Journal of Experimental Psychology:
Learning, Memory, and Cognition
2005, Vol. 31, No. 4, 654–669

0278-7393/05/$12.00 DOI: 10.1037/0278-7393.31.4.654

Discontinuous Categories Affect Information-Integration but Not Rule-Based Category Learning

W. Todd Maddox
University of Texas at Austin

J. Vincent Filoteo
Veterans Affairs San Diego Healthcare System and University of California, San Diego

J. Scott Lauritzen
University of Texas at Austin

Emily Connally
Harvard University

Kelli D. Hejl
University of Texas at Austin

Three experiments were conducted that provide a direct examination of within-category discontinuity manipulations on the implicit, procedural-based learning and the explicit, hypothesis-testing systems proposed in F. G. Ashby, L. A. Alfonso-Reese, A. U. Turken, and E. M. Waldron's (1998) competition between verbal and implicit systems model. Discontinuous categories adversely affected information-integration but not rule-based category learning. Increasing the magnitude of the discontinuity did not lead to a significant decline in performance. The distance to the bound provides a reasonable description of the generalization profile associated with the hypothesis-testing system, whereas the distance to the bound plus the distance to the trained response region provides a reasonable description of the generalization profile associated with the procedural-based learning system. These results suggest that within-category discontinuity differentially impacts information-integration but not rule-based category learning and provides information regarding the detailed processing characteristics of each category learning system.

Keywords: category learning, striatum, procedural memory, working memory, multiple systems

Categorization is an important skill that is critical to the survival of all organisms. For example, knowledge of category membership provides information about how an object should be used or manipulated or whether it should be approached or avoided. As such, categorization can be used to help guide many behaviors under a variety of circumstances. Category learning involves laying down a memory trace that improves the efficiency (i.e., accuracy and speed) of responding. It is now widely accepted that mammals have multiple memory systems (Poldrack & Packard, 2003; Schacter, 1987; Squire, 1992), and thus it is reasonable to postulate that multiple category learning systems might also exist. The convergence of evidence in support of multiple category learning systems is growing and comes from a wide range of research areas including animal learning (McDonald & White, 1993, 1994; Packard & McGaugh, 1992), neuropsychology (Filoteo, Maddox, & Davis, 2001a, 2001b; Maddox & Filoteo, 2001, in press; Myers et al., 2003), functional neuroimaging (Poldrack, Prabhakaran, Seger, & Gabrieli, 1999; Reber, Stark, & Squire, 1998; E. E. Smith, Patalano, & Jonides, 1998), and cognitive psychology (for reviews, see Keri, 2003, and Maddox & Ashby, 2004)[1]. Most multiple-systems theorists argue for at least one explicit system that is tied to conscious awareness and for at least one implicit system that does not have full access to conscious awareness. One of the most successful multiple systems models of category learning, and the only one that specifies the underlying neurobiology, is the competition between verbal and implicit systems model (COVIS; Ashby, Alfonso-Reese, Turken, & Waldron,

W. Todd Maddox, J. Scott Lauritzen, and Kelli D. Hejl, Department of Psychology, University of Texas at Austin; J. Vincent Filoteo, Veterans Affairs San Diego Healthcare System, and Department of Psychiatry, University of California, San Diego; Emily Connally, Department of Psychology, Harvard University.

This research was supported in part by National Institute of Health Grant R01 MH59196 to W. Todd Maddox, National Institute of Health Grant R01 NS41372 to J. Vincent Filoteo, and a McDonnell-Pew Consortium grant. We thank Les Cohen, Brad Love, Matt Jones, Art Markman, and Brian Stankiewicz for useful discussions that influenced this work and Derek Koehler for helpful comments on a draft of this article.

Correspondence concerning this article should be addressed to W. Todd Maddox, 1 University Station A8000, University of Texas at Austin, TX 78712. E-mail: maddox@psy.utexas.edu

[1] In some ways the multiple-systems approach grew out of the multiple-process approach to category learning that suggests that observers have available different processing modes that can be used during category learning (Allen & Brooks, 1991; Erickson & Kruschke, 1998; Kemler-Nelson, 1984; Nosofsky, Palmeri, & McKinley, 1994; Pickering, 1997; Reber & Squire, 1994; Rehegr & Brooks, 1993; J. D. Smith & Shapiro, 1989).

1998; Ashby & Waldron, 1999). COVIS postulates two systems that compete throughout learning—an explicit, hypothesis-testing system that uses logical reasoning and depends on working memory and executive attention, and a procedural-based learning system that relies more on incremental and feedback-learning processes.

In COVIS, the explicit, hypothesis-testing system is assumed to dominate the learning of rule-based (RB) tasks, whereas the implicit, procedural-based learning system dominates the learning of information-integration (II) tasks. *RB 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 et al., 1998). For example, if the stimulus is a line that varies in length and orientation, the observer might set a criterion along line length to determine whether the line is short or long and a criterion along orientation to determine if the angle is shallow or steep. The decision along each dimension might then be integrated to determine category membership (e.g., short, shallow angle lines are assigned to Category A; all others are assigned to Category B). This integration is postdecisional because a decision is first made about the value along each dimension, and that information is explicitly integrated to generate a response. *II category learning tasks,* on the other hand, are those in which accuracy is maximized only if information from two or more stimulus components is integrated at some predecisional stage that occurs outside of conscious awareness, such as when observers adopt a weighted linear combination of the dimensional values (Ashby & Gott, 1988). In many cases, the optimal rule in II tasks is difficult or impossible to describe verbally (Ashby et al., 1998). Postdecisional integration rules, like the one described above for RB categories, can be applied to II conditions, but they generally lead to suboptimal performance levels.

COVIS assumes that learning in RB tasks is dominated by an explicit system that uses working memory[2] and executive attention and is mediated primarily by the anterior cingulate, the prefrontal cortex, and the head of the caudate nucleus. This system learns through a conscious process of hypothesis generation and testing. Thus, this system should be flexible with respect to the placement and timing of the feedback and the response requirements but should be affected by manipulations that increase the working memory load or executive attention demands of the task. In contrast, learning in II tasks is assumed to be dominated by an implicit procedural-based learning system that is mediated largely within the tail of the caudate nucleus (Ashby et al., 1998; Ashby & Ell, 2001; Willingham, 1998). 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). A procedural-based 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.[3] As a result, learning in this system should depend heavily on the placement and timing of the feedback and, because of its procedural nature, should be affected by the response requirements. On the other hand, under acceptable feedback and response requirement conditions, II learning should be essentially automatic and should not be affected by working memory load or attentional demands.

Each of these predictions has been tested empirically (see Maddox & Ashby, 2004 for a detailed review). To examine the effects of the placement and timing of the feedback, researchers compared RB and II category learning across an observational training condition (in which observers were informed before stimulus presentation of what category the ensuing stimulus was from) and a traditional feedback training condition (in which the category label followed the response; Ashby, Maddox, & Bohil, 2002) and across an immediate feedback condition (in which corrective feedback was provided immediately following the response) and a delayed feedback condition (in which corrective feedback was delayed by 2.5, 5.0 or 10.0 s following the response; Maddox, Ashby, & Bohil, 2003; see also Maddox & Ing, 2005). In line with predictions based on COVIS, observational training and delayed feedback negatively impacted II category learning but had little effect on RB category learning.

To examine the effects of the response requirements, Ashby, Ell, and Waldron (2003; see also Maddox, Bohil, & Ing, 2004) incorporated a procedure originally developed by Willingham, Wells, Farrell, and Stemwedel (2000) into a category-learning task to study serial reaction time. Willingham et al. (2000) showed that changing the location of the response keys interferes with SRT learning, even when the sequence of stimulus positions is unchanged. In addition, they showed that SRT learning is unaffected when the sequence of finger movements is changed as long as the location of the response keys remains fixed. If the implicit system in COVIS is procedural-based learning, then it should be the case that changing the location of the response keys would adversely affect learning in this system, and thus II category learning, whereas changing the finger press associated with each category response should not. Ashby et al. (2003) tested and found support for this prediction.

To examine the effect of working memory load and executive attention, researchers examined RB and II learning when the participant was asked to perform a second working memory demanding task. Waldron and Ashby (2001) showed that RB category learning was disrupted more than II category learning by the simultaneous performance of a numerical Stroop task, and Maddox, Ashby, Ing, and Pickering (2004) showed that RB category learning was disrupted by a sequential memory scanning task, whereas II category learning was not.

Taken together, these studies provide support for the existence of separate hypothesis-testing and procedural-based learning systems of category learning. Even so, they tell us little about the detailed properties of these systems. The overriding aim of the present research is to begin this more detailed examination. The

[2] Many argue for the existence of an implicit form of working memory that may not be available to conscious awareness. However, when we use the term working memory, we refer to a conscious, verbalizable process.

[3] Crick and Koch (1990, 1995, 1998) offered a cognitive neuroscience theory of consciousness that states that one can have conscious awareness only of activity in brain areas that project directly to the prefrontal cortex. The caudate nucleus does not project to the prefrontal cortex (it first projects through the thalamus), so the Crick–Koch hypothesis predicts that we are not aware of activity within the caudate nucleus.

primary focus of this research is on the procedural-based learning system, although we examined RB category learning in Experiment 1. Our approach was to examine II category learning across a series of category structures that differ systematically in ways that should have affected procedural-based learning. These more direct manipulations will allow us to rigorously test predictions regarding the nature of learning within each system.

The three experiments presented in this report provide a first step toward the more ambitious goal of understanding the qualitative properties of each system through systematic manipulation of the category structures. All three studies used the same stimuli, consisting of a line that varies in length and orientation and the randomization technique developed by Ashby and Gott (1988). In the randomization technique, each category is represented by a single bivariate normal distribution or a collection of bivariate normal distributions (bivariate because the stimuli are two-dimensional). The optimal decision bound can also be defined and constructed in such a way that either a RB or II strategy is optimal (i.e., maximizes accuracy). Each category exemplar is generated by taking a random sample from the relevant bivariate normal distribution and constructing a stimulus with the associated length and orientation value. Because of the two-dimensional nature of the stimuli, each stimulus can be represented by a point in a two-dimensional length–orientation space. The collection of stimuli used in a particular experimental condition can be displayed in a two-dimensional scatter plot. The scatter plots and optimal decision bounds used in Experiments 1–3 are shown later in this article.

