

# Observational versus feedback training in rule-based and information-integration category learning

F. GREGORY ASHBY

*University of California, Santa Barbara, Santa Barbara, California*

and

W. TODD MADDOX and COREY J. BOHIL

*University of Texas, Austin, Texas*

The effects of two different kinds of categorization training were investigated. In observational training, observers are presented with a category label and then shown an exemplar from that category. In feedback training, they are shown an exemplar, asked to assign it to a category, and then given feedback about the accuracy of their response. These two types of training were compared as observers learned two types of category structures—those in which optimal accuracy could be achieved via some explicit rule-based strategy, and those in which optimal accuracy required integrating information from separate perceptual dimensions at some predecisional stage. There was an overall advantage for feedback training over observational training, but most importantly, type of training interacted strongly with type of category structure. With rule-based structures, the effects of training type were small, but with information-integration structures, accuracy was substantially higher with feedback training, and people were less likely to use suboptimal rule-based strategies. The implications of these results for current theories of category learning are discussed.

We learn about categories in a variety of ways. Sometimes we are told the category label prior to viewing the object—for example, when the excited parent spots a deer crossing the road and exclaims, “Look, there’s a deer.” Other times, we view objects, make categorization responses, and receive feedback. For example, the child might spot some four-legged furry object and exclaim, “Look, there is a deer.” The parent then verifies whether or not this observation is correct.

Although we learn about categories in a number of different ways, nearly all research on category learning uses an approach in which a single stimulus is presented, the observer is required to generate a response, and corrective feedback is provided (hereafter called “feedback training”). A few studies have examined a different approach, in which the stimulus and the category label are presented simultaneously and no response is required (hereafter called “observational training”; e.g., Estes, 1976, 1994; Izawa, 1967; Reber & Millward, 1968). Performance in these ob-

servational conditions differed from performance in standard feedback training conditions in the following ways. First, observational training led to superior performance early in learning, but later in learning performance was about the same or there was a slight advantage for feedback training (from Estes, 1994). Unfortunately, however, it is impossible to determine the locus of the performance difference in these studies since the feedback and observational conditions differed in two important ways. First, during standard feedback training a response is required, whereas in observational training no response is required. Second, during feedback training the identity of the correct category is revealed after the response is made, whereas in observational training the category label is presented before the stimulus appears.

A major goal of the present study was to disentangle this confound by combining factorially two levels of motor response (response vs. no response) with two levels of training type (category label before stimulus vs. category label after stimulus). Note that traditional feedback training is equivalent to our response/after condition, whereas observational training is equivalent to our no response/before condition. To our knowledge, no previous studies have included response/before or no response/after conditions.

A variety of recent empirical and theoretical results suggest that the effects of these experimental manipulations might be different depending on the specific category structures that are chosen. In this article, we focus on two differ-

---

This research was supported in part by National Science Foundation Grant SBR-9796206 and National Institutes of Health Grant R01 MH59196. We thank Lorilei Cardenas and Amy Kaderka for help with data collection. Correspondence should be addressed to F. G. Ashby, Department of Psychology, University of California, Santa Barbara, CA 93106 (e-mail: ashby@psych.ucsb.edu), or to W. T. Maddox, Department of Psychology, Mezes Hall 330, University of Texas, Austin, TX 78712 (e-mail: maddox@psy.utexas.edu).

ent types of category structure (Ashby & Ell, 2001). *Rule-based category learning tasks* are those in which the category structures can be learned via some explicit reasoning process. Frequently, the rule that maximizes accuracy (i.e., the optimal rule) is easy to describe verbally (Ashby, Alfonso-Reese, Turken, & Waldron, 1998). In the most common applications, only one stimulus dimension is relevant, and the observer's task is to discover this relevant dimension and then to map the different dimensional values to the relevant categories. Rule-based tasks have a long history in cognitive psychology, and not surprisingly they have been popular with proponents of the so-called classical theory of categorization, which assumes category learning is the process of discovering the set of necessary and sufficient conditions that determine category membership (e.g., Smith & Medin, 1981).

*Information-integration category learning tasks* are those in which accuracy is maximized only if information from two or more stimulus components (or dimensions) is integrated at some predecisional stage (Ashby & Gott, 1988).

Perceptual integration could take many forms—from treating the stimulus as a Gestalt to computing a weighted linear combination of the dimensional values. However, a conjunction rule (e.g., “Respond A if the stimulus is small on dimension  $x$  and small on dimension  $y$ ”) is a rule-based task rather than an information-integration task because separate decisions are first made about each dimension (e.g., small or large) and then the outcome of these decisions is combined (integration is not predecisional). In many cases, the optimal rule in information-integration tasks is difficult or impossible to describe verbally (Ashby et al., 1998). Information-integration tasks have been favorites of exemplar theorists, who argue that categorization requires accessing the memory representations of every previously seen exemplar from each relevant category (Estes, 1994; Medin & Schaffer, 1978; Nosofsky, 1986), and of decision bound theorists, who argue that category learning is a process of associating category labels with regions of perceptual space (Ashby & Gott, 1988; Ashby & Maddox, 1992, 1993). In typical applications, however,

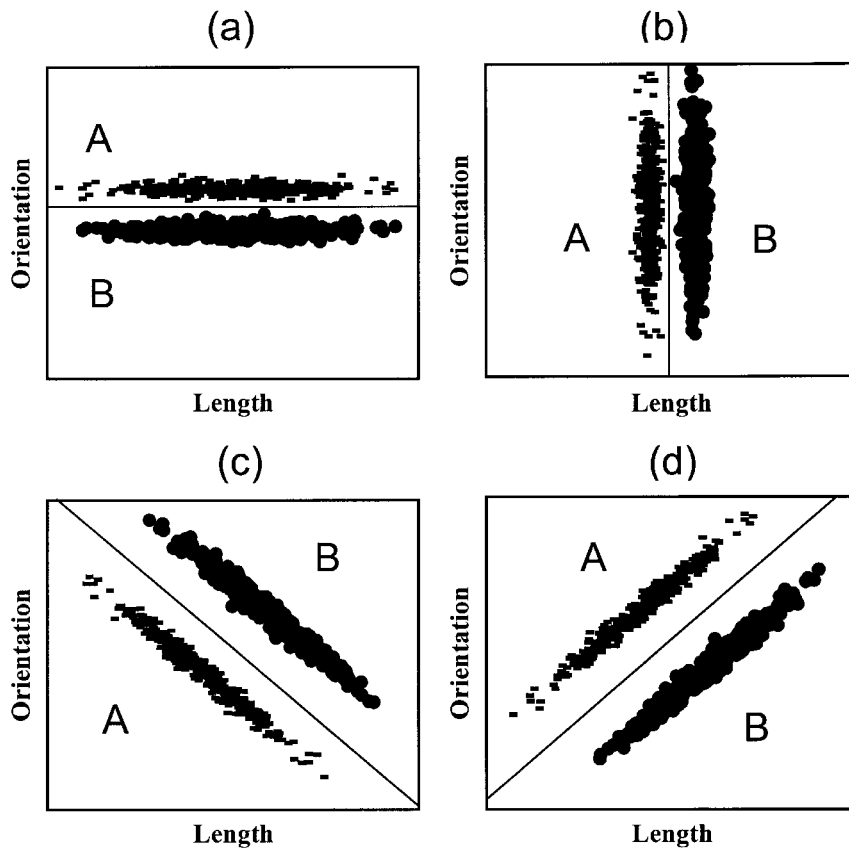


Figure 1. Category structures used in Experiment 1: (a) unidimensional-orientation (Uni-O), (b) unidimensional-length (Uni-L), (c) diagonal-negative (Diag-Neg), and (d) diagonal-positive (Diag-Pos). Each dash denotes the length and orientation of a line from Category A. Each dot denotes the length and orientation of a line from Category B. The solid line in each panel denotes the optimal decision bound.

exemplar theorists have used information-integration tasks with few exemplars per category, whereas decision bound theorists have used many exemplars per category.

