Evidence for a procedural-learning–based system in perceptual category learning

W. TODD MADDOX, COREY J. BOHIL, and A. DAVID ING University of Texas, Austin, Texas

The consistency of the mapping from category to response location was investigated to test the hypothesis that abstract category labels are learned by the hypothesis testing system to solve rule-based tasks, whereas response position is learned by the procedural-learning system to solve information-integration tasks. Accuracy rates were examined to isolate global performance deficits, and model-based analyses were performed to identify the types of response strategies used by observers. A–B training (consistent mapping) led to more accurate responding relative to yes—no training (variable mapping) in the information-integration category learning task. Model-based analyses indicated that the yes—no accuracy decline was due to an increase in the use of rule-based strategies to solve the information-integration task. Yes—no training had no effect on the accuracy of responding or distribution of best-fitting models relative to A–B training in the rule-based category learning tasks. These results both provide support for a multiple-systems approach to category learning in which one system is procedural-learning—based and argue against the validity of single-system approaches.

A growing body of research suggests that the learning of different types of category structures is mediated by different categorization systems (Ashby & Ell, 2001; Erickson & Kruschke, 1998; Maddox, Filoteo, Hejl, & Ing, 2004; Reber & Squire, 1994; Smith, Patalano, & Jonides, 1998). Whereas most multiple-systems theorists agree that one system is explicit and another is implicit, there is disagreement about the nature of the implicit system. Some argue for an exemplar-based system (e.g., Erickson & Kruschke, 1998), some for a perceptual representation system (e.g., Reber, Stark, & Squire, 1998), and others for a procedural-learning—based system (e.g., Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Ashby, Ell, & Waldron, 2003).

One intriguing aspect of the procedural-learning—based system is its association with motor performance (e.g., Hazeltine & Ivry, 2002; Willingham, 1998; Willingham, Nissen, & Bullemer, 1989). Whereas there is no a priori reason to expect that exemplar- or perceptual-representation—based category learning should be closely linked to associated motor responses, categories learned via a procedural-learning—based system should have a close link to the motor response. In support of this claim, Ashby et al. (2003) found an empirical dissociation between the learning of a rule-based category structure and an information-

This research was supported in part by National Institutes of Health Grant R01 MH59196 to WTM. We are indebted to Peter MacNeilage for providing the initial inspiration for this study. We also thank Maribel Chavez, Kelli Hejl, Randy Jupio, Amy Kaderka, Mina Wilcox-Ghanoonparvar, and Jessica Valverde for help with data collection and three anonymous reviewers for helpful comments on an earlier draft. Correspondence concerning this article should be addressed to W. T. Maddox, University of Texas, 1 University Station A8000, Department of Psychology, Austin, TX 78712 (e-mail: maddox@psy.utexas.edu).

integration category structure. Rule-based category learning tasks are those in which the category structures can be learned via some explicit reasoning process that is generally verbalizable. In some applications, 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. In other applications, two dimensions are relevant, and the observer's task is to learn a decision criterion along each dimension and to map the different dimensional values to the relevant categories. *Information-integration* category learning tasks are those in which accuracy is maximized only if information from two or more stimulus dimensions is integrated at some predecisional stage (Ashby & Gott, 1988). In many cases, the optimal rule in informationintegration tasks is difficult or impossible to describe verbally (i.e., is nonverbalizable).

Ashby et al. (2003) postulated that verbalizable, rulebased category learning is mediated by an explicit system, whereas nonverbalizable, information-integration category learning is mediated by an implicit procedurallearning system. Observers were trained on either rulebased or information-integration category structures. There were three conditions: control, hand-switch, and button-switch. In the control condition, the response key assigned to category A was pressed with the left index finger and the response key assigned to category B was pressed with the right index finger during both training and transfer. In the hand-switch condition, the hands were crossed during training so that the response key assigned to category A was pressed with the right index finger and the response key assigned to category B was pressed with the left index finger. During transfer, the hands were uncrossed on the response keys. In the buttonswitch condition, training was identical to that in the control condition, but during transfer the location of the buttons was switched. Ashby et al. (2003) found no effect of hand switching or button switching on rule-based category learning. However, for information-integration category learning button switching led to a large performance deficit, whereas hand switching did not. In agreement with results from the procedural-learning literature, the researchers concluded that response positions are learned by the procedural-learning—based system, whereas abstract category labels are learned by the rule-based system.

