

Learning categories by making predictions: An investigation of indirect category learning

JOHN PAUL MINDA and BRIAN H. ROSS

University of Illinois at Urbana-Champaign, Urbana, Illinois

Categories are learned in many ways, but studies of category learning have generally focused on classification learning. This focus may limit the understanding of categorization processes. Two experiments were conducted in which participants learned categories of animals by predicting how much food each animal would eat. We refer to this as *indirect category learning*, because the task and the feedback were not directly related to category membership, yet category learning was necessary for good performance in the task. In the first experiment, we compared the performance of participants who learned the categories indirectly with the performance of participants who first learned to classify the objects. In the second experiment, we replicated the basic findings and examined attention to different features during the learning task. In both experiments, participants who learned in the prediction-only condition displayed a broader distribution of attention than participants who learned in the classification-and-prediction condition did. Some participants in the prediction-only group learned the family resemblance structure of the categories, even when a perfect criterial attribute was present. In contrast, participants who first learned to classify the objects tended to learn the criterial attribute.

Learning about categories and making categorical decisions are important for intelligent behavior. Knowledge about an object's category membership provides information about how that object should be used, manipulated, approached, or avoided. Psychologists have often studied categorization by training participants in a classification-learning task in which the participants classify exemplars one at a time, with feedback on every trial. In this article, we argue that a focus on classification learning can lead to a restricted view of category learning. Although other means of category learning have been examined recently, much still remains to be learned about how different ways of learning might affect the category representation. The goal of this article is to explore this issue by introducing an indirect way of learning categories that we believe is common, provides a clear contrast to classification learning, and allows a better understanding of the theoretical implications of nonclassification category learning.

Classification Learning

Research in which the classification paradigm has been used has supported several different theories of categorization, including the idea that people represent categories

as prototypes (Blair & Homa, 2001; Minda & Smith, 2001; E. E. Smith & Medin, 1981; J. D. Smith & Minda, 1998), as collections of exemplars (Medin & Schaffer, 1978; Nosofsky, 1987; Nosofsky & Zaki, 1998), and as rules and exceptions (Nosofsky, Palmeri, & McKinley, 1994; Palmeri & Nosofsky, 1995; E. E. Smith, Patalano, & Jonides, 1998). These theories emphasize categories as something that participants use to classify a new item, and they emphasize the importance of a contrast category in making a decision. In other words, in order to be useful for making classification decisions, two prototypes should be distinguishable from each other or two clusters of exemplars should be distinguishable from each other. Participants learn to attend to the features that maximize this distinction (Kruschke, 1992; Nosofsky, 1987).

Although these properties are important in categorization, it is possible that their primary importance in these theories comes from a reliance on the classification-learning paradigm. Classification explicitly requires participants to distinguish one family of exemplars from another and may affect how a category is represented. In situations outside the lab, people may not interact with an item solely by classifying it. Rather, people may use classification as a step in attaining some goal for which knowledge of an object's category is useful. A thorough examination of category learning must include learning tasks that require participants to do more than simply classify an item and must specify how different learning environments shape category representations. This examination would better reflect real-world learning situations and would result in a better understanding of how different category-learning tasks affect what is learned about a category.

The research was supported in part by the Arnold and Mabel Beckman Foundation. We thank Gregory L. Murphy and three anonymous reviewers for their helpful comments on earlier versions of this article. We also thank Seth Chin-Parker, Jane Erickson, and other members of the Categorization lab at the Beckman Institute, as well as Bradley Love for suggestions and questions about this project. Correspondence concerning this article should be addressed to J. P. Minda, Department of Psychology, University of Western Ontario, London, ON, N6A 5C2 Canada (e-mail: jpminda@uwo.ca).

Category Use and Inference Learning

Although research on categorization has favored the classification paradigm, some researchers have examined alternative learning tasks. This research has generally shown that what people learn about a category depends on how they learn it (Markman & Ross, 2003). Different learning tasks introduce different goals to participants. These goals, in turn, can affect the category representation. Classification learning, for example, is a supervised task with a clear goal: to learn to determine the category membership of the items. The resulting category representation reflects that goal by including the information that is most diagnostic of category membership. However, other category-learning tasks with different goals may lead to differences in the representation.

Category use. In some research, category learning has been examined via inferences and problem solving (Ross, 1996, 1997, 1999, 2000). Participants in these experiments classify the objects and then use information related to category membership to attain a goal. The category use task requires participants to attend to more information than the classification task alone, and this additional information is incorporated into a category representation. For example, Ross (1996) found that participants who learned to classify and solve equations were more likely to sort new equations on the basis of underlying mathematical structure than were participants who learned only to classify the equations. The participants learned about the category structure because that category structure was useful in arriving at the solution to the equation. In related research, Ross (1997) asked participants to first classify a problem and then apply an operator based on the classification. Later, when the information that originally permitted classification was removed, participants were still able to classify the test items, suggesting that they had attended to other information that allowed categorization. These studies suggest that category use, even after a classification is made, affects what information is learned about a category.

Inference learning. Inference learning does not include a classification task. Rather, the category membership is given, and participants supply the value of a missing feature (e.g., Anderson, Ross, & Chin-Parker, 2002; Yamauchi & Markman, 1998, 2000). Inference and classification are both supervised category-learning tasks and are equated in the amount of information presented, but they differ in terms of what information is queried and given feedback (the missing feature or category membership). There is a variety of evidence that participants in these tasks learn different information about the category (see Markman & Ross, 2003, for a full review).

For example, Yamauchi and Markman (1998) found that participants in an inference condition learned prototypical feature values, whereas participants in a classification condition learned exemplar information. Anderson et al. (2002) showed that classification learners performed bet-

ter than inference learners on a whole-exemplar classification task but that inference learners performed better than classification learners on a single-feature classification task. Participants in the inference-learning condition had better knowledge of prototypical features than did participants in the classification-learning condition.

In summary, both the category use and the inference research suggest that a category-learning task can have an effect on what information people acquire about a category. Classification is not the only way to learn a category and may impose its own specific demands on processing. Classification promotes the learning of diagnostic features, whereas inference promotes the learning of prototypical features (Chin-Parker & Ross, 2004).

Unsupervised Category Learning

Much of the research in the field has examined supervised tasks, such as the work on classification learning, category use, and inference learning. The goals are fairly well specified with respect to the categories, and the feedback reinforces the goals of the task to participants. However, in many real-world situations, categories are acquired in the absence of direct feedback, and the goals may not include direct references to the categories. Some researchers have examined unsupervised category learning in which the participants are not aware that there are different categories and good performance on the task does not require learning the categories (Billman & Heit, 1988; Billman & Knutson, 1996; Brooks, 1978; Fried & Holyoak, 1984; Kemler Nelson, 1984; Love, 2002, 2003). This research suggests that when people learn categories in unsupervised tasks, their representations are different from when they learn by classification.

For example, Kemler Nelson (1984) had participants acquire categories of schematic faces that could be learned well by either a criterial attribute (CA) strategy or a family resemblance (FR) strategy. Participants that were given explicit classification-learning instructions tended to focus on CA information. However, other participants were given an unsupervised task. In one experiment, they were asked to view the exemplars for a memorization task. In another experiment, the participants were asked to indicate which faces best fit their stereotype of a firefighter or a police officer, but no feedback was given. Participants in these unsupervised conditions focused more on FR structure.