The four category structures used in the present study are described in Figure 1. Every stimulus was a single line that varied across trials in length and orientation. Each symbol in Figure 1 denotes the length and orientation of a single stimulus. Category A exemplars are denoted by dashes and Category B exemplars are denoted by dots. In each condition, there were two distinct categories that did not overlap, so perfect accuracy was always possible.

The category structures in the four conditions were generated by successively rotating the categories shown in the upper left panel of Figure 1 by increments of 45°. Thus, optimal accuracy, within-category scatter, and category coherence are all identical in the four conditions (Ashby, Queller, & Berretty, 1999). Also shown in Figure 1 are the decision bounds that maximize categorization accuracy. In two conditions the optimal bound is unidimensional and in two conditions it is diagonal. In the two diagonal conditions, the most accurate unidimensional rule yields a response accuracy of about 80%. In addition, because of the continuous-valued stimulus dimensions, it would be difficult or impossible to respond optimally in the diagonal conditions by using a unidimensional rule and memorizing exceptions.

Both unidimensional conditions are rule-based tasks. In both cases, there is a simple explicit rule that separates the contrasting categories. For example, the vertical bound in the upper right panel of Figure 1 corresponds to the following rule: "Respond A if the line is short and B if it is long." In contrast, the two diagonal conditions are examples of information-integration tasks. In both of these conditions, perfect accuracy requires integrating length and orientation information, and there is no simple verbal description of the optimal decision bound.

In every condition of our experiments, each observer completed a series of 10 blocks. The odd-numbered blocks served as training and were used to instantiate one of the four different training procedures that were constructed from the factorial combination of the two levels of motor response (response vs. no response) and the two levels of training (observational vs. feedback<sup>1</sup>). The even-numbered blocks contained transfer trials, during which observers made categorization responses, but no feedback was provided. Each observer received only one type of training and only one level of motor response.

The stimuli and category structures shown in Figure 1 were used by Ashby, Queller, and Berretty (1999) in a study of unsupervised categorization. In a series of experiments, observers were told the number of categories (i.e., two) and that perfect accuracy was possible, and they were given extensive experience in the task (i.e., 800 trials), but they were never given any feedback about the accuracy of their responses. Thus, participants in the Ashby, Queller, and Berretty study received no training of any kind, in contrast to our participants, who all received some form of train-

ing (either observational or feedback). Despite the lack of training in the Ashby, Queller, and Berretty study, however, in both unidimensional conditions, all participants learned to respond optimally. However, in the diagonal conditions, none of the participants responded optimally. Instead they all used some sort of unidimensional rule, even when explicitly encouraged to use both stimulus dimensions. When trial-by-trial feedback was provided in a separate experiment, all participants responded optimally in the diagonal conditions. These results provide clear evidence that the nature of the feedback interacts with the type of category structure.

A wide variety of evidence now indicates that the learning of rule-based and information-integration category structures might be mediated by different neural circuits, and this literature provides further reason to expect that the experimental manipulations studied in this article might have different effects on these two types of tasks (Ashby et al., 1998; Ashby & Ell, 2002; Erickson & Kruschke, 1998; Smith, Patalano, & Jonides, 1998; Waldron & Ashby, 2001). In particular, Ashby and his colleagues have proposed that with rule-based structures, learning is mediated by a circuit that includes the anterior cingulate, the prefrontal cortex, and the head of the caudate nucleus, whereas in information-integration tasks, learning is mediated largely within the tail of the caudate nucleus (with visual stimuli) (Ashby et al., 1998; Ashby & Ell, 2001; Ashby, Isen, & Turken, 1999; Ashby & Waldron, 1999).

A review of all the evidence supporting this hypothesis is beyond the scope of this article. Instead, we mention only a few of the more prominent supporting results. We begin with a few results supporting the role of the anterior cingulate, prefrontal cortex, and head of the caudate nucleus in rule-based category learning. First, patients with lesions of the prefrontal cortex are well known to be impaired in rule-based tasks (e.g., such as the Wisconsin Card Sorting Test; e.g., Robinson, Heaton, Lehman, & Stilson, 1980), but not in information-integration tasks (Knowlton, Mangels, & Squire, 1996). Second, an fMRI study of a rule-based task showed activation in the right dorsal-lateral prefrontal cortex, the anterior cingulate, and the head of the right caudate nucleus (Rao et al., 1997). Third, many studies implicate these structures as key components of executive attention (e.g., Posner & Petersen, 1990) and working memory (e.g., Goldman-Rakic, 1987, 1995), both of which are likely to be critically important to the explicit processes of rule formation and testing that are assumed to mediate rule-based category learning. Fourth, a recent neuroimaging study identified the (dorsal) anterior cingulate as the site of hypothesis generation in a rule-based category-learning task (Elliott & Dolan, 1998). Finally, lesion studies in rats implicate the dorsal caudate nucleus in rule switching (Winocur & Eskes, 1998).

A prominent role for the tail of the caudate nucleus in (visual) information-integration category learning is supported by several studies that have reported information-integration category learning deficits in patients with dis-

eases of the basal ganglia (i.e., Parkinson's or Huntington's disease; Filoteo, Maddox, & Davis, 2001a; Knowlton et al., 1996; Maddox & Filoteo, 2001). In contrast, medial temporal lobe amnesic patients are normal (Filoteo, Maddox, & Davis, 2001b). Other evidence for a basal ganglia contribution to category learning comes from a long series of lesion studies in rats and monkeys that show that the tail of the caudate nucleus<sup>2</sup> is both necessary *and* sufficient for normal visual discrimination learning. In primates, all of extrastriate visual cortex projects directly to the tail of the caudate nucleus, and the cells in this area then project, via the globus pallidus (the output portion of the basal ganglia) and thalamus, to the prefrontal and premotor cortices. These projections place the caudate in an ideal position to link percepts and actions, and many researchers have hypothesized that this is its primary role (e.g., Rolls, 1994; J. Wickens, 1993). Many studies have shown that lesions of the tail of the caudate nucleus impair the ability of animals to learn visual discriminations that require one response to one stimulus and a different response to some other stimulus (e.g., McDonald & White, 1993, 1994; Packard, Hirsh, & White, 1989; Packard & McGaugh, 1992). Since the visual cortex is intact in these animals, it is unlikely that their difficulty is in perceiving the stimuli. Rather, it appears that their difficulty is in learning to associate an appropriate response with each stimulus alternative. Technically, such studies are categorization tasks with one exemplar per category. It is difficult to imagine how adding more exemplars to each category could alleviate the deficits caused by caudate lesions, and it is for this reason that the caudate lesion studies support the hypothesis that the caudate contributes to normal category learning. The sufficiency of the caudate nucleus for visual discrimination learning was shown in a series of studies by Gaffan and colleagues that lesioned all pathways out of visual cortex except into the tail of the caudate (e.g., projections into prefrontal cortex were lesioned by Eacott & Gaffan, 1992, and Gaffan & Eacott, 1995; projections to the hippocampus and amygdala were lesioned by Gaffan & Harrison, 1987). None of these lesions prevented visual discrimination learning.

