FOOTBALL	26. MASTER CHEF	CUB SCOUT
Object big enough to hold X cups.
Object big enough to hold X cups that has laces along its side.		Someone X years old.
Someone X years old with badges all over their shirt.	
13. NEW TV	DIME	27. BALLET DANCER	SOCIAL SECURITY RECIPIENT
Object worth X dollars.
Object worth X dollars that says IN GOD WE TRUST on it.		Someone X years old.
Someone X years old who has retired.	
14. WRIST WATCH	LOCAL PHONE CALL	28.* CAR	BOXING RING
Something worth X dollars.
Something worth X dollars that makes communication possible.		Object that is X feet wide.
Object that is X feet wide surrounded by rope.	
15. CALCULATOR	CAN OF PEPSI	29.* LIVING ROOM	FOOTBALL FIELD
Object worth X dollars.
Object worth X dollars that is cool and refreshing.		Object that is X yards wide.
Object that is X yards wide and covered with grass.	
16. ROCK GROUP	US SENATE	30.* BOOKSHELF	AMERICAN EXPRESS CREDIT CARD
Group with X members.
Group with X members, two from each state.		Something X inches wide.
Something X inches wide that is usually mostly green.	
17. COLLEGE FACULTY	US SUPREME COURT		
Group with X members.
Group with X members which makes legal decisions.			
18. JAR OF PICKLES	DECK OF PLAYING CARDS		
Something with X things.
Something with X things which are flat.			

(Manuscript received June 14, 1993;
revision accepted for publication October 15, 1993.)