Memory & Cognition
2003, 31 (5), 790-799

The nonindependence of stimulus properties in human category learning

BRADLEY C. LOVE and ARTHUR B. MARKMAN
University of Texas, Austin, Texas

Typically, models of category learning are verified through behavioral experiments with stimuli consisting of putatively independent dimensions such as shape, size, and color. The assumption of independence is critical in both the design of behavioral experiments and the development of models and theories of learning. Using the standard classification learning paradigm and a common stimulus set, the present work demonstrates that the assumption of independence is unwarranted. Systematic relations span stimulus dimensions and govern learning performance. For example, shape is not independent of size and color, because humans quantify size and color over shape when shape is relevant to the categorization. This quantification is reflected in natural language use (e.g., "blue triangle" as opposed to "triangle and blue"). In this example, color and size are predicates and shape is the argument. Across four experiments, the difficulty of mastering a classification rule can be predicted by the number of predicates that must be unbound in order to free rule-relevant stimulus dimensions.

An important step in developing a theory or model of categorization consists in specifying how stimuli are represented. Categorization researchers have begun to appreciate the complexity of this task. Partitioning a stimulus into feature dimensions depends on relevant domain knowledge (Wisniewski & Medin, 1994) and the comparison context (Medin, Goldstone, & Gentner, 1993). In some cases, the main problem facing learners may be that of learning what the features are, rather than learning what the mapping from features to categories is (Schyns, Goldstone, & Thibaut, 1998). Developing the appropriate feature vocabulary may partially explain what differentiates experts from novices (Biederman & Shiffrar, 1987). Matters can be further complicated when stimulus feature dimensions are relationally connected to one another (e.g., cause/effect, part/whole).

Many researchers have tried to make headway in understanding category learning by constructing laboratory situations that eliminate many of these representational complexities. In service of this goal, the canonical stimulus set used by category learning researchers can be easily and clearly partitioned into a set of seemingly orthogonal and independent stimulus dimensions (see, e.g., Medin & Schaffer, 1978; Nosofsky, Gluck, Palmeri, McKinley, & Glauthier, 1994; Shepard, Hovland, & Jenkins, 1961; Yamauchi, Love, & Markman, 2002). This stimulus set consists of three binary-valued dimensions: size (e.g., small/large), shape (e.g., triangle/square), and color (e.g., blue/red). Few other stimulus sets (e.g., random dot patterns) rival the popularity of the canonical stimulus set.

In this article, we argue that complexities remain even in this highly constrained stimulus set. In particular, relations exist across stimulus dimensions, and these relations play an important role in determining learning performance. Our results suggest that color and size are quantified over shape, which leads to difficulties in forming conjunctive rules involving shape and either color or size. Our findings are in accord with previous research that grants an important role to shape in object representations. For example, young children tend to generalize new count nouns to objects of the same shape as a sample's (Imai, Gentner, & Uchida, 1994; Landau, Smith, & Jones, 1988). Theories of object recognition give a central role to shape (Edelman, 1999; Ullman, 1996). Our findings qualify the interpretation of numerous experimental studies utilizing similar stimulus sets, as well as the models that have been developed to account for such studies. We propose a representational account that makes allowances for the relations among stimulus dimensions (cf. Norman & Rumelhart, 1975), and we offer linguistic tests and empirical methods for inferring the nature of people's stimulus representations.

Relational Independence

The canonical stimulus set is usually assumed to consist of three independent sets:

$$\begin{aligned} color &= \{red, blue\} \\ size &= \{small, large\} \\ shape &= \{triangle, square\}. \end{aligned} \tag{1}$$

A stimulus is constructed by selecting one member from each set. The three selected properties are unified in

This work was supported by AFOSR Grant F49620-01-1-0295 to B.C.L. and NSF Grant SBR-9905013 to A.B.M. We thank Safa Zaki and an anonymous reviewer for their helpful comments. Correspondence concerning this research should be addressed to B. C. Love, Department of Psychology, University of Texas, Austin, TX 78712 (e-mail: love@psy.utexas.edu; http://love.psy.utexas.edu/).

Copyright 2003 Psychonomic Society, Inc.

a single object, which can be seen as a container for the three independent properties. One way to conceptualize this kind of representation is to assume that each dimension is a function that takes the object identifier as an argument and evaluates to the value that the object has on that dimension. The following example is for an object whose size is large, whose color is red, and whose shape is triangle.

$$\begin{aligned} &object \\ &color(object) = red \\ &size(object) = large \\ &shape(object) = triangle. \end{aligned} \tag{2}$$

The functions on the left side in Equation 2 return the member of each set in Equation 1 that is present. Although there is a binding between the entity and each dimension's predicate, these bindings have no influence on learning because they do not vary across predicates or stimuli. In other words, learning would progress in the same fashion if the learner simply knew which three properties were present. Importantly, there are no relations (i.e., predicates) that span stimulus dimensions in Equation 2. For instance, information about shape does not serve as an argument for a predicate concerning size. We refer to representational schemes that lack cross-dimensional relations as being *relationally independent.*

Relational independence makes it easy to construct categorization models and to design experiments. The assignment of stimulus dimensions and their values is arbitrary. Often, such assignments are counterbalanced, and the results are aggregated in statistical analyses. As a result, instead of being consumed with issues related to stimulus representation, researchers can focus on how the abstract structure of the categories (i.e., the assignment of stimuli to categories) affects performance. For example, Shepard et al. (1961) enumerated the six possible partitions (Type I–VI) of the eight stimulus items defined in Equation 1 into two contrastive categories of equal size. People learned to classify the eight items over numerous trials with corrective feedback, and the relative difficulty of the six problems was considered.

