Memory & Cognition
2006, 34 (2), 387-398

Dual-task interference in perceptual category learning

DAGMAR ZEITHAMOVA and W. TODD MADDOX
University of Texas, Austin, Texas

The effect of a working-memory–demanding dual task on perceptual category learning was investigated. In Experiment 1, participants learned unidimensional rule-based or information integration category structures. In Experiment 2, participants learned a conjunctive rule-based category structure. In Experiment 1, unidimensional rule-based category learning was disrupted more by the dual working memory task than was information integration category learning. In addition, rule-based category learning differed qualitatively from information integration category learning in yielding a bimodal, rather than a normal, distribution of scores. Experiment 2 showed that rule-based learning can be disrupted by a dual working memory task even when both dimensions are relevant for optimal categorization. The results support the notion of at least two systems of category learning: a hypothesis-testing system that seeks verbalizable rules and relies on working memory and selective attention, and an implicit system that is procedural-learning based and is essentially automatic.

Humans live in a world of categories, rather than unique instances. Categories divide the world into meaningful pieces. Humans categorize in order to reach cognitive economy of memory, to communicate and understand, and to explain and predict properties and actions of new stimuli on the basis of older experiences. Because categorization is essential for higher level cognition, much attention in cognitive research has been paid to category learning (see, e.g., Ashby & Maddox, 2005; Estes, 1994; Kruschke, 1992; Love, Medin, & Gureckis, 2004; Medin & Schaffer, 1978; Nosofsky, 1986).

A large and growing body of research suggests that participants have available multiple processing modes that can be used during categorization. Well established in the literature is a distinction between categorization according to a rule and categorization based on overall similarity (Allen & Brooks, 1991; Erickson & Kruschke, 1998; Folstein & Van Petten, 2004; Kemler Nelson, 1984; Nosofsky, Palmeri, & McKinley, 1994; Regehr & Brooks, 1993). Building upon this work on multiple processing modes is a recent interest in understanding the neurobiological underpinnings of category learning and examining the possibility of multiple systems of category learning (Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Poldrack, Prabhakaran, Seger, & Gabrieli, 1999; Reber, Stark, & Squire, 1998; E. E. Smith, Patalano, & Jonides, 1998; for reviews, see Kéri, 2003, and Maddox & Ashby, 2004). Relevant to this work are studies of multiple memory systems (Poldrack & Packard, 2003; Schacter & Tulving, 1994; Squire, 1992) and multiple reasoning systems (Sloman, 1996).

One multiple systems model of perceptual category learning, and the only one that specifies the underlying neurobiology, is the competition between verbal and implicit systems (COVIS) model proposed by Ashby et al. (1998; Ashby & Waldron, 2000). COVIS postulates two systems that compete throughout learning: an explicit hypothesis-testing system, which uses logical reasoning and depends on working memory and executive attention, and an implicit procedural-learning–based system. (Relations between COVIS and the multiple process [rule vs. overall similarity] approach are reserved for the General Discussion section.)

At the implementation level, the explicit hypothesis-testing and the implicit procedural-learning systems have distinct but partially overlapping neurobiological underpinnings. The key neural structures for the hypothesis-testing system are the prefrontal cortex, the anterior cingulate, and the head of the caudate nucleus. The key neural structures for the procedural-learning system are the inferotemporal cortex and the tail of the caudate nucleus. A dopamine-mediated reward signal from the substantia nigra is critical for learning in this system. Both systems attempt to acquire and solve every categorization task encountered. However, the relative weight of each system in the category judgment depends on the relative success of each system in category learning, which, in turn, depends on the type of category structure to be acquired.

The hypothesis-testing system searches for and applies explicit rules that are typically easy to verbalize (the hypothesis-testing system is often called *verbal*, although such a description may not be appropriate in all cases and is not appropriate for nonhumans). One example is

This research was supported in part by National Institutes of Health Grant R01 MH59196 to W.T.M. We thank Evan Heit, Sarah Queller, and two anonymous reviewers for useful comments on an earlier version of the manuscript. We are especially indebted to Larry Cormack for his guidance on the statistical analyses. Correspondence concerning this article should be addressed to D. Zeithamova or W. T. Maddox, Department of Psychology, University of Texas, 1 University Station A8000, Austin, TX 78712 (e-mail: zeithamova@mail.utexas.edu or maddox@psy.utexas.edu).

Copyright 2006 Psychonomic Society, Inc.

a unidimensional (UD) rule that allows categorization on the basis of a criterion along a single separable stimulus dimension. For example, if the stimulus is a Gabor patch (a sinusoidal luminance grating windowed by a Gaussian envelope) that varies across trials in spatial frequency and spatial orientation, one UD rule may be "respond A if the spatial frequency is low and respond B if the spatial frequency is high." Category structures for which the optimal rule is likely to be explicitly verbalized as a categorization strategy by a responder are called *rule based*.

The implicit procedural-learning–based system learns to associate a category response with a region of perceptual space without deriving any explicit rule. Stimuli are represented perceptually in higher order visual areas, such as the inferotemporal (IT) cortex. It is well established that there is a many-to-one convergence of IT cells onto each cell in the tail of the caudate (Wilson, 1995). The striatum learns to associate subregions in the perceptual space with category assignments (Packard & Knowlton, 2002). Category structures acquired by the implicit system may be very complex (see, e.g., Ashby & Maddox, 1992, 2005). The categorization rule typically combines two or more stimulus dimensions expressed in different units. Such a rule may be "respond A if the spatial frequency is greater than the spatial orientation; otherwise, respond B." A person that categorizes according to such a rule is not likely to be able to verbalize it in this form, because it compares values expressed in different units and, thus, is not easy to comprehend logically. We say that the optimal rule is not verbalizable. Category structures in which the optimal rule is of this form are called *information integration*.[1]

As a consequence of the proposed underlying neurobiology, the nature of the feedback and response mapping should affect the implicit procedural-learning–based system but not the hypothesis testing system and, thus, should affect information-integration (II) but not rule-based (RB) category learning. Indeed, several studies have reported results consistent with this prediction (Ashby, Queller, & Berretty, 1999; Maddox, Ashby, & Bohil, 2003; Maddox, Bohil, & Ing, 2004; Maddox & Ing, 2005). On the other hand, working memory load should affect the explicit hypothesis-testing system but not the implicit procedural-learning–based system and, thus, should affect RB but not II category learning. Waldron and Ashby (2001) found support for this prediction with a working memory–demanding dual task. Specifically, they found a large dual-task interference on UD rule-based category learning but only a small dual-task interference on (multidimensional) II category learning when a small number of highly discriminable binary-value dimension stimuli were used.