Love (2002) asked participants to learn four of the six category types from Shepard, Hovland, and Jenkins (1961). With classification instructions, participants did best on Type I (single-dimensional rule), second best on Type II (an XOR problem), third best on Type IV (FR), and worst on Type VI (nonlinearly separable). This rank order is the same as that found by Shepard et al. However, other participants learned the categories by being told to look for patterns (intentional unsupervised) or by simply rating the stimuli for pleasantness (incidental unsupervised). These participants performed better on Type IV than on Type II. This result suggests that super-

vised learning (including classification) encourages explicit rule formation, whereas unsupervised learning encourages the learning of FRs.

More recent research has suggested that different modes of unsupervised learning can affect what information is acquired, as can different modes of supervised learning (classification, use, and inference). Love (2003) compared several unsupervised tasks that varied in terms of goals and how they related to the stimulus set. When participants were explicitly told to learn how the categories were defined (but without corrective feedback), they learned the set of categories well. When participants performed an incidental learning task that emphasized the physical dimensions of the stimuli (pleasantness ratings), they also performed well. However, participants who performed an incidental task that did not emphasize the relevant dimensions did not learn the categories as well. In short, what people learn about a category reflects information that is relevant to the goals of the learning task.

Indirect Learning

Given the importance of unsupervised learning in real-world tasks and the relative paucity of research examining it, further research is necessary for understanding category learning. Given the evidence that different unsupervised tasks may lead to different representations (Love, 2003), it is important to consider what task to examine and how. We believe that much unsupervised category-learning research suffers from at least two problems. First, many unsupervised category-learning studies consist of stimulus-viewing tasks or variations on memorization tasks (e.g., Clapper & Bower, 1994, 2002; Kemler Nelson, 1984; Love, 2002). The goals in these tasks, at least with respect to category learning, are not very clear. Certainly these tasks do not have the same kinds of category-related goals that real-world situations might present, in which knowing the category of an object allows for useful and adaptive behaviors. Second, it is difficult to compare the results of unsupervised learning experiments with classification (or even inference and use), because the tasks are so different. If we could compare a task more directly with a supervised task, such as classification learning, it would be an aid for understanding how and why a learning task affects category representations. Most examinations of unsupervised learning have no supervised task against which to make a clear comparison.

In this article, we investigate category learning using an *indirect learning* task. We refer to it as *indirect* because the participants are not told that there are categories, nor is learning about categories given as the goal (similar to unsupervised learning), but learning the categories will lead to improved performance on the task on which they are being given feedback (similar to supervised learning). We believe that indirect learning reflects some common real-world category-learning scenarios. For example, faculty who teach undergraduates often form categories of the different types of students that they encounter. One category might include students who always come to office hours and always ask ques-

tions but who perform poorly on exams. Another category might include students who miss class and always need an extension on papers. These categories are often used when deciding how to interact with students. However, these categories are not learned through explicit classification training, and classification of the students is never the goal. These categories are acquired, instead, by interacting with students; feedback comes in the form of seeing how different students respond to different interactions. The goal might be to decide whether to make an exception on a make-up exam policy. The processes that operate during learning to make these predictions may not be the same ones as those used in classification learning and may result in different category representations.

In the present experiments, participants were shown a series of imaginary animals of various sizes and were asked to predict how much food they thought the animal should receive. Size could be used as predictor, but in order to perform well in the task, participants had to learn about the underlying category structure. Correct responses were a function of size and category membership. Feedback was provided in the form of the correct amount of food for each animal. The food prediction and the accompanying feedback were only indirectly related to category membership. This task was designed to capture the indirect process of learning about categories when predictions about stimuli were made. Although this task was supervised, the goal was not to learn the categories but, rather, to learn how much to feed the animals.

In Experiment 1 we contrasted the performance of participants who learned categories by doing the food prediction task with the performance of participants who first classified each exemplar before predicting the correct amount of food. Since both groups completed the food prediction task, both groups were equated in terms of this primary goal of the task. The only difference between the groups was that the classification-and-prediction (CP) group first made a classification. We analyzed the data to ascertain what participants had learned about the categories and how they had allocated attention to various dimensions in the task. Experiment 2 replicated the key effects of Experiment 1 but also served to investigate participants' attentional strategies during the learning phase.

EXPERIMENT 1

In Experiment 1, we compared the performance of participants in the prediction task with the performance of participants who were first asked to classify each stimulus and then make a prediction. Since both groups of participants performed the prediction task, the only differences between the two groups were that the CP group knew that the objects belonged to two categories, they made a classification decision, and they received feedback on the classification prior to making a food prediction. To examine potential differences in how the participants in the two conditions might learn the same category, we used a category set that could be learned either by at-

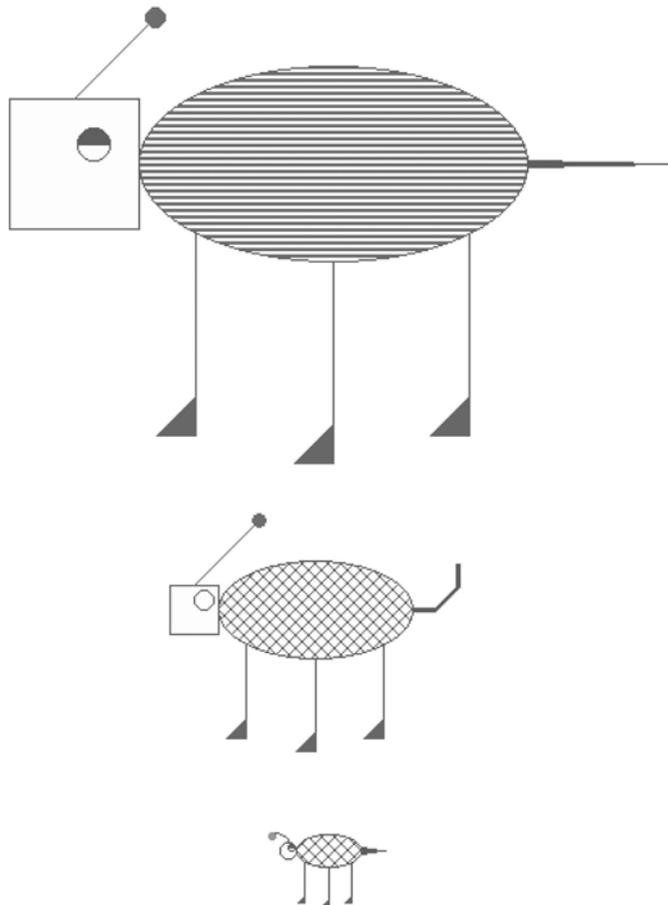


Figure 1. Some examples of the stimuli that were used in Experiments 1 and 2. On the screen, the body and the lines for most of the features were in red, except for the feet (which were gray) and the ball on the antenna (which was either green or maroon). Details of the category set and the stimuli are noted in the text.

tending to a single CA or by attending to overall FR. Classification learning emphasizes the partitioning of the objects into two categories and might encourage more CA performance. The prediction task emphasizes the connection between some set of features and an outcome and might encourage attention to a greater number of features and perhaps more FR performance. Given prior evidence that other non-classification-learning tasks may allow FR learning (Kemler Nelson, 1984; Love, 2002) and prior evidence that participants in classification tasks often prefer single-dimensional rules (Nosofsky et al., 1994), we expected that many participants who learned in the CP task would tend to learn the CA, whereas participants who learned in the prediction-only (PO) group would not.