To our knowledge Ashby et al. (2003) is the first study to provide direct evidence in support of a procedurallearning-based component of category learning. This article extends their research by examining rule-based and information-integration category learning under two different training procedures. In all conditions, observers learned two categories of Gabor patches (sine-wave gratings) that varied across trials in spatial frequency and orientation. In the A-B training condition, the stimulus was displayed on each trial, along with the query "Is this an A or B?" The observer pressed one key for category A and a separate key for category B. The response was followed by a 500-msec randomly generated Gabor pattern mask, 750 msec of corrective feedback ("Correct" or "Error"), a 500-msec intertrial interval, and the next trial. In the *ves-no* training condition, the stimulus was displayed on each trial, along with one of the following two queries (selected randomly, but with equal probability), "Is this an A?" or "Is this a B?" The observer pressed one key to respond "no" and a separate key to respond "yes." The response was followed by a 500 msec randomly generated Gabor pattern mask, 750 msec of corrective feedback ("Correct" or "Error"), a 500-msec intertrial interval, and the next trial. Notice that in the A-B condition, each category label is associated with a unique response location (consistent category to response location mapping), whereas in the yes-no condition, each category label is not associated with a unique response location (variable category to response location mapping). If rule-based categories are solved via an explicit reasoning process that learns abstract category labels, then performance should not differ across A-B and yes—no training procedures. If information-integration categories are solved via a procedural-learning-based process that learns response positions, then learning should be poor in the yes-no conditions relative to the A-B conditions.

The Present Experiment

One information-integration category structure and two rule-based structures were used in the experiment and are described in Figure 1. The distribution parameters are outlined in Table 1. Each symbol in Figure 1 denotes the spatial frequency and orientation of a single Gabor patch. Also shown in Figure 1 are the decision bounds that maximize categorization accuracy. In the unidimensional rule-based task (Figure 1A), the optimal

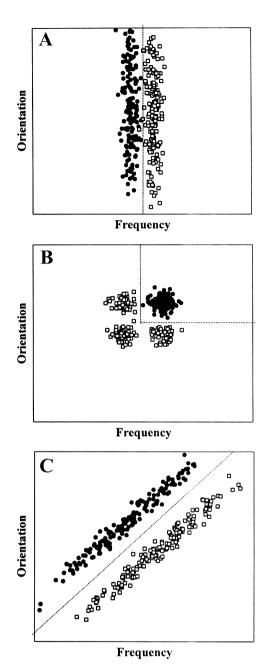


Figure 1. (A) Unidimensional, (B) conjunctive, and (C) information-integration category structures from the experiment. Each circle denotes the spatial frequency and orientation of a Gabor pattern from category A. Each square denotes the spatial frequency and orientation of a Gabor pattern from category B. The broken line(s) in each panel denote(s) the location of the optimal decision bound.

bound requires observers to attend to spatial frequency and ignore orientation, a process that corresponds to the rule "Respond A if the bars are thick and B if they are thin." In the conjunctive rule-based task (Figure 1B), the optimal bound requires observers to attend to spatial frequency *and* orientation, a rule that could be stated thus:

Table 1
Distribution Parameters

Condition	Category	μ_{f}	$\mu_{\rm o}$	$\sigma_{ m f}^2$	$\sigma_{\rm o}^2$	cov _{f,o}
Unidimensional	A	285	125	75	9,000	0
	В	315	125	75	9,000	0
Conjunctive	Α	317	152	75	75	0
•	B_1	283	98	75	75	0
	\mathbf{B}_{2}	283	152	75	75	0
	\mathbf{B}_{3}^{-}	317	98	75	75	0
Information-integration	A	272	153	4,538	4,538	4,463
	В	327	97	4,538	4,538	4,463