The goal of the present research was twofold. First and foremost, we wished to test the generality of Waldron and Ashby's (2001) results when applied to a unidimensional rule-based and a (two-dimensional) information-integration category-learning task, using a large number of perceptually similar continuous-value dimension stimuli—that is, Gabor patches that varied across trials in spatial frequency and spatial orientation. Stimuli of this sort have been used extensively to study category learning (see Maddox & Ashby, 2004, for a review). Second, we explored the dual-task interference phenomenon in more detail and provided a critical test of Nosofsky and Kruschke's (2002; see also Ashby & Ell, 2002) single-system explanation of the original Waldron and Ashby results by examining two-dimensional, conjunctive rule-based category learning in a dual-task setting.

In the next sections, we will briefly review a number of empirical studies in which a priori predictions from COVIS have been tested and will provide a more detailed review of Waldron and Ashby (2001). Then we will present the results from two experiments. We will conclude with some general comments that will include a discussion of alternative approaches to categorization and how they may or may not account for the experimental results.

Brief Review of COVIS and the Dissociation Studies

COVIS assumes that, regardless of the nature of the category structures (i.e., rule based or information integration), both the hypothesis-testing system and the procedural-learning system attempt to learn. The two systems then compete to determine the response. COVIS assumes an initial bias for the explicit system. If an explicit rule exists that yields good performance, the hypothesis-testing system is likely to be successful and dominate the implicit system. If no such explicit rule exists, the hypothesis-testing system will continuously fail to discover the correct rule, and the implicit system will eventually dominate. To study the properties of each system, different category structures are therefore used: a rule-based category structure for studying the explicit system and an information-integration category structure for studying the implicit system.

Evidence that multiple processes are involved in category learning has come from a number of sources (see Maddox & Ashby, 2004, for a review). Several experiments have shown that II, but not RB, category learning may be disrupted by feedback or instruction manipulation. First, category learning is qualitatively different with trial-by-trial feedback than it is without feedback. Without supervision, people typically use simple unidimensional rules (Ashby et al., 1999), whereas with trial-by-trial feedback they are able to learn complex nonlinear decision bounds that are difficult to describe verbally (Ashby & Maddox, 1992). Second, when the feedback is delayed, processing in the implicit system is affected, so that learning of an II category structure may be impossible (Maddox, Ashby, & Bohil, 2003; Maddox & Ing, 2005). Third, in most experiments, consistent stimulus–response mappings have been used. For example, when a stimulus is presented, the participant is asked to press Button A with the left hand and Button B with the right hand (A–B training). Maddox, Bohil, and Ing (2004; see also Ashby, Ell, & Waldron, 2003) used a variable stimulus–response mapping. The participants were asked to press either a *yes* button or a *no* button to a stimulus in response to a question "Is this an A?" or "Is this a B?" (*yes–no* training). As is predicted by COVIS, *yes–no* training impaired II category learning, as compared with A–B training, but had no effect on RB category learning.

The previous studies show that the implicit system differs from the hypothesis-testing system in that it requires immediate feedback and a consistent stimulus–response mapping. When feedback or a consistent stimulus–response mapping is not provided, learning by the implicit system is adversely affected. COVIS postulates that this is due to the fact that learning in the implicit system is dopamine mediated. Positive feedback induces dopamine to be released from the substantia nigra into the tail of the caudate nucleus, strengthening recently activated synapses. When the feedback is not provided or is substantially delayed, synaptic activation within the striatum decays, and learning does not occur (Arbuthnott, Ingham, & Wickens, 2000; Kerr & Wickens, 2001).

One may argue that the disruption of information-integration category learning, but not of rule-based learning, is due to differences in the complexity and, therefore, to the difficulty of simple (e.g., one-dimensional) verbalizable rules in the RB condition versus complex (multidimensional) nonverbalizable integration rules in the II condition.[2] To provide evidence for the existence of two alternative systems, double dissociation should be demonstrated. Recent studies have introduced manipulations that impair RB category learning but not II category learning (Maddox, Ashby, Ing, & Pickering, 2004; Maddox, Filoteo, Hejl, & Ing, 2004). Waldron and Ashby (2001) provided empirical evidence of that kind by introducing a second task that was to be performed concurrently with category learning.

Review of Waldron and Ashby (2001)

Recall that COVIS postulates that the hypothesis-testing system relies on working memory and selective attention to solve rule-based category tasks, whereas learning in the procedural-learning system is essentially automatic. Waldron and Ashby (2001) provided an empirical test of this prediction by comparing rule-based and information-integration category learning under dual-task conditions with that in a single-task control. They chose a numerical analogue of the Stroop task (for a detailed review of the Stroop task, see MacLeod, 1991) to serve as a dual task. The Stroop task is known to require working memory and selective attention and to strongly activate the anterior cingulate and prefrontal cortex (Bench et al., 1993), neural structures associated with the explicit hypothesis-testing system, but not with the implicit procedural-learning system proposed in COVIS.

Waldron and Ashby (2001) had participants learn to categorize colored geometric figures presented on a colored background that varied on four binary dimensions. In the unidimensional (UD) rule-based condition, one dimension was relevant, and the remaining three were irrelevant. In the information-integration condition, information from three dimensions had to be integrated, and one dimension could be ignored (see Waldron & Ashby, 2001, for details). Under control conditions, the participant simply categorized each stimulus on every trial. In the dual-task conditions, the participant had to perform a numerical analogue of the Stroop task during each trial of categorization. The Stroop task stimulus was presented simultaneously with the categorization stimulus for 200 msec. The Stroop stimulus was then masked, and the categorization stimulus remained on the screen until the participant categorized it. After categorization feedback, the participant was to respond to the Stroop stimulus that he or she had seen at the beginning of the trial. Therefore, the participant was required to hold a representation of the Stroop stimulus in working memory during the process of categorization. Performance in the Stroop task was emphasized over that in the categorization task.

Waldron and Ashby (2001) found that the dual task produced greater interference for the UD rule-based task than for the II task. These findings support the COVIS prediction that a dual working memory task impairs rule-based but not information-integration category learning and argues against the *complexity* arguments offered against multiple-systems theories.

EXPERIMENT 1

The main aim of Experiment 1 was to test the generalizability of Waldron and Ashby's (2001) results in an experiment using a large number of unique continuous-valued dimension stimuli. The stimuli were Gabor patches that varied across trials in spatial frequency and spatial orientation. UD rule-based and II category learning were examined under control and dual Stroop conditions. Scatterplots of the stimuli used in the UD rule-based and II category-learning conditions are shown in Figure 1, along with the optimal decision bound. Each point in the scatterplot denotes the spatial frequency and spatial orientation of a single stimulus. In the UD rule-based condition, spatial frequency was relevant and spatial orientation was irrelevant, and the optimal rule required the participants to respond A when the spatial frequency was low and to respond B when the spatial frequency was high. Both dimensions were relevant in the II condition. The optimal rule required the participants to respond A when the difference of the value on spatial frequency dimension and the value on the spatial orientation dimension was low and to respond B when the difference of the values on the two dimensions was high. Such a rule is not easy to comprehend logically, because it compares values in different units. The category discriminabilities (d') were 4.5 for the UD and 10.3 for the II category structures.[3]

Method

Participants. One hundred seventy students at the University of Texas at Austin participated in the experiment in partial fulfillment of a class requirement or for pay. All the observers were tested for 20/20 vision, and no observer completed more than one experimental condition. Each participant completed one of four experimental conditions: UD rule-based control (UDC), UD rule-based dual Stroop (UDS), II control (IIC), and II dual Stroop (IIS).