Of interest in the present work are three of these problems: Type I, II, and IV. These category structures are shown in Table 1. Type I is a simple, unidimensional structure in which all items with a particular value on one dimension (the first dimension listed in Table 1) belong in one category and all items with the other value on that dimension are in the other category. Categories following this rule are easiest for people to acquire. Type II requires attending to two dimensions (the first two dimensions listed in Table 1), and categories with this structure are typically harder to learn than categories with the Type I structure. Finally, Type IV requires attending to all three dimensions, and is more difficult to learn than Types I and II. An important preliminary test of any new model of classification learning has often been to ensure that it captures this pattern of data.

Shepard et al. (1961) has often been used as a benchmark against which models of category learning are compared (see, e.g., Anderson, 1991; Kruschke, 1992; Love, Medin, & Gureckis, in press; Nosofsky, Gluck, et al., 1994; Nosofsky, Palmeri, & McKinley, 1994). Shepard et al. has also served as the conceptual blueprint for countless empirical studies (some of which have been mentioned above). In this article, we present data that qualify the interpretation of these studies and models. We argue that the field's characterization of the canonical stimulus set is incorrect and that people's stimulus representations are not relational independent. In particular, we argue that shape serves as an argument for size and color. This relational dependency correctly predicts how the assignment of dimensions to the abstract category structures shown in Table 1 modulates the difficulty order of Shepard et al.'s problems. The predictions will be specified below.

Table 1
Types I, II, and IV Category Structures

Item	I	II	IV
1 1 1	A	A	A
1 1 2	A	A	A
1 2 1	A	B	A
1 2 2	A	B	B
2 1 1	B	B	A
2 1 2	B	B	B
2 2 1	B	A	B
2 2 2	B	A	B

Note—The category assignment (either A or B) varies for the eight stimuli (consisting of three binary dimensions) across the three problems.

Relational Dependence

We propose that a stimulus such as a large red triangle is represented as

$$\begin{aligned} &triangle \\ &color(triangle) = red \\ &size(triangle) = large. \end{aligned} \tag{3}$$

In contrast with Equation 2, the value of the shape dimension serves as the argument for the color and size predicates. In other words, color and size are quantified over shape. The stimulus is not a red object, but a red triangle. The bindings between the shape value and the color and size predicates are not trivial. The shape value changes across situations and trials. Thus, attention is required to bind the shape value to its argument (Kahneman, Treisman, & Gibbs, 1992). Disrupting spatial attention should lead to a representation that is more like that shown in Equation 2, which is functionally equivalent to a bundle of independent properties.

A simple linguistic test can determine whether a stimulus set is relationally dependent. The dimensions in relationally dependent stimulus sets can enter into adjective/noun relationships. In particular, when one stimulus dimension can describe other stimulus dimensions, but the reverse cannot be done, then the stimulus set is relationally dependent. For example, if a stimulus can be described as a "large triangle" or a "red triangle," but not as

a "triangle large" or "triangle red," then the stimulus set is relationally dependent. In the present case, size and color behave as adjectives (i.e., one-place predicates or functions), whereas shape acts as the noun (i.e., entity or argument). For relationally independent dimensions, descriptions can contain simple conjunctions of properties. For example, it is quite natural to say "The object is large and red," whereas it is less natural to say "The object is triangle and red." That is, which dimensions are nouns and adjectives is diagnostic of the dimensions' roles in the representation. In relationally independent stimulus sets, all dimensions are adjectives.

Importantly, the relationally independent and relationally dependent stimulus accounts differ in their predictions with regard to classification-learning performance. The predictions are intimately tied to the task demands of the classification-learning task. Classification learning can be characterized as a hypothesis-testing procedure in which people explicitly test rules that are constructed over independent Boolean literals (Bruner, Goodnow, & Austin, 1956; Feldman, 2000; Nosofsky, Palmeri, & McKinley, 1994). Relationally independent representations are ideally suited to this type of induction operation. The values (e.g., red, large, and triangle) are readily accessible as inputs to a rule-construction process. In relationally independent representations, each dimension is related to each other dimension only by virtue of being properties of the same object. Therefore, it should be just as hard to learn a Type II rule involving shape and color as it is to learn a rule[1] involving color and size.

In contrast, relationally dependent representations are not well suited for acquiring conjunctive rules. The difficulty of acquiring a rule depends on which dimensions are relevant. With the canonical stimulus set, conjunctive rules involving color and size should be readily acquired because these dimensions are relationally independent, whereas rules combining shape and either size or color should be more difficult because shape serves as an argument to these predicates. Rule formation requires treating the Boolean literals, such as *triangle* and *red*, as independent. Thus, the cognitive system must either re-represent the stimulus set or engage a different set of learning processes that are not compatible with the task demands of classification learning when dimensions sharing a relation are relevant to a conjunctive classification rule. Either way, we predict that rules involving interrelated dimensions will be more difficult to acquire.

It is important to note that we are not arguing that relationally dependent stimulus sets always lead to poorer performance. For example, acquisition of one-dimensional rules (e.g., the Type I problem) should not be affected by relations across stimulus dimensions, because these rules do not integrate information across dimensions. Similarly, perceptual decision tasks that involve selectively attending to one dimension while ignoring the other dimensions (e.g., tests of perceptual separability—Garner, 1974; Lockhead, 1966) should not be affected by the relations considered above. We are not claiming that shape and size or color are integral.

Some tasks may even benefit from relationally dependent representations. For example, people prefer analogies between two representations that contain matching relational structures that span dimensions (Gentner, 1983, 1989). It is often quite difficult to form analogies in the absence of interconnecting relations (Keane, Ledgeway, & Duff, 1994). Though not considered in this paper, many learning problems may also benefit from relationally dependent representations. For example, in case-based reasoning, previously solved examples are stored in memory and are retrieved and adapted (perhaps through an analogical process) to fit the current situation (Kolodner, 1993). Situations that encourage this type of learning may not share the task demands of classification learning and, therefore, may be more like analogy and benefit from interstimulus structure, which can facilitate inference. Given the presumption that most real-world stimulus sets are relationally dependent, it would be sensible for many real-world tasks to be well matched with this type of representation.