Note—Each random sample from the relevant bivariate normal distribution was converted to a stimulus by deriving the frequency, $f=.25+(x_1/50)$, and orientation, $o=x_2(\pi/500)$. For example, the category A mean for the unidimensional category structure was converted to a Gabor pattern with frequency f=.25+(285/50)=5.95 cycles/degree and orientation $o=125(\pi/500)=0.785$ radians counterclockwise from horizontal. The scaling factors were chosen in an attempt to equate the salience of frequency and orientation.

"Respond A if the bars are thin *and* if the orientation is steep; otherwise respond B." In the information-integration task (Figure 1C), both dimensions are relevant, but in contrast to the conjunctive task, dimensional integration is predecisional; thus, there is no simple verbal description of the optimal decision bound.

It is important to equate rule-based and informationintegration category learning in the A–B task so that task difficulty differences cannot be used to explain any differential effects of yes-no training on category learning. To achieve this goal, we conducted a number of pilot studies using A-B training in which we adjusted category discriminability across the three tasks to ensure equivalent category learning. Notice that category discriminability is lower in the unidimensional rule-based task than in the conjunctive rule-based or information-integration tasks. In fact, there is a small amount of category overlap in the unidimensional task. Although this overlap was necessary in order to equate performance with the other two conditions, it does introduce a potential problem—namely, that performance in the unidimensional task might be affected by perceptual discriminability differences across categories. Fortunately, including the conjunctive rulebased task alleviates this problem, since category discriminability is high in this condition. One additional advantage of including the conjunctive rule-based task is that it requires that both dimensions be processed, similar to the information-integration task.

METHOD

Observers and Design

One hundred thirty-nine observers were solicited from the University of Texas community and received course credit for participation. Twenty-seven, 19, and 19 observers received yes—no training in the information-integration (II), unidimensional (UD), and conjunctive (CJ) conditions, respectively. Thirty, 25, and 19 observers received A—B training in the II, UD, and CJ conditions, respectively. 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 randomization technique introduced by Ashby and Gott (1988). Forty category A and 40 category B stimuli from the II categories were generated by sampling randomly from two bivariate normal distributions. The same procedure was used to generate the UD stimuli. Each set of 80 stimuli was displayed in a random order in each of four blocks of trials. For the CJ categories, 39 category A and 39 category B stimuli (13 from each of the three category B distributions: $B_1,\,B_2,\,$ and $B_3;\,$ see Table 1) were generated by random sampling. Each set of 78 stimuli was displayed in a random order in each of four blocks of trials. Each Gabor patch was generated and displayed on a 21-in. monitor with $1,360\times1,024$ resolution.

Procedure

Each observer was tested individually in a dimly lit room and informed that there were two categories of equal probability. They were informed that nearly perfect performance was possible and told to emphasize accuracy without worrying about speed of responding.

RESULTS AND THEORETICAL ANALYSES

Analyses were performed separately on each block of data. First we will present the accuracy rates using an analysis of variance (ANOVA); then we will introduce the model-based analyses.

ANOVA Results

The averaged accuracy rates are displayed in Figure 2A, and the accuracy rates for the final block only are in Figure 2B. To verify that performance under A–B training procedures was equivalent across the three category structures, we conducted a 3 (category structure: II vs. UD vs. CJ) \times 4 (block) ANOVA. The category structure effect was nonsignificant (F < 1), suggesting equivalent performance across all three category structures. The block effect was significant [F(3,213) = 40.98, p < .001], suggesting learning with experience. The interaction was significant [F(6,213) = 2.32, p < .05] but was due only to a performance difference during the first block of trials.