Experiment 1 examined the effects of within-category discontinuity on RB and II category learning using structurally equivalent categories. Within-category discontinuity was manipulated by varying the location of some stimulus clusters relative to others within the same category. We predicted that within-category discontinuity would differentially impact II and RB category learning. Where there are discontinuous clusters of stimuli belonging to the same category, observers are required to learn to associate clusters of stimuli that are very different in terms of their perceptual appearance to the same category. Because procedural-based learning systems appear to be highly dependent on stimulus similarity (e.g., Cohen, Poldrack, & Eichenbaum, 1997), and II category learning is thought to be mediated by a procedural-based learning system, we predicted that II category learning would be adversely affected by discontinuous clusters. Such a prediction was also based on other lines of research. For example, previous studies have shown that perceptual learning (a possible form of procedural learning) appears to be highly stimulus specific (e.g., Schoups, Vogels, & Orban, 1995; Shiu & Pashler, 1992; Vogels & Orban, 1985), suggesting that the acquisition of certain experience-dependent perceptual effects do not readily generalize to changes in even basic perceptual features such as orientation. Thus, II category learning should be adversely affected by the need for observers to learn to categorize clusters of different looking stimuli into the same category. In contrast, we predicted that the impact of within-category discontinuity would have minimal effects on RB learning given that it is a decision rule that is learned by the hypothesis-testing system. This decision rule is not derived from the specific values of the stimulus attributes as is the case in II category learning, but rather the rule that is learned is more abstract and therefore should be impacted to a lesser degree by discontinuous category clusters. Thus, as long as the discontinuous clusters all follow the same rule, then RB category learning should be unaffected. To anticipate, we confirmed the prediction. Experiments 2 and 3 were designed to examine the procedural-based learning system in more detail. Experiment 2 reexamined the effect of within-category discontinuity when the amount of stimulus space trained was controlled. Experiment 3 examined three levels of within-category discontinuity to determine whether increasing within-category discontinuity led to a consistent decrease in performance or whether any amount of within-category discontinuity led to a performance decrement.

Experiment 1

In Experiment 1, we constructed II and RB categories for which either a single cluster of stimuli was associated with each category (no spread [NS] condition) or two discontinuous clusters of stimuli (i.e., perceptually dissimilar clusters) were associated with each category (discontinuous spread [DS] condition). Scatter plots of the stimuli and optimal bounds are displayed in the top four panels of Figure 1. The bottom two panels display the stimuli used in the transfer phase and will be described later. In the II-DS condition,

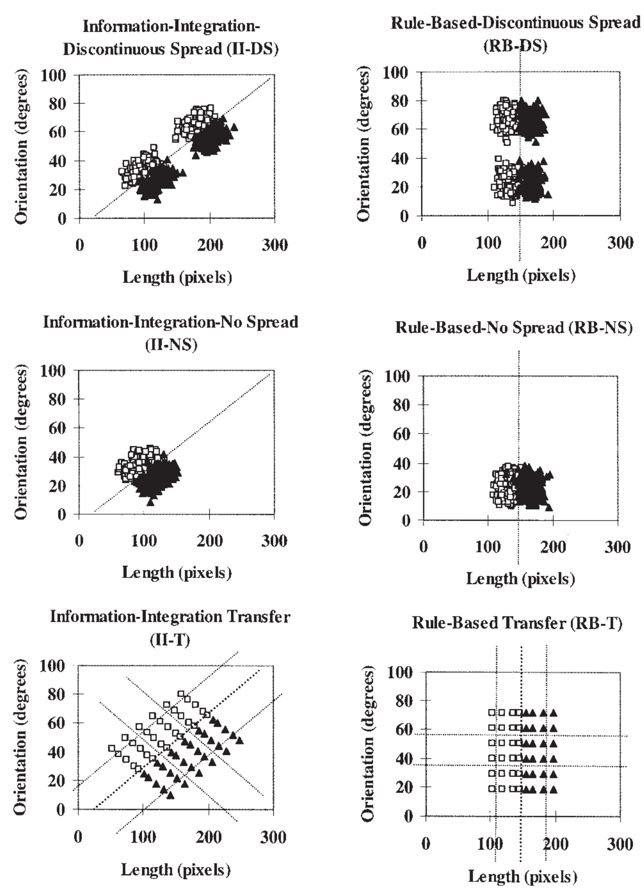

Figure 1. Scatter plots of the stimuli along with the optimal decision bounds from the four conditions of Experiment 1 and the transfer items used in the rule-based and information-integration conditions. Open squares denote stimuli from Category A. Filled triangles denote stimuli from Category B.

two discontinuous clusters of stimuli (each composed of two subclusters; see Table 1) are associated with Category A and two with Category B, but the same decision rule can be used to classify each pair of clusters within a category. In the II-NS condition, a single cluster of stimuli is associated with Category A and a separate cluster with Category B. It is important to note that the same decision rule is optimal in both conditions.

Structurally equivalent RB conditions were also constructed by rotating the II condition stimuli 45° around the center of the length–orientation space.[4] In the RB-DS condition, we again have two discontinuous clusters of stimuli associated with Category A and a separate pair of discontinuous clusters associated with Category B. A single, linear decision bound can be used to accurately classify all clusters. In the RB-NS condition, we have a single cluster associated with Category A and a single cluster associated with Category B. Again, the same decision bound can be used to optimally classify the stimuli across both conditions.

If the procedural-based learning system is dependent on within-category similarity, II category learning will be adversely affected

Table 1
Category Distribution Parameters From Experiment 1

Category	μ_l	μ_o	σ_l	σ_o	cov_{lo}
Information-integration discontinuous spread					
A_1	86	83	10	10	0
A_2	107	104	10	10	0
B_1	171	168	10	10	0
B_2	192	189	10	10	0
A_3	108	61	10	10	0
A_4	129	82	10	10	0
B_3	193	146	10	10	0
B_4	214	167	10	10	0
Information-integration no spread					
A_1	86	83	10	10	0
A_2	107	104	10	10	0
B_1	171	168	10	10	0
B_2	192	189	10	10	0
Rule-based discontinuous spread					
A_1	134	50	10	10	0
A_2	134	80	10	10	0
B_1	166	50	10	10	0
B_2	166	80	10	10	0
A_3	134	170	10	10	0
A_4	134	200	10	10	0
B_3	166	170	10	10	0
B_4	166	200	10	10	0
Rule-based no spread					
A_1	134	50	10	10	0
A_2	134	80	10	10	0
B_1	166	50	10	10	0
B_2	166	80	10	10	0

Note. Optimal accuracy was held constant at 95% in all conditions. μ_l = the population mean on the length dimension; μ_o = the population mean on the orientation dimension; σ_l = the population standard deviation on the length dimension; σ_o = the population standard deviation on the orientation dimension; cov_{lo} = the population covariance between length and orientation.

by discontinuous clusters. If the hypothesis-testing system learns an abstract rule and that rule is identical across DS and NS conditions, then RB category learning should be unaffected by discontinuous clusters. We tested these predictions in Experiment 1. For exploratory purposes, we also included at the end of each experimental session a set of transfer trials in which no corrective feedback was provided. The transfer stimuli (see Figure 1) included stimuli from the portion of the length–orientation space presented during training, as well as items from novel (untrained) portions of the space. The transfer trials were included because it was of interest to examine performance generalization to portions of the stimulus space that were not encountered during training. If an abstract rule is learned by the hypothesis-testing system, then learning should generalize fairly well to novel (untrained) items in the RB condition. In fact, untrained items that are better described by the rule should yield superior performance. On the other hand, if learning involves linking regions of the stimulus space with a particular category response, as predicted by the procedural-based learning system, then generalization should not be as good in the II condition and should be worse for transfer items coming from regions of the stimulus space that are farther from the training items, even if they are well described by the optimal rule.

Method

Observer. We solicited 16 observers (8 women and 8 men) from the University of Texas community and paid them $25 for participating in this study. Each observer completed all four experimental conditions with the condition order being determined from a Latin square. Only one of the four conditions (approximately 60 min) was completed during a single test day, and 1 rest day was required between testing sessions. Visual acuity was tested in each observer, and all observers had 20/20 vision or vision corrected to at least 20/20.

Stimuli and stimulus generation. The experiment used the randomization technique introduced by Ashby and Gott (1988). The category structures are displayed in Figure 1 along with the optimal decision bound(s). The category distribution parameters are outlined in Table 1 and optimal accuracy was 95%. In the RB-NS and II-NS conditions, 120 stimuli were sampled randomly from each of the four distributions for a total of 480 stimuli. In the RB-DS and II-DS conditions, 60 stimuli were sampled randomly from each of the eight distributions for a total of 480 stimuli. The resulting 480 stimuli were randomized and divided into five 96-trial blocks. These were presented during categorization training. Sixty stimuli (30 from the A response region and 30 from the B response region) were used during the transfer phase (see Figure 1). Each of these stimuli was presented twice for a total of 120 transfer trials.

Procedure. Each observation was run individually in a dimly lit testing room with an approximate viewing distance of 35 cm. The observers were informed that there were two equally likely categories. They were informed that perfect performance was impossible but that high levels of accuracy could be achieved. They were instructed to learn about the categories, to be as accurate as possible, and to not worry about speed of responding. At the start of each training trial, a fixation point was displayed for 1 s and then the stimulus appeared. The stimulus remained on the screen until the observer generated a response by pressing one of two keys. The correct category label was then presented on the screen for 1 s along with the word *wrong* if their response was incorrect or *right* if their response was correct.

[4] By structurally equivalent, we mean category structures for which the optimal accuracy, number of stimulus clusters, within-cluster scatter, and cluster coherence is equivalent across RB and II conditions.

Once feedback was given, the next trial was initiated. The procedure for the transfer trials was identical except that feedback was omitted.