These results are important because the mechanisms that mediate learning-related changes in synaptic efficacy within these two neural circuits are qualitatively different, and such differences suggest that providing the category label before versus after the stimulus may have different effects on rule-based and information-integration tasks.

The rule-based category learning system proposed above is under conscious control and has full access to working memory and executive attention. As a result, the timing of the category label should not matter too much in rule-based tasks. In contrast, an information-integration category learning system that is mediated within the tail of the caudate nucleus would not be accessible to conscious awareness and is far removed from working memory. As a result, it would depend more heavily on local learning mechanisms.

Within the tail of the caudate nucleus, a reward-mediated feedback signal is thought to be provided by dopamine released from the substantia nigra (e.g., J. Wickens, 1993). Specifically, dopamine is released into the tail of the caudate (among other regions) from the substantia nigra (pars compacta) shortly after the animal receives an unexpected reward (Hollerman & Schultz, 1997; Schultz, 1992), and the presence of this dopamine is widely thought to strengthen recently active synapses (which presumably are responsible for the animal obtaining the reward) (e.g., Arbuthnott, Ingham, & Wickens, 2000; Calabresi, Pisani, Centonze, & Bernardi, 1996). However, observing a category label before seeing the stimulus is not inherently rewarding, and even if it were, it is too removed in time from the categorization response for such a procedural-learning mechanism to operate efficiently. According to this argument, learning in information-integration tasks should be more difficult when the category label is provided before the stimulus than when it occurs after the response.

## EXPERIMENT 1

### Method

**Observers and Design.** Eighty observers were solicited from the University of Texas community and received course credit for participation. Five observers participated in each of 16 experimental conditions constructed from the factorial combination of two levels of motor response (response vs. no response), two levels of training type (observational vs. feedback), and four category structures. No observer participated in more than one experimental condition. All observers reported 20/20 vision or vision corrected to 20/20. Each observer completed one session of approximately 60 min duration.

**Stimuli and stimulus generation.** The experiment used the so-called randomization technique introduced by Ashby and Gott (1988). The four pairs of category structures are displayed in Figure 1. Each point in the space represents the line length and line orientation for a single stimulus. Category A stimuli are denoted by the dashes and Category B stimuli are denoted by dots. The stimuli from each category were generated by sampling randomly from a bivari-

**Table 1**  
Category Distribution Parameters Used in Experiment 1

Condition	Category A					Category B				
	$\mu_x$	$\mu_y$	$\sigma_x^2$	$\sigma_y^2$	$cov_{xy}$	$\mu_x$	$\mu_y$	$\sigma_x^2$	$\sigma_y^2$	$cov_{xy}$
Uni-O	300	165	9,000	75	0	300	85	9,000	75	0
Uni-L	260	125	75	9,000	0	340	125	75	9,000	0
Diag-Neg	243	68	4,538	4,538	-4,463	357	182	4,538	4,538	-4,463
Diag-Pos	243	182	4,538	4,538	4,463	357	68	4,538	4,538	4,463

Note—Uni-O, unidimensional-orientation; Uni-L, unidimensional-length; Diag-Neg, diagonal-negative; Diag-Pos, diagonal-positive.

ate normal distribution. Each category distribution is specified by a mean and a variance on each dimension and by a covariance between dimensions. For each of the four category structures it was always the case that the covariance matrix for Category A was identical to the covariance matrix for Category B. The categories differed only in the location of the means. Under these conditions the optimal decision bound will be linear. The optimal decision bound is displayed for each of the four conditions in Figure 1. The exact parameter values are displayed in Table 1. The category structures were taken from a study by Ashby, Queller, and Berretty (1999).

To begin, we generated the stimuli for the unidimensional-orientation condition (Uni-O; Figure 1a) by randomly sampling 400 stimuli from Category A and 400 stimuli from Category B. The order of these 800 stimuli was randomized separately for each observer and then divided into 10 blocks of 80 trials each. The odd-numbered blocks served as the training blocks and were used to instantiate the four different training procedures that were constructed from the factorial combination of two levels of motor response (response vs. no response) and two levels of training type (observational vs. feedback). The even-numbered blocks served as the transfer blocks, during which observers generated categorization responses but no feedback was provided. The stimuli for the unidimensional-length category structures (Uni-L; Figure 1b) were generated by rotating the Uni-O categories 90° counterclockwise. The stimuli for the diagonal-negative category structures (Diag-Neg; Figure 1c) and diagonal-positive category structures (Diag-Pos; Figure 1d) were generated by rotating the Uni-O categories 135° counterclockwise and 45° counterclockwise, respectively, around a central point located at 300 pixels in length and 45° in orientation, and by shifting each stimulus 40 units away from the optimal decision bound along the relevant dimension, thereby increasing the discriminability of the categories. The categories used in the unidimensional conditions were less discriminable than those used in the diagonal conditions because a pilot study indicated that equal discriminability led to ceiling effects for the unidimensional conditions. Importantly, though, note that perfect performance is possible in all four conditions.

The stimuli were computer generated and displayed on a 21-in. monitor with  $1,360 \times 1,024$  resolution. Each line was presented in white on a black background. To minimize line jaggedness, Alfonso-Reese's (1997) anti-aliasing routine, developed for use with Brainard's (1997) Psychophysics Toolbox, was applied. Each stimulus was created by converting the  $x$  value into a line length and the  $y$  value (after applying the scaling factor  $\pi/500$ ) into a line orientation. The scaling factor  $\pi/500$  was chosen in an attempt to equate the salience of line length and line orientation.

**Procedure.** Each observer was run individually in a dimly lit testing room. The observers were informed that there were two categories denoted "A" and "B," and that each category was equally likely. They were instructed to learn about the two categories during the training blocks so that they could accurately separate the stimuli into two categories during the transfer blocks. They were informed that perfect performance was possible. The motor response and training type manipulations were instantiated during the training blocks. The procedure for a typical training trial in each of the four training conditions was as follows: (1) No response/before: category label (500 msec), blank screen (500 msec), stimulus (500 msec); response/before: category label (500 msec), blank screen (500 msec), stimulus (500 msec), response; response/after<sup>3</sup> stimulus (500 msec), blank screen (500 msec), response-category label (500 msec); no response/after: stimulus (500 msec), blank screen (500 msec), category label (500 msec), where the category label was a simple "A" or "B," depending on which category was presented.