To determine how yes—no training affected learning across the three category structures, we conducted another 3 (category structure: II vs. UD vs. CJ) \times 4 (block) ANOVA. For this training condition, the category structure effect was significant [F(2,62) = 4.74, p < .05]. Post hoc analyses suggested that performance in the information-integration task was worse than performance in the unidimensional and conjunctive rule-based tasks and also that performance in the latter two conditions was equivalent. The block effect was again significant [F(3,186) = 22.65, p < .001], suggesting learning with experience, and the interaction was nonsignificant [F(6,186) = 1.56, p > .05].

As outlined earlier, we predicted a priori that relative to A–B training, yes–no training should adversely affect information-integration category learning but should not affect rule-based category learning—that is, either uni-

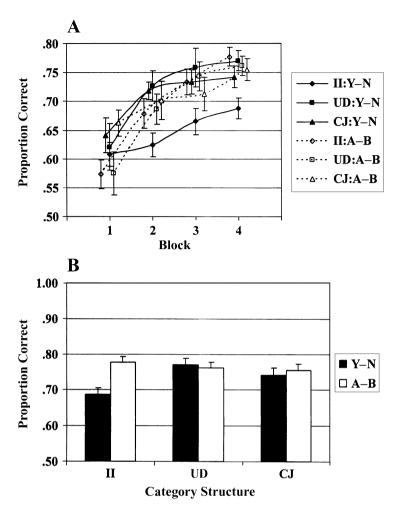


Figure 2. (A) Proportion correct for the information-integration (II) category structure and the unidimensional (UD) and conjunctive (CJ) rule-based category structures under the A–B (broken line) and yes—no (Y–N; solid line) training procedures by block (standard error bars included). Note that the horizontal locations of the six symbols are "jittered" to facilitate a comparison of the performance differences across conditions. (B) Proportion correct during the final block of trials.

dimensional or conjunctive learning. We tested each of these predictions directly by conducting a 2 (training procedure: yes-no vs. A-B) × 4 (block) ANOVA separately on the data from each of the three category structures. For the information-integration category structure, the block effect was significant [F(3,165) = 33.13, p <.001] and the training effect was marginally significant [F(1,55) = 3.38, p = .07]. These effects were qualified by a significant interaction [F(3,165) = 6.31, p < .001]. To determine the locus of this effect, we examined performance separately for each block of trials. Yes-no and A-B performance was equivalent in Blocks 1 and 2 (p >.05), but by Blocks 3 and 4, yes-no performance was significantly worse than A–B performance. For the unidimensional category structure, the block effect was significant [F(3,126) = 21.20, p < .001], but the training effect and interaction were both nonsignificant (Fs < 1).

A similar pattern held for the conjunctive category structure, with a significant block effect [F(3,108) = 14.64, p < .001] but a nonsignificant training effect and interaction (Fs < 1).

Two important conclusions can be drawn from these data. First, for the unidimensional and conjunctive rule-based category structures in all learning blocks, there was no performance difference across yes—no and A—B training procedures. Second, a performance advantage for the A—B relative to the yes—no training procedure emerged over blocks for the information-integration category structures. These results support the claim that information-integration category learning is mediated by a procedural-learning—based system that associates category labels with response locations, whereas rule-based category learning is mediated by an explicit hypothesis testing system in which abstract category labels are learned.

Modeling Results

The accuracy-based analyses provide information regarding overall performance, but they tell us little about the types of strategies that observers might use to solve these tasks. An understanding of strategy use and how these strategies might be affected by the different training procedures is of critical importance to a complete understanding of category learning. One intriguing hypothesis suggested by the procedural-learning literature, and particularly by Ashby et al. (2003), is that observers in the information-integration condition will be forced to resort to rule-based categorization strategies under the yes—no training procedure, because the procedural-learning system cannot associate each categorization response with a separate location. To gain insight into these processes, a number of different decision-bound models were fit to the final block of trials from each observer. 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.