However, both groups had enough potentially important differences from conditions that had been studied before that alternative results were certainly possible. First, for the CP group, the prediction task forced them to consider additional features beyond what might be needed for classification and provided an opportunity on each trial to look beyond the CA that they might exclu-

sively focus on if the task were simply to classify. That is, they could not predict the amount of food by using only the CA that they might have used to classify, so in trying to find other predictors for this food task, they might well see more of the FR structure within the category. This paradigm is similar in some respects to the category use paradigm of Ross (1997), and in those cases, the use of the category after the classification did influence the category representations formed, including the incorporation of additional features. Second, although a memorization unsupervised learning task had led to learning the FR structure rather than the CA (Kemler Nelson, 1984), the prediction task had a very definite goal for which determining category membership was necessary. Given that, it was very possible that PO participants would end up focusing on the feature that helped them do the task most easily (the CA), in addition to the size.

Method

Participants. Forty University of Illinois undergraduates participated to fulfill course requirements.

Table 1
Abstract Notation for Training Items Used in
Experiments 1 and 2

Stimulus	Dimension					Size	Food
	1	2	3	4	5		
Category A							
A1	0	0	0	0	0	0	4
A2	0	1	0	0	0	0	4
A3	0	0	1	0	0	0	4
A4	0	0	0	1	0	0	4
A5	0	0	0	0	1	0	4
A6	0	0	0	0	0	1	7
A7	0	1	0	0	0	1	7
A8	0	0	1	0	0	1	7
A9	0	0	0	1	0	1	7
A10	0	0	0	0	1	1	7
A11	0	0	0	0	0	2	10
A12	0	1	0	0	0	2	10
A13	0	0	1	0	0	2	10
A14	0	0	0	1	0	2	10
A15	0	0	0	0	1	2	10
Category B							
B1	1	1	1	1	1	0	8
B2	1	0	1	1	1	0	8
B3	1	1	0	1	1	0	8
B4	1	1	1	0	1	0	8
B5	1	1	1	1	0	0	8
B6	1	1	1	1	1	1	11
B7	1	0	1	1	1	1	11
B8	1	1	0	1	1	1	11
B9	1	1	1	0	1	1	11
B10	1	1	1	1	0	1	11
B11	1	1	1	1	1	2	14
B12	1	0	1	1	1	2	14
B13	1	1	0	1	1	2	14
B14	1	1	1	0	1	2	14
B15	1	1	1	1	0	2	14

Note—Each dimension consisted of two values, 0 and 1. For example, for some participants, Dimension 1 might be head shape, and a value of 0 would be round and 1 would be square. The full set of dimensions and values are given in the text.

Materials. The stimuli were line drawings of fictional animals constructed of five binary features: antenna (forward or backward), eye (open or half closed), head (circle or square), body (striped or cross-hatched), and tail (straight or curved). An example of these animals at three different size levels is shown in Figure 1.

The binary notation for the category set we used is shown in Table 1. Each category contained 15 stimuli. There were three different sizes and five exemplars of each size. Note that the five exemplars at each size were identical (except for size) to the five exemplars at the other sizes. For example, Stimuli A1, A6, and A11 contained the same features but were different sizes. Each of the training items shared a majority of its features with the category prototype, so this category set could be learned well with an FR strategy. However, the first dimension provided perfect evidence of category membership, so the categories could also be learned well with a CA strategy. Since the items could be learned with either strategy, test items (Table 2) differentiated between FR learning and CA learning. The evidence provided by the CA (the value on Dimension 1) conflicted with the FR evidence provided by the other dimensions. The actual features that corresponded to each dimension were rotated across participants, so that some participants saw the body as the CA, others saw the head as the CA, and so forth. Both the training items (Table 1) and the test items (Table 2) were shown during the transfer phase.

Procedure. Participants were randomly assigned to one of the two learning conditions. Participants in the PO condition were told to imagine that they were curators of a zoo that had just acquired animals from another planet and that their job was to figure out how much food the animals should receive. They were also told that some features or aspects of the animals might help them predict the correct amount, which would be in the range of 1–14 pounds. On each trial, the participants saw the animal, and they used the number pad to type in how many pounds of food the animal should be fed. The word *correct* was shown on the screen if they were correct. The word *incorrect* was shown if they were incorrect, along with the correct amount of food that the animal should have received.

Participants in the CP condition were given prediction instructions as above but were first given instructions that the animals belonged to two different categories, A and B. On each trial, they made a classification decision and received feedback. After the classification feedback, they indicated how much food to give the animal, just as participants in the PO group did.

Participants in both conditions performed the training task until they had finished at least four blocks and had responded correctly to 12 prediction questions in a row or until they had performed 600 trials, whichever came first. This criterion was chosen to be conservative enough to ensure that participants had learned the categories but to be liberal enough to allow most participants to reach the criterion. The data of participants who did not reach the criterion were not analyzed further.

After reaching the criterion, both groups received two postlearning tests. First, in the prediction test, participants were shown each of

Table 2
Abstract Notation for Test Items Used in Experiments 1 and 2

Stimulus	Dimension					Size
	1	2	3	4	5	
T1	0	1	1	1	1	0
T2	0	0	1	1	1	0
T3	0	1	0	1	1	0
T4	0	1	1	0	1	0
T5	0	1	1	1	0	0
T6	0	1	1	1	1	1
T7	0	0	1	1	1	1
T8	0	1	0	1	1	1
T9	0	1	1	0	1	1
T10	0	1	1	1	0	1
T11	0	1	1	1	1	2
T12	0	0	1	1	1	2
T13	0	1	0	1	1	2
T14	0	1	1	0	1	2
T15	0	1	1	1	0	2
T16	1	0	0	0	0	0
T17	1	1	0	0	0	0
T18	1	0	1	0	0	0
T19	1	0	0	1	0	0
T20	1	0	0	0	1	0
T21	1	0	0	0	0	1
T22	1	1	0	0	0	1
T23	1	0	1	0	0	1
T24	1	0	0	1	0	1
T25	1	0	0	0	1	1
T26	1	0	0	0	0	2
T27	1	1	0	0	0	2
T28	1	0	1	0	0	2
T29	1	0	0	1	0	2
T30	1	0	0	0	1	2

Note—Each dimension consisted of two values, 0 and 1. For example, for some participants, Dimension 1 might be head shape, and a value of 0 would be round and 1 would be square. The full set of dimensions and values are given in the text.

Table 3
Summary of Experiment 1 Results

Measure	Training Condition	
	PO	CP
Trials to criterion	245	160
Old items	.86	.93
Conflict items—CA	.58	.88
Conflict items—FR	.31	.04
Model fit (RMS_e)	0.90	0.41
Model fit (PVAF)	.92	.97
Average heaviest weight	.75	.87
Variance among weights	.09	.12
Number of ignored dimensions	2.17	2.53

Note—Information to help interpret the variety of dependent measures: old items, proportion of correct predictions; conflict items, proportion consistent (with CA or FR); measures using weights, sum of all weights is 1.0. PO, prediction only; CP, classification and prediction; CA, criterial attribute; FR, family resemblance; RMS_e , root-mean square error; PVAF, percentage of variance accounted for.

the training and test stimuli shown in Tables 1 and 2, and they indicated how much food they thought each animal should receive. There was no feedback, and participants saw each animal only once, in a random order. Second, participants completed a single-feature prediction test that was designed to measure how well participants had attended to each dimension in the stimulus set. Participants saw single features at various sizes (along with a label, such as *small head* or *large body*) and indicated how much food they thought an animal with that feature should receive.