The remainder of this article is organized as follows: Experiments 1 and 2 test our account of the canonical stimulus set within a standard classification-learning task. In Experiment 3, the classification-learning task is coupled with a secondary task that demands spatial attention in order to test our prediction that maintaining relations across stimulus dimensions requires attentional resources. We predict that under a load, the disadvantage of forming conjunctive rules over shape and either size or color will disappear. In Experiment 4, a new stimulus set involving part relations is considered in order to evaluate the potential of applying our analyses to other domains.

EXPERIMENT 1

The canonical stimulus set was used (see Equation 1). Small shapes were 1.5 × 1.5 cm (subtending 2.1° of visual arc), and large shapes were 3.0 × 3.0 cm. In Experiment 1, we tested the claim that the canonical stimulus set is relationally dependent. We predicted that acquiring the Type II category structure when shape is one of the two relevant stimulus dimensions will be more difficult than when shape is irrelevant and color and size are the two relevant dimensions (i.e., shape is assigned to the third stimulus dimension in Table 1). If the three stimulus dimensions are relationally independent, no differences between these two conditions should be observed. The Type IV problem is also included for comparison. We expect to replicate previous results and find Type II (overall) to be easier to learn than Type IV.

Method

Design and Subjects. The subjects in the experiments were 78 undergraduates at the University of Texas who received course credit in exchange for their participation. The subjects were randomly assigned to conditions. Three groups of 26 subjects were run to test the critical hypotheses. One group was given Type IV categories, and the other two were given Type II categories. One Type II group—the *shape-irrelevant* group—was given categories in which color and size were relevant to categorizing the stimuli, and the other

group—the *shape-relevant* group—was given categories in which shape and one of the other two dimensions (either color or size) were relevant.

Stimuli and Procedure. Each subject completed 128 learning trials. On each trial, a stimulus was presented on a CRT monitor with a dark background, and the subject indicated whether the item was in Category A or B by pressing the corresponding key. The subject was then provided with corrective feedback (both auditory and visual) for 1,500 msec, followed by a 1,000-msec intertrial interval. Trial order was randomized for each subject within blocks. The first two blocks of trials consisted of one presentation of each stimulus. The subsequent seven blocks consisted of two presentations of each stimulus. The assignment of dimensions (when not constrained by the experimental design) and values to the abstract category structures shown in Table 1 was randomized for each subject.

Results and Discussion

Two dependent measures were used to compare subjects' learning performance across conditions: overall accuracy (the typical performance measure used in learning experiments) and the proportion of subjects reaching a learning criterion of correct trials in a row. The latter criterion was used as a dependent variable, because the Type II and Type IV category structures differed in the degree to which particular suboptimal rules yielded accurate performance. For example, a one-dimensional rule can yield 75% accuracy for the Type IV problem, but it yields only chance performance (i.e., 50% accuracy) for the Type II problem. In light of this observation, the proportion of subjects reaching the criterion offers a purer measure of mastery than does overall accuracy.

Table 2 displays the results from Experiment 1. In consistency with our prediction, the subjects given Type II categories defined using color and size attained higher accuracy levels than did the subjects given such categories defined using shape and one of the other dimensions [$t(50) = 2.27, p < .05$]. This prediction also holds when the proportion of subjects reaching criterion is analyzed [$\chi^2(1) = 4.46, p < .05$], though not all expected cell counts are above 5. Accuracy was higher for the shape-irrelevant Type II problem than for the Type IV problem [$t(50) = 3.56, p < .001$], whereas the Type II shape-relevant and Type IV problems were not reliably different ($t < 1$). This pattern of results was mirrored by the criterion measure [$\chi^2(1) = 11.34, p < .001$, and $\chi^2(1) = 2.00, p \approx .16$, respectively].

Following learning, the subjects completed a questionnaire in which they indicated whether they used a rule to categorize the stimuli. The subjects who responded affirmatively stated their classification rule for Categories A and B. Of particular interest is the language used by the subjects given the Type II categories in which shape was not relevant. Nine subjects provided a classification rule that involved quantification (e.g., "a small, blue object or a large, red object"), whereas 13 subjects did not explicitly quantify (e.g., "small and blue or large and red"). Those who described their categorization rule by quantifying over an entity were significantly less accurate than those who did not [.84 vs. .91; $t(20) = 2.57, p < .05$].

Evaluating Alternative Accounts

One alternative explanation for the relatively poor performance of the Type II shape-relevant subjects is that shape might be less salient than the other dimensions, and, thus, people may first attempt to construct rules on the more salient dimensions. To rule out this possibility, we had three groups of 16 subjects complete the Type I problem with the category-relevant dimension being shape, color, or size. The subjects in all three groups reached the learning criterion, and no reliable difference in accuracy was observed across groups ($F < 1$). In fact, all three groups had an overall accuracy level of .96. This study had the virtue of being procedurally identical to Experiment 1, but may lack the sensitivity necessary to reveal small differences in dimensional salience.

To rule out this possibility, pairwise similarity ratings ($8 \times 7 = 56$ ratings on a scale of 1–9) were collected from 36 subjects. The salience of each dimension was calculated as the difference in mean similarity of stimulus pairs that matched on the dimension minus the mean similarity of pairs that mismatched on the dimension. There was a significant effect of dimension [$F(2,70) = 11.97, MS_e = 10.34, p \approx 0$]. Shape matches increased similarity more than did color matches [1.94 vs. 0.92; $t(35) = 4.47, p \approx 0$] and size matches [1.94 vs. 1.13; $t(35) = 3.03, p < .01$].

In order to confirm that the operable dimensions were shape, color, and size, the mean similarity ratings were transformed into dissimilarities (i.e., 9 minus the rated similarity) and the 8×8 dissimilarity (i.e., distance) matrix was subjected to multidimensional scaling (MDS). Metrical MDS reduced the proximity data to three dimensions corresponding to shape, color, and size, verifying that people interpreted the stimulus set as having the expected dimensions. The coordinates of the eight stimuli are shown in Table 3. The eigenvalues for the dimensions of shape, size, and color are 42.27, 30.30, and 27.79, re-

Table 2
Mean Accuracy and Proportion of Subjects Reaching the Learning Criterion for Each Category Type as a Function of Task

	Type II					
	Shape Irrelevant		Shape Irrelevant		Type IV	
Task	Accuracy	Proportion	Accuracy	Proportion	Accuracy	Proportion
Single task (Experiment 1)	.85	24/26	.77	18/26	.74	13/26
Single task, text (Experiment 2)	.77	31/40	.71	23/40	.75	20/41
Dual task (Experiment 3)	.68	13/29	.67	15/29	.69	4/29

Table 3
MDS Recovery of the Three Dimensions of Shape, Size, and Color From Pairwise Similarity Ratings of the Eight Stimulus Items

	Dimension		
Item	1	2	3
111	−2.69	−1.44	2.03
112	−2.63	−1.78	−1.70
121	−1.92	2.32	1.59
122	−1.86	1.97	−2.07
211	2.06	−1.95	2.22
212	1.97	−2.48	−1.71
221	2.49	1.96	1.57
222	2.58	1.40	−1.91

Note—Dimension 1, shape; Dimension 2, size; Dimension 3, color.

spectively. The eigenvalue for the fourth dimension was not distinguishable from zero.

The MDS analysis of the pairwise similarity ratings confirmed that the subjects were using the three intended dimensions. One interesting question is whether any effect of relational dependence on the similarity ratings can be seen. Shape plays a central role in the relationally dependent interpretation of the stimulus set (see Equation 3). One prediction from this observation is that shape matches will weigh more than size and color matches. The analyses above support this prediction. A more subtle prediction based on the alignment of object representations is that shape matches will highlight size and color matches more than size matches will highlight color matches and vice versa. This prediction was also confirmed. The effect of size and color matches on similarity conditioned on shape matches {i.e., [(size or color match − size or color mismatch) | shape match] − (size or color match − size or color mismatch) | shape mismatch} is greater (i.e., less negative) than the effect of size matches conditioned on color matches and vice versa [−0.64 vs. −1.21; $t(35) = 2.02, p \approx .051$]. Ceiling effects led to the negative means. Notice that this ceiling effect works strongly against the borderline effect found in the preceding analysis, because shape matches led to higher ratings than did size or color matches (see the analyses above).

EXPERIMENT 2

Experiment 1 yielded a dramatic result: Despite the fact that shape is more salient than color or size, categories defined by a conjunction of shape and either size or color are more difficult to learn than categories defined by a conjunction of size and color. Experiment 2 replicates Experiment 1 in order to rule out the possibility that Experiment 1's surprising result was a statistical anomaly. Experiment 2 also extends Experiment 1 in that a different stimulus set is considered. The procedure and design of Experiment 2 are identical to those of Experiment 1, but the stimuli in Experiment 2 are verbal descriptions of the stimuli shown in Experiment 1. For example, in Experiment 1 a large red triangle was displayed as a picture, but in Experiment 2 the text "large," "triangle," "red" was shown (displayed vertically with one stimulus dimension per line). Because color and size should be quantified over shape, we expect the same pattern of results in Experiment 2 as we observed in Experiment 1, despite the fact that the stimulus presentation encourages participants to treat the three values for each stimulus independently.

Method

Design and Subjects. The design of Experiment 2 was identical to that of Experiment 1. One hundred twenty-one undergraduates at the University of Texas participated in Experiment 2 in exchange for course credit.

Procedure. The procedure was identical to that of Experiment 1. The sole difference between the experiments was the stimulus set. In Experiment 2, stimuli consisted of verbal descriptions of the geometric stimuli shown in Experiment 1. Each stimulus dimension was displayed on consecutive lines (e.g., "blue" above "large" above "triangle"). The vertical order in which the dimensions were displayed was randomized for each subject and was held constant across trials.

Results and Discussion

Table 2 displays the results from Experiment 2. The results mimic those of Experiment 1. In accordance with our prediction, the subjects given Type II categories defined using color and size attained higher accuracy levels than did those given such categories defined using shape and one of the other dimensions [$t(78) = 2.20, p < .05$]. This prediction also held (as a trend) when the proportion of subjects reaching criterion was analyzed [$\chi^2(1) = 3.65, p \approx .06$]. Accuracy for the shape-irrelevant Type II problem was not significantly better than that for the Type IV problem ($t < 1$). However, a significant advantage for the Type II shape-irrelevant problem over the Type IV problem was observed when the proportion of subjects reaching criterion was considered [$\chi^2(1) = 3.65, p < .01$]. The Type II shape-relevant displayed lower (though not reliably so) accuracy than did the Type IV problem [$t(79) = -1.53, p \approx .13$]. Similarly, a nonreliable disadvantage was observed in the criterion data [$\chi^2 < 1$].

As in Experiment 1, at the end of the experiment the subjects completed a questionnaire in which they indicated whether they used a rule to categorize the stimuli. The subjects who responded affirmatively stated their classification rule for Categories A and B. In Experiment 1, the subjects in the shape-irrelevant condition who used quantifying language to describe the rule (e.g., "small or blue *object*") were less accurate than those who did not. Interestingly, no subjects in Experiment 2's Type II shape-irrelevant condition used quantifying language. This result is in accord with the idea that the text stimuli encourage the subjects to view the stimulus dimensions as independent. Despite this manipulation, the difficulty of forming rules involving shape and another dimension was still observed.

EXPERIMENT 3

Our explanation of the difference in performance for the shape-relevant and shape-irrelevant Type II problems in Experiments 1 and 2 is that the dimensions of color and

size were quantified over shape. Object file theories of visual object representation suggest that attention (and in particular spatial attention) is required to bind together the attributes of objects (Kahneman et al., 1992). Thus, if spatial attention were disrupted during category acquisition, people might have difficulty binding the color and size dimensions with the shape dimension. Thus, this manipulation would eliminate the observed difference between performance on Type II shape-relevant and -irrelevant problems. In effect, disrupting spatial attention should lead to relational independence.

Method

Design and Subjects. To test this possibility, three additional groups of 29 subjects (undergraduates at the University of Texas) were run. As in Experiments 1 and 2, one group learned Type IV categories, one learned Type II categories in which the rule was defined over color and size, and one learned Type II categories in which the rule was defined over shape and one of the other dimensions. Subjects in all three groups learned the categories in a dual-task setting in which they had to repeatedly tap the four arrow keys on the keyboard in a fixed order on each trial (the order was varied across trials). This manipulation was expected to reduce the available spatial working memory, which we anticipated would increase the likelihood that shape would be treated as independent of the other dimensions.

Procedure. The subjects completed a secondary task during the learning task. The subjects were shown a sequence of four keys at the beginning of a trial (e.g., up, left, down, right) and had to repeat this sequence by pressing the arrow keys. After completing the sequence correctly three times at a sufficient pace (less than 2,500 msec for each sequence), the geometric stimulus was shown. The subjects continued repeating the sequence correctly at the requisite speed until they were ready to respond A or B, at which point they replaced the last member of the sequence with either A or B. After the subjects responded, the stimulus was no longer shown on the screen and the subjects received corrective feedback. Trials in which the subject did not follow this procedure were aborted.[2] All the subjects received training in this procedure (without category stimuli) prior to data collection. Each key sequence was randomly generated for each trial.

Results and Discussion

The results are shown in Table 2. The main prediction was confirmed. Performance on the Type II shape-irrelevant and shape-relevant problems did not differ significantly. In fact, pairwise comparisons of the accuracy levels for all three conditions did not differ significantly ($t < 1$), although accuracy levels above chance ($p \approx 0$) were displayed in all three conditions. The criterion data also revealed little difference between the two Type II conditions [$\chi^2(1) = 0.27, p > .10$]. This result was probably not due to both groups' being at floor level of performance as a result of the dual task, because the subjects in both the shape-relevant and the shape-irrelevant groups were more likely to reach the learning criterion than were the subjects in the Type IV group [$\chi^2(1) = 6.74$ and 9.47, respectively, $p < .01$].

Experiments 1 and 2 suggest that quantifying over shape (which should require attention to establish the binding) led to poorer performance for the Type II shape-relevant subjects than for the Type II shape-irrelevant subjects. In Experiment 3, a manipulation that reduced the likelihood that size and color would be quantified over shape eliminated the difference in difficulty of learning of Type II categories in which shape was relevant to the classification rule. In contrast with Experiments 1 and 2, the shape-relevant subjects actually outperformed the shape-irrelevant subjects in terms of reaching criterion. Taken together, the experiments provide strong support for the hypothesis that shape is not independent of size and color in the representations developed during category learning. Experiment 3 also offers a new methodology for evaluating acquired representations.

As in Experiment 2, the subjects completed postexperiment questionnaires. Because we assume that the subjects could not form bindings involving shape during the course of learning, we predicted that differences in language would simply reflect word choice and not be indicative of underlying differences in the representations acquired during learning. As we predicted, there was almost no difference in performance for the 6 subjects who described the proper rule using quantifying language in the shape-irrelevant Type II condition and the 14 subjects who used nonquantifying language in this condition (.76 vs. .75; $t < 1$).

EXPERIMENT 4

Before discussing the implications of these findings for the exploration of category learning, we examine one further issue. The experiments described thus far suggest that predicting the difficulty of a classification learning task requires an analysis of stimulus representation. When a category structure involves a rule that considers a pair of properties that are logically independent (i.e., they are members of different sets), the rule is easier to learn when relational dependencies do not have to be broken in order to separate the two properties.

In the first three experiments, the relational dependency involved properties that were quantified over shape. In Experiment 4, we considered the case of parts. In particular, the stimuli were constructed from the following sets:

$$\begin{aligned} &drive = \{CD\ ROM, DVD\} \\ &processor = \{budget, performance\} \\ &computer\ type = \{desktop, laptop\}. \end{aligned} \tag{4}$$

This stimulus set also fails our linguistic test for relational independence (e.g., it is awkward to say "The object is laptop and budget processor"). The test for relational dependence, though not as awkward, does not sound as natural as it did for the shape stimuli (e.g., "This is a DVD laptop" vs. "This is a red square"). If we represent this stimulus set in a manner analogous to our relationally dependent scheme for shape, we get:

$$\begin{aligned} &laptop \\ &drive\ (laptop) = DVD \\ &processor\ (laptop) = budget. \end{aligned} \tag{5}$$

This scheme would predict that learning Type II rules involving drive and processor should be easier than learn-

ing Type II rules involving computer type and one of the other two dimensions. However, the awkwardness of the adjective–noun sentence above suggests we should try a different approach. In computers, drives and processors are not properties, but parts. In order to capture this intuition, we simply added a part relation to the representation:

$$\begin{aligned} &laptop \\ &part\ of\ (laptop, drive\ (laptop) = DVD) \\ &part\ of\ (laptop, processor\ (laptop) = budget). \end{aligned} \quad (6)$$

The linguistic test for this particular form of relational dependence naturally states that one dimension "has" another dimension, but not vice versa. For example, it is natural to say "The laptop has a DVD" or "The laptop with a DVD," but it is not natural to say "The DVD has a laptop" or "The DVD with a laptop." Beyond our intuitions, Markman and Makin (1998) found that this is the preferred phrasing for referring to parts and entities. The test excludes relationally dependent items involving properties. For example, it is unnatural to state "The triangle with a red."

In the computer stimulus set, computer type plays a role analogous to shape in the canonical stimulus set, in that both dimensions are entities and appear as arguments in all predicates. In the canonical stimulus set, Type II rules were more difficult to learn when shape was relevant, because the subjects had to adopt a learning strategy inconsistent with the task demands of classification learning or had to re-represent the stimulus set (e.g., unbind the shape value). In contrast, conjunctive rules could easily be formed over size and color, because the values of these dimensions were readily accessible and not bound to any predicate.

In the computer stimulus set, all dimensions are bound to predicates. As with the canonical stimulus set, we predicted that the relative difficulty level of a Type II rule should be related to the number of predicates that must be unbound in order to "free" the rule-relevant literals. In contrast to the canonical stimulus set, rules involving the base dimension (i.e., computer type) and one of the other dimensions (either drive or processor) should be easier to learn than those involving both of the other dimensions. To form a rule involving computer type and either drive or processor, one *part_of* predicate must be decomposed, whereas both *part_of* predicates must be unbound to learn a rule involving the two parts (see Equation 6).

Thus, the introduction of the *part_of* relation leads to a set of predictions that contrast sharply with those of Experiments 1 and 2. Learning conjunctive rules involving the base dimension should now be easier than learning rules that involve the other two dimensions. A cross-experiment prediction (which should be evaluated with caution) is that performance levels for Type II (overall) should be lower in Experiment 4 than in Experiments 1 and 2, because more relations must be unbound.

Method

Design, Stimuli, Subjects, and Procedure. The design of Experiment 4 followed from the previous experiments. The stimuli in Experiment 4 differed from those in the previous experiments. The stimuli were varied along three binary dimensions: computer type (laptop or desktop), drive type (DVD or CD ROM), and processor type (budget or performance). The computer type dimension served a role analogous to that of the shape dimension in the previous experiments. In the Type II computer-type-relevant condition, computer type and one of the other two dimensions (either drive type or processor type) were relevant to categorizing the stimuli, whereas in the Type II computer-type-irrelevant condition the relevant dimensions were drive and processor type. As in the previous experiment, a third group of subjects were trained on the Type IV structure. The subjects assigned stimuli to Company A or Company B instead of to Category A or Category B. Ninety-four undergraduates at the University of Texas participated in Experiment 4 in exchange for course credit. The procedure was identical to that of Experiment 2. Experiment 4 followed the methods from Experiment 2 for displaying and randomizing the text-based stimuli.

Results and Discussion

Table 4 displays the results from Experiment 4. In accordance with our prediction, the subjects given Type II categories defined using computer type and one of the other dimensions attained higher accuracy levels than the subjects given such categories defined using drive and processor type [$t(60) = 2.25, p < .05$]. Although not significant, this prediction also held when the proportion of subjects reaching criterion was analyzed [$\chi^2(1) = 2.14$, $p \approx .14$]. The Type IV problem also showed an advantage over the computer-irrelevant Type II problem [$t(63) = 3.54, p < .01$ and $\chi^2(1) = 2.64, p \approx .10$]. No differences between the computer-type-relevant Type II and Type IV problems approached significance.

One prediction was that the Type II problems involving the computer stimulus set would be more difficult overall than the Type II problems involving the shape stimulus set. An informal comparison of Experiment 2's data that involved the shape stimulus set in verbal form supported this prediction, although care should be taken in making cross experiment comparisons, especially given the number of differences between the two stimulus sets.

Unlike in the previous experiments, neither Type II problem led to performance superior to that of the Type IV problem. One explanation is that the increased difficulty level of Experiment 4 (indicated by the overall performance levels) favored category structures such as that of

Table 4
Mean Accuracy and Proportion of Subjects Reaching the Learning Criterion for Each Category Type in Experiment 4

Type II (Computer Type Irrelevant)		Type II (Computer Type Relevant)		Type IV	
Accuracy	Proportion	Accuracy	Proportion	Accuracy	Proportion
.61	11/33	.71	15/29	.72	10/32

the Type IV problem, in which suboptimal rules can lead to above-chance performance.

Questionnaire analysis. As in the previous experiments, following learning, the subjects completed a questionnaire in which they indicated whether they had used a rule to categorize the stimuli. The subjects who responded affirmatively stated their classification rule for Companies A and B. Of particular interest is the language used by the subjects given the Type II categories in which computer type was relevant. Six subjects provided a classification rule that involved treating the noncomputer-type dimension as a part (e.g., "a laptop with a DVD or a desktop with a CD ROM"), whereas six subjects did not express any relationship (e.g., "a laptop and DVD or desktop and CD ROM"). Those who described their categorization rule with part language tended to perform better than those who did not [.87 vs. .72; $t(10) = 1.76, p \approx .11$]. Notice that this predicted finding runs counter to the findings of Experiment 1. In Experiment 1, the subjects who used independent language performed better than those who used the quantifying language. The difference between part and property language appears critical.

Similarity ratings. In Experiment 1, an analysis of pairwise similarity ratings indicated that shape was more salient than color or size, despite the fact that the shape-relevant Type II problem was more difficult than the shape-irrelevant Type II problem. Following Experiment 1, pairwise similarity judgments ($8 \times 7 = 56$ ratings on a scale of 1–9) were collected from 31 subjects. The salience of each dimension was calculated as the difference in mean similarity of stimulus pairs that matched on the dimension minus the mean similarity of pairs that mismatched on the dimension. There was a significant effect of dimension [$F(2,60) = 7.71, MS_e = 4.75, p < .01$]. As was predicted, computer-type matches increased similarity more than did drive-type matches [2.16 vs. 1.40; $t(30) = 3.49, p < .01$] and processor-type matches [2.16 vs. 1.62; $t(30) = 2.38$, $p < .05$]. Considering the combined results from Experiments 1 and 4, saliency proves to be a poor predictor of the difficulty of acquiring Type II rules. This is not to say that saliency did not affect the results (cue saliency surely played a role in determining performance), but, rather, that its effect was insufficient to determine the pattern of results.

In order to confirm that the operable dimensions were computer type, drive type and processor type, the mean similarity ratings were transformed into dissimilarities (i.e., 9 minus the rated similarity), and the 8×8 dissimilarity (i.e., distance) matrix was subjected to MDS. Metrical MDS reduced the proximity data to three dimensions corresponding to computer, drive, and processor type, verifying that people interpreted the stimulus set as having the expected dimensions. The coordinates of the eight stimuli are shown in Table 5. The eigenvalues for the dimensions of computer type, processor type, and drive type are 43.98, 36.66, and 34.32, respectively. The eigenvalue for the fourth dimension was not distinguishable from zero.

Like shape in its stimulus set, computer type plays a central role according to the relationally dependent interpretation of the stimulus set (see Equation 6). One prediction that can be derived from this observation is that computer-type matches will weigh more than processor- and drive-type matches. The analyses above supported this prediction (as they did for shape in the analyses of Experiment 1). An analysis of the pairwise similarity data of Experiment 1 also revealed a relational dependency effect in which shape matches highlighted size and color matches more than size matches highlighted color matches and vice versa. Although the classification-learning predictions for Experiments 1 and 4 differed, the predictions for the highlighting effect were the same. As we predicted, computer-type matches highlighted processor- and drive-type matches more than processor-type matches highlighted drive-type matches and vice versa. The effect of processor- and drive-type matches conditioned on computer-type matches {i.e., [(processor or drive match − processor or drive mismatch) | computer type match] − (processor or drive match − processor or drive mismatch) | computer type mismatch} is greater (i.e., less negative) than the effect of processor-type matches conditioned on drive-type matches and vice versa [−0.80 vs. −1.24; $t(30) = 2.13$, $p < .05$]. Ceiling effects led to the negative means. Notice that this ceiling effect works strongly against the effect found in the preceding analysis because computer-type matches led to higher ratings than did processor- or drive-type matches (see analyses above).

Table 5
MDS Recovery of the Three Dimensions of Computer Type, Processor, and Drive From Pairwise Similarity Ratings of the Eight Stimulus Items

	Dimension		
Item	1	2	3
111	−2.54	−2.09	1.98
112	−2.39	−1.92	−2.23
121	−2.26	2.16	2.02
122	−2.17	2.34	−2.00
211	2.16	−2.41	2.00
212	2.25	−2.11	−2.23
221	2.35	1.87	2.26
222	2.60	2.16	−1.81

Note—Dimension 1, computer type; Dimension 2, processor type; Dimension 3, drive type.

GENERAL DISCUSSION

There are three main issues in these experiments that we want to highlight. First, even with very simple stimuli such as those that have been used in previous studies of classification, there are relational dependencies among dimensions that significantly influence ease of learning. These dependencies are evident in the language that participants use to describe the items. In Experiments 1 and 2, with the stimulus set involving shapes, two of the dimensions were described by adjectives and one by a noun. We suggested that this phrasing reflected that the adjectives predicated the entity described by the noun. In accordance with this analysis, rules involving pairs of predicates were easier to learn than rules involving a predicate

and the entity it described. Further evidence for this binding came from Experiment 3, in which a second task, which imposed a spatial memory load, appeared to allow the subjects to treat the shape separately from the predicates.

Finally, in Experiment 4 we focused on a stimulus set in which a particular entity had parts. In this case, the language that people used to describe the parts involved a relation between the part and the entity. We suggested that if people represent a part as a relation between the part and the entity, then conjunctive rules involving the two arguments of a part relation should be easier to learn than conjunctive rules involving arguments from different relations. This prediction was also supported. As in Experiments 1–3, the difficulty of forming conjunctive rules was predicted by the number of relations (i.e., predicates) that needed to be decomposed (i.e., unbound) in order to free the literals necessary for rule formation.

Previous critiques of category-learning research have focused on the generalizability of the results. In an attempt to understand category learning, people have selected stimulus sets that are assumed to have relationally independent dimensions. In practice, there are few sets of dimensions that people treat as relationally independent. In fact, interactions are observed even for sensory stimulus dimensions of different modalities (Marks, 1989). Thus, models that account for the standard findings in Shepard et al.'s (1961) study do not generalize to most natural stimuli. For example, it is difficult to identify the independent dimensions that characterize complex stimuli such as birds, tables, and faces. The present results suggest that category-learning research also faces issues of internal validity. In particular, the present work argues that the search for principles of category learning that are independent of the particular way that stimuli are represented is ill-advised.

In addition to issues of stimulus representation, the line of research following from Shepard et al. (1961) has focused on one induction task—namely, classification learning—despite evidence that comparable induction tasks (e.g., inference learning and unsupervised learning) lead to contrasting patterns of acquisition (Love, 2002; Yamauchi et al., 2002; Yamauchi & Markman, 1998) This line of research has been productive, but has addressed only limited aspects of human categorization (Love, 2001; Schank, Collins, & Hunter, 1986). These issues of ecological validity loom large; rarely do humans learn about easily dimensionalized stimuli with ever-present corrective feedback.

Taken together, these results provide an important caution about developing models of classification that are too strongly based on the structure of the task typically given to subjects. The stimuli in the current study were developed along the lines of those used in previous research. They have a small number of clearly separable dimensions with binary values. However, the natural way that the perceptual system provides representations for these items does not treat them as independent. This observation suggests that there may be few cases in which there are feature dimensions that are treated as logically independent. This observation argues strongly that general principles that govern ease of category learning (e.g., Feldman, 2000) cannot be defined without consideration of the principles that govern how representations are formed (Markman, 1999; Shanon, 1988). In fact, the latter set of principles may play a larger role in determining category learning performance than does the logical structure of categories (cf., Love, 2002, 2003).

Thus, we advocate an approach to the study of categorization that extends current research in two directions. First, we recommend that researchers pay more careful attention to the ways in which people represent the stimuli that are being presented. Models that require relationally independent dimensions should include some test to ensure that the dimensions are in fact independent. Furthermore, some acknowledgment of this independence assumption would be useful in constraining the range of generalization of the studies.

Second, researchers must also be more careful about the relation between the task and the representation. In the present experiments, the difficulty of classification was an interaction between the particular rule that defined the category and the nature of the stimulus representation. Because category learning is also influenced by the way categories are used, it is also crucial that we explore the interactions between types of representations and category uses (Markman & Ross, in press).

REFERENCES

Anderson, J. (1991). The adaptive nature of human categorization. *Psychological Review*, **98**, 409-429.

Biederman, I., & Shiffrar, M. (1987). Sexing day-old chicks: A case study and expert systems analysis of a difficult perceptual learning task. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 640-645.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking*. New York: Wiley.

Edelman, S. (1999). *Representation and recognition in vision*. Cambridge, MA: MIT Press.

Feldman, J. (2000). Minimization of Boolean complexity in human concept learning. *Nature*, **407**, 630-633.

Garner, W. R. (1974). *The processing of information and structure*. Potomac, MD: Erlbaum.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, **7**, 155-170.

Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical processing* (pp. 199-241). Cambridge: Cambridge University Press.

Imai, M., Gentner, D., & Uchida, N. (1994). Children's theories of word meaning: The role of shape similarity in early acquisition. *Cognitive Development*, **9**, 45-75.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object specific integration of information. *Cognitive Psychology*, **24**, 175-219.

Keane, M. T., Ledgeway, T., & Duff, S. (1994). Constraints on analogical mapping: A comparison of three models. *Cognitive Science*, **18**, 387-438.

Kolodner, J. (1993). *Case-based reasoning*. San Mateo, CA: Kaufmann.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, **99**, 22-44.

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, **3**, 299-321.

Lockhead, G. R. (1966). Effects of dimensional redundancy on visual discrimination. *Journal of Experimental Psychology*, **72**, 95-104.

Love, B. C. (2001). Three deadly sins of category learning modelers. *Behavioral & Brain Sciences*, **24**, 687-688.

Love, B. C. (2002). Comparing supervised and unsupervised category learning. *Psychonomic Bulletin & Review*, **9**, 829-835.

Love, B. C. (2003). The multifaceted nature of unsupervised category learning. *Psychonomic Bulletin & Review*, **10**, 190-197.

Love, B. C., Medin, D. L., & Gureckis, T. (in press). SUSTAIN: A network model of human category learning. *Psychological Review*.

Markman, A. B. (1999). *Knowledge representation*. Mahwah, NJ: Erlbaum.

Markman, A. B., & Makin, V. S. (1998). Referential communication and category acquisition. *Journal of Experimental Psychology: General*, **127**, 331-354.

Markman, A. B., & Ross, B. H. (in press). Category use and category learning. *Psychological Bulletin*.

Marks, L. E. (1989). On cross-modal similarity: The perceptual structure in pitch, loudness, and brightness. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 586-602.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review*, **100**, 254-278.

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

Norman, D. A., & Rumelhart, D. E. (1975). *Explorations in cognition*. San Francisco: Freeman.

Nosofsky, R. M., Gluck, M. A., Palmeri, T. J., McKinley, S. C., & Glauthier, P. (1994). Comparing models of rule-based classification learning: A replication and extension of Shepard, Hovland, and Jenkins (1961). *Memory & Cognition*, **22**, 352-369.

Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review*, **101**, 53-79.

Schank, R. C., Collins, G. C., & Hunter, L. E. (1986). Transcending inductive category formation in learning. *Behavioral & Brain Sciences*, **9**, 639-686.

Schyns, P. G., Goldstone, R. L., & Thibaut, J.-P. (1998). The development of features in object concepts. *Behavioral & Brain Sciences*, **21**, 1-54.

Shanon, B. (1988). The similarity of features. *New Ideas in Psychology*, **6**, 307-321.

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

Ullman, S. (1996). *High-level vision*. Cambridge, MA: MIT Press.

Wisniewski, E. J., & Medin, D. L. (1994). On the interaction of theory and data in concept learning. *Cognitive Science*, **18**, 221-281.

Yamauchi, T., Love, B. C., & Markman, A. B. (2002). Learning nonlinearly separable categories by inference and classification. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 585-593.

Yamauchi, T., & Markman, A. B. (1998). Category learning by inference and classification. *Journal of Memory & Language*, **39**, 124-149.

NOTES

1. We use the term "rule" loosely. Our theoretical claims would not differ if rule-like behavior arose by pairing a selective attention mechanism with stored clusters or exemplars.

2. On average, 13.02 trials were aborted for each subject, with no strong error patterns emerging across conditions.

(Manuscript received June 12, 2002;
revision accepted for publication February 21, 2003.)