Results

Training block analyses. Analyses were performed separately on each of the five 96-trial blocks of data. A nature of the Optimal Decision Strategy (RB vs. II) × Within-Category Discontinuity (DS vs. NS) × Block (five 96-trial blocks) within-observer design analysis of variance (ANOVA) was conducted on the accuracy rates. The accuracy rates averaged across observers are presented in Figure 2. The main effects of nature of the optimal rule, $F(1, 15) = 16.53$, $p < .01$, within-category discontinuity, $F(1, 15) = 49.02$, $p < .001$, and block, $F(4, 60) = 27.15$, $p < .001$, were significant. The nature of the Optimal Rule × Within-Category discontinuity, $F(1, 15) = 18.35$, $p < .01$, nature of the Optimal Rule × Block, $F(4, 60) = 7.24$, $p < .001$, and the three-way interaction, $F(4, 60) = 4.88$, $p < .01$, were all significant. To determine the locus of the three-way interaction, we conducted follow-up analyses that examined the effects of the nature of the optimal rule and within-category discontinuity on a block-by-block basis. In all five blocks of trials, there was a significant decline in performance for the II-DS relative to the II-NS condition, yielding performance drops of .10, .10, .10, .10, and .17 in Blocks 1–5, respectively (all *p*s $< .001$). On the other hand, there was no significant performance difference across the RB-DS and RB-NS conditions (all *p*s $> .05$ except in Block 3), yielding performance differences of only .04, .01, .06, .01, and .00 in Blocks 1–5, respectively.

Transfer block analyses. To provide an initial examination of transfer performance, we conducted a nature of the Optimal Decision Strategy (RB vs. II) × Within-Category Discontinuity (DS vs. NS) within-observer design ANOVA on the transfer block accuracy rates. The accuracy rates averaged across observers are presented in Figure 2 and are denoted by the block marked *T* for transfer. Only the main effect of nature of the optimal rule, $F(1, 15) = 8.99$, $p < .01$, was significant. As suggested by Figure 2, RB category learning led to better transfer (0.83) relative to transfer following II category learning (0.70). To determine how performance changed from the final training block to the transfer block, we computed a *difference score* by subtracting transfer performance from performance in the final training block and subjecting this score to the same ANOVA. The main effect of within-category discontinuity, $F(1, 15) = 20.16$, $p < .001$, and the interaction, $F(1, 15) = 10.98$, $p < .01$, were both significant, whereas the effect of the nature of the optimal rule was marginally significant, $F(1, 15) = 3.63$, $p = .08$. To determine the locus of the interaction, we compared the NS and DS difference scores separately for the RB and II category structures. For the RB structures, the effect of within-category discontinuity was nonsignificant, $t(15) = 1.34$, *ns*, with no change in accuracy from the final training to the transfer block in the NS condition and a slight increase of .03 in the DS condition. The effect of within-category discontinuity in the II condition, on the other hand, was highly significant, $t(15) = 4.60$, $p < .001$, and suggested a small increase in accuracy for the DS condition of .05 but a large drop of .13 for the NS condition.

Recall that the transfer items were sampled from trained and untrained regions of the stimulus space, and that participants in the DS conditions were presented with a wider range of items. These initial transfer results suggest a strong interaction between the nature of the optimal rule and within-category discontinuity. In the RB conditions, in which observers are hypothesized to learn an explicit, verbalizable rule, the within-category discontinuity manipulation has little effect on transfer performance. In other words, whether observers were trained on a small or larger portion of the stimulus space had no effect on global transfer accuracy. On the other hand, in the II conditions, in which observers are assumed to assign responses to regions of the stimulus space and a decision bound is not learned directly, transfer performance was better in the DS condition, in which more of the stimulus space was trained, than in the NS condition, in which less of the stimulus space was trained. It is worth mentioning that task difficulty cannot be used to explain this finding, as II-NS performance was significantly better than DS performance by the final block of training.

To gain a more detailed understanding of the transfer results, we partitioned the transfer stimuli into 12 groups of stimuli, 6 groups on the A side of the optimal decision bound and 6 groups on the B side of the decision bound. The idea was to examine performance for (a) items sampled from the trained region of the stimulus space, (b) items sampled from untrained regions of the space that were equidistant to the optimal bound with trained items, and (c) items sampled from an untrained region of the space that were farther from the optimal bound. The partitioning is outlined in Figure 1. For ease of exposition we focus the description on the RB transfer items. Three vertical lines and two horizontal lines partition the 60 transfer stimuli into the 12 groups. The two stimuli in the lower left portion are from an untrained region of the stimulus space, are far from the optimal bound, and are on the A side of the bound. These items are directly to the left of the transfer items that were sampled from the trained portion of the stimulus in both the DS and NS conditions. Similarly, the two stimuli in the lower right portion are from an untrained region of the stimulus space, are far from the optimal bound, and are on the B side of the bound. These items are directly to the right of transfer items that were sampled from the trained portion of the stimulus in both the DS and NS conditions. This collection of four items (two from the A side and two from the B side of the bound) is referred to as *far/DS–NS items* because they are far from the bound (on either the A or B side) and

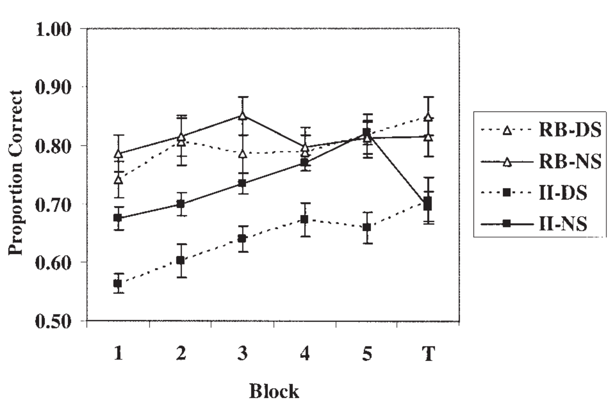

Figure 2. Proportion correct (averaged across observers, ± *SE*) from Experiment 1. RB-DS = rule-based discontinuous-spread condition; RB-NS = rule-based no-spread condition; II-DS = information-integration discontinuous-spread condition; II-NS = information-integration no-spread condition; T = transfer.

are directly next to items trained in the DS and NS conditions. The four stimuli directly above the far/DS–NS items (two from the A and two from the B side of the bound) are also far from the bound, but they are next to an untrained region of the space and thus are referred to as *far/untrained items*. The four items directly above these again are far from the bound and next to items from the trained region of the space but only in the DS condition. These are referred to as *far/DS items*. The column of items to the right (on the A side of the bound) and to the left (on the B side of the bound) of these far items are near the bound and so are referred to as *near/DS–NS*, *near/untrained*, and *near/DS* items going from the bottom to the top, respectively. The proportion correct (averaged across observers) for these six types of items in the II-DS, II-NS, RB-DS, and RB-NS conditions are displayed in Table 2. The groups of stimuli that were trained in each condition are highlighted in bold type.

First, we compared transfer performance across the near/DS–NS, near/untrained, and near/DS items separately for each condition using a one-way ANOVA. We were interested in determining whether categorization performance differed for items equidistant from the decision bound but from different regions of the space (trained or untrained). We first examined performance in the near RB-NS and II-NS conditions because (a) this provided the best test of our hypotheses given that trained items were confined to one cluster of stimulus space for each category (whereas this was not the case for the RB-DS and II-DS conditions), thereby allowing us to determine whether transfer accuracy decreases at the furthest points from the trained region, and (b) by examining this effect in the near transfer items only, we were able to hold constant the impact of the distance to the bound. Thus, these comparisons should provide the best insight in the nature of what is learned by the hypothesis-testing and procedural-learning systems. For the II-NS conditions the effect of stimulus type was significant, $F(2, 30) = 18.24$, $p < .001$, whereas for the RB-NS condition the effect was nonsignificant, $F(2, 30) = 2.13$, $p > .05$. Post hoc analyses of the II-NS results indicated that performance was significantly worse for the near/DS (untrained items in this condition) than for the near/untrained items, and that performance for the near/untrained items was significantly worse than for the near/DS–NS items. In other words, as distance from the training items increased for items equidistant from the optimal decision bound, II transfer performance declined, whereas RB transfer performance remained statistically unchanged. This is a strong piece of evidence in support of the prediction that it is a rule that is learned by the hypothesis-testing system, and that perceptually similar stimulus-category label assignments are learned by the procedural-learning system.

We also conducted the analogous set of analyses for the II-DS and for the RB-DS conditions, although these analyses could potentially provided somewhat less of a stringent test of the effects of distance to the trained regions because more of the stimulus space was trained in the II-DS and RB-DS conditions. For both the II-DS and RB-DS conditions, the effect of stimulus type was significant, II-DS, $F(2, 30) = 5.16$, $p < .05$; RB-DS, $F(2, 30) = 4.02$, $p < .05$. Post hoc analyses of the II-DS results indicated that performance was significantly worse for the near/untrained items than for the near/DS–NS or near/DS items, and that performance was equivalent across the near/DS–NS and near/DS items. This pattern was expected, as both the near/DS-NS and near/DS items were trained in the II-DS condition. Post hoc analyses of the RB-DS results led to the same conclusion, but, it is important to note that the effects were about half the magnitude of those from the II-DS condition. The finding that untrained items as equidistant to the decision bound as trained items yield worse transfer performance in II conditions provides strong support for the hypothesis that the procedural-learning system learns to assign perceptually similar items to the same category label and does not learn a decision bound directly. The finding that this same pattern does not occur in RB tasks (e.g., the RB-NS condition), and that when it does occur, as in the RB-DS condition, the effect is much smaller (5% decline for near/untrained relative to near/trained for RB-DS, but an 11% decline in the II-DS condition), suggests that it is a decision rule that is being learned by the hypothesis-testing system. It is important to note that these effects were observed despite the fact that more of the stimulus space was trained in the II-DS and the RB-DS conditions as compared with the II-NS and the RB-NS conditions.

To examine potential differences in the two systems more fully, we replicated these analyses for the far items. In this case, none of the items were from a trained region of the space, but they did differ in distance from trained items. Again, we predicted no effect of stimulus type for the RB conditions because a decision rule is learned and all far items are equally distant from the rule. On the other hand, we did predict an effect of stimulus type for the II conditions, especially the II-NS condition, as items farther from the trained region of the space should yield lower performance. Both predictions were born out in the analyses of the far transfer

Table 2
Probability Correct for the Transfer Trials From Experiment 1

	Distance to bound	
Transfer trial	Near	Far
II-DS		
DS	**0.75**	0.78
Untrained	0.65	0.78
DS-NS	**0.73**	0.79
II-NS		
DS	0.59	0.72
Untrained	0.71	0.88
DS-NS	**0.77**	0.92
RB-DS		
DS	**0.85**	0.91
Untrained	0.79	0.96
DS-NS	**0.84**	0.95
RB-NS		
DS	0.76	0.89
Untrained	0.78	0.93
DS-NS	**0.82**	0.96

Note. Groups of stimuli that were trained in each condition are in boldface. II-DS = information-integration discontinuous-spread condition; II-NS = information-integration no-spread condition; RB-DS = rule-based discontinuous-spread condition; RB-NS = rule-based no-spread condition.

items. The effect of stimulus type was significant for the II-NS condition, $F(2, 30) = 13.28$, $p < .001$, yielding significantly worse performance for the far/DS items relative to the far/untrained items and relative to the far/DS-NS items, although performance did not differ between the far/untrained, and far/DS-NS items. The effect of stimulus type was nonsignificant in the other three conditions, II-DS, $F(2, 30) < 1$; RB-NS, $F(2, 30) = 2.83$, $p > .05$; RB-DS, $F(2, 30) = 1.93$, $p > .05$. One excellent example of the performance difference between the RB and II tasks can be seen when one compares the RB-NS and II-NS conditions. In the RB-NS condition, the far/DS items yield an accuracy rate of .89, which is only slightly lower than the accuracy rate of .96 for the far/DS-NS items. In the II-NS condition, on the other hand, accuracy for the far/DS items is .72, whereas accuracy for the far/DS-NS items is .93. This is a drop of .21, whereas in the analogous RB condition, the drop was only .07. Taken together these results suggest that a rule is extracted in the RB conditions and items that better follow the rule (i.e., items more distant from the bound), regardless of whether they are near or close to trained items, yield higher accuracy rates. In contrast, II category learning is characterized by a strengthening of the relationship between responses and regions of the stimulus space. In this case, as the distance between a transfer item and a trained region of the space increases, accuracy generally decreases.

It is important to make clear that we are not arguing that distance to the bound has no effect on II category learning. An examination of Table 2 suggests that distance to the bound affects both RB and II category learning. Rather, our claim is that distance to the bound alone seems to characterize performance of the hypothesis-testing system, whereas there is a large effect of distance to the training items (along with distance to the bound effects) on II category learning. To our knowledge, this is the first study to directly and rigorously test this prediction.

Discussion

The results from Experiment 1 suggest that RB category learning is unaffected by within-category discontinuity when the optimal decision bound remains constant. On the other hand, II category learning is adversely affected by within-category discontinuity. The transfer results suggest a qualitative difference in the nature of the learning in the two systems. RB category learning is abstract in that a rule is learned directly and learning is less tied to the specific stimulus regions trained, resulting in a strong distance-to-the-bound effect. In contrast, II learning is more directly tied to the specific regions of the stimulus space that were trained, a rule is not learned directly, and distance from the trained regions strongly affects performance.

These data suggest that the procedural-based categorization learning system is highly dependent on the continuity of stimulus similarity within a category, whereas the hypothesis-testing system is not. There is, however, one alternative explanation for the effect on II category learning that cannot be ruled out at this point. It is possible that the poor II learning in the DS condition relative to the NS condition was due to the fact that more of the stimulus space had to be learned in the former condition, and, as such, the differences observed were not due to the fact that the category clusters were discontinuous. Recall that twice as many within-category clusters were trained in the DS condition than in the NS condition. Perhaps within-category discontinuity is irrelevant, and instead, II category learning decreases as the amount of the stimulus space to be trained increases. Experiment 2 was conducted to address this shortcoming.

Experiment 2

Experiment 2 continued our examination of the effect of within-category discontinuity on II category learning, but in this experiment, we equated the amount of the stimulus space to be trained across conditions. In addition, we replaced the within-observer design used in Experiment 1 with a between-observer design. RB conditions were not included because no effect of within-category spread was observed in Experiment 1. The NS condition from Experiment 1 was replaced with two *continuous spread* (CS) conditions. Scatter plots of the stimuli from the continuous spread A (CS-A) and continuous spread B (CS-B) conditions, along with the optimal bounds, are displayed in Figure 3. Figure 3 also displays the stimuli for the DS condition. Notice that the optimal decision bound is identical across all three conditions. If the within-category discontinuity effect observed in Experiment 1 was due to differences in the amount of stimulus space trained, then there would be no performance difference observed across the three conditions. On the other hand, if within-category spread had been the mediating factor, then we predicted worse performance in the DS condition than in either CS condition.

To continue our exploratory analysis of generalization, we presented a block of transfer stimuli at the end of the study. These stimuli are presented in Figure 3. Notice that a larger range of the stimulus space is examined during transfer in Experiment 2 as compared with Experiment 1. We took this approach to better characterize transfer performance.

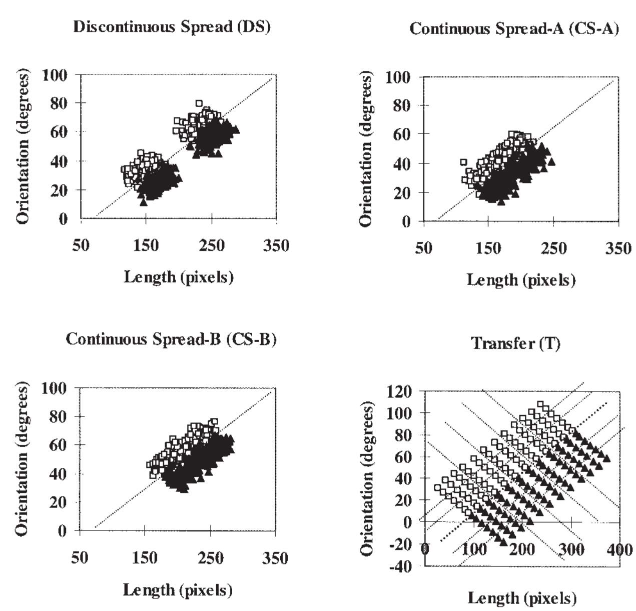

Figure 3. Scatter plots of the stimuli along with the optimal decision bounds from the three conditions of Experiment 2 and the transfer items. Open squares denote stimuli from Category A. Filled triangles denote stimuli from Category B.

Method

Observer. We solicited 70 observers (24 in the DS condition and 23 in each of the CS conditions) from the University of Texas community. The observers received course credit or were paid for participating in this study. Visual acuity was tested in each observer, and all observers had 20/20 vision, or vision corrected to at least 20/20.

Stimuli and stimulus generation. The stimuli and stimulus generation procedure was identical to that used in Experiment 1, except that the number of transfer stimuli was increased to 168, with each item presented once. The category structures are displayed in Figure 3 along with the optimal decision bounds. The category distribution parameters are outlined in Table 3, and optimal accuracy was 95%.

Procedure. The procedure was identical to that used in Experiment 1.

Results

Training block analyses. Analyses were performed separately on each of the five 96-trial blocks of data. A Within-Category Discontinuity (DS vs. CS-A vs. CS-B) × Block (five 96-trial blocks) mixed-design ANOVA was conducted on the accuracy rates. The accuracy rates averaged across observers are presented in Figure 4. The main effects of within-category discontinuity, $F(2, 67) = 3.31$, $p < .05$, and block, $F(4, 268) = 7.54$, $p < .001$, were significant, whereas the interaction was nonsignificant, $F(8, 268) = 1.22$, $p > .05$. Post hoc analyses of the within-category discontinuity effect suggested that performance was significantly worse in the DS condition (.55) than in both CS conditions ($p < .05$). Performance in the CS-A (.59) and CS-B (.59) conditions did not differ.

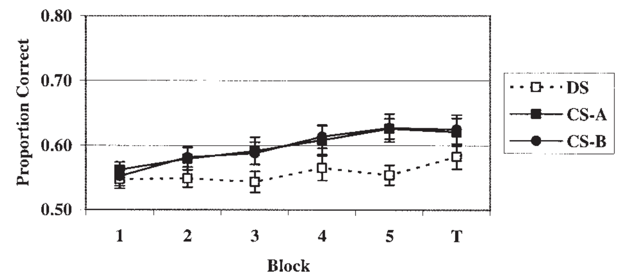

Figure 4. Proportion correct (averaged across observers, ± *SE*) from Experiment 2. DS = discontinuous-spread condition; CS-A = continuous-spread condition A; CS-B = continuous-spread condition B; T = transfer.

Table 3
Category Distribution Parameters From Experiment 2

Category	μ_l	μ_o	σ_l	σ_o	cov_{lo}
		Discontinuous spread			
A_1	136	83	10	10	0
A_2	157	104	10	10	0
A_3	221	168	10	10	0
A_4	242	189	10	10	0
B_1	158	61	10	10	0
B_2	179	82	10	10	0
B_3	243	146	10	10	0
B_4	264	167	10	10	0
		Continuous spread–A			
A_1	136	83	10	10	0
A_2	157	104	10	10	0
A_3	178	126	10	10	0
A_4	199	147	10	10	0
B_1	158	61	10	10	0
B_2	179	82	10	10	0
B_3	201	103	10	10	0
B_4	222	124	10	10	0
		Continuous spread–B			
A_1	178	126	10	10	0
A_2	199	147	10	10	0
A_3	221	168	10	10	0
A_4	242	189	10	10	0
B_1	201	103	10	10	0
B_2	222	124	10	10	0
B_3	243	146	10	10	0
B_4	264	167	10	10	0

Note. Optimal accuracy was held constant at 95% in all conditions. μ_l = the population mean on the length dimension; μ_o = the population mean on the orientation dimension; σ_l = the population *SD* on the length dimension; σ_o = the population *SD* on the orientation dimension; cov_{lo} = the population covariance between length and orientation.