In each condition the intertrial-interval was 500 msec. The procedure was identical for all transfer trials. During a typical transfer trial the stimulus was presented for 1 sec followed by a blank screen. The observer generated a categorization response by pressing either the "Z" key, for A responses, or the "F" key, for B responses. The ob-

server's response was followed by initiation of the next trial. An observer-paced break followed each block of trials.

## Results and Theoretical Analyses

Analyses were performed only on data from the transfer blocks. In the first section (entitled ANOVA results) we analyze the accuracy rates using analysis of variance (ANOVA). In the second section (entitled Modeling results) we introduce the model-based analyses.

**ANOVA results.** We begin with a set of analyses to determine whether performance differed between the two unidimensional conditions (Uni-O and Uni-L) and between the two diagonal conditions (Diag-Neg and Diag-Pos). The main effect of category structure was significant for the unidimensional conditions [ $F(1,32) = 7.873$ ,  $p < .01$ ], yielding higher accuracy in the Uni-O condition (95.5%) than in the Uni-L condition (90.4%), but not for the diagonal conditions [ $F(1,32) = 1.544$ , n.s.]. For the unidimensional conditions there was also a significant condition  $\times$  block interaction [ $F(4,128) = 2.827$ ,  $p < .05$ ], revealing different learning rates for the two conditions. Post hoc analyses indicated that performance in the Uni-O and Uni-L conditions differed significantly during the first two blocks of trials (Uni-O = 95.4%, 96.8%, and Uni-L = 84.1%, 89.1%, For Blocks 1 and 2, respectively), but did not differ across the last three blocks (Uni-O = 94.98%, 95.5%, 94.8%, and Uni-L = 90.8%, 93.4%, 94.4% for Blocks 3–5, respectively). Because our focus is on asymptotic performance, we collapsed across the two unidimensional and two diagonal conditions throughout the remainder of the analyses.

The accuracy during the last transfer block in each of the eight experimental conditions is shown in Figure 2. A 2 (motor response; no/yes)  $\times$  2 (training types; observational/feedback)  $\times$  2 (category structures; unidimensional/diagonal) ANOVA revealed a main effect of training type [ $F(1,72) = 8.595$ ,  $p < .01$ ] and category structure [ $F(1,72) = 27.908$ ,  $p < .001$ ], suggesting that performance was superior with feedback training (accuracy: before = 85.5%; after = 92.0%), and in the unidimensional conditions (accuracy: Uni = 94.6%; Diag = 82.9%). The main effect of motor response was not significant [ $F(1,72) = 1.81$ ,  $p > .10$ ]. The ANOVA also revealed an interaction between training type and category structure [ $F(1,72) = 7.785$ ,  $p < .01$ ]. As predicted by the neuropsychological hypothesis discussed above, feedback training provided a significant increase in performance in the diagonal conditions, but not in the unidimensional conditions. This same interaction occurred for both the no response and response data separately, although it was only marginally significant in the response conditions [no response:  $F(1,36) = 5.067$ ,  $p < .05$ ; response:  $F(1,36) = 3.172$ ,  $p < .10$ ]. The response  $\times$  training type [ $F(1,72) = 0.65$ ,  $p > .10$ ], response  $\times$  category structure [ $F(1,72) = 0.34$ ,  $p > .10$ ], and the three-way interaction [ $F(1,72) = 0.01$ ,  $p > .10$ ] were all nonsignificant.

The ANOVA results can be summarized as follows. First, performance was superior with rule-based category struc-

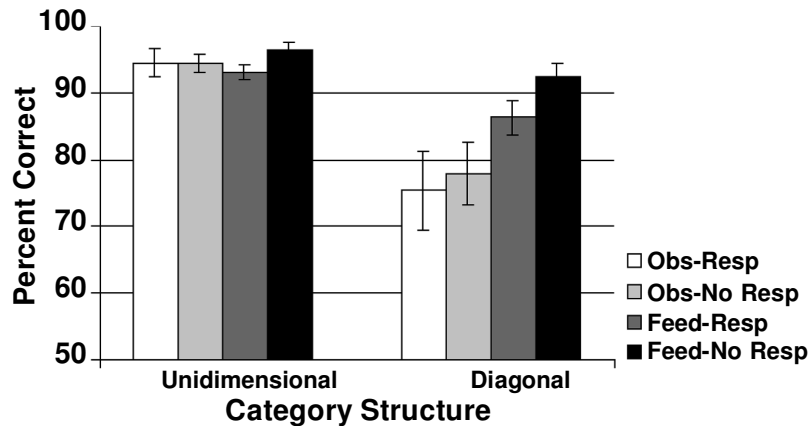


Figure 2. Accuracy during the final block of trials for each of the eight conditions in Experiment 1. Standard errors are also indicated. Obs, observational training; Feed, feedback training; Resp, motor response required; No Resp, no motor response required.

tures. Second, performance was better with feedback training than with observational training. Third, accuracy did not depend on whether observers made a motor response during training. Finally, training type interacted with the nature of the categorization task. Specifically, feedback training was no better than observational training with the unidimensional categories, but it led to a substantial performance improvement with the diagonal categories.

**Modeling results.** To get a more detailed description of how observers categorized the stimuli, a number of different decision bound models (Ashby, 1992a; Maddox & Ashby, 1993) were fit to each observer's responses. Decision bound models are derived from general recognition theory (GRT; Ashby & Townsend, 1986), which is a multivariate generalization of signal detection theory (e.g., Green & Swets, 1966). The fundamental assumption of GRT is that there is trial-by-trial variability in the perceptual information obtained from every stimulus, no matter what the viewing conditions (Ashby & Lee, 1993). On each trial, however, it is assumed that the percept can be represented as a point in a multidimensional psychological space. Decision bound theory assumes each observer partitions the perceptual space into response regions by constructing a decision bound. On each trial, the observer determines which region the percept is in and then emits the associated response. Despite this deterministic decision rule, decision bound models predict probabilistic responding because of trial-by-trial perceptual and criterial noise.

Two different types of decision bound models were fit to each observer's responses (see Ashby, 1992a; Maddox & Ashby, 1993, for a more formal treatment of these models). One type assumed observers used an explicit rule-based strategy and one type assumed an information-integration strategy. With feedback training, of course, we predict observers will use a rule-based strategy in the unidimensional conditions and an information-integration

strategy in the diagonal conditions. If observational training disrupts information-integration category learning, then we might expect an increase in rule-based responding under these conditions.

*Rule-based models.* Two models assumed observers used an explicit rule-based strategy. The *unidimensional model* assumes the observer sets a criterion on a single perceptual dimension and then makes an explicit decision about the level of the stimulus on that dimension (Ashby & Gott, 1988; Shaw, 1982). For example, in the present experiments, the observer might use the following rule:

Respond A if the line is short and B if it is long.