Results

Two participants in the PO condition and 1 participant in the CP condition failed to reach the criterion before 600 trials. These 3 participants were dropped from the analysis, leaving 18 participants in the PO condition and 19 in the CP condition. The results are shown in Table 3. Participants in the PO condition took longer to reach the criterion [$t(35) = 2.51, p < .05$]. Despite the difference in reaching the criterion, performance on the old items in the test phase was only marginally better for the CP group [$t(35) = -1.74, p < .10$], although the possibility of ceiling effects makes this a weak conclusion.

The main result of interest was whether participants would make CA responses or FR responses and whether the basis of the response was related to the training condition. For the test items in Table 2, we calculated the proportion of responses for each training condition that were consistent with either the CA prediction or the FR prediction. Sometimes, a response did not agree with either prediction, so the two proportions do not add to 1.0. We then subtracted the proportion of FR-consistent predictions from the proportion of CA-consistent predictions to get an overall measure of sensitivity on the test (a higher number means more predictions were based on the CA). As can be seen from Table 3, this difference was .27 for participants in the PO group, as compared with .84 for the CP participants [$t(35) = -3.12, p < .01$]. We also examined each proportion separately. Participants in the PO condition made fewer CA responses than participants in the CP condition did [$t(35) = -3.09, p <$

.01] and made more FR responses than participants in the CP condition did [$t(35) = 2.96, p < .01$].

Although the analysis of CA and FR performance suggests a difference in what participants learned, we also wanted to assess more finely their attention to dimensions. Some participants may have attended to a single non-CA dimension; others may have attended to two or three dimensions. The simple CA versus FR analysis might miss these subtle distinctions. To investigate participants' attentional spread more generally, we turned to categorization models.

We fit the transfer data with a modified version of an exemplar model, the GCM (Nosofsky, 1987). We also fit a prototype model, which led to similar results but with worse fits, so we will report only the results of fitting the GCM. The details of the model and the model fitting are discussed in the Appendix. To summarize, the GCM has five attention weight parameters that sum to 1.0. Attention to size was held constant. The parameters correspond to attention to each of the five nonsize dimensions in the task and can be arranged to correspond to full attention on one dimension (ignoring the others), to equal attention on all five dimensions, and to every case in between. We used a hill-climbing algorithm (also discussed in the Appendix) to search for the weights that minimized the error between each participant's prediction data in the transfer phase and the model's predictions. The optimal weights from the model fitting are assumed to correspond to the attention distribution adopted by each participant in the task.

The model did a good job of fitting the data for each condition and accounted for over 90% of the variance. The fit was better for the CP condition ($RMS_e = 0.41$) than for the PO condition [$RMS_e = 0.90; t(35) = 3.56, p < .01$].

We analyzed the attention parameters in three ways. First, we calculated the heaviest weight of the five for each participant. A high number means that the participant's data were best described by a heavy weight on a single dimension and corresponded to a narrow attentional focus; a smaller number corresponds to a broader attentional focus. Second, we calculated the variance among the five weights. If the model estimated all of its attention on one dimension, the variance should be relatively high (.16). If the model estimated equal attention to all dimensions, the variance would be zero. Finally, we calculated the number of dimensions with zero weight to estimate how many dimensions were ignored by each participant. These three measures provide a converging description of attention.

Across the three measures, the PO group showed a wider allocation of attention. First, the average heaviest weight was lower in the PO condition [$t(35) = -2.87, p < .01$], suggesting that these participants were less focused on a single dimension than were participants in the CP condition. Figure 2A shows the distribution of the heaviest weights for each participant. Most participants in the CP condition attended to a single dimension, whereas participants in the PO condition did not. Second, the

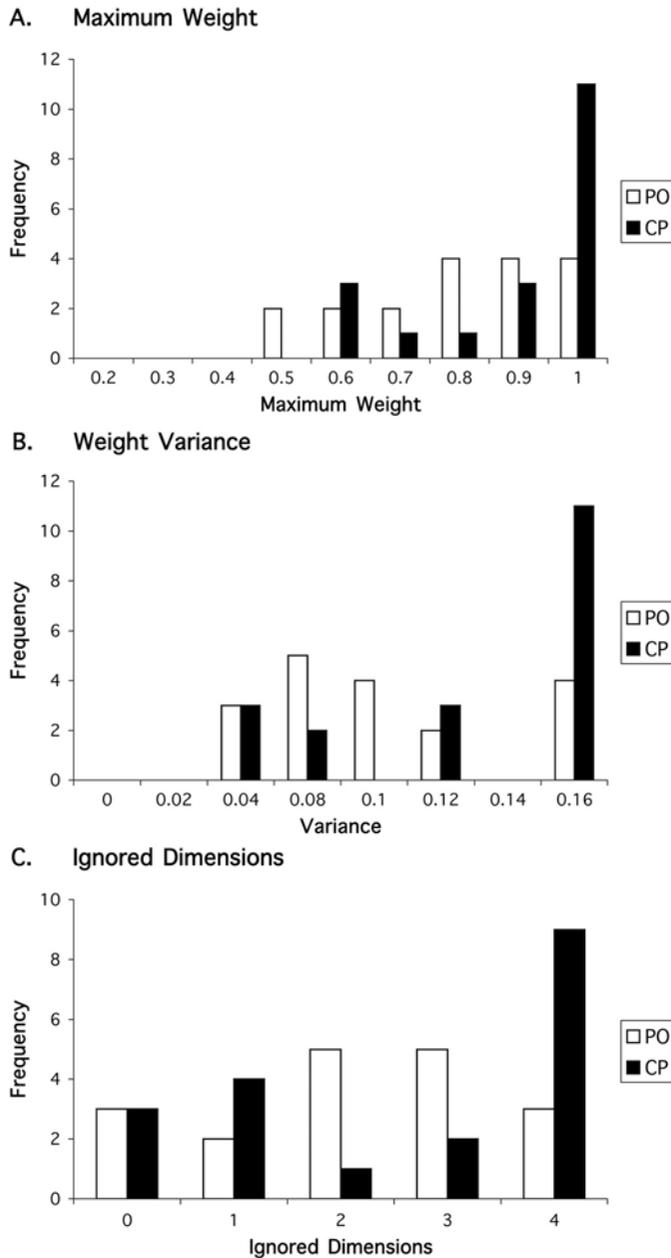


Figure 2. (A) The distribution of average heaviest weights for the participants in Experiment 1. (B) The distribution of weight variance for the participants in Experiment 1. (C) The distribution of ignored dimensions (given a weight of zero by the GCM) for participants in Experiment 1. PO, prediction-only group; CP, classification-and-prediction group.

variance among the five attention weights was lower in the PO condition [$t(35) = 2.03, p = .05$], indicating that these participants adopted a broader allocation of attention than participants in the CP condition did. Figure 2B shows the distribution of these variances and reinforces this conclusion. Third, we analyzed how many dimensions had a weight of zero. There was a small nonsignificant difference between the groups [2.17 for the PO condition vs. 2.53 for the CP condition; $t(35) = -0.73,$

n.s.]. However, the distribution of ignored dimensions shown in Figure 2C does look as if there might be some difference not captured by this statistic. In particular, many participants in the CP condition ignored four out of the five dimensions, but relatively few participants in the PO group adopted this strategy.