Stimuli and Apparatus. The categorization stimuli were Gabor patches that varied across trials in spatial frequency and spatial orientation. The experiment used the randomization technique introduced by Ashby and Gott (1988). Forty Category A and 40 Category B

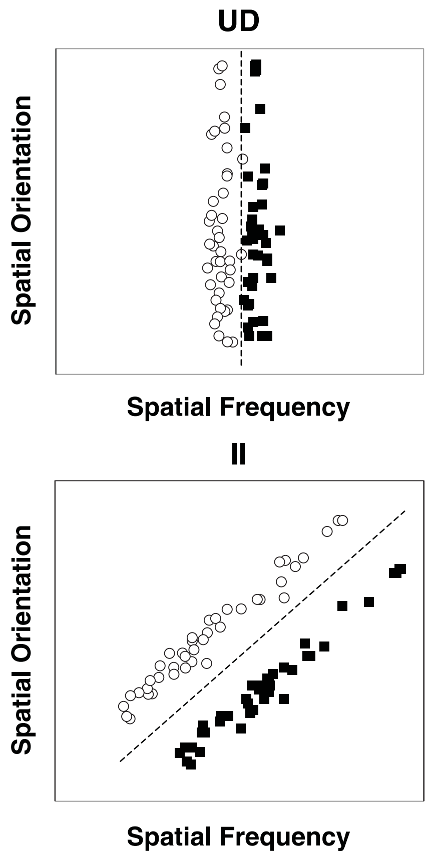

Figure 1. Unidimensional rule-based (UD; upper panel) and information integration (II; lower panel) category structures used in Experiment 1. Open circles denote Category A; filled squares denote Category B. The dashed lines represent the optimal decision bounds.

stimuli from the UD categories were generated by sampling randomly from two bivariate normal distributions. Each random sample (x_1, x_2) was converted to a stimulus by deriving the frequency [$f = .25 + (x_1/50)$] and the orientation [$o = x_2(\pi/500)$]. The stimuli for the II categories were generated by rotating the 80 rule-based stimuli clockwise by 45º around the center of the spatial-frequency–spatial-orientation space and then shifting the spatial frequency and spatial orientation by an amount that resulted in the appropriate d', using linear algebra method. The category distribution parameters for both structures are listed in Table 1.

Each Gabor patch was generated using MATLAB (MathWorks, Natick, MA) routines from the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). The size of each stimulus was 200 × 200 pixels, covering about 4º of visual angle, and was centered on the computer screen. Following Waldron and Ashby (2001), the Stroop task stimuli used in the dual task were two whole numbers sampled without replacement from the range of 2–8. On 85% of the trials, the numerically larger number was physically smaller (95 pixels tall vs. 180 pixels tall). The stimuli were presented on a gray background.

Procedure. Each condition consisted of five 80-trial blocks. In the control conditions, the participants were told that there were two categories of stimuli and that these were to be learned via corrective feedback. On each trial, a categorization stimulus was presented on the screen and remained there until the participant categorized the stimulus into one of the two categories by pressing either the "Z" button with the left hand or the "?" button with the right hand on the computer keyboard. Corrective feedback was then provided for 1,000 msec, followed by a 1,000-msec delay and a 1,000-msec intertrial interval.

In the dual Stroop task conditions, a categorization stimulus was presented centered on the screen, with the Stroop task stimuli presented concurrently to the left and right of the categorization stimulus for 200 msec, followed by a rectangular white mask for another 200 msec. The categorization stimulus remained on the screen until the participant categorized it into one of the two categories by pressing "Z" or "?" on the keyboard. The categorization response was followed by 1,000 msec of corrective feedback and a 1,000-msec blank screen delay. Then either the word "value" or the word "size" appeared on the screen. The participant then indicated on which side the number with the larger value or larger size was presented. The response was followed by 1,000 msec of corrective feedback and a 1,000-msec intertrial interval. The timing of each trial was identical to that used in Waldron and Ashby (2001).

Results

Stroop task performance. Fifty and 45 participants completed the UD Stroop task and the II Stroop task conditions, respectively. The overall proportion correct for the Stroop task was .84. There was no difference in Stroop task accuracy between the UDS ($M = .831$, $SE = .022$) and the IIS ($M = .849$, $SE = .019$) groups [$t(93) = 0.582$, $p = .562$], suggesting that the effort and cognitive resources allocated to the Stroop task were equal in both groups. Fifteen participants in the UDS condition and 13 participants in the IIS condition did not reach the 80% required accuracy minimum on the Stroop task, and their data were excluded from further analyses.

Category-learning performance. For each participant, we computed the proportion correct for each block and the overall proportion correct. We began by examining the shape of the II and UD overall score distributions collapsed across control and Stroop task conditions. The distribution of overall scores for the UD category structure deviated significantly from normality [Kolmogorov–Smirnov (KS) $D(76) = .212$, $p = .002$], whereas the distribution of overall scores for the II category structure did not [KS $D(66) = .097$, $p = .557$]. This pattern held in each block as well. To illustrate, histograms of the overall accuracy distributions for the UDC and IIC conditions are shown in Figure 2. Whereas the II distribution is unimodal and close to normal, the UD distribution is bimodal, with one modus close to the chance level of accuracy (.5) and another at a much higher level of performance.