A different version of this model assumes the observer attends selectively to orientation, rather than length. The unidimensional models have two free parameters: a decision criterion on the relevant perceptual dimension and the variance of internal (perceptual and criterial) noise (i.e.,  $\sigma^2$ ). In the unidimensional conditions, a special case of the unidimensional model assumes observers use the unidimensional decision bound that maximizes accuracy (i.e., the vertical and horizontal bounds shown in Figure 1). This special case has only one free parameter (i.e., noise variance).

The *conjunction model* assumes observers use a conjunction rule in which separate decisions are made about the levels on the two dimensions and then a response is selected based on the outcome of these two decisions. For example, in the Diag-Neg condition, an observer might use the following explicit rule:

Respond A if the line is short and its orientation is small, otherwise respond B.

This rule partitions the perceptual space into four regions, one of which is assigned to Category A and three of which are assigned to Category B. As mentioned in the introduction, such a strategy is rule based because it is easy to

describe verbally and it does not require perceptual integration of length and orientation.

Although a number of distinct conjunction rules can be constructed, only two of these were reasonable for each of the diagonal conditions. Specifically, for the Diag-Neg condition, the following two rules are plausible:

Respond A if length and orientation are both small,  
otherwise respond B,

and

Respond B if length and orientation are both large,  
otherwise respond A.

For the Diag-Pos condition, the following two rules make sense:

Respond A if the line is short and its orientation is large,  
otherwise respond B,

and

Respond B if the line is long and its orientation is small,  
otherwise respond A.

Conjunction models have three parameters (a criterion on each dimension, and  $\sigma^2$ ).

*Information-integration models.* The general linear classifier (GLC) assumes that the decision bound between each pair of categories is linear. This produces an information-integration decision strategy because it requires linear integration of perceived length and orientation. The GLC has three parameters (slope and intercept of the linear bound and  $\sigma^2$ ). In the diagonal conditions, a special case of the GLC assumes observers use the linear bound that maximizes accuracy (i.e., the diagonal bounds shown in Figure 1). This model has only one free parameter (noise variance).

*Model fits.* Each of these models was fit separately to the data from each of the five transfer blocks for every observer. The model parameters were estimated using max-

imum likelihood (Ashby, 1992b; T. D. Wickens, 1982) and the goodness-of-fit statistic was

$$AIC = 2r - 2\ln L,$$

where  $r$  is the number of free parameters and  $L$  is the likelihood of the model given the data (Akaike, 1974; Takane & Shibayama, 1992). The AIC statistic penalizes a model for extra free parameters such that the smaller the AIC, the closer a model is to the “true model,” regardless of the number of free parameters. Thus, to find the best model among a given set of competitors, one simply computes an AIC value for each model and chooses the model associated with the smallest AIC value.

For each of 400 data sets (four training conditions  $\times$  four category structures  $\times$  five observers  $\times$  five blocks), we determined which of the two model types provided the best account of the data (i.e., rule-based or information integration). The results of these fits were similar for the Diag-Neg and Diag-Pos conditions, so these were collapsed across conditions. In addition, because the results were also similar in the two unidimensional conditions, the modeling results were collapsed across the Uni-O and Uni-L conditions. The percentages of data sets for which each model type provided the best account by motor response, training type, category structure, and block are presented in Table 2.

Table 2 indicates the following: First, as expected, with traditional feedback training, observers were most likely to use a rule-based strategy in the unidimensional conditions and an information-integration strategy in the diagonal conditions [ $F(1,16) = 38.21, p < .001$ ], especially during the later blocks. The tendency for feedback training to cause an eventual shift to information-integration strategies in the diagonal conditions is especially evident when a response was required. During the first two transfer blocks, most observers used a rule-based strategy, but by Block 4 the numbers reverse, showing that most observers used an information-integration strategy. Second, with ob-

**Table 2**  
Percentage of Observers in Experiment 1 (by Category Structure and Block) Whose Data Were Best Fit by a Rule-Based Model or by an Information-Integration (Inf-Int) Model

Block	Observational Training						Feedback Training					
	1	2	3	4	5	Ave.	1	2	3	4	5	Ave.
Response												
Uni												
Rule-based	80	40	80	80	70	70	70	90	100	70	80	82
Inf-int	20	60	20	20	30	30	30	10	0	30	20	18
Diag												
Rule-based	30	60	50	80	60	56	70	70	50	30	30	50
Inf-int	70	40	50	20	40	44	30	30	50	70	70	50
No Response												
Uni												
Rule-based	70	80	70	80	90	78	70	60	90	90	60	74
Inf-Int	30	20	30	20	10	22	30	40	10	10	40	26
Diag												
Rule-based	40	50	40	40	40	42	30	30	20	30	10	24
Inf-int	60	50	60	60	60	58	70	70	80	70	90	76

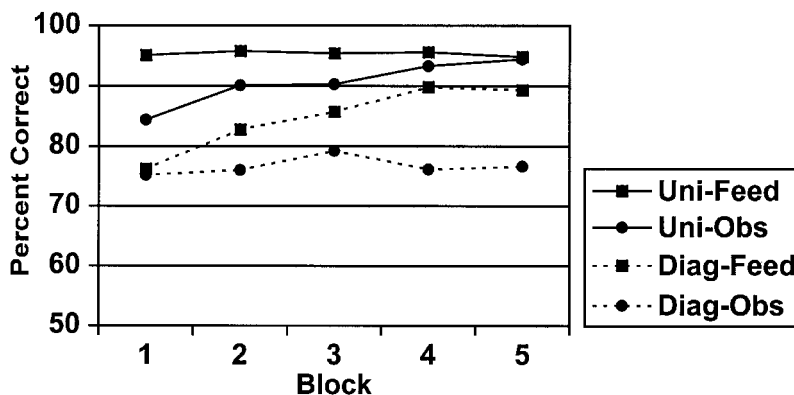


Figure 3. Average accuracy in Experiment 1 during each of the five transfer blocks. Uni, unidimensional category structure; Diag, diagonal category structure; Obs, observational training; Feed, feedback training.

servational training, most observers again used a rule-based strategy in the unidimensional conditions [ $F(1,16) = 17.36, p < .001$ ], and in fact training type had no effect on this preference [ $F(1,16) = 0.42, p > .1$ ]. Third, and perhaps most importantly, observers were more likely to use a rule-based strategy<sup>4</sup> in the diagonal conditions when they had observational training (49%) than when they had feedback training (37%) [ $F(1,16) = 3.47, p = .08$ ].

## EXPERIMENT 2

One possibility, which is important to consider, is that observational training was as effective as feedback training with the unidimensional category structures because these tasks were so easy for observers to learn. According to this idea, the observational training data displays a ceiling effect wherein performance is so good that there is little room for improvement with feedback training. One way to test this hypothesis is to examine performance in the early transfer blocks. These data are shown in Figure 3, collapsed over rule type (i.e., Uni-O and Uni-L for the unidimensional rules and Diag-Pos and Diag-Neg for the diagonal rules) and motor response (yes and no).