Two different types of decision-bound models were fit to each observer's responses (Ashby, Maddox, & Bohil, 2002): One type was compatible with the assumption that observers used an explicit rule-based strategy, and one type assumed an information-integration strategy. Page constraints preclude a detailed overview of the models. The details can be found in numerous articles (e.g., Ashby, 1992a; Maddox, Filoteo, et al., 2004). The following models were fit to each observer's responses (Ashby, 1992a).

Rule-based models. Two models were compatible with the assumption that observers used an explicit rule-based strategy.

The *unidimensional model* assumes that observers set a criterion on a single perceptual dimension and then make an explicit decision about the level of the stimulus on that dimension (Ashby & Gott, 1988). For example, in a unidimensional category structure, observers might use the rule "Respond A if the spatial frequency is low and B if it is high." The unidimensional model has two free parameters: a decision criterion on the relevant perceptual dimension and the variance of internal (perceptual and criterial) noise (i.e., σ^2).

The *conjunction model* assumes that observers use a conjunction rule in which they make separate decisions about the levels on the two dimensions and then select a response based on the outcome of these two decisions. Two conjunction rules were examined for the information-integration task (1 and 2), and a third for the conjunctive task (3):

- (1) Respond A if spatial frequency is low and orientation is large, otherwise respond B.
- (2) Respond B if spatial frequency is high and orientation is small, otherwise respond A.

(3) Respond A if spatial frequency is high and orientation is large, otherwise respond B.

These strategies are rule-based because they are easy to describe verbally and do not require perceptual integration of spatial frequency and orientation. Conjunction models have three parameters (a criterion on each dimension and σ^2).

Information-integration models. The *general linear* classifier (GLC) assumes that the decision bound between each pair of categories is linear; this model was applied to the unidimensional and information-integration category learning data. The GLC has three parameters (slope and intercept of the linear bound and σ^2). The *striatal* pattern classifier (SPC; Ashby & Waldron, 1999) assumes that the observer constructs four decision bounds to separate the A and B categories in the conjunctive rule-based condition. (The SPC contains the conjunction model as a special case.) In fitting this model, we have assumed that there are four "units" in the frequencyorientation space. On each trial, the observer determines which unit is closest to the perceptual effect and gives the associated response. When fitting the SPC to the conjunctive rule-based data, we assumed that one unit was associated with category A and three units with category B (since category B consisted of three separate clusters of stimuli), which yielded four linear decision bounds. Because the location of one of the units can be fixed and a uniform expansion or contraction of the space will not affect the location of the resulting (minimum distance) decision bounds, in this case the SPC contains six free parameters (i.e., five that determine the location of the units, and thus the decision bounds, and one noise variance σ^2).

Model fits. Each of these models was fit separately to the data from the final block of trials for every observer. The model parameters were estimated using maximum likelihood (Ashby, 1992b), 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). The AIC statistic penalizes a model for extra free parameters in such a way that the smaller the AIC, the closer a model is to the "true model," regardless of the number of free parameters. Thus, the best model is the one with the smallest AIC value.

For each observer's final block of trials, we determined which model type (i.e., rule-based or information-integration) provided the best account of the data. The proportion of observers for which a rule-based model provided the best account of the data by training procedure (A–B and yes–no) and category structure (UD, CJ, and II) is displayed in Figure 3.

Two findings are of interest. First, the yes—no training procedure led to a large increase in the use of rule-based strategies to solve the information-integration category

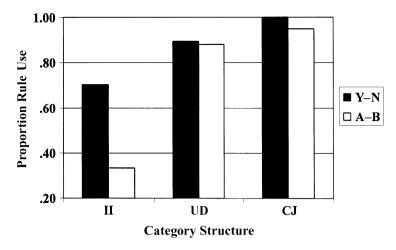


Figure 3. Proportion of participants' final block data that were best fit by a rule-based model for the information-integration (II), unidimensional (UD), and conjunctive (CJ) category structures under the A–B and yes–no training procedures.

learning task. Specifically, during the final block of trials, 70% of the observers' data was best fit by a rulebased model in the yes-no training condition (30% using a unidimensional rule and 40% using a conjunction rule), whereas only 33% was best fit by a rule-based model in the A–B training condition (10% using a unidimensional rule and 23% using a conjunction rule). Second, the training procedure manipulation did not affect the proportion of observers whose data was best fit by a rule-based model in the unidimensional or conjunctive category learning tasks. For both training procedures, 90%–100% of the observers' data was best fit by a rule-based model in the rule-based category learning tasks. Across all conditions, the best-fitting model provided a good account of the data, accounting for 75%-85% of the responses in the data on average.