Table 1
Category Distribution Parameters for the Unidimensional and Information Integration Category Structures Used in Experiment 1

Category Structure	μ_x	μ_y	σ_x^2	σ_y^2	cov_{xy}
Unidimensional					
Category A	280	125	75	9,000	0
Category B	320	125	75	9,000	0
Information integration					
Category A	268	157	4,538	4,538	4,351
Category B	332	93	4,538	4,538	4,351

Figure 3 presents the mean accuracy scores (proportions correct) for each group. The experimental hypothesis predicts that the dual task will have a bigger impact on UD rule-based than on II category learning. Assessing the effect of the dual task on underlying distributions with such different shapes provided a challenge, since the ANOVA, like most standard statistical methods, assumes normal distributions with equal variance. We used a *bootstrapping*[4] procedure to compare the drop in mean performance across the control and Stroop conditions, to determine whether this drop was larger for the UD than for the II category structures. Specifically, the test was designed to verify that the 95% confidence interval for the difference in performance drops [(UDC − UDS) − (IIC − IIS)] was reliably bigger than zero. We found that the 15.2% drop in overall performance in the UD rule-based category learning dual task, relative to the control task, was reliably bigger than the 6.1% drop in the overall performance observed in the II dual task, relative to the control task.[5]

We found a substantial drop in mean performance for the Stroop task condition, relative to the control condition, for UD rule-based category learning but a smaller drop for II category learning. We were interested in examining the effect that the Stroop task had on the distribution of scores—whether it was a shift in the peaks of the two modes or a change in the shape. To address this issue, we computed the index of asymmetry (skewness) of the overall accuracy distribution for each group. The results

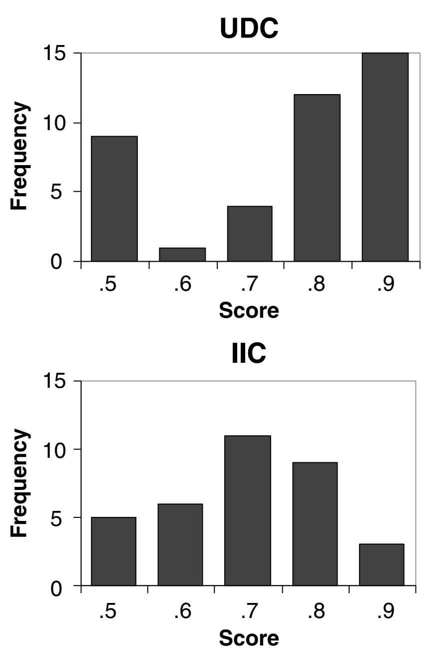

Figure 2. Distribution of the overall scores (proportions correct) for unidimensional control (UDC; upper panel) and information-integration control (IIC; lower panel) groups. Numbers along the abscissa denote the midpoints of the bins, except that the .5 bin includes all participants below .55. No participant reached an accuracy above .95.

are shown in Figure 4. The UD rule-based control group is significantly skewed negative, suggesting that the majority of the participants are in the high-accuracy modus, with a few participants doing poorly on the task. Under the dual Stroop task condition, the distribution of scores became significantly skewed positive, suggesting that the majority of the participants are in the chance modus, with a few participants learning the task well. The distribution of scores in the II category-learning condition was symmetrical for both control and dual Stroop task groups, suggesting that the distribution shifted slightly to the left but did not substantially change the shape.

Discussion

Experiment 1 yielded several interesting results. First and foremost, including the dual Stoop task had a large effect on UD rule-based, but not II, category learning. This finding replicates that observed in Waldron and Ashby (2001) and extends it to a situation in which a large number of normally distributed continuous-valued dimension stimuli were used, providing further support for COVIS. Importantly, this pattern holds even though performance was best in the UD control condition and worst in the UD Stroop condition, ruling out a complexity explanation of the results. Second, the results suggested that UD rule-based category learning (under control and dual-task conditions) differs qualitatively from II category learning. Specifically, whereas the distribution of scores observed in the II conditions was unimodal and close to normal, the distribution of scores observed in the rule-based conditions was bimodal, suggesting an all-or-none character to category acquisition (for a similar result, see J. D. Smith, Minda, & Washburn, 2004).

The qualitative difference in the performance profiles across UD rule-based and II conditions is predicted by COVIS. Rule-based category learning involves the explicit system. In this system, different rules are tested and are either accepted or rejected. This system relies on working memory and executive attention processes (Ashby et al., 1998). When the correct categorization rule is identified, categorization accuracy improves dramatically. When incorrect rules are applied, categorization accuracy is often near chance, resulting in a bimodal performance distribution. Information-integration category learning involves an implicit procedural learning-based system. The implicit system learns gradually, incrementally, and automatically, leading to a normal unimodal distribution of scores.

EXPERIMENT 2

One potential weakness of Experiment 1 and Waldron and Ashby (2001) is that the number of dimensions relevant for optimal categorization differs across conditions. Indeed, as Nosofsky and Kruschke (2002) pointed out, the results of Waldron and Ashby are consistent with a single-system approach that operates on a single exemplar representation with normal (control) or limited (Stroop) selective attention. To elaborate, Nosofsky and Kruschke argued that the Stroop task will disrupt ALCOVE's (Kruschke,

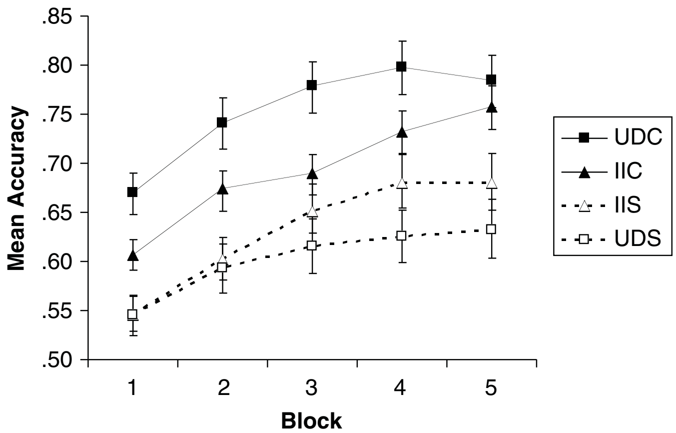

Figure 3. Mean categorization block accuracies (proportion correct) for each group in Experiment 1. The control groups are denoted with solid lines and filled marks; dual Stroop task groups are denoted with broken lines and open marks. Unidimensional rule-based category observers are marked with squares; information integration category observers are marked with triangles. Error bars denote bootstrapped 68% confidence intervals (equivalent to standard errors of the mean). UDC, unidimensional rule-based control; IIC, information integration control; IIS, information integration dual Stroop; UDS, unidimensional rule-based Stroop.

1992) selective attention learning parameter. Failure to attend to the single relevant dimension in the UD rule-based task will cause strong interference, because attending to the three irrelevant dimensions will waste vast amounts of processing capacity. In the complex, II category structure, three dimensions are relevant and only one irrelevant, and thus a wide variety of attentional weights will lead to reasonable performance and only a little processing capacity is wasted on the one irrelevant dimension.

Ashby and Ell (2002) demonstrated that ALCOVE, although able to account for the qualitative pattern found in Waldron and Ashby (2001), could not account for the quantitative pattern. ALCOVE either underestimates the observed difference between UD control and II control category learning, or assumes no attention learning in the Stroop task condition (leading the participants in the UD Stroop task condition to be unaware that a single dimension was relevant).