Note that there is a trend for observational training to be less effective than feedback training with the unidimensional rules in the early transfer blocks, although only the Block 1 difference is significant ( $p < .05$ ). As observers gain practice, this difference disappears. In contrast, with the diagonal categories, performance is initially identical with observational and feedback training, but with more experience, observers given feedback training continually improve, whereas there is no evidence of additional learning in the group given observational training. Although the Figure 3 results indicate a clear dissociation between the learning of the unidimensional and diagonal rules, they also raise several important questions. First, is the early advantage of feedback training on unidimensional rule learning real? And second, is there a ceiling effect with unidimensional rule learning? In particular, if the unidi-

dimensional tasks had been more difficult, perhaps the feedback training group would have maintained its initial advantage over the observational group. Experiment 2 was designed to answer these questions. Essentially, we replicated the design of the unidimensional rule conditions of Experiment 1, but increased the difficulty.

## Method

**Observers and Design.** Twenty observers were solicited from the University of Texas community; they received course credit for participation. Five observers participated in each of four experimental conditions constructed from the factorial combination of two levels of training (observational vs. feedback), and two rule-based category structures (Uni-O vs. Uni-L). Observers generated a motor response in all conditions. All observers reported 20/20 vision or vision corrected to 20/20. Each observer completed one session of approximately 60 min duration.

**Stimuli and stimulus generation.** The stimulus dimensions were the same as those from Experiment 1. The two category structures were identical to the Uni-O and Uni-L category structures from Experiment 1 except that the Category A and Category B means were each displaced ten units toward the optimal category boundary, thus lowering category discriminability. In fact, the stimuli used in Experiment 2 were generated by moving every stimulus from the unidimensional conditions of Experiment 1, ten units toward the optimal decision bound. Thus, the parameter values shown in Table 1 were identical to those used in Experiment 2, except that in Experiment 2 mean orientations for the two categories in the Uni-O conditions were 95 and 155, respectively, and mean lengths in the Uni-L conditions were 270 and 330, respectively. All other aspects of the stimuli were identical to those from Experiment 1.

**Procedure.** The procedures were identical to those for the response/before and response/after conditions from Experiment 1.

## Results and Theoretical Analyses

An ANOVA revealed a main effect of rule type [ $F(1,16) = 10.23, p < .01$ ]. In each of the five transfer blocks, accuracy was higher for the unidimensional rule on orientation than for the rule on length (average accuracy was 94.3% on orientation and 80.3% on length). There was also a main effect of block [ $F(4,64) = 2.78, p < .05$ ], indicating that performance did significantly improve from the first

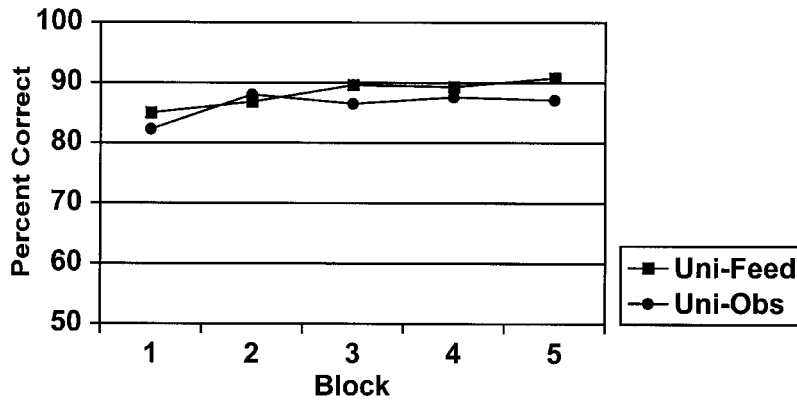


Figure 4. Average accuracy in Experiment 2 during each of the five transfer blocks. Uni, unidimensional category structure; Obs, observational training; Feed, feedback training.

transfer block to the last, and therefore that learning did occur. Most importantly, there was no effect of training type ( $p > .50$ ). This is easily seen in Figure 4, which shows the percentage of correct responses in each of the five transfer blocks (averaged across rule type). Note that, as expected, accuracy is lower than in Experiment 1. Most importantly, in contrast to Experiment 1, the feedback training group gradually increases its accuracy across the transfer blocks. Despite this fact, accuracy is essentially the same in the observation and feedback groups, and this equivalence persists from the first transfer block through the last, unlike what occurred in Experiment 1.

Table 3 shows the results of the modeling analyses (averaged across rule type). Note that there were no significant differences between decision strategies used by observers in the observational and feedback conditions [ $F(1,8) = 1.22, p > .20$ ]. In particular, the majority of observers used a rule-based strategy in both conditions.<sup>5</sup>

The Experiment 2 results strongly argue that the differences observed in Experiment 1 between unidimensional and diagonal rule learning were not due to a ceiling effect in the unidimensional conditions. Even when difficulty is increased, there is no difference between observational and feedback training when observers are learning unidimensional categorization rules. Second, the lack of any difference in Experiment 2 between observational and feedback training during the early transfer blocks suggests that the trend for such a difference that we observed in Experiment 1 may not be real.

## GENERAL DISCUSSION

The most important result reported in this article is the interaction that we found between category structure and type of categorization training. With our rule-based category structures, it made very little difference whether observers received observational or feedback training. In particular, when accuracy was maximized via some explicit rule that is easy to verbalize, observational training was nearly as effective as traditional feedback training. However, with category structures that required information integration for optimal responding, traditional feedback training was more effective than observational training. More specifically, with observational training people were less accurate and were more likely to use suboptimal rule-based strategies than with feedback training.

These results are generally consistent with the hypothesis that learning in these two types of categorization tasks is mediated by separate systems. According to the version of this hypothesis described in the introduction, learning in rule-based tasks is dominated by an explicit system that is mediated by frontal cortical structures (and the head of the caudate nucleus). This system is assumed to have access to working memory and executive attention. Learning occurs through a process of hypothesis generation and testing, and there is no close relationship with motor output systems. For these reasons, whether or not a response is required should have little effect on rule-based learning; nor should the position of the category label. In fact,

Table 3  
Percentage of Observers in Experiment 2 Whose Data Were Best Fit by a Rule-Based Model (i.e., the Unidimensional, Conjunctive, or Optimal) or by an Information-Integration Model (i.e., the GLC)

Block	Observational Training					Feedback Training						
	1	2	3	4	5	Ave.	1	2	3	4	5	Ave.
Rule-based	70	40	80	80	60	66	60	50	70	40	60	56
Information-integration	30	60	20	20	40	34	40	50	30	60	40	44

Ashby, Queller, and Berretty (1999) showed that with these same stimuli and category structures, people learn to respond optimally in the rule-based conditions *even when no category label is ever shown*.

In contrast, according to the hypothesis reviewed earlier in this article, optimal responding in the information-integration conditions requires the participation of an implicit system, in which learning is mediated largely within the tail of the caudate nucleus. This system is assumed to use a form of procedural learning that is more closely tied to motor systems and that depends on a dopamine-mediated reward signal for learning. The idea is that an unexpected reward causes dopamine to be released from the substantia nigra into the tail of the caudate nucleus, and that the presence of this dopamine strengthens recently active synapses (e.g., Schultz, 1992; J. Wickens, 1993). This is a reward-mediated version of Hebbian stimulus–response association learning. Of course, it is not clear what sorts of internal events occur following stimulus presentation during observational training (e.g., caused by confirming or failing to confirm expectations), but it seems clear that this system should learn best when feedback follows a motor response. A slightly stronger prediction is that learning should continue to deteriorate the more severely these conditions are violated. In the present experiment, the conditions least conducive to this form of implicit learning occur when the label comes before the stimulus and no response is required.

Our results indicate that whether or not a response was required had little effect on accuracy or on the type of learning that occurred in the diagonal conditions. This, at first, might seem surprising, but of course we have no way of knowing what observers in the no response conditions were doing—we only know that they did not fully depress the response keys. In particular, they could have been making a subvocal or submanual response. They could have been imagining a response, or perhaps they were consciously associating the label with the response position. Since we can't rule out any of these possibilities, it is difficult to draw strong conclusions from our failure to find differences between the response and no response data.

On the other hand, our results indicate a substantial difference between the decision strategies used in the diagonal conditions depending on whether observers received observational or feedback training. In both the response and no response conditions, observers were less accurate and were more likely to use a rule-based strategy with observational training. In the no response conditions, this difference in decision strategy persisted throughout all transfer blocks, but it was equally dramatic in the later blocks of the response conditions. In particular, during the last two transfer blocks of the response conditions, 70% of the observers receiving observational training were using a rule-based strategy, whereas only 30% of the observers receiving feedback training were using a rule-based strategy.

According to the COVIS theory of category learning (Ashby et al., 1998), explicit and implicit systems compete during category learning. Initially, however, the explicit

system dominates. So, COVIS predicts that in the absence of significant learning in the implicit system, people will persist in using explicit, rule-based strategies. This prediction is supported by our results, and also by the results of Ashby, Queller, and Berretty (1999), who used these same categories in a study of unsupervised categorization. In particular, in the absence of feedback, Ashby, Queller, and Berretty found that all observers in the diagonal conditions used rule-based strategies.

An important alternative hypothesis that must be considered is that a single system mediates learning in the unidimensional and diagonal conditions, and that our results occurred only because the unidimensional categories are easier to learn and therefore more resistant to the interference caused by observational training. Although this hypothesis has intuitive appeal, it has several serious weaknesses. First, because the unidimensional and diagonal categories are simple rotations of each other, an ideal observer would perform identically in these conditions (even in the presence of perceptual and criterial noise). Second, Experiment 2 was designed specifically to test this hypothesis. In this study, to increase difficulty, the unidimensional categories were moved closer together. As a consequence, asymptotic accuracy in the unidimensional/feedback conditions of Experiment 2 was about equal to the asymptotic accuracy of the diagonal/feedback conditions of Experiment 1. Despite controlling for difficulty in this manner, only unidimensional category learning was unaffected by type of training (i.e., observational vs. feedback).

Third, note that a critical assumption of the difficulty hypothesis is that rule-based category learning is more susceptible to interference than information-integration category learning. However, this assumption is not always true. For example, Waldron and Ashby (2001) showed that unidimensional category learning is disrupted more by a simultaneous task that activates frontal cortex (a numerical Stroop task) than is learning of information-integration structures. If difficulty is the most important factor, then simultaneously performing a second task should interfere more strongly with learning the more difficult information-integration structures. Since Waldron and Ashby found the opposite pattern of results, it seems likely that factors other than difficulty are at work. As another example, Ashby, Noble, Filoteo, Waldron, and Ell (2001) found that the same group of Parkinson's disease patients were much more impaired at rule-based category learning than at information-integration category learning. If a single system mediates learning in these two types of categorization tasks, and if Parkinson's disease damages this system, then we would expect the more serious deficits to occur in the more difficult information-integration tasks.

## REFERENCES

- AKAIKE, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**, 716-723.
- ALFONSO-REESE, L. A. (1997). *On the dangers of ignoring noise in high-level perception experiments* (Tech. rep. no. 196). Bloomington: Indiana University, Department of Psychology.

- ARBUTHNOTT, G. W., INGHAM, C. A., & WICKENS, J. R. (2000). Dopamine and synaptic plasticity in the neostriatum. *Journal of Anatomy*, **196**, 587-596.
- ASHBY, F. G. (1992a). Multidimensional models of categorization. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 449-483). Hillsdale, NJ: Erlbaum.
- ASHBY, F. G. (1992b). Multivariate probability distributions. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 1-34). Hillsdale, NJ: Erlbaum.
- 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.
- ASHBY, F. G., & ELL, S. W. (2001). The neurobiological basis of category learning. *Trends in Cognitive Sciences*, **5**, 204-210.
- ASHBY, F. G., & ELL, S. W. (2002). Single versus multiple systems of learning and memory. In J. Wixted & H. Pashler (Eds.), *Stevens' handbook of experimental psychology: Vol. 4. Methodology in experimental psychology* (3rd ed., pp. 655-692). New York: Wiley.
- ASHBY, F. G., & GOTT, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **14**, 33-53.
- ASHBY, F. G., ISEN, A. M., & TURKEN, A. U. (1999). A neuropsychological theory of positive affect and its influence on cognition. *Psychological Review*, **106**, 529-550.
- ASHBY, F. G., & LEE, W. W. (1993). Perceptual variability as a fundamental axiom of perceptual science. In S. C. Masin (Ed.), *Foundations of perceptual theory* (pp. 369-399). Amsterdam: Elsevier.
- ASHBY, F. G., & MADDOX, W. T. (1992). Complex decision rules in categorization: Contrasting novice and experienced performance. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 50-71.
- ASHBY, F. G., & MADDOX, W. T. (1993). Relations between prototype, exemplar, and decision bound models of categorization. *Journal of Mathematical Psychology*, **37**, 372-400.
- ASHBY, F. G., NOBLE, S., FILOTEO, J. V., WALDRON, E. M., & ELL, S. W. (2001). *Category learning deficits in Parkinson's disease*. Manuscript submitted for publication.
- ASHBY, F. G., QUELLER, S., & BERRETTY, P. M. (1999). On the dominance of unidimensional rules in unsupervised categorization. *Perception & Psychophysics*, **61**, 1178-1199.
- ASHBY, F. G., & TOWNSEND, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, **93**, 154-179.
- ASHBY, F. G., & WALDRON, E. M. (1999). The nature of implicit categorization. *Psychonomic Bulletin & Review*, **6**, 363-378.
- BRAINARD, D. H. (1997). Psychophysics software for use with MATLAB. *Spatial Vision*, **10**, 433-436.
- CALABRESI, P., PISANI, A., CENTONZE, D., & BERNARDI, G. (1996). Role of Ca<sup>2+</sup> in striatal LTD and LTP. *Seminars in the Neurosciences*, **8**, 321-328.
- EACOTT, M. J., & GAFFAN, D. (1992). Inferotemporal-frontal disconnection: The uncinate fascicle and visual associative learning in monkeys. *European Journal of Neuroscience*, **4**, 1320-1332.
- ELLIOTT, R., & DOLAN, R. J. (1998). Activation of different anterior cingulate foci in association with hypothesis testing and response selection. *NeuroImage*, **8**, 17-29.
- ERICKSON, M. A., & KRUSCHKE, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General*, **127**, 107-140.
- ESTES, W. K. (1976). The cognitive side of probability learning. *Psychological Review*, **83**, 37-64.
- ESTES, W. K. (1994). *Classification and cognition*. Oxford: Oxford University Press.
- FILOTEO, J. V., MADDOX, W. T., & DAVIS, J. D. (2001a). A possible role of the striatum in linear and nonlinear categorization rule learning: Evidence from patients with Huntington's disease. *Behavioral Neuroscience*, **115**, 786-798.
- FILOTEO, J. V., MADDOX, W. T., & DAVIS, J. D. (2001b). Quantitative modeling of category learning in amnesic patients. *Journal of the International Neuropsychological Society*, **7**, 1-19.
- GAFFAN, D., & EACOTT, M. J. (1995). Visual learning for an auditory secondary reinforcer by macaques is intact after uncinate fascicle section: Indirect evidence for the involvement of the corpus striatum. *European Journal of Neuroscience*, **7**, 1866-1871.
- GAFFAN, D., & HARRISON, S. (1987). Amygdectomy and disconnection in visual learning for auditory secondary reinforcement by monkeys. *Journal of Neuroscience*, **7**, 2285-2292.
- GOLDMAN-RAKIC, P. S. (1987). Circuitry of the prefrontal cortex and the regulation of behavior by representational knowledge. In F. Plum & V. Mountcastle (Eds.), *Handbook of physiology* (pp. 373-417). Bethesda, MD: American Physiological Society.
- GOLDMAN-RAKIC, P. S. (1995). Cellular basis of working memory. *Neuron*, **14**, 477-485.
- GREEN, D. M., & SWETS, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- HOLLERMAN, J. R., & SCHULTZ, W. (1997). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature Neuroscience*, **1**, 304-308.
- IZAWA, C. (1967). Function of test trials in paired-associate learning. *Journal of Experimental Psychology*, **75**, 194-209.
- KNOWLTON, B. J., MANGELS, J. A., & SQUIRE, L. R. (1996). A neostriatal habit learning system in humans. *Science*, **273**, 1399-1402.
- MADDOX, W. T., & ASHBY, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception & Psychophysics*, **53**, 49-70.
- MADDOX, W. T., & FILOTEO, J. V. (2001). Striatal contribution to category learning: Quantitative modeling of simple linear and complex nonlinear rule learning in patients with Parkinson's disease. *Journal of the International Neuropsychological Society*, **7**, 710-727.
- MCDONALD, R. J., & WHITE, N. M. (1993). A triple dissociation of memory systems: Hippocampus, amygdala, and dorsal striatum. *Behavioral Neuroscience*, **107**, 3-22.
- MCDONALD, R. J., & WHITE, N. M. (1994). Parallel information processing in the water maze: Evidence for independent memory systems involving dorsal striatum and hippocampus. *Behavioral & Neural Biology*, **61**, 260-270.
- MEDIN, D. L., & SCHAFFER, M. M. (1978). Context theory of classification learning. *Psychological Review*, **85**, 207-238.
- NOSOFSKY, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, **115**, 39-57.
- PACKARD, M. G., HIRSH, R., & WHITE, N. M. (1989). Differential effects of fornix and caudate nucleus lesions on two radial maze tasks: Evidence for multiple memory systems. *Journal of Neuroscience*, **9**, 1465-1472.
- PACKARD, M. G., & MCGAUGH, J. L. (1992). Double dissociation of fornix and caudate nucleus lesions on acquisition of two water maze tasks: Further evidence for multiple memory systems. *Behavioral Neuroscience*, **106**, 439-446.
- POSNER, M. I., & PETERSEN, S. E. (1990). Attention systems in the human brain. *Annual Review of Neuroscience*, **13**, 25-42.
- RAO, S. M., BOBHOLZ, J. A., HAMMEKE, T. A., ROSEN, A. C., WOODLEY, S. J., CUNNINGHAM, J. M., COX, R. W., STEIN, E. A., & BINDER, J. R. (1997). Functional MRI evidence for subcortical participation in conceptual reasoning skills. *NeuroReport*, **8**, 1987-1993.
- REBER, A. S., & MILLWARD, R. B. (1968). Event observation in probability learning. *Journal of Experimental Psychology*, **77**, 317-327.
- ROBINSON, A. L., HEATON, R. K., LEHMAN, R. A. W., & STILSON, D. W. (1980). The utility of the Wisconsin Card Sorting Test in detecting and localizing frontal lobe lesions. *Journal of Consulting & Clinical Psychology*, **48**, 605-614.
- ROLLS, E. T. (1994). Neurophysiology and cognitive functions of the striatum. *Revue Neurologique*, **150**, 648-660.
- SCHULTZ, W. (1992). Activity of dopamine neurons in the behaving primate. *Seminars in Neuroscience*, **4**, 129-138.
- SHAW, M. L. (1982). Attending to multiple sources of information. I: The integration of information in decision making. *Cognitive Psychology*, **14**, 353-409.
- SMITH, E. E., & MEDIN, D. L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.
- SMITH, E. E., PATALANO, A., & JONIDES, J. (1998). Alternative strategies of categorization. *Cognition*, **65**, 167-196.
- TAKANE, Y., & SHIBAYAMA, T. (1992). Structures in stimulus identifica-

- tion data. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 335-362). Hillsdale, NJ: Erlbaum.
- WALDRON, E. M., & ASHBY, F. G. (2001). The effects of concurrent task interference on category learning. *Psychonomic Bulletin & Review*, **8**, 168-176.
- WICKENS, J. (1993). *A theory of the striatum*. New York: Pergamon.
- WICKENS, T. D. (1982). *Models for behavior: Stochastic processes in psychology*. San Francisco: W.H. Freeman.
- WINOCUR, G., & ESKES, G. (1998). Prefrontal cortex and caudate nucleus in conditional associative learning: Dissociated effects of selective brain lesions in rats. *Behavioral Neuroscience*, **112**, 89-101.

#### NOTES

1. Note that in the no response conditions, the term *feedback training* is a bit of a misnomer. Nevertheless, in most applications, providing a category label after the stimulus is presented does provide feedback to the observer. So we prefer to call the two types of training "observational" and "feedback."

2. In the rat, the caudate and putamen merge into a single entity, so for the rat studies, it is more proper to refer to the dorsal striatum.

3. Observers were allowed to generate a categorization response during the 500-msec blank interval. A variant on this condition was also run in which the procedure was as follows: stimulus (500 msec), blank screen (500 msec), category label (500 msec), respond. The results were similar across these two variants.

4. In most of the cases in which a rule-based model provided the best fit to data from a diagonal condition, the conjunction model fit better than the unidimensional model.

5. If we repeat this analysis separately for the two rule types, we find no differences between observational and feedback training on decision strategy for the unidimensional rule on orientation [ $F(1,8) = 0.29, p > .50$ ], but we do find significance for the rule on length [ $F(1,8) = 7.54, p < .05$ ]. In this case, rule-based responding was more common with observational training (60%) than with feedback training (32%).

(Manuscript received March 26, 2001;  
revision accepted for publication March 7, 2002.)