We also analyzed performance on the feature prediction task, in which participants saw each feature in isolation and made a food prediction for an animal that pos-

sessed that feature. We defined a correct response as the number of pounds of food that was typically given to animals with that specific feature. This amounted to the prototypical feature values. Overall performance was not very good for either condition (.48 for the PO condition and .54 for the CP condition). There was no difference between the two conditions [$t(35) = -1.07$, n.s.]. However, performance on the CA feature was lower for the participants in the PO condition than for participants in the CP condition [.59 and .78, respectively; $t(35) = -2.11$, $p < .05$]. This finding reinforces the results of the other analyses that the CP group learned CA information but the PO group did not. Performance on non-CA dimensions was poor for both conditions [.46 for participants in the PO condition and .48 for participants in the CP condition; $t(35) = -0.37$, n.s.].

Discussion

The results of Experiment 1 suggest two conclusions. First, although participants in both conditions learned to make accurate predictions, they did not learn equivalent information about the categories, indicating that category learning may be quite flexible and is sensitive to the specific goals and demands of a task. Simply requiring participants to first make a classification changed what participants learned about the category and seemed to encourage attention to a single dimension.

Second, participants in the PO condition were more likely than participants in the CP condition to spread attention over several dimensions in the task and to learn the FR structure of the categories, consistent with research in which other nonclassification modes of learning have been examined (Kaplan & Murphy, 1999; Kehler Nelson, 1984; Love, 2002). Participants in the CP condition often attended solely to the CA and adopted narrower attentional policies in general.

But in Experiment 1, we examined performance on the transfer data only after learning had taken place. We do not know how these attentional strategies developed during the learning task, and the understanding of how attention develops is crucial to explaining how these performance differences emerge. For example, it is possible that participants in both conditions started out with heavy attention to a single dimension but that only participants in the PO condition broadened this attentional policy, due to the prediction task. Alternatively, perhaps both groups began by attending to many dimensions, but only the CP group narrowed their focus due to the constraints of making classifications. The purpose of Experiment 2 was to replicate the basic results of Experiment 1 and also to examine how participants might be processing the features in the task as they were learning to make correct predictions.

EXPERIMENT 2

Method

Participants. Forty University of Illinois undergraduates participated to fulfill course requirements.

Materials. Experiment 2 used the same category set as that in Experiment 1. The stimuli were similar as well, but the animals now had a short neck between the head and the body. This feature, like the feet, did not vary.

Procedure. The procedure in Experiment 2 was similar to that in Experiment 1, with only a few differences. Participants were assigned to one of the two learning conditions, the PO condition or the CP condition. Both groups were given the same instructions as those in Experiment 1, and the structure of the task was nearly the same as well. In addition, about every 30 trials (after a correct prediction), participants were given a short rating task. They were asked to rate on a scale of 1 (*not at all important*) to 7 (*very important*) how important each feature (e.g., head shape) had been in making the food prediction on the previous trial. The rating task asked about each feature individually. Participants were also asked to supply a rating for the neck and the feet, which did not vary and were not relevant for the prediction task.

Participants in both groups performed the training task until they had finished at least four blocks and had responded correctly to 12 prediction trials in a row or until they had performed 600 trials, whichever came first. The data of participants who did not reach the criterion were not analyzed further. After reaching the criterion, participants were shown each of the stimuli in Tables 1 and 2 and indicated how much food they thought each animal should receive. There was no feedback, and participants saw each animal only once, in a random order.

Results

Two participants in the PO condition failed to reach the criterion before 600 trials and were dropped from the analysis, leaving 18 participants in the PO condition and 20 in the CP condition. The results are shown in Table 4. As before, participants in the PO condition took more trials to reach the criterion [$t(36) = 2.21$, $p < .05$]. Performance on the old items at time of transfer did not differ between the two conditions [$t(36) = -0.56$, n.s.].

Overall, the results of this study were similar to those in Experiment 1, but the differences between conditions were attenuated, presumably due to the intrusive rating task during learning. As before, one issue was whether the pattern of CA and FR responses was related to the training condition. We determined this by examining performance on the test items in Table 2, using the dif-

Table 4
Summary of Experiment 2 Results

Measure	Training Condition	
	PO	CP
Trials to criterion	245	159
Old items	.84	.88
Conflict items-CA	.61	.72
Conflict items-FR	.35	.16
Model fit (RMS_e)	1.36	0.79
Model fit (PVAf)	.86	.91
Average heaviest weight	.64	.72
Variance among weights	.07	.10
Number of ignored dimensions	1.78	2.70

Note—Information to help interpret the variety of dependent measures: old items, proportion of correct predictions; conflict items, proportion consistent (with CA or FR); measures using weights, sum of all weights is 1.0. PO, prediction only; CP, classification and prediction; CA, criterion attribute; FR, family resemblance; RMS_e , root-mean square error; PVAf, percentage of variance accounted for.

ference score (CA – FR) as in Experiment 1. The PO group was more sensitive to FR and less to CA (a difference score of .26). The CP group did not show as strong a CA performance as in Experiment 1 [a difference score here of .56; $t(36) = -1.41$, n.s.]. There was no significant difference between the groups on CA performance

[$t(36) = -0.94$, n.s.]. However, participants in the PO condition made somewhat more FR predictions than the participants in the CP condition did [$t(36) = 1.89$, $p < .10$]. One possible reason for this difference from Experiment 1 is that the addition of the rating task changed how participants processed each feature, thus leading to