Although a number of previous studies provided sufficient evidence for the existence of at least two modes of category learning and resulting category representations (Kemler Nelson, 1984; Seger & Cincotta, 2002; Shanks & St. John, 1994; E. E. Smith et al., 1998; J. D. Smith & Shapiro, 1989), we decided to investigate the notion of Nosofsky and Kruschke (2002), using a conjunctive rule-based category structure in which both dimensions were relevant for optimal categorization (see the Method section for details). Nosofsky and Kruschke would predict no or very little dual-task interference, because no dimension is irrelevant in this task and a wide range of attentional weights provides a high level of performance. Also, because the attention learning mechanism is disrupted, attention will be spread over both dimensions throughout the course of learning. COVIS, however, would predict stronger dual-task interference, because the conjunctive task, unlike the II task, is solved under control condition by the hypothesis-testing system. Under the dual Stroop task condition, use of conjunctive rules (attending to both dimensions) is less likely, and use of suboptimal UD rules (selective attention to one dimension while ignoring the other) is more likely, because conjunctive rules require more working memory capacity than do UD rules. The aim of Experiment 2 was to provide a test of these two alternatives.

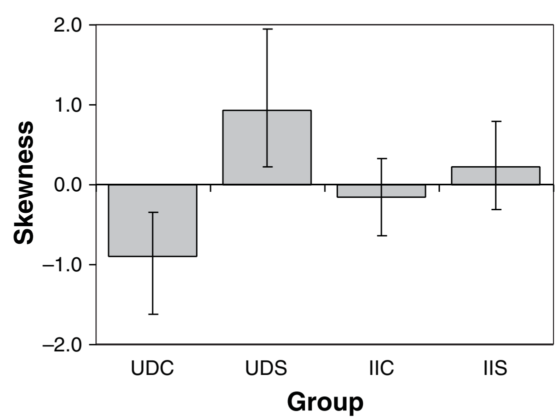

Figure 4. Index of asymmetry (skewness) of the overall accuracy scores for each group. Error bars denote 95% confidence intervals. UDC, unidimensional rule-based control; UDS, unidimensional rule-based dual Stroop; IIC, information integration control; IIS, information integration dual Stroop.

Method

Participants. Sixty students at the University of Texas at Austin participated in the experiment in partial fulfillment of a class requirement or for pay. Thirty completed the conjunctive control (CJC) condition, and 30 completed the conjunctive Stroop task (CJS) condition. All the participants were tested for 20/20 vision.

Stimuli and Apparatus. The stimuli, stimulus generation procedure, and apparatus were identical to those used in Experiment 1. The only difference was in the nature of the category structures. Eighty stimuli were generated by sampling randomly from four bivariate normal distributions. Three were assigned to Category A and one to Category B. The four distribution parameters and the number of stimuli generated from each are displayed in Table 2. A scatterplot of the stimuli and the optimal rule is presented in Figure 5. The optimal rule required the participants to respond B when the spatial frequency was high *and* the orientation was steep and to respond A otherwise. Note that both dimensions are relevant for correct categorization. The number of stimuli generated from each distribution was chosen to equate the number of stimuli in both categories and in an attempt to reduce the usage of UD rules to solve the task.

Procedure. The procedure was identical to that in Experiment 1, except that there were four, rather than five, blocks of 80 trials. The participants were told that perfect performance was possible and that they should certainly achieve above 80% correct before the end of training.

Results

Stroop task performance. Mean Stroop task accuracy was .862 ($SEM = .026$). Five participants did not reach required 80% Stroop task accuracy, and their data were excluded from further analyses.

Categorization task performance. We first inspected the distribution of scores in order to compare them with those in Experiment 1 (data not shown). Although there was a tendency toward bimodality, the KS test did not show a significant deviation from normality for either condition or when collapsed across control and dual-task conditions in any block [KS $D(55) = .153$, $p = .153$, for the collapsed data and overall score distribution].

Mean categorization accuracy for each block of trials is shown in Figure 6. Overall categorization accuracy was 70.2% in the control group and 60.6% in the Stroop task group. Thus, the Stroop task produced a 9.6% drop in categorization accuracy that was significant [bootstrapped 95% confidence interval for the drop (CJC − CJS)].

For comparison with Experiment 1, we computed asymmetry (skewness) of the distributions in the control and dual-task groups. The distribution of scores for the control group was slightly skewed negative (skew = −0.345, nonsignificant; bootstrapped 95% confidence interval, [−0.916; 0.197]), suggesting that there were about equal numbers of participants doing well as doing poorly on the task. In the dual-task group, the distribution of scores was significantly skewed positive (skew = 0.860; bootstrapped 95% confidence interval, [0.118; 1.614]), suggesting that the majority of participants did poorly on the task.

Table 2
Four Distribution Parameters and the Number of Stimuli Derived From Each Distribution for the Conjunctive Category Structure Used in Experiment 2

Distribution	μ_x	μ_y	σ^2	cov_{xy}	N
A_1	283	98	75	0	8
A_2	317	98	75	0	16
A_3	283	152	75	0	16
B	317	152	75	0	40

Note—Stimuli from the A_1, A_2, and A_3 distributions were all members of Category A.

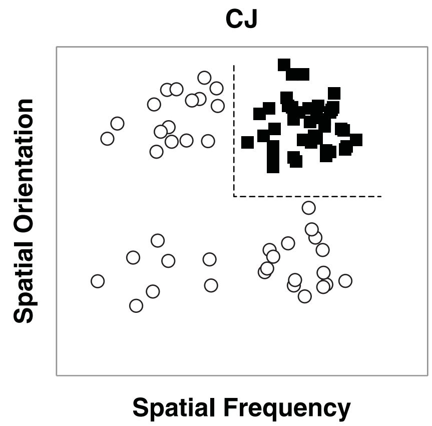

Figure 5. Conjunctive (CJ) category structure used in Experiment 2. Open circles denote Category A; filled squares denote Category B. The dashed line represents the optimal decision bound.

To examine response strategies, we fit conjunctive and UD decision bounds to each participant's responses in the last block (the details of these models can be found in numerous articles; e.g., Maddox et al., 2003). We found that the proportion of participants who used a rule employing both dimensions (either a conjunctive rule-based strategy or a decision based on linear combination of the two dimensions' values) dropped from 77% in the control condition to 44% in the dual-task condition and that the proportion of participants using a UD rule for categorization increased from 7% in the control condition to 17% in the dual-task condition.[6]

Discussion

The results from Experiment 2 supported the COVIS prediction that categorization based on the combination of both dimensions is less likely and using UD strategies more likely under a dual-task condition than under a control condition—a prediction that is opposite of that in Nosofsky and Kruschke's (2002) account of Waldron and Ashby's (2001) results. To compare the observed drop in performance in the conjunctive rule-based condition with those in Experiment 1, we computed the average performance across the first four blocks of trials (since Experiment 2 included four, not five, blocks of trials) in each condition. The results are displayed in Figure 7. Figure 7 suggests that the impact of the dual Stroop task was indeed larger on conjunctive rule-based than on II category learning, a result that is predicted by COVIS and is opposite that predicted by Nosofsky and Kruschke.