DISCUSSION

This article reports the results from an experiment that examined the differential effects of A–B versus yes–no training procedures on the learning of unidimensional rule-based, conjunctive rule-based, and information-integration category structures. If an explicit hypothesis testing system that learns abstract labels mediates rule-based learning, there should be no performance difference across training procedures in the unidimensional and conjunctive rule-based tasks. If a procedural-learning system that associates category labels with response positions mediates information-integration learning, there should be a performance decrement for yes–no training relative to A–B training.

For the unidimensional and conjunctive rule-based category structures, the yes-no training procedure had no effect on the accuracy of rule-based category learning or on the type of decision strategy used by observers rela-

tive to the A–B training procedure.² For the information-integration category structure, the yes–no training procedure led to both a significant decrease in the accuracy of responding relative to the A–B training procedure and an increase in the use of rule-based response strategies. This dissociation between rule-based and information-integration category learning under the yes–no training procedure held even when A–B training performance was equated across category structures. This result provides evidence in support of a procedural-learning–based category learning system.

Single Versus Multiple Systems

These findings add to the body of literature in support of multiple memory and category learning systems (e.g., Ashby & Ell, 2001; Erickson & Kruschke, 1998; Maddox & Filoteo, 2001; Pickering, 1997) and suggest that rule-based and information-integration category learning are mediated by separate systems. The findings are also consistent with the hypothesis that learning in rulebased tasks is dominated by an explicit system that uses a conscious process of hypothesis generation and testing, and learns abstract category labels. This system should not require that each category be associated with a unique response location, and thus it should be (and was) unaffected by the training procedure manipulation. This hypothesis also suggests that other aspects of the training, such as the timing and placement of the feedback, should have no effect. Ashby et al., (2002) compared an observational learning procedure in which the category label was presented prior to the stimulus with a feedback learning procedure in which the category label was presented after the stimulus, and they found no effect of this manipulation on rule-based category learning. Maddox, Ashby, and Bohil (2003) compared an immediate-feedback condition with a delayed-feedback

condition in which the feedback followed the response by 5 sec and likewise found no effect on rule-based category learning.

In contrast, it has been hypothesized that learning in information-integration tasks is dominated by an implicit procedural-learning-based system that is mediated largely within the tail of the caudate nucleus (Ashby et al., 1998; Ashby & Ell, 2001). It has been proposed that a dopamine-mediated reward signal is critical for learning in this system. 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; Wickens, 1993). Because of its link to procedural learning, this system should prefer situations in which each category is associated with a unique response location, and thus it should be (and was) adversely affected by the yes-no training procedure. In addition, because of this system's reliance on dopamine-mediated reward, the timing and placement of the feedback should be critical. As predicted, informationintegration category learning was worse for observational relative to feedback learning (Ashby et al., 2002) and for delayed relative to immediate feedback (Maddox et al., 2003). In addition, across both of these studies the performance decrement observed in informationintegration category learning was associated with an increase in the use of rule-based strategies to solve the information-integration problem.