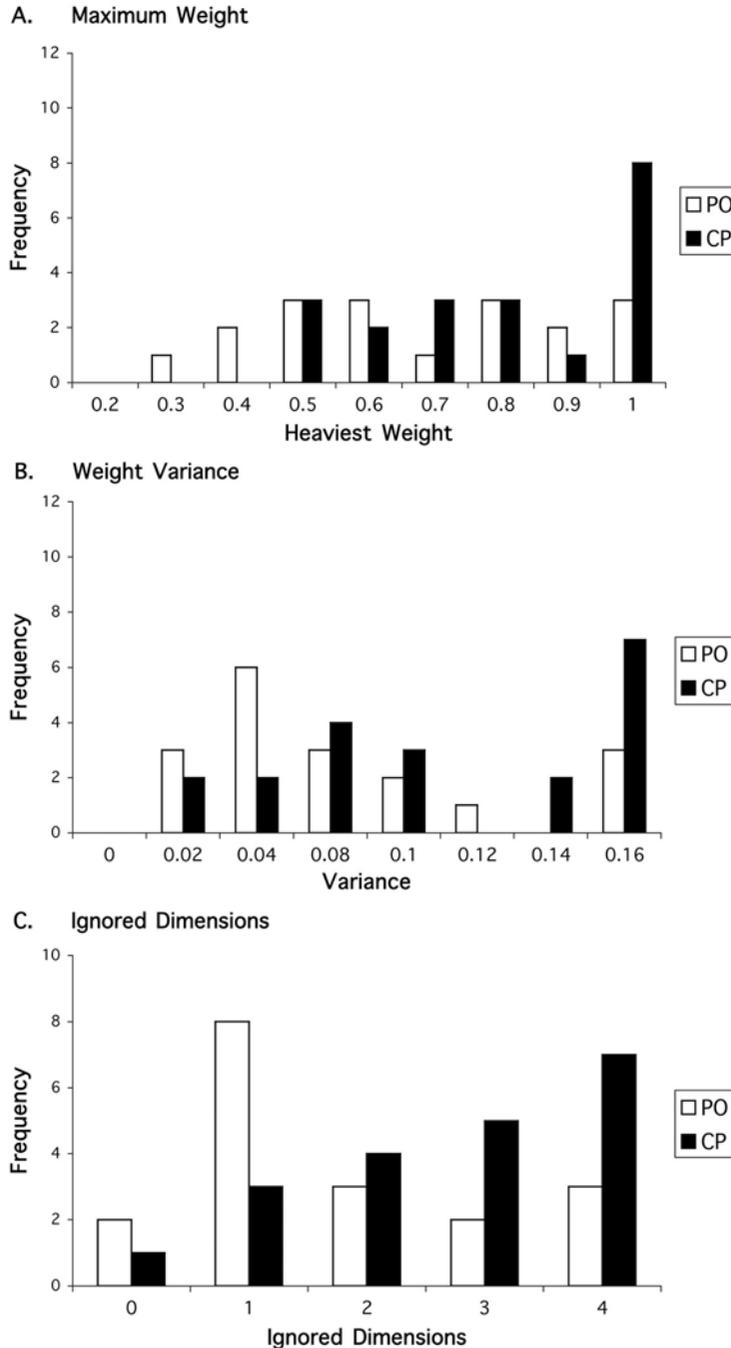


Figure 3. (A) The distribution of average heaviest weights for participants in Experiment 2. (B) The distribution of weight variance for participants in Experiment 2. (C) The distribution of ignored dimensions (given a weight of zero by the GCM) for participants in Experiment 2. PO, prediction-only group; CP, classification-and-prediction group.

slightly more CA performance in the PO group and slightly more FR and less CA performance in the CP group. We will address this possibility further in the Discussion section. However, modeling should reveal any attentional differences.

As in Experiment 1, we fit the transfer data with the GCM. The model again did a good job of fitting the data for each condition and accounted for over 85% of the variance across the two conditions. The model fit better in the CP condition ($RMS_e = 0.79$) than in the PO condition [$RMS_e = 1.36$; $t(36) = -2.45$, $p < .05$].

We now will turn to the three interrelated measures of attention. Although the differences were attenuated here as well, the pattern was consistent with that in Experiment 1. The average heaviest weight was somewhat higher in the CP condition, although the advantage was only marginally significant [$t(36) = -1.77$, $p < .10$]. This difference suggests that participants in the CP condition paid more attention to a single dimension than the participants in the PO condition did. Figure 3A shows the distribution of the heaviest weights for each participant. Most participants in the CP condition attended to a single dimension, whereas participants in the PO condition did not. The variance among the five attention weights was somewhat higher in the CP condition [$t(36) = -1.87$, $p = .10$], suggesting that participants in the PO condition adopted a broader allocation of attention than participants in the CP condition did. Figure 3B shows the distribution of these variances and reinforces this conclusion. Finally, participants in the CP condition ignored significantly more dimensions, as compared with the PO group [$t(35) = -2.21$, $p < .05$], as is shown in Figure 3C. Many participants in the CP condition ignored four out of the five dimensions, but relatively few participants in the PO condition adopted this strategy. This result suggests that participants in the PO condition did attend to more dimensions in the task than participants in the CP condition did. These three measures provide a consistent picture of participants in the PO condition showing wider attentional allocation than the CP group.

Participants in Experiment 2 also completed a rating task periodically during learning. They rated all of the individual dimensions in the task (plus two noninformative dimensions) for how important they were in making a food prediction. We analyzed the five relevant dimensions and excluded the two irrelevant dimensions. The minimum number of rating trials was 4, which happened when participants reached criterion in 120 trials, and the maximum was 17. In order to compare all the participants, we analyzed only 4 trials per person. In order that these 4 trials might reflect any changes that occurred with time, we examined the first 2 trials and the last 2 trials. The average ratings for the CA showed no significant difference, with 5.32 for the PO condition and 6.00 for the CP condition [$t(36) = -1.59$, n.s.]. The ratings for the four non-CA dimensions were higher in the PO condition (3.48) than in the CP condition [2.79; $t(36) = 2.11$, $p < .05$].

We considered two measures of attentional spread for the rating data during learning. First, we calculated how many weights were given a rating of 1, or *not at all informative*. If participants in the CP condition were more likely to base their decisions on a single dimension, the number of uninformative dimensions should be higher. Second, we calculated how many features were given a rating of 5 or higher, as a measure of how many dimensions were considered at least somewhat important. A higher number would indicate a broader attention spread, because it would mean that several dimensions were considered at least somewhat important. An interaction with trial would suggest a change in attention policy from the beginning to the end of the training phase.

The results of the feature rating task can be seen in Table 5. The *not at all informative* ratings were roughly equivalent between the two groups, although participants in the CP condition rated slightly more dimensions as not informative. An analysis of variance (ANOVA) with trial (1–4) as a within-subjects variable and condition (PO or CP) as a between-subjects variable showed no main effect for condition [$F(1,36) = 2.13$, $MS_e = 4.54$, n.s.], and no interaction between trial and condition [$F(3,108) = 0.21$, $MS_e = 0.94$, n.s.]. However, participants in the PO condition rated more dimensions as at least *somewhat important* [5 or higher; $F(1,36) = 4.32$, $MS_e = 2.19$, $p < .05$]. Furthermore, although the number of dimensions given these rating was similar between groups on the first trial, this number increased for participants in PO condition, but not for participants in the CP condition, as was confirmed by the significant interaction between trial and condition [$F(3,108) = 2.65$, $MS_e = 0.56$, $p < .05$].

Discussion

The goals in Experiment 2 were to replicate the indirect learning differences and to examine the learning process. First, the findings generally replicated the results in Experiment 1. The differences were attenuated due to the rating task intrusion, but the overall pattern was clearly very similar. Even with the attenuation, participants in the PO condition paid attention to more dimensions than participants in the CP group did and ignored fewer dimensions. Second, the results of the feature rating task suggest that the PO participants may

Table 5
Frequencies of Feature Ratings for the Five Relevant Dimensions

Trial	Number Rated 1		Number Rated 5 or Higher	
	PO	CP	PO	CP
1	1.39	2.05	1.67	1.90
2	1.67	2.00	2.28	1.60
3	1.50	1.95	2.39	1.85
4	1.72	2.30	2.67	1.65
Average	1.57	2.08	2.25	1.75

Note—Ratings were from 1 (*not important*) to 7 (*very important*). PO, prediction only; CP, classification and prediction.

have developed this strategy over time and expanded their attention focus only after some initial exposure to the stimuli.

One result that did not entirely replicate was that participants in the CP condition did not make significantly more CA predictions than participants in the PO condition did. This change was largely due to the lower CA performance in the CP condition, from .89 in Experiment 1 to only .72 in Experiment 2. One reason for this decrease is that several participants in the CP condition in Experiment 2 based their predictions on a single non-CA dimension. Thus, although they were choosing only one dimension, the CA/FR measure we used scored this as an FR-based performance, but the modeling still picked up a difference in attention allocation between the two groups. This decrease may have occurred because the addition of the rating task interfered with some participants' search for the correct CA. For example, even if performance was not perfect for a single non-CA strategy, it could be that rating all the stimuli, including that non-CA dimension, encouraged some participants to stick with that strategy.

GENERAL DISCUSSION

In the research described here, we examined a novel paradigm for studying category learning. Participants learned categories by prediction with indirect feedback, rather than by classification and direct feedback. Both experiments showed that participants in the PO condition paid attention to more dimensions and participants in the CP condition paid attention to fewer dimensions. People who learned in the PO condition were more likely to learn the FR than were participants who learned them by first making a classification.

One of the primary motivations for this research was the idea that people learn categories in many different ways with many different goals. These different goals can affect what is learned about a category, but experimental research has focused predominantly on classification learning. Classification learning may encourage participants to rely on single-dimensional rules and to learn the optimal attention weights for classification (Chin-Parker & Ross, 2004; Nosofsky, 1987; Nosofsky et al., 1994). However, a small but growing body of research has examined other ways in which people learn categories (Love, 2003; Markman & Ross, 2003). These other ways of learning categories may reflect real-world categorization situations better than the classification task does. For example, it is possible that categories can be learned by simply observing perceptual similarity or behavioral similarities (Fried & Holyoak, 1984; Kaplan & Murphy, 1999; Wattenmaker, 1991). Children may learn that some animals are friendly and like affection (such as dogs) and others are less friendly and do not always like to be touched (such as cats). This learning need not be the result of classification training. In fact, the dog and cat categories are probably learned before the child has

acquired labels for them (Quinn & Eimas, 1996). Alternatively, categories are sometimes learned by people interacting with objects or predicting or inferring things about objects (Anderson et al., 2002; Ross, 1996, 1997; Yamauchi & Markman, 1998, 2000). An amateur photographer might learn about different categories of light conditions and camera settings without ever receiving classification feedback, simply by seeing the good or bad pictures that result. In these cases, classification is not the goal of the task, and classification is not the method of training. It is reasonable to think that what participants learn in these cases would not necessarily mirror what they would learn in a classification task.

The present research suggests that people learn categories by predicting things about stimuli, even when category membership is not known. Furthermore, participants might learn the same category differently, depending on whether they learned in the prediction task or whether they learned to make classifications first. These differences were shown by the attentional strategies of the participants in each group—by participants in the PO group often adopting a broader distribution of attention and by participants in the CP group often focusing on a single dimension. The extent to which our results can be generalized to other category learning depends on the differences between classification and prediction and prompts a closer look at the processing demands of each task.

The present data suggest that the category-learning process operating during the prediction task may encourage participants to attend to more dimensions than the classification task does. In this case, although some participants in the PO condition learned the CA, some participants also spread their attention over many dimensions and learned FRs. They did not use a single-feature rule, even though a perfect one existed. This result is consistent with some results of incidental and unsupervised learning studies (Kaplan & Murphy, 1999; Kemler Nelson, 1984; Love, 2002, 2003) and also with the results of category use experiments, where participants attended to information other than the defining characteristics (Ross, 1997).

The results from the PO condition are inconsistent with classification research that has argued for rules as the default approach to category learning. For example, some classification theories have assumed that participants will prefer unidimensional rules over other strategies when learning categories (Ashby, Alfonso-Reese, Turken, & Waldron, 1998). Although most of our CP participants seemed to follow this strategy, by attending to a single dimension and size, many of our PO participants did not. Other research has suggested that participants will search for the best unidimensional or conjunctive rule and then memorize exceptions to those rules (Nosofsky et al., 1994). Again, our CP participants seemed to follow this strategy, but many of our PO participants did not. Finally, some researchers have argued that humans learn categories by imposing the simplest

possible structure on the set of exemplars (see Feldman, 2000, 2003a, 2003b, for a more in-depth description of category simplicity). Many of the CP participants did learn the simplest way because they learned the CA, but many of the PO participants did not. We are not claiming that these theories of classification are incorrect: Our CP participants behaved as predicted by these theories. Our claim is that these models and theories of classification may not generalize to other nonclassification cases of category learning, such as prediction learning.

Although the GCM provided good fits to the prediction data, there are at least two reasons why this finding is not evidence that classification models are sufficient to account for indirect learning results. First, we chose a simple contrast of CA versus FR, both of which the GCM can account for, as a way of showing how the representation was affected by the CP versus PO manipulation. Second, although the GCM can fit the prediction data, it provides no account of why the two conditions led to these different attention weight patterns, which is the main result of the experiments.

Because this is an initial investigation of indirect learning, several alternative possibilities have not been ruled out. We mention two here. One possibility is that the difference in performance might have been due to differences in task constraints. Perhaps participants in the CP task used a single feature to classify the exemplars and then simply stuck with that attentional allocation when performing the prediction task. Since participants in the PO condition did not have the additional information given by the classification task, they were less likely to find the correct CA (although some did). Although this explanation is consistent with our general claim above, it is possible that the effects we observed are tied to the specific task that we used. Further research in which different indirect learning tasks are used might help to clarify how and why many of our PO participants behaved as they did. However, note that the PO condition is usually (always?) going to require more complex learning, because it requires learning both the classification and the category-based judgment simultaneously, rather than getting feedback on a classification response before making the other judgment.

A second possibility is that some PO participants learned six discrete food categories rather than the two *light eater* and *heavy eater* categories. This learning might have caused some participants to attend to a broader distribution of features than did participants who first made a two-category classification, although the CA (with size) might still be the simplest solution. We are investigating this possibility.

Nonetheless, given the agreement between our data and those in other studies of non-classification-learning tasks, one sees that classification learning may be a specialized and more constrained version of category learning. Of course, participants can attend to more than one dimension and learn FRs in a classification paradigm as

well, but the evidence here and elsewhere suggests that participants in classification-learning tasks tend to look for single-dimensional rules first and then broaden their attention only if those single-dimensional rules are insufficient.

Classification makes clear to the participants that there are definite categories. Participants may have expectations or beliefs about categories as rule defined, and classification may introduce a single-dimension constraint on performance. All things being equal, participants may seek the simplest way to divide the stimuli into two categories (Feldman, 2000, 2003a, 2003b). But in the prediction task, there is no such knowledge of categories, and the goal of the task does not directly relate to category membership. Some participants may find the CA and may find this to be a workable strategy for doing well in the prediction task. Others may develop a category representation that includes many of the dimensions, perhaps because they seek several sources of information to enter into a prediction. Something about the prediction task allows this processing to happen, or something about the classification task prevents it. But why should participants in the CP condition have displayed this constraint, since they also made food predictions during their training?

One possibility is that CP participants have additional information in the form of the category label when they have to make a food prediction, so processing beyond the CA may be unnecessary for the generation of the food response. Participants may simply learn the feature that is best for making the classification and rely on the results of their classification for making predictions. This possibility seems consistent with some intuitions about categorization outside the lab. Although classification is an important part of categorization, making a classification decision is often the precursor to what the observer might actually do with the stimulus. Simply knowing that an object is a chair may be helpful when referring to that object, but in most cases one would also need to know that chairs afford sitting.

In summary, our research suggests that category learning is not just classification learning. Rather, people learn categories in different ways, by predicting things about objects, by learning to use objects, and so forth. The study of classification learning is one important part of categorization learning, but clearly not the only part. Theories and models of categorization that rely solely on classification data may not be describing category learning as a whole. When given alternative learning tasks, such as our prediction task, participants adopt different attentional strategies and form different category representations than participants who are asked to classify the exemplars do. A complete account of category learning and categorization must consider these alternative learning tasks and the psychological processes that result from them.

REFERENCES

- ANDERSON, A. L., ROSS, B. H., & CHIN-PARKER, S. (2002). A further investigation of category learning by inference. *Memory & Cognition*, **30**, 119-128.
- ASHBY, F. G., ALFONSO-REESE, L. A., TURKEN, A. U., & WALDRON, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, **105**, 442-481.
- BILLMAN, D., & HEIT, E. (1988). Observational learning from internal feedback: A simulation of an adaptive learning method. *Cognitive Science*, **12**, 587-625.
- BILLMAN, D., & KNUTSON, J. (1996). Unsupervised concept learning and value systematicity: A complex whole aids learning the parts. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 458-475.
- BLAIR, M., & HOMA, D. (2001). Expanding the search for linear separability constraint on category learning. *Memory & Cognition*, **29**, 1153-1164.
- BROOKS, L. R. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 169-211). Hillsdale, NJ: Erlbaum.
- CHIN-PARKER, S., & ROSS, B. H. (2004). Diagnosticity and prototypicality in category learning: A comparison of inference learning and classification learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **30**, 216-226.
- CLAPPER, J. P., & BOWER, G. H. (1994). Category invention in unsupervised learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 443-460.
- CLAPPER, J. P., & BOWER, G. H. (2002). Adaptive categorization in unsupervised learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 908-923.
- FELDMAN, J. (2000). Minimization of Boolean complexity in human concept learning. *Nature*, **407**, 630-633.
- FELDMAN, J. (2003a). A catalog of Boolean complexity. *Journal of Mathematical Psychology*, **47**, 75-89.
- FELDMAN, J. (2003b). The simplicity principle in human concept learning. *Current Directions in Psychological Science*, **12**, 227-232.
- FRIED, L. S., & HOLYOAK, K. J. (1984). Induction of category distributions: A framework for classification learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **10**, 234-257.
- KAPLAN, A. S., & MURPHY, G. L. (1999). The acquisition of category structure in unsupervised learning. *Memory & Cognition*, **27**, 699-712.
- KEMLER NELSON, D. K. (1984). The effect of intention on what concepts are acquired. *Journal of Verbal Learning & Verbal Behavior*, **23**, 734-759.
- KRUSCHKE, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, **95**, 471-484.
- 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.
- MARKMAN, A. B., & ROSS, B. H. (2003). Category use and category learning. *Psychological Bulletin*, **129**, 592-613.
- MEDIN, D. L., & SCHAFFER, M. M. (1978). Context theory of classification learning. *Psychological Review*, **85**, 207-238.
- MINDA, J. P., & SMITH, J. D. (2001). Prototypes in category learning: The effects of category size, category structure, and stimulus complexity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 775-799.
- NOSOFSKY, R. M. (1987). Attention and learning processes in the identification and categorization of integral stimuli. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 87-108.
- NOSOFSKY, R. M., PALMERI, T. J., & MCKINLEY, S. K. (1994). Rule-plus-exception model of classification learning. *Psychological Review*, **101**, 53-79.
- NOSOFSKY, R. M., & ZAKI, S. R. (1998). Dissociations between categorization and recognition memory in amnesic and normal individuals: An exemplar-based interpretation. *Psychological Science*, **9**, 247-255.
- PALMERI, T. J., & NOSOFSKY, R. M. (1995). Recognition memory for exceptions to the category rule. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 548-568.
- QUINN, P. C., & EIMAS, P. D. (1996). Perceptual cues that permit categorical differentiation of animal species by infants. *Journal of Experimental Child Psychology*, **63**, 189-211.
- ROSS, B. H. (1996). Category representations and the effects of interacting with instances. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 1249-1265.
- ROSS, B. H. (1997). The use of categories affects classification. *Journal of Memory & Language*, **37**, 240-267.
- ROSS, B. H. (1999). Postclassification category use: The effects of learning to use categories after learning to classify. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **25**, 743-757.
- ROSS, B. H. (2000). The effects of category use on learned categories. *Memory & Cognition*, **28**, 51-63.
- SHEPARD, R. N., HOVLAND, C. I., & JENKINS, H. M. (1961). Learning and memorization of classifications. *Journal of Experimental Psychology*, **65**, 94-102.
- SMITH, E. E., & MEDIN, D. L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.
- SMITH, E. E., PATALANO, A. L., & JONIDES, J. (1998). Alternative strategies of categorization. *Cognition*, **65**, 167-196.
- SMITH, J. D., & MINDA, J. P. (1998). Prototypes in the mist: The early epochs of category learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 1411-1430.
- WATTENMAKER, W. D. (1991). Learning modes, feature correlations, and memory-based categorization. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 908-923.
- YAMAUCHI, T., & MARKMAN, A. B. (1998). Category learning by inference and classification. *Journal of Memory & Language*, **39**, 124-148.
- YAMAUCHI, T., & MARKMAN, A. B. (2000). Learning categories composed of varying instances: The effect of classification, inference, and structural alignment. *Memory & Cognition*, **28**, 64-78.

APPENDIX

Model Specifications

The GCM is an exemplar model first described by Medin and Schaffer (1978) and generalized by Nosofsky (1987). To categorize an item, the model first finds the distance between an item and an exemplar, as shown in Equation A1:

$$d_{ij} = c \left[\sum_{k=1}^N w_k |x_{ik} - x_{jk}| \right]. \quad (\text{A1})$$

In Equation A1, x_{ik} and x_{jk} are the values of the item and exemplar on dimension k , and w_k is the attentional weight for dimension k . Each attentional weight varies between 0.0 and 1.0, and attentional weights are constrained to sum to 1.0 across dimensions. The c parameter is a sensitivity parameter that varies from 0.0 to 20.0. Larger values magnify psychological space and increase the differentiation among stimuli. The similarity η_{ij} between the item and the exemplar is calculated from the scaled psychological distance d_{ij} between the stimuli, as shown in Equation A2:

$$\eta_{ij} = e^{-d_{ij}}. \quad (\text{A2})$$

Overall similarity to Category A is divided by the sum of overall A and B similarity to obtain the probability of a Category A response for stimulus i , as shown in Equation A3.

$$P(R_A | S_i) = \frac{\sum_{j \in C_A} \eta_{ij}}{\sum_{j \in C_A} \eta_{ij} + \sum_{j \in C_B} \eta_{ij}}. \quad (\text{A3})$$

Since this model was originally designed to make classification predictions, we needed to add one step to make the food predictions. We assumed that the participants used a combination of size and category membership when making their predictions. Since size was perceptually obvious, size information was given to the model. Food prediction was then a function of category membership (as given by Equation A3) and size, as shown in Equation A4:

$$\text{food} = [P(R_A | S_i) \times G] + (\text{size} \times H). \quad (\text{A4})$$

In the final equation, G is the category factor constant and was set to 4, and H is the size factor constant and was set to 3. These two values simply scaled the result to the range of values that we were fitting. This algorithm is the same one that we used in determining the correct amount of food for each stimulus in designing the experiment.

Model Fitting

To find the best-fitting parameter settings of each model, a single parameter configuration (attention weights and c) was chosen, and the food prediction for each stimulus in an experiment was calculated according to that configuration. We then minimized the root-mean square error (RMS_e) between the observed data and the prediction of the model with a hill-climbing algorithm that made a small adjustment to the provisional best-fitting parameter settings and chose the new settings if they produced a better fit (i.e., a lower RMS_e). During each iteration of the hill-climbing algorithm, a parameter and a directional change were chosen at random. These changes were small: gradations of 1/100 for the attentional weights and 1/200 for the sensitivity parameter. These changes always respected the upper and the lower bounds of the parameters. To ensure that attentional weights always summed to 1.0, the weight parameters were adjusted in randomly chosen complementary pairs. To ensure that local minima were not a problem, the fitting procedure was repeated by choosing five more, different configurations of the model and hill climbing from there. The variance among the six fits tended to be zero (or very small), indicating that the minima found were close to global ones. We report both the fitting error (RMS_e) and the percentage of variance accounted for (PVAF), which was the r^2 between observed and predicted food amounts. The fits reported in the present experiments represent the best RMS_e and the r^2 from the six different fittings.

(Manuscript received June 18, 2002;
revision accepted for publication April 19, 2004.)