Both the shape of the score distributions and the skewness values from the conjunctive task were intermediate

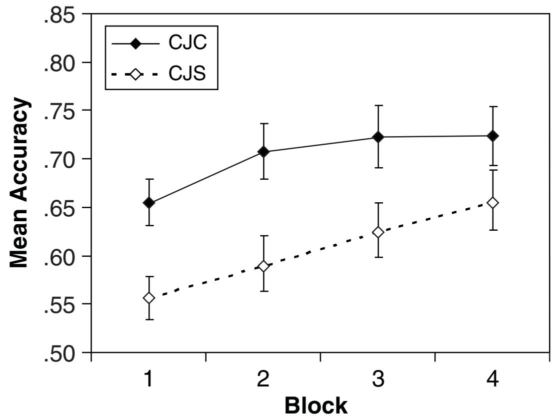

Figure 6. Mean block accuracies in the Experiment 2 conjunctive (CJ) category-learning task. The control group (CJC) is denoted with a solid line and filled diamonds; the dual Stroop task group (CJS) is denoted with a broken line and open diamonds. Error bars denote bootstrapped 68% confidence intervals (equivalent to standard errors of the mean).

between those found in Experiment 1, suggesting that a number of strategies may be used to resolve the conjunctive task and each of these strategies may be influenced by the dual task differently. Detailed discussion of the dual-task interference for the three category structures examined in this article is reserved for the General Discussion section.

GENERAL DISCUSSION

The theoretical framework that gave rise to the experiments reported in this article was the COVIS model of category learning. COVIS builds upon a body of research that has identified alternative strategies of category learning and has extended it by identifying the underlying neurostructures. This line of research contrasts with theories in which a single system of category learning is assumed. In this discussion, we will first focus on the COVIS account of the observed pattern of data, then will review alternative multiple-system approaches to categorization, and finally will ask whether a single-system approach to categorization may be sufficient to account for the results observed here and elsewhere.

COVIS

COVIS assumes the existence of at least two category-learning systems: an initially favored hypothesis-testing system that seeks explicit rules and relies on working memory and selective attention and an implicit system that is procedural-learning based and essentially automatic. Two predictions result from this notion. First, category learning by the hypothesis-testing system when a simple correct categorization rule exists that yields nearly perfect performance (such as our UD rule-based category structure) should have an all-or-none character, whereas learning by the procedural-learning–based system is gradual and incremental. Second, a dual task requiring limited cognitive resources, working memory, and selective attention should impair the hypothesis-testing system, but not the II system.

The three category structures used in the two experiments reported in this article differed in their level of attainability by the two systems. UD rule-based category learning resulted in a bimodal, all-or-none distribution of scores and was affected most by the dual task, suggesting a strong reliance on the hypothesis-testing system in solving the task. The UD category structure is indeed well acquired by the hypothesis-testing system, because a simple rule can yield almost perfect accuracy. However, if the correct rule is not found, alternative rules yield performance at chance levels of accuracy. The implicit procedural-learning–based system may exhibit poor acquisition of such a structure because the variance along the relevant dimension is small, whereas the variance along the irrelevant dimension is high. The high convergence of connections from the IT cortex to the tail of the caudate nucleus may cause the same striatal units to be activated by stimuli coming from different categories but sharing similar values on the irrelevant dimension, making the stimulus–response mapping within the caudate difficult. Thus, although this task was easiest under the control condition, the need to find the one correct rule by the hypothesis-testing system with limited resources and unreliable responses from the implicit system made it most difficult under the dual-task condition.

A neuroimaging study by Bench et al. (1993) showed that the anterior cingulate and frontal cortex are structures strongly activated while a Stroop task is performed. The fact that the presence of the Stroop task affected UD rule-based category learning the most provides an empirical test of the COVIS proposition that the explicit hypothesis-testing system, but not the implicit system, relies on working memory and attentional processes and on these same underlying brain structures (i.e., the anterior cingulate and the frontal cortex). The dual Stroop task may influence

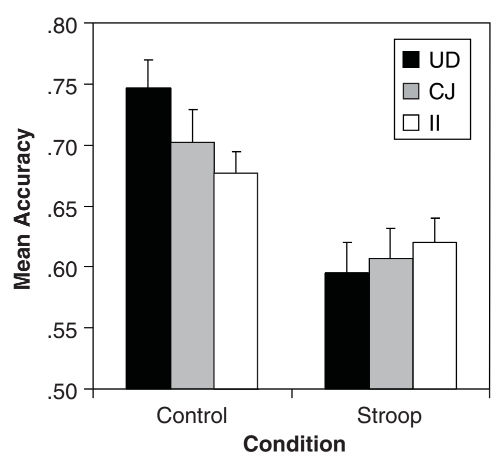

Figure 7. Comparison of mean categorization accuracies across the first four blocks in Experiments 1 and 2. UD, unidimensional; CJ, conjunctive; II, information integration.

several stages of the hypothesis-testing system. It may make selective attention to the relevant dimension more difficult to achieve, because selective attention is needed for the Stroop task. Its working memory load may make it harder to remember the current rule to be tested and which rules did not work previously. It may impair the ability to detect conflict and evaluate performance and to select and switch to a new rule (anterior cingulate functions).

II category learning was most difficult in the control condition but exhibited the smallest decrement in performance in the dual-task condition, becoming the easiest. The II category structure is better acquired by the implicit system than by the hypothesis-testing system. COVIS predicts that after trying unsuccessfully all salient rules, the weight on the hypothesis-testing system decreases, and the responses are dominated more often by the implicit system, which learns the stimulus–response mapping gradually and incrementally, yielding a normal distribution of scores. The stimulus–response mapping in the caudate is facilitated by the larger distance of the stimuli from the two categories in the stimulus space ($d' = 10.3$, as compared with 4.5 for the UD structure). The dual Stroop task may influence II category learning in two opposite ways: It may hurt performance because it reduces the cognitive resources needed for initially biased hypothesis testing and/or slows down the shift of the overall system in favor of the implicit system, or it may facilitate performance because the limited capacity hypothesis-testing system becomes less initially biased and/or the overall system shifts more quickly toward the implicit system. We found a slight performance drop in the dual-task condition, as compared with the control condition, suggesting that the first type of influence or a combination of both types is more likely. Because the implicit system itself is unaffected by the dual task, once the weight of that system increases sufficiently, accuracy will be expected to be essentially the same under the control as under the dual-task condition.