Despite the strong evidence in support of multiple systems, a single-system theorist might argue against such claims by suggesting that information-integration categories are inherently more difficult than rule-based categories and that any manipulation that affects learning will disrupt the more difficult (information-integration) task. There are a number of problems with this reasoning. First, there is no a priori reason why the informationintegration task should be more difficult. In fact, the unidimensional and conjunctive rule-based tasks are of lower category discriminability, and thus by all relevant objective measures they are more difficult. In addition, pigeons have no more difficulty learning information-integration categories than rule-based categories (Herbranson, Fremouw, & Shimp, 1999). A second problem with the difficulty argument is that it predicts that any experimental manipulation disrupting category learning should have greater effects with information-integration than with rule-based tasks. This prediction was disconfirmed in a recent study by Waldron and Ashby (2001; see also Maddox, Ashby, Ing, & Pickering, 2004), who showed that rule-based category learning was disrupted more by a simultaneous task that requires working memory and executive attention (a numerical Stroop task) than was information-integration category learning.

Another argument that one might make against our analysis is that a yes—no task places greater demands on working memory than does an A—B task, and that this additional working memory load adversely affects information-integration, but not rule-based, category

learning. Taken at face value, this is a multiple-systems proposal in which information-integration category learning is mediated by working memory processes, whereas rule-based learning is not. This approach faces a number of problems as well. First, as reviewed above, Waldron and Ashby (2001; see also Maddox, Ashby, et al., 2004) showed that rule-based category learning was disrupted more by a simultaneous task that requires working memory than was information-integration category learning. This pattern is predicted from our dual-systems model, because we postulate that an explicit hypothesis testing system (which relies on working memory) mediates rulebased but not information-integration category learning. Second, both the button-switch and hand-switch conditions of Ashby et al. (2003) required a change in the participant's response characteristics during transfer that likely placed greater demands on working memory. According to the working memory hypothesis, an informationintegration performance decrement should have been observed in both conditions. Instead, performance in the hand-switch condition was unaffected—only performance in the button-switch condition was adversely affected. Taken together, these findings argue against a working memory explanation of the results.

Even so, the yes—no procedure likely is more difficult to process and may require more working memory in order to interpret the feedback, since the feedback is present only after the stimulus display has been removed. As a direct test of this hypothesis, we conducted a second experiment using the Figure 1A (rule-based) and 1C (informationintegration) category structures. Observers were trained using either a yes-no or an A-B procedure.³ The trial procedure was identical to that outlined in the Method section, with one critical exception. Instead of presenting only the word "Correct" or "Error" during the feedback phase of each trial, we also presented the stimulus, the query ("Is this an A?" "Is this a B?" or "Is this an A or B?") and the observer's response ("yes," "no," "A," or "B"). In this way, we reduced the memory load difference across yes-no and A-B procedures required to process the feedback. During the final block of trials, accuracy was .74, .77, .73, and .66 for the rule-based with A-B, rulebased with ves-no, information-integration with A-B, and information-integration with yes-no procedures, respectively. Performance did not differ significantly across the two A–B training conditions (p > .05), nor did performance differ across the A–B and yes–no training procedures for the rule-based category structures (p > .05). However, performance was significantly worse with yes-no training relative to A-B training for the information-integration condition (p < .05), in line with the results we outlined above. These data argue strongly against the hypothesis that an increased working memory load for the yes-no procedure can account for our original results.

It should be stressed that the dissociation reported in this article, between A–B and yes–no training and the type of category structure, was predicted a priori from the dual-systems model, and further that these predictions were parameter free. Once the difficulty hypothesis has been rejected, it is hard to see how a single-system model could predict our results in a similar a priori fashion. Of course, it is possible that a single-system model could account for our results via post hoc manipulation of its parameters, but post hoc accounts are less parsimo nious than parameter-free a priori predictions, regardless of the number of systems assumed by the two models.

REFERENCES

- AKAIKE, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**, 716-723.
- ASHBY, F. G. (1992a). Multidimensional models of categorization. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition*. 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 neurobiology of human category learning. *Trends in Cognitive Sciences*, **5**, 204-210.
- ASHBY, F. G., ELL, S. W., & WALDRON, E. M. (2003). Procedural learning in perceptual categorization. *Memory & Cognition*, 31, 1114-1125.
- 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., MADDOX, W. T., & BOHIL, C. J. (2002). Observational versus feedback training in rule-based and information-integration category learning. *Memory & Cognition*, **30**, 666-677.
- ASHBY, F. G., & WALDRON, E. M. (1999). On the nature of implicit categorization. *Psychonomic Bulletin & Review*, **6**, 363-378.
- ERICKSON, M. A., & KRUSCHKE, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General*, 127, 107-140.
- HAZELTINE, E., & IVRY, R. B. (2002). Motor skill. In V. S. Ramachandran (Ed.), *Encyclopedia of the brain*. San Diego: Academic Press.
- HERBRANSON, W. T., FREMOUW, T., & SHIMP, C. P. (1999). The randomization procedure in the study of categorization of multidimensional stimuli by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, 25, 113-134.
- MADDOX, W. T., ASHBY, F. G., & BOHIL, C. J. (2003). Delayed feedback effects on rule-based and information-integration category learning. <u>Journal of Experimental Psychology: Learning, Memory, & Cognition</u>, 29, 650-662.
- MADDOX, W. T., ASHBY, F. G., ING, A. D., & PICKERING, A. D. (2004).
 Disrupting feedback processing interferes with rule-based but not information-integration category learning. <u>Memory & Cognition</u>, 32, 582-591.
- MADDOX, W. T., & FILOTEO, J. V. (2001). Striatal contributions to cat-

- egory learning: Quantitative modeling of simple linear and complex nonlinear rule learning in patients with Parkinson's disease. <u>Journal</u> of the International Neuropsychological Society, 7, 710-727.
- MADDOX, W. T., FILOTEO, J. V., HEJL, K. D., & ING, A. D. (2004). Category number impacts rule-based but not information-integration category learning: Further evidence for dissociable category learning systems. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 30, 227-235.
- PICKERING, A. D. (1997). New approaches to study of amnesic patients: What can a neurofunctional philosophy and neural network methods offer? *Memory*, 5, 255-300.
- REBER, P. J., & SQUIRE, L. R. (1994). Parallel brain systems for learning with and without awareness. *Learning & Memory*, 1, 217-229.
- REBER, P. J., STARK, C. E. L., & SQUIRE, L. R. (1998). Cortical areas supporting category learning identified using functional MRI. Proceedings of the National Academy of Sciences, 95, 747-750.
- SCHULTZ, W. (1992). Activity of dopamine neurons in the behaving primate. Seminars in the Neurosciences, 4, 129-138.
- SMITH, E. E., PATALANO, A. L., & JONIDES, J. (1998). Alternative strategies of categorization. *Cognition*, **65**, 167-196.
- Waldron, E. M., & Ashby, F. G. (2001). The effects of concurrent task interference on category learning: Evidence for multiple category learning systems. *Psychonomic Bulletin & Review*, **8**, 168-176.
- WICKENS, J. (1993). A theory of the striatum. New York: Pergamon.WILLINGHAM, D. B. (1998). A neuropsychological theory of motor skill learning. Psychological Review, 105, 558-584.
- WILLINGHAM, D. B., NISSEN, M. J., & BULLEMER, P. (1989). On the development of procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 1047-1060.

NOTES

- 1. 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 only then are the outcomes of these decisions combined (integration is postdecisional, not predecisional).
- 2 It is worth mentioning that yes—no and A—B training were compared in three additional unidimensional category structures and one additional conjunctive category structure. In each of these four cases, the categories were more discriminable than in the study outlined here, and thus A—B performance was better than in the present study. In every case, A—B category learning was statistically equivalent to yes—no category learning, and the proportion of rule use was both high and unaffected by the training procedure manipulation.
- 3. Twenty-five, 14, 20, and 17 observers participated in the rule-based with A–B, rule-based with yes–no, information-integration with A–B, and information-integration with yes–no procedures, respectively.

(Manuscript received January 16, 2003; revision accepted for publication November 10, 2003.)