These results suggest that within-category discontinuity does have an adverse effect on II category learning when the amount of stimulus space trained is held constant. To quantify the magnitude of this effect, we computed the average probability correct across the two CS conditions minus that in the DS condition. For Blocks 1–5, respectively, these values are .01, .03, .05, .05, and .07.

Transfer block analyses. As in Experiment 1, we began by examining overall transfer performance. We conducted a within-category discontinuity (DS vs. CS-A vs. CS-B) ANOVA on the transfer block accuracy rates. The accuracy rates averaged across observers are presented in Figure 4 and are denoted by the block marked *T* for transfer. The effect of within-category discontinuity was nonsignificant, $F(2, 67) = 1.25$, $p > .05$. To determine how performance changed from the final training block to the transfer block, we computed a difference score by subtracting transfer performance from performance in the final training block and subjected this score to the same ANOVA. The within-category discontinuity effect was nonsignificant, $F(2, 67) = 1.20$, $p > .05$. It is interesting that the difference scores were all near 0 (DS = −.03, CS-A = .01, CS-B = 0).

To gain a more detailed understanding of the transfer results, we partitioned the transfer stimuli into 30 groups of stimuli, 15 groups on the A side of the optimal decision bound and 15 groups on the B side of the decision bound. The partitions are outlined in the bottom panel of Figure 3. We took the same general approach that we did in Experiment 1, examining transfer items at different distances from the optimal bound and different distances from trained areas of the stimulus space. Because we had a much larger set of transfer items, the partitioning was more fine grained. The 15 partitions on each side of the bound were constructed from the factorial combination of three distances to the optimal bound: near, medium, or far from the bound with 5 partitions orthogonal to the optimal decision bound. With respect to distance to the bound, only near items were potentially from trained regions of the stimulus space. Whether they were from trained or untrained regions depended on the condition. With respect to the 5 partitions orthogonal to the optimal decision bound, the two most extreme levels were always untrained. For ease of presentation, we averaged across analogous partitions on the A and B side of the bound, as we did in Experiment 1. The proportion correct (averaged across observers) for these 15 partitions separately by condition are displayed in Table 4. The groups of stimuli from trained regions of the space are highlighted in bold type.

Our approach was to compare trained and untrained region item accuracy separately for each distance to the bound (near, medium, and far). We conducted three mixed-design ANOVAs (one for each distance to the bound), all with within-category discontinuity as a between-subjects factor and trained versus untrained region as a within-subjects factor. For near items, the effect of training was significant, $F(1, 67) = 4.71$, $p < .05$, but the main effect of condition and the interaction were both nonsignificant, $Fs(1, 67) < 1$. As expected, items from the untrained region yielded lower accuracy than items from the trained region. The same pattern held for medium and far items. Specifically, the main effect of training was significant, medium, $F(1, 67) = 34.72$, $p < .001$; far, $F(1, 67) = 9.13$, $p < .01$, whereas the main effect of condition, medium, $F(2, 67) = 1.67$, $p > .05$; far, $F(2, 67) = 1.27$, $p > .05$, and the interaction were nonsignificant, medium, $F(2, 67) = 1.20$, $p > .05$; far, $F(2, 67) < 1$.

Two comments are in order. First, these data suggest that items from untrained regions yield lower accuracy than items from trained regions even when distance to the bound is controlled. This replicates the effects observed in Experiment 1. Second, there was no effect of condition, whereas there were performance differences across the II-DS and II-NS conditions in Experiment 1. Most likely, the fact that the amount of space trained is held constant here partially explains this result. In summary, as in Experiment 1, II transfer performance appears to be governed by both distance to the bound and distance from the trained regions.

Table 4
Probability Correct for the Transfer Trials From Experiment 2

Transfer trial	Distance to optimal bound: Near	Medium	Far
Discontinuous spread			
Untrained	0.54	0.57	0.62
Trained	**0.56**	0.65	0.65
Untrained	0.55	0.61	0.70
Trained	**0.54**	0.66	0.65
Untrained	0.50	0.58	0.57
Continuous spread-A			
Untrained	0.52	0.64	0.66
Untrained	0.60	0.69	0.77
Trained	**0.59**	0.74	0.74
Trained	**0.55**	0.68	0.74
Untrained	0.54	0.52	0.58
Continuous spread-B			
Untrained	0.52	0.62	0.68
Trained	**0.60**	0.74	0.70
Trained	**0.61**	0.76	0.77
Untrained	0.58	0.67	0.76
Untrained	0.52	0.61	0.57

Note. Groups of stimuli from trained regions of the space are in boldface.

Discussion

The results from Experiment 2 suggest that within-category spread adversely affects II category learning even when the amount of stimulus space trained is controlled across conditions. In contrast, discontinuity did not influence RB category learning (see Experiment 1). It is worth mentioning that overall accuracy levels were lower in Experiment 2 than in Experiment 1. This is likely due to the fact that Experiment 1 was run within-observers, which would give observers more exposure to the nature of the stimuli and tasks. Even so, this does not diminish the fact that discontinuity had a strong effect on II category learning when the amount of stimulus space trained was equated. Experiment 3 takes the next step by examining the effects of within-category discontinuity parametrically. Specifically, we examined three levels of within-category discontinuity and their effects on II category learning.

Experiment 3

Experiment 3 compared performance across three within-category discontinuity conditions to determine whether increases in the magnitude of within-category discontinuity are associated with decreases in performance, or whether any level of within-category spread is associated with a decrease in performance. To achieve this goal, we compared performance in a CS condition with that in a low DS (LDS) and a high DS (HDS) condition. As in Experiment 2, we held the amount of stimulus space trained constant across conditions, and the conditions were such that the optimal decision bound was identical. Scatter plots of the stimuli along with the optimal decision bound for all three conditions are displayed in Figure 5. If the magnitude of the within-category discontinuity affected performance, then we would have expected a monotonic decline in performance across the CS, LDS, and HDS conditions. On the other hand, if any within-category discontinuity had an equivalent effect, then we would have predicted worse, but equivalent, performance in the two DS conditions relative to the CS condition. We again include a block of transfer stimuli at the end of the study (see Figure 5).

Method

Observer. We solicited 69 observers (29 in the continuous condition, 21 in the low spread condition, and 19 in the high spread condition) from the University of Texas community. Observers received course credit or pay for participating in this study. Visual acuity was tested in each observer, and all observers had 20/20 vision, or vision corrected to at least 20/20.

Stimuli, stimulus generation, and procedures. The stimuli, stimulus generation, and experimental procedures were identical to those used in Experiment 2. The category structures are displayed in Figure 5 along with the optimal decision bound(s). The category distribution parameters are outlined in Table 5, and optimal accuracy was 95%.

Results

Training block analyses. Analyses were performed separately on each of the five 96-trial blocks of data. A Within-Category Discontinuity (CS vs. LDS vs. HDS) × Block (five 96-trial

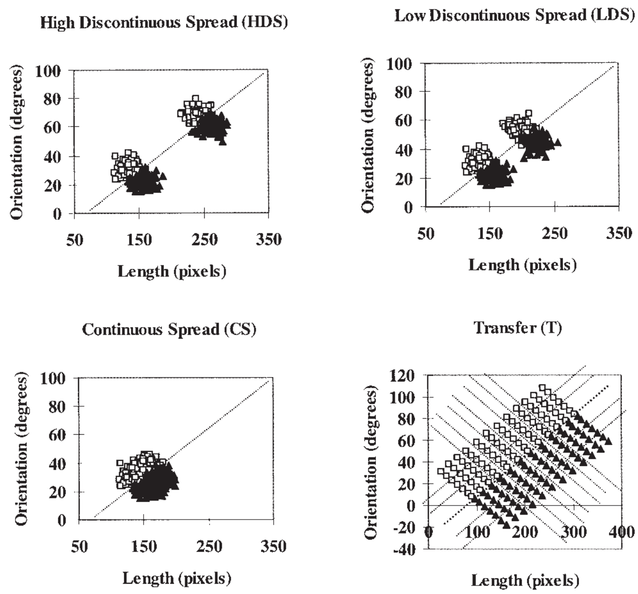

Figure 5. Scatter-plots of the stimuli along with the optimal decision bounds from the three conditions of Experiment 3 and the transfer items. Open squares denote stimuli from Category A. Filled triangles denote stimuli from Category B.

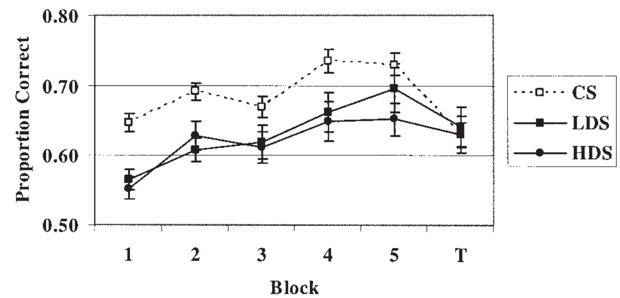

Figure 6. Proportion correct (averaged across observers, ± *SE*) from Experiment 3. CS = continuous-spread condition; LDS = low discontinuous-spread condition; HDS = high discontinuous-spread condition; T = transfer.

blocks) mixed-design ANOVA was conducted on the accuracy rates. The accuracy rates averaged across observers are presented in Figure 6. The main effects of within-category discontinuity, $F(2, 66) = 7.14$, $p < .01$, and block, $F(4, 264) = 28.33$, $p < .001$, were significant. The interaction was nonsignificant, $F(8, 264) = 1.01$, $p > .05$. Post hoc analyses of the within-category discontinuity main effect suggested better performance in the CS condition than in both the LDS and HDS conditions ($p < .05$) and equivalent performance in the LDS and HDS conditions. This result suggests that increasing the magnitude of the spread does not lead to a monotonic decline in performance, but rather that any spread adversely affects II learning. Although there was a trend toward LDS performance to be better than HDS performance during the final block of training, this effect was nonsignificant based on a *t* test ($p > .05$). Even so, the results do suggest that, with additional training, the monotonic decline in performance with increased DS observed during the final training block might increase in magnitude.

Table 5
Category Distribution Parameters From Experiment 3

Category	μ_l	μ_o	σ_l	σ_o	cov_{lo}
Continuous spread					
A_1	136	83	10	10	0
A_2	157	104	10	10	0
B_1	158	61	10	10	0
B_2	179	82	10	10	0
Low discontinuous spread					
A_1	136	83	10	10	0
A_2	199	147	10	10	0
B_1	158	61	10	10	0
B_2	222	124	10	10	0
High discontinuous spread					
A_1	136	83	10	10	0
A_2	242	189	10	10	0
B_1	158	61	10	10	0
B_2	264	167	10	10	0

Note. Optimal accuracy was held constant at 95% in all conditions. μ_l = the population mean on the length dimension; μ_o = the population mean on the orientation dimension; σ_l = the population *SD* on the length dimension; σ_o = the population *SD* on the orientation dimension; cov_{lo} = the population covariance between length and orientation.

Transfer block analyses. As in the previous two experiments, we conducted a within-category discontinuity (CS vs. LDS vs. HDS) ANOVA on the transfer block accuracy rates. The accuracy rates averaged across observers are presented in Figure 6 and are denoted by the block marked *T* for transfer. The effect of within-category discontinuity was nonsignificant, $F(2, 67) < 1$. To determine how performance changed from the final training block to the transfer block, we computed a difference score by subtracting transfer performance from performance in the final training block and subjected this score to the same ANOVA. The within-category discontinuity effect was significant, $F(2, 66) = 4.93$, $p < .01$. Post hoc analyses indicated that performance declined from the final training to the transfer block in all conditions, with the decline being largest in the CS condition (.098), intermediate in the LDS condition (.055), and smallest in the HDS condition (.020).

We performed the same type of partitioning in these data that we did in Experiment 2, except that the partitioning was more fine-grained. Specifically, instead of examining five partitions orthogonal to the optimal decision bound we examined eight. This change was needed because the trained subregions were smaller in Experiment 3 (compare Figures 3 and 5). The probability correct (averaged across observers) for these 24 partitions separated by condition are displayed in Table 6. The groups of stimuli from trained regions of the space are highlighted in bold type.

We followed the approach in Experiment 2 and conducted three mixed-design ANOVAs (one for each distance to the bound: near, medium, and far) all using within-category spread as a between-subjects factor and trained versus untrained region as a within-subjects factor. For near items, the effect of training was signifi-

Table 6
Probability Correct for the Transfer Trials From Experiment 3

Transfer trial	Distance to optimal bound: Near	Medium	Far
High discontinuous spread			
Untrained	0.61	0.57	0.68
Trained	**0.68**	0.75	0.72
Untrained	0.59	0.71	0.67
Untrained	0.61	0.63	0.79
Untrained	0.56	0.61	0.76
Untrained	0.50	0.61	0.64
Trained	**0.63**	0.70	0.78
Untrained	0.51	0.54	0.53
Low discontinuous spread			
Untrained	0.53	0.50	0.58
Untrained	0.60	0.58	0.61
Untrained	0.63	0.58	0.75
Trained	**0.67**	0.69	0.79
Untrained	0.64	0.76	0.82
Untrained	0.63	0.67	0.76
Trained	**0.63**	0.70	0.76
Untrained	0.58	0.58	0.65
Continuous spread			
Untrained	0.54	0.53	0.57
Untrained	0.60	0.62	0.55
Untrained	0.59	0.63	0.61
Untrained	0.59	0.65	0.75
Untrained	0.64	0.72	0.87
Trained	**0.63**	0.72	0.74
Trained	**0.61**	0.68	0.80
Untrained	0.52	0.60	0.64

Note. Groups of stimuli from the trained regions of the space are in boldface.

cant, $F(1, 66) = 8.88$, $p < .01$, but the main effect of condition and the interaction were both nonsignificant, (Fs < 1). As expected, items from the untrained region yielded lower accuracy than items from the trained region. The same pattern held for medium and far items. Specifically, the main effect of training was significant, medium, $F(1, 66) = 29.64$, $p < .001$; far, $F(1, 66) = 19.41$, $p < .01$, whereas the main effect of condition (medium, $F < 1$; far, $F < 1$), and the interaction were nonsignificant, medium, $F(2, 66) = 1.90$, $p > .05$; far, $F(2, 66) = 2.13$, $p > .05$. These findings mimic those from Experiment 2 and suggest that items from untrained regions of the space yield lower transfer performance than items from trained regions of the space, with some evidence that performance declines with distance from a trained region (when distance to the bound is held fixed). In addition, transfer performance increases as distance to the bound increases. In summary, II transfer performance appears to be governed by both distance to the bound and distance from the trained regions.

Discussion

The results from Experiment 3 converge with the results from Experiment 2, suggesting that within-category discontinuity adversely affects II category learning even when the amount of stimulus space trained is controlled across conditions. Although only a single study, Experiment 3 also suggests that both a low and high within-category DS lead to approximately the same performance decrement suggesting that the existence of within-category discontinuity, and not the magnitude of such discontinuity, determines performance. Even so, during the final block of training there was a monotonic decline in performance as the magnitude of the spread increased, although this effect was nonsignificant.

General Discussion

The aim of the studies outlined in this report was to examine the qualitative properties of the implicit, procedural-based learning and explicit, hypothesis-testing systems proposed in COVIS. Information regarding the processing characteristics of the two systems has been obtained in a number of previous studies (for a review, see Maddox & Ashby, 2004), but in each case the examination was indirect in the sense that a single II and RB category structure was selected and some external manipulation was introduced. The current studies take the next, more direct, step by introducing systematic manipulations of the II and RB category structures. In COVIS, the procedural-learning system is involved in learning II categories. Processing in this system is not directly accessible to conscious awareness and is mediated largely within the tail of the caudate nucleus (Ashby et al., 1998; Ashby & Ell, 2001; Willingham, 1998). Learning in this system is incremental and requires a dopamine reward signal. In COVIS, the hypothesis-testing system learns RB categories. This system uses working memory and executive attention and is mediated primarily by the anterior cingulate, the prefrontal cortex, and the head of the caudate nucleus. This system appears to learn through a conscious process of hypothesis generation and testing.

The focus of this report was on the effects of within-category discontinuity on II and RB category learning. Because learning in the procedural-based categorization learning system involves associating perceptually similar clusters of stimuli with the same categorization response, we predicted that within-category discontinuity should affect the learning of II category structures. In particular, increasing within-category discontinuity should adversely affect II category learning. Because learning in the hypothesis-testing categorization system requires abstracting a rule, we postulated that within-category discontinuity should have no effect on RB learning as long as the optimal rule remains constant across within-category discontinuity conditions.

Experiment 1 examined the effects of within-category discontinuity on II and RB category learning. In the NS condition, the items from each category were perceptually similar and the perceptual variation across a category was gradual. In the DS condition, two distinct and perceptually dissimilar clusters of stimuli were associated with each category (see Figure 1). Both the DS and NS conditions were run with II and RB category structures. It is important to note that the optimal decision bound was equivalent across discontinuous and NS conditions. As predicted, the DS manipulation led to a large performance decrement in II category learning but did not lead to a performance decrement in RB category learning.

Although we argue that the results of Experiment 1 were due to within-category discontinuity, another possible explanation was that the effects observed could have been due to the amount of stimulus space that was trained. Specifically, the DS condition trained a larger portion of the stimulus space, and this could have led to the decreased learning in that condition. Experiment 2 addressed this possibility by equating the amount of stimulus space trained. Within-category discontinuity continued to adversely affect II category learning when the amount of stimulus space trained was equated. Experiment 3 examined the effects of within-category discontinuity parametrically by examining three levels of discontinuity: CS, LDS, and HDS. Performance in both DS conditions was worse than that observed in the CS condition, but the LDS and HDS conditions yielded the same level of performance. However, during the final training block, HDS performance was worse than low CS performance, suggesting that perhaps with additional training a monotonic decline in performance with increasing spread would emerge. Future research should address this important issue.

All three experiments included a block of transfer trials at the end of each session that included items of differing distances from the optimal decision bound as well as items from trained and untrained regions of the stimulus space. The results were clear. Transfer performance in the RB conditions was mediated by distance to the bound regardless of whether items were sampled from trained or untrained regions. Transfer performance in the II conditions, on the other hand, was mediated by both distance to the bound and distance from the trained-regions of the stimulus space. Specifically, as the distance to the bound increased, transfer performance generally increased. However, for items equidistant from the decision bound, transfer performance was better for items sampled from the trained rather than the untrained regions of the space.

Within-Category Discontinuity Effects

These data add to the growing body of research in support of the existence of functionally and neurobiologically distinct category learning systems. Ashby and his colleagues proposed that with RB structures, learning is mediated by a circuit that includes the anterior cingulate, the prefrontal cortex, and the head of the caudate nucleus, whereas in II 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; and the body of the caudate for auditory stimuli; Maddox, Molis, & Diehl, 2002). The mechanisms that mediate learning-related changes in synaptic efficacy within these two neural circuits are qualitatively different, and such differences suggest that within-category discontinuity may have different effects on RB and II tasks. The hypothesis-testing category learning system proposed above is under conscious control and has full access to working memory and executive attention, and accurate RB category learning depends on good use of these processes. As long as the optimal rule remains constant, the working memory load and executive attention resources necessary to learn the rule should remain constant, and performance should be unaffected regardless of the degree of within-category discontinuity. On the other hand, the procedural-based learning system learns to associate clusters of visual cortical cells with a specific response location. Ashby and Waldron (1999; see also Ashby, Waldron, Lee, & Berkman, 2001; Maddox, 2001, 2002; Waldron & Ashby, 2001) proposed a neurobiologically plausible model of learning in this system. In short, visual stimuli are represented perceptually in higher level visual areas, such as inferotemporal cortex (IT). Given the many-to-one convergence of information from IT into the striatum (Wilson, 1995), it is assumed that a low-resolution map of the perceptual space is represented among the striatal units. As the observer gains experience with the task, each unit becomes associated with a particular response through a gradual incremental learning process. Thus, the striatum can be thought of as associating a categorization response with a cluster of visual cortical cells that are associated with perceptually similar stimuli. When accurate performance requires that perceptually dissimilar clusters of stimuli be associated with the same response, this system is required to train and disambiguate more striatal units, which ultimately leads to a performance decrement.

When designing Experiment 3, our expectation was that we would observe a monotonic decline in accuracy with increased DS. Our thinking was that perceptual information about the stimuli might be spatio-topically represented in the tail of the caudate in such a way that items that look more alike will activate cells that are closer in proximity. With continuous categories, caudate cells that are spatially close will be activated for items from the same category, whereas in the high discontinuous condition, caudate cells that are spatially separated will be activated for items from the same category, with an intermediate situation holding in the low discontinuous condition. For the continuous condition, when dopamine is released into the caudate, the neurotransmitter only has to be distributed within a continuous group of cells, whereas for the other conditions, dopamine has to be distributed to clusters of spatially separate cells. As the cluster distance increases, the impact of the dopamine lessens because the same amount of dopamine is released regardless of cluster continuity. At this stage, this is still a viable hypothesis especially given the results from the final training block. One avenue for future research would be to increase the amount of training to determine whether the accuracy trend observed in Block 5 increases in magnitude. Another avenue would be to increase the discontinuity. It is possible that the discontinuity difference across the low and high discontinuous conditions was too small to yield a significant effect.

Effects of the Amount of Stimulus Space Trained

These data provide strong support for the prediction that the procedural-learning based system should be adversely affected by within-category discontinuity, but there is also some evidence that the amount of stimulus space trained during learning also affects performance. Notice that the amount of stimulus space trained in the Experiment 3 CS condition was half of that trained in the two CS conditions in Experiment 2. Performance was better in the Experiment 3 CS condition (.70) than in the two CS conditions in Experiment 2 (CS-A = .59, CS-B = .59) that trained twice as much of the stimulus space.

The interpretation of these results is very similar to that outlined above for discontinuous category spread. II category learning by the procedural-learning based system involves learning to associate units in the striatum with a categorization response. Each striatal unit is associated with a cluster of visual cortical cells that are associated with perceptually similar stimuli. When accurate performance requires that more clusters of stimuli be associated with the same response, this system is required to train more striatal units, which ultimately leads to a performance decrement. This results when more of the stimulus space is trained regardless of whether the spread is continuous.

Generalization Profiles for the Two Systems

The transfer results from Experiment 1 provide some of the strongest and perhaps most striking evidence in support of distinct category learning systems. These results suggest that the nature of performance generalization might be very different in the two category learning systems. An abstract rule appears to be learned in the RB conditions that leads to a distance-to-the-bound effect whereby stimuli that are farther from the bound yield higher accuracy rates, even though stimuli from this region were never presented during training and might be quite distant from trained items (e.g., far/DS items in the RB-NS condition). This rule is abstract in the sense that it can be applied to novel items and is not tied directly to the trained items. Items that are more representative of the rule (e.g., very short lines relative to a short line vs. long line categorization problem) yield higher accuracy rates. This result is very much in line with that predicted from the explicit, hypothesis-testing system proposed in COVIS, and in following, strongly suggests that those brain structures involved in this system (i.e., prefronto-cortical-thalamic loops and the head of the caudate) represent this rule. II learning is also partially affected by distance to the bound, but in contrast to the RB system, it appears to be more closely linked to the regions of the stimulus space that are trained. Transfer performance for items equidistant to the decision bound declined with distance from the training items in the II condition but not in the RB condition, suggesting that generalization in II tasks is strongly affected by the distance to the trained response region, with more distant items yielding worse performance.

It is important to be clear that we are not arguing that specific training item effects on RB category learning do not exist. On the contrary, a large body of work (e.g., Allen & Brooks, 1991; Sakamoto & Love, 2004) suggests that specific exemplars can have a strong effect on RB category learning. Although more work is needed, our belief is that specific training items have less effect on RB category learning than II category learning, especially when the categories are composed of large numbers of stimuli constructed from simple perceptual dimensions. An important focus of future research should be on attempting to merge these two different, but related, approaches.

Model-Based Analyses of Participant's Strategies

The conclusions drawn from the accuracy results are suggestive of two unique category learning systems. Even so, it is also important to determine what strategies an observer might use when solving these tasks. An understanding of strategy use and how these strategies might be affected by the nature of the category structures and the within-category discontinuity manipulation is of central importance to a more complete understanding of category learning. In addition, and perhaps more important, it is critical to determine whether participants are using a strategy similar to the optimal or are using a qualitatively different strategy. To determine the strategies used by observers in the present study, we fit a number of different decision bound models (Ashby, 1992; Maddox & Ashby, 1993) to the individual participant's block by block data. The details of our model-based approach are outlined in numerous published articles (e.g., Maddox, Filoteo, Hejl, & Ing, 2004) and will not be elaborated here. The important point is the two different classes of models that were applied. One class is compatible with the assumption that observers used an explicit hypothesis-testing strategy and one class assumes an II strategy. In general, the results support the assumption that participants were using hypothesis-testing strategies to solve the RB task and II strategies to solve the II tasks. These results support our conclusions that discontinuous category clusters impact II category learning but not RB category learning.

Discontinuous Category Learning Under Extended Training

In all three experiments we found that assigning discontinuous clusters of stimuli to the same category label led to poor II category learning relative to NS conditions. We attributed this effect to the fact that more striatal units were required to learn the II discontinuous category structures. One obvious question to ask is whether participants can learn discontinuous, II categories if given enough experience. We recently conducted a study in which 4 observers completed two consecutive sessions in the Experiment 1, II-DS condition and 4 observers completed two consecutive sessions in the Experiment 1, II-NS condition. By the final block of training in Session 1, the average proportion correct in the II-DS and II-NS conditions was .71 and .82, respectively. These values are in line with those from Experiment 1 (II-DS = .66; II-NS = .82), and suggest a large performance decrement for discontinuous category clusters. By the final block of training in Session 2, however, the average proportion correct in the II-DS and II-NS conditions was .80 and .83, respectively. In addition, the final block of data for all observers in Session 2 was best accounted for by an II model. These data suggest that observers can in fact learn to assign discontinuous clusters of stimuli to the same category nearly as well as they learn to assign a single coherent cluster of stimuli to the same category, but that extended training is required to do so. This finding supports our claim that more striatal units are necessary to learn the II-DS structures and that this process requires additional experience with the category exemplars.

An Alternative Complexity Explanation

One might argue that discontinuity affects II and not RB category learning because II categories are more difficult to learn. The idea being that any increase in difficulty, such as the discontinuity manipulation, will affect a potentially more complex task, such as an II task, more so than a potentially less complex task, such as an RB task. The data from the NS conditions from Experiment 1 do suggest that the unidimensional RB task was learned more quickly than the multidimensional II task, so this is a reasonable alternative explanation. However, there are a number of results that argue against this complexity explanation. First, the complexity argument would predict that requiring participants to perform a categorization task and a second task should adversely affect II category learning more than RB category learning. COVIS, on the other hand, predicts the opposite, as the second task places a greater demand on working memory. Consistent with predictions based on COVIS, the results from two studies (Maddox, Ashby, Ing, & Pickering, 2004; Waldron & Ashby, 2001) indicated that the secondary task impacted RB category learning more so than II category learning, even though the RB task was less difficult than the II task. Second, Maddox and Ing (2005; see also Maddox, Bohil, & Ing, 2004) examined the effects of delayed feedback on II and RB category learning, but in their study, the RB category learning task required that both stimulus dimensions be attended. In their study, this multidimensional, conjunctive RB task was more complex than the analogous multidimensional II task (on the basis of the category-level discriminability required to equate control condition performance). Because the conjunctive RB task was more complex, the complexity explanation predicts that delayed feedback should affect RB category learning more than it affects II category learning, as delaying the feedback increases the complexity of the task. COVIS, on the other hand, predicts that delayed feedback should affect II but not RB category learning. In line with the predictions from COVIS, delayed feedback adversely affected the simpler II task but not the more complex RB task. Thus, taking these findings into account, it does not appear that task complexity can entirely explain why category discontinuity impacted II but not RB category learning.

Alternative Theories of Category Learning

Despite the growing interest in multiple systems approaches to category learning, a number of single system models of category learning remain popular. Two single system alternatives are exemplar-similarity models, such as Nosofsky's (1986) generalized context model (GCM) and Kruschke's (1992) attention learning coverage map. To our knowledge, neither of these single system approaches could provide an a priori accounting of the differential impact of category discontinuity on the II and RB tasks in the present study. Even so, it is likely that either single-system model could account for the observed data by postulating different parameter settings across conditions. For example, the estimate of the psychological scaling parameter (c) in the GCM might be smaller in the DS condition relative to the NS condition resulting in a performance decrement. This could account for the discontinuity effect on II category learning. To account for the lack of an effect of discontinuity on RB category learning, the GCM would likely adjust the attention weight parameter (w) so that selective attention was operative. Although one might argue that selective attention constitutes a separate cognitive system in and of itself (and thus does not adhere to the concept of a single system), adjustments in selective attention could effectively wash out any performance differences due to the use of different c parameters in the two discontinuity conditions. This post hoc adjustment of parameter values might account for the performance dissociation we observed in the II and RB conditions.

It is also important to point out that multiple system models other than COVIS have also been developed. In fact, there is a long and rich tradition in category learning that focuses on the distinction between rule application and similarity-based processing (e.g., Allen & Brooks, 1991; Brooks, 1978; Folstein & Van Petten, 2004; Kemler-Nelson, 1984; J. D. Smith & Shapiro, 1989; see also Shanks & St. John, 1994). Several of these other multiple systems models have similar components as COVIS. For example, in the model proposed by E. E. Smith et al. (1998), rule application involves a high working memory load and requires analytic, serial processing of criterial attributes with differential weighting of attributes, much like the hypothesis-testing system in COVIS. Similarly, one system in Erickson and Kruschke's (2002) attention to rules and instances in a unified model instantiates rules in a similar manner as the hypothesis-testing system in COVIS. Where these other multiple-system models differ from COVIS is in their proposal of a second system that is a similarity-based process. In contrast, the second system in COVIS is a procedural-based learning system that is involved in associating category labels (or responses) to regions of perceptual space. As noted above, similarity-based models have a difficult time providing an a priori accounting of the findings from the present study. Even so, like the single-system approaches outlined above, it is likely that multiple process models, such as ATRIUM, could account for the observed data by postulating different parameter settings across conditions.

Another important difference between COVIS and other multiple systems models is the detail in which COVIS proposes the neurobiological underpinnings of the different category learning systems. Although the neurobiology of category learning has been elaborated in the context of other multiple-systems models (e.g., Patalano, Smith, Jonides, & Koeppe, 2001), COVIS has the added advantage of incorporating biological constraints into its architecture that have enabled more detailed predictions of the impact of various experimental manipulations on category learning. For example, COVIS hypothesizes that the II system relies on the many-to-one convergence of cells from the IT cortex onto cells within the tail of the caudate. Such a funneling of information onto the caudate results in the prediction that II category learning will be better when stimuli from a category are perceptually similar (i.e., when there is no discontinuity in the distribution of stimuli within a category), which was supported by the findings in the present study. Other biological constraints have also been incorporated into predictions made by COVIS. For example, based on COVIS, the dopamine-reward signal associated with feedback has to occur in close temporal proximity to the response in order for efficient II category learning to occur. This prediction has been supported by the finding that delayed feedback negatively impacts II category learning (Maddox et al., 2003; Maddox & Ing, 2005). Thus, COVIS attempts to provide computationally and biologically plausible accounts of multiple category learning systems.

Conclusions

To summarize, the studies outlined in this report provide a direct examination of the qualitative properties of the implicit, procedural-based and explicit, hypothesis-testing systems proposed in COVIS by systematically manipulating within-category discontinuity while holding a number of important structural properties fixed. Experiment 1 showed that within-category discontinuity adversely affects II but not RB category learning. Experiments 2 showed that the DS effect observed in Experiment 1 continued to hold when the amount of stimulus space trained was controlled. Experiment 3 showed that increasing the magnitude of the within-category discontinuity does not lead to a significant decline in performance, except perhaps during the final block of training. Although preliminary at this stage, we also found evidence that the distance to the bound provides a reasonable description of the generalization profile associated with the hypothesis-testing system, whereas the distance to the bound plus the distance to the trained response region provides a reasonable description of the generalization profile associated with the procedural-learning system. These data provide useful information regarding the detailed processing characteristics of each category learning system.

References

Allen, S. W., & Brooks, L. R. (1991). Specializing the operation of an explicit rule. *Journal of Experimental Psychology: General, 120,* 3–19.

Ashby, F. G. (1992). Multidimensional models of categorization. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 449–484). 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, and 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., 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.

Ashby, F. G., Waldron, E. M., Lee, W. W., & Berkman, A. (2001). Suboptimality in human categorization and identification. *Journal of Experimental Psychology: General, 130,* 77–96.

Brooks, L. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B. L. Lloyd (Eds.), *Cognition and categorization* (pp. 169–211). Hillsdale, NJ: Erlbaum.

Cohen, N. J., Poldrack, R. A., & Eichenbaum, H. (1997). Memory for items and memory for relations in the procedural/declarative memory framework. *Memory, 5,* 131–178.

Crick, F., & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in Neuroscience, 2,* 2263–2275.

Crick, F., & Koch, C. (1995, May 11). Are we aware of neural activity in primary visual cortex? *Nature, 375,* 121–123.

Crick, F., & Koch, C. (1998). Consciousness and neuroscience. *Cerebral Cortex, 8,* 97–107.

Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General, 127,* 107–140.

Erickson, M. A., & Kruschke, J. K. (2002). Rule-based extrapolation in perceptual categorization. *Psychonomic Bulletin & Review, 9,* 160–168.

Filoteo, J. V., Maddox, W. T., & Davis, J. D. (2001a). A possible role of the striatum in linear and nonlinear category learning: Evidence from patients with Huntington's disease. *Behavioral Neuroscience, 115,* 786–798.

Filoteo, J. V., Maddox, T. W., & Davis, J. D. (2001b). Quantitative modeling of category learning in amnesic patients. *Journal of the International Neuropsychological Society, 7,* 1–19.

Folstein, J. R., & Van Petten, C. V. (2004). Multidimensional rule, unidimensional rule, and similarity strategies in categorization: Event-related brain potential correlates. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 1026–1044.

Kemler-Nelson, D. G. (1984). The effect of intention on what concepts are acquired. *Journal of Verbal Learning and Verbal Behavior, 10,* 734–759.

Keri, S. (2003). The cognitive neuroscience of category learning. *Brain Research Reviews, 43,* 85–109.

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

Maddox, W. T. (2001). Separating perceptual processes from decisional processes in identification and categorization. *Perception & Psychophysics, 63,* 1183–1200.

Maddox, W. T. (2002). Learning and attention in multidimensional identification, and categorization: Separating low-level perceptual processes and high-level decisional processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 99–115.

Maddox, W. T., & Ashby, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception & Psychophysics, 53,* 49–70.

Maddox, W. T., & Ashby, F. G. (2004). Dissociating explicit and procedural-learning based systems of perceptual category learning. *Behavioral Processes, 66,* 309–332.

Maddox, W. T., Ashby, F. G., & Bohil, C. J. (2003). Delayed feedback effects on rule-based and information-integration category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 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. *Memory & Cognition, 32,* 582–591.

Maddox, W. T., Bohil, C. J., & Ing, A. D. (2004). Evidence for a procedural learning-based system in category learning. *Psychonomic Bulletin & Review, 11,* 945–952.

Maddox, W. T., & Filoteo, J. V. (2001). Striatal contributions 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.

Maddox, W. T., & Filoteo, J. V. (in press). Multiple systems of category learning: Neuropsychological tests. In C. Lefebre & H. Cohen (Eds.), *Handbook of categorization in cognitive science*. New York: Elsevier.

Maddox, W. T., Filoteo, J. V., Hejl, K. D., & Ing, A. D. (2004). Category numerosity impacts rule-based but not information-integration category learning: Further evidence for dissociable category learning systems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 227–235.

Maddox, W. T., & Ing, A. D. (2005). Delayed feedback disrupts the procedural-learning system but not the hypothesis testing system in perceptual category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 100–107.

Maddox, W. T., Molis, M. R., & Diehl, R. L. (2002). Generalizing a neuropsychological model of visual categorization to auditory categorization of vowels. *Perception & Psychophysics, 64,* 584–597.

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.

Myers, C. E., Shohamy, D., Gluck, M. A., Grossman, S., Onlaor, S., & Kapur, N. (2003). Dissociating medial temporal and basal ganglia memory systems with a latent learning task. *Neuropsychologia, 41,* 1919–1928.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General, 115,* 39–57.

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

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.

Patalano, A. L., Smith, E. E., Jonides, J., & Koeppe, R. A. (2001). PET evidence for multiple strategies of categorization. *Cognitive, Affective, & Behavioral Neuroscience, 1,* 360–370.

Pickering, A. D. (1997). New approaches to the study of amnesic patients: What can a neurofunctional philosophy and neural network methods offer? *Memory, 5,* 255–300.

Poldrack, R. A., & Packard, M. G. (2003). Competition among multiple memory systems: Converging evidence from animal and human brain studies. *Neuropsychologia, 41,* 245–251.

Poldrack, R. A., Prabhakaran, V., Seger, C. A., & Gabrieli, J. D. E. (1999). Striatal activation during acquisition of a cognitive skill. *Neuropsychology, 13,* 564–574.

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 magnetic resonance imaging. *Proceedings of the National Academy of Sciences, USA, 95,* 747–750.

Rehegr, G., & Brooks, L. R. (1993). Perceptual manifestations of an analytic structure: The priority of holistic individuation. *Journal of Experimental Psychology: General, 122,* 92–114.

Sakamoto, Y., & Love, B. C. (2004). Schematic influences on category learning and recognition memory. *Journal of Experimental Psychology: General, 133,* 534–553.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 501–518.

Schoups, A. A., Vogels, R., & Orban, G. A. (1995). Human perceptual learning in identifying the oblique orientation: Retinotopy, orientation specificity and monocularity. *Journal of Physiology, 483,* 797–810.

Schultz, W. (1992). Activity of dopamine neurons in the behaving primate. *Seminars in Neuroscience, 4,* 129–138.

Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences, 17,* 367–447.

Shiu, L. P., & Pashler, H. (1992). Improvement in line orientation discrimination is retinally local but dependent on cognitive set. *Perception & Psychophysics, 52,* 582–588.

Smith, E. E., Patalano, A., & Jonides, J. (1998). Alternative strategies of categorization. *Cognition, 65,* 167–196.

Smith, J. D., & Shapiro, J. H. (1989). The occurrence of holistic categorization. *Journal of Memory & Language, 28,* 386–399.

Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys and humans. *Psychological Review, 99,* 195–231.

Vogels, R., & Orban, G. A. (1985). The effect of practice on the oblique effect in line orientation judgments. *Vision Research, 25,* 1679–1687.

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 Press.

Willingham, D. B. (1998). A neuropsychological theory of motor skill learning. *Psychological Review, 105,* 558–584.

Willingham, D. B., Wells, L. A., Farrell, J. M., & Stemwedel, M. E. (2000). Implicit motor sequence learning is represented in response location. *Memory & Cognition, 28,* 366–375.

Wilson, C. J. (1995). The contribution of cortical neurons to the firing pattern of striatal spiny neurons. In J. C. Houk, J. L. Davis, & D. G. Beiser (Eds.), *Models of information processing in the basal ganglia* (pp. 29–50). Cambridge, MA: Bradford.

Received September 29, 2004
Revision received December 20, 2004
Accepted January 4, 2005 ■