The conjunctive rule-based category structure used in Experiment 2 has properties intermediate between those structures used in Experiment 1, yielding intermediate difficulty and performance drop under the two conditions. The optimal conjunctive rule yields the highest accuracy (100% possible); however, UD rules on either dimension can provide an accuracy of up to 80%, and II strategies may be successful as well, due to the relatively high separability of the four underlying distributions. The proportion of the participants using a combination of both dimensions (spreading attentional weights in terms of ALCOVE [Kruschke, 1992]) for categorization decision decreased and the proportion of the participants using values on a single dimension increased under the dual-task condition, contrary to the ALCOVE prediction (Nosofsky & Kruschke, 2002) and in agreement with the COVIS prediction. Experiment 2's results also argue for dual Stroop task interference on performance evaluation and rule switching, in addition to working memory load, because the participants were more likely to stick with suboptimal UD rules despite corrective feedback.

Rule Versus Similarity

There has been a long tradition in cognitive psychology research of focusing on a distinction between perceptual categorization that is based on rule application and that based on overall similarity to previously seen instances (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). Rule versus similarity distinction provides an alternative theory of multiple strategies of categorization. Rule application involves a high working memory load and requires analytic, serial processing of criterial attributes with differential weighting of attributes, whereas similarity-based processing involves a low working memory load and holistic, parallel, automatic processing with equal weighting of attributes (E. E. Smith et al., 1998).

The theories assuming alternative strategies of categorization involving qualitatively distinct processes of rule application and similarity judgment are strikingly similar to the computational-level description of the COVIS model. Rule application is assumed to involve working memory and selective attention to criterial attributes, like the explicit hypothesis-testing system. The similarity judgment is an automatic, holistic process that does not have a high working memory load, comparable to the implicit procedural-learning–based system. A dual task reduces the likelihood of using analytical rules in categorization (J. D. Smith & Shapiro, 1989).

Rule versus similarity distinction theories would, therefore, predict a pattern of results similar to that obtained here, because the holistic strategies promoted over the analytic strategies under the dual task have different relative utilities for correct categorization. However, direct application of these theories to the results of our two experiments is complicated by the dissimilarity of the experimental paradigms. In experiments illustrating the dissociation between rules and similarity, a unitary category structure has often been used in which category membership could be determined perfectly from rule application *or* similarity-based processes and induction of either process was achieved by instruction manipulation (e.g., explicit formulation of the rule vs. feedback training only, in Allen & Brooks, 1991). Alternatively, real-world categories have been used for which the existence or nonexistence of necessary attributes (rules) was known to the participants (e.g., size of a quarter, in Rips, 1989). In our experiments, the participants had no prior knowledge about the nature of the category structure, and training was based only on feedback for all the category structures. The category structures themselves, rather than instruction or prior knowledge, promoted or inhibited the use of either system. The processes of rule discovery and testing are of equal importance to rule application in the COVIS model, and the interaction and relative weighting of the two systems is explicitly stated. On the other hand, studies in which use of a rule versus a similarity judgment is addressed in conditions in which both strategies are available on any given trial may help to shed more light on how the competition between the hypothesis-testing and the implicit systems is resolved. Also, although the rule versus similarity distinction may be widely valid across modalities and extend to higher level cognition, such as language, COVIS has a narrower focus on visual perceptual categorization and, because of the specified underlying neurobiology, cannot be automatically applied outside its original domain. In sum, despite some methodological and terminological differences, the neuropsychological COVIS model and the cognitive-psychology–based theories of alternative rule and similarity strategies of categorization are more likely to complement than to oppose one another.

Single System Versus Multiple Systems in Category Learning

The alternative to the notion of multiple systems in categorization is the notion of a single categorization system. First we have to make clear what *a system* means. In COVIS, the two category-learning systems operating in parallel differ in both the computational and the implementation levels of description—one system coding for explicit rules in frontal structures using selective attention and working memory, the other encoding instances in the IT cortex and procedural-learning–based stimulus–response mapping in the striatum (see the introduction or Ashby et al., 1998, for more details). Both systems then compete (or cooperate) to determine the response of the overall system (the organism). When arguing against the multiple-system account of Waldron and Ashby's (2001) results, Nosofsky and Kruschke (2002) accepted that other processes, such as selective attention to the relevant dimension, may take place in category learning. However, they emphasized that different processes operate on a single exemplar category representation. What seems to be a distinction between categorization based on a rule application and that based on overall similarity evaluation is, then, a distinction between exemplar-based categorization when all attention weight is on one diagnostic dimension and the same exemplar-based categorization when attention weight is spread about equally across many dimensions. A similar idea has recently been presented by Pothos (2005), who argued that rules and similarity represent two extremes on a single continuum of similarity operations, with no need to model rule and similarity processes separately. Rule application is a similarity evaluation process in which only a single or small number of an object's features are involved.

These are compelling ideas, and the imperative of parsimony requires accepting the single-system notion (a single representation with a single process), unless we have a sufficient body of evidence that a single-system explanation cannot account for the empirical data available.

Several lines of evidence lead us to believe that a single-system explanation is not sufficient to account for the data observed in our two experiments and, most important, in the complex of a broad range of other studies. First, we have already discussed how Nosofsky and Kruschke's (2002) account of Waldron and Ashby (2001) is inconsistent with the results of our Experiment 2 (see also Ashby & Ell, 2002, for an evaluation of Nosofsky & Kruschke, 2002).

Second, single-system models, such as ALCOVE (Kruschke, 1992), do not specifically address the underlying neural substrate for exemplar storage and the neural mechanism of categorization. A number of studies have focused on differential activation of the brain in different categorization paradigms and have suggested that humans, at least, have available more than one system, involving different neural circuits and category representations (Seger & Cincotta, 2002; E. E. Smith et al., 1998; see Kéri, 2003, for a review of studies including clinical neuropsychology findings, functional neuroimaging, and single-cell research). To address specifically the single- versus multiple-representations issue, Kéri has summarized a number of studies showing that the IT cortex is responsible for category instances representation, whereas the prefrontal cortex encodes abstract rules. Such a finding supports at least a twofold category representation; one is the representation of specific instances (exemplars), the other of rules. Because an organism behaves as an integral system, different representations and processes will interact and act in concordance in order to produce meaningful behavior. Pothos's (2005) putative continuum from rules to similarity may thus reflect a relative contribution of each system to the overall response of an organism, such as that postulated in COVIS by the relative weighing of the two subsystems' responses in producing the final decision.

Third, even if a single-system model can account for the pattern of data observed in our two experiments, we may still question whether that provides us with a valid and more parsimonious explanation. First, exemplar models are often viewed as highly flexible. Recently, Olsson, Wennerholm, and Lyxzèn (2004) showed that exemplar and other mathematical models often suffer from overfitting—that is, accounting perfectly for noise, as well as actual variance due to cognitive processes. Second, with respect to the issue of parsimony, it is unclear whether a single-system model that requires different sets of assumptions (and parameter values) about the underlying processes to account for the wide array of "multiple-systems" data[7] is more parsimonious than a multiple-systems model that a priori predicts the patterns observed in the multiple-systems data.

REFERENCES

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

Arbuthnott, G. W., Ingham, C. A., & Wickens, J. R. (2000). Dopamine and synaptic plasticity in the neostriatum. *Journal of Anatomy*, **196**, 587-596.

Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, **105**, 441-481.

Ashby, F. G., & Ell, S. W. (2002). Single versus multiple systems of category learning: Reply to Nosofsky and Kruschke (2002). *Psychonomic Bulletin & Review*, **9**, 175-180.

Ashby, F. G., Ell, S. W., & Waldron, E. M. (2003). Procedural learning in perceptual categorization. *Memory & Cognition*, **31**, 1114-1125.

Ashby, F. G., & Gott, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **14**, 33-53.

Ashby, F. G., & Maddox, W. T. (1992). Complex decision rules in categorization: Contrasting novice and experienced performance. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 50-71.

Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology*, **56**, 149-178.

Ashby, F. G., Queller, S., & Berretty, P. M. (1999). On the dominance of unidimensional rules in unsupervised categorization. *Perception & Psychophysics*, **61**, 1178-1199.

Ashby, F. G., & Waldron, E. M. (2000). The neuropsychological bases of category learning. *Current Directions in Psychological Science*, **9**, 10-14.

Bench, C. J., Frith, C. D., Grasby, P. M., Friston, K. J., Paulesu, E., Frackowiak, R. S. J., & Dolan, R. J. (1993). Investigations of the functional anatomy of attention using the Stroop test. *Neuropsychologia*, **31**, 907-922.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, **10**, 433-436. Online record: color.psych.upenn.edu/brainard/pubs.html.

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

Efron, B. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall.

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

Estes, W. K. (1994). *Classification and cognition*. Oxford: Oxford University Press.

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

Kemler Nelson, D. G. (1984). The effect of intention on what concepts are acquired. *Journal of Verbal Learning & Verbal Behavior*, **23**, 734-759.

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

Kerr, J. N., & Wickens, J. R. (2001). Dopamine D-1/D-5 receptor activation is required for long-term potentiation in the rat neostriatum in vitro. *Journal of Neurophysiology*, **85**, 117-124.

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

Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A network model of category learning. *Psychological Review*, **111**, 309-332.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, **109**, 163-203.

Maddox, W. T., & Ashby, F. G. (2004). Dissociating explicit and procedural-learning based systems of perceptual category learning. *Behavioural 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, & Cognition*, **29**, 307-320.

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 perceptual category learning. *Psychonomic Bulletin & Review*, **11**, 945-952.

Maddox, W. T., Filoteo, J. V., Hejl, K. D., & Ing, A. D. (2004). Category number impacts rule-based but not information-integration category learning: Further evidence for dissociable category-learning systems. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **30**, 227-245.

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, & Cognition*, **31**, 100-107.

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

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

Nosofsky, R. M., & Kruschke, J. K. (2002). Single-system models and interference in category learning: Commentary on Waldron and Ashby (2001). *Psychonomic Bulletin & Review*, **9**, 169-174.

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

Olsson, H., Wennerholm, P., & Lyxzèn, U. (2004). Exemplars, prototypes, and the flexibility of classification models. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **30**, 936-941.

Packard, M. G., & Knowlton, B. J. (2002). Learning and memory functions of the basal ganglia. *Annual Review of Neuroscience*, **25**, 563-596.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, **10**, 437-442. Online record: vision.nyu.edu/VideoToolbox/VideoToolbox.html.

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.

Pothos, E. M. (2005). The rules versus similarity distinction. *Behavioral & Brain Sciences*, **28**, 1-49.

Reber, P. J., Stark, C. E. L., & Squire, L. R. (1998). Contrasting cortical activity associated with category memory and recognition memory. *Learning & Memory*, **5**, 420-428.

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

Rips, L. J. (1989). Similarity, typicality, and categorization. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 21-59). Cambridge: Cambridge University Press.

Schacter, D. L., & Tulving, E. (Eds.) (1994). *Memory systems 1994*. Cambridge, MA: MIT Press.

Seger, C. A., & Cincotta, C. M. (2002). Striatal activity in concept learning. *Cognitive, Affective, & Behavioral Neuroscience*, **2**, 149-161.

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

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, **119**, 3-22.

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

Smith, J. D., Minda, J. P., & Washburn, D. A. (2004). Category learning in rhesus monkeys: A study of the Shepard, Hovland, and Jenkins (1961) tasks. *Journal of Experimental Psychology: General*, **133**, 398-414.

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.

Waldron, E. M., & Ashby, F. G. (2001). The effects of concurrent task interference on category learning: Evidence for multiple category learning systems. *Psychonomic Bulletin & Review*, **8**, 168-176.

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: MIT Press, Bradford Books.

NOTES

1. Using the term *implicit* for the procedural-learning–based system does not imply an unconscious nature for category learning by this

system (see Shanks & St. John, 1994, for a discussion of unconscious learning). Rather, we mean that the optimal categorization rule characterizing the II category structure is not directly represented in the system but is only implicitly present in the stimulus–response mapping within striatum. The participants are not likely to be able to express the rule explicitly after the training, even when their response strategies suggest that they are able to employ such a rule.

2. Some of the studies reported above utilized multidimensional rule-based tasks that addressed this potential shortcoming and continued to show the predicted results (e.g., Maddox et al., 2004; Maddox & Ing, 2005).

3. These discriminabilities were chosen to avoid ceiling effects in the UD conditions and floor effects in the II conditions. In addition, we hoped to approximately equate performance across these two category structures in the control condition. To anticipate, we were not successful in equating control condition performance. The II control condition performance was worse than the UD control condition performance. However, if the two conditions differed only in difficulty, as is sometimes suggested, we would expect a larger dual-task interference effect on II than on UD category learning. COVIS, on the other hand, predicts a larger dual-task interference effect on UD category learning.

4. Bootstrap analysis is a statistical method for obtaining an estimate of reliability or error, such as confidence intervals, without a priori assumptions about population distribution. The sample distribution and variability are used as a model for the population distribution, and simulations carried out on actual samples are used to draw inference. Bootstrapping is appropriate to use when the distribution shape is unknown (Efron, 1993).

5. The difference between the median drop in performance across the two category structures was even stronger. The median performance drop for the UD-category observers was 29.0% in the Stroop task condition, as compared with the control condition, which is reliably bigger (bootstrapped 95% confidence interval) than the 6.8% median performance drop for the II category observers. The same category structure × condition interaction was detected using parametric methods [ANOVA interaction, $F(1,138) = 4.006$, $MS_e = 0.367$, $p = .047$].

6. The responses of the rest of the participants were best accounted for by a random response strategy.

7. For instance, Nosofsky and Kruschke (2002) claim that "as long as the sensitivity parameter c is not too high, ALCOVE predicts far greater interference on the simple one-dimensional task than on the complex three-dimensional task" (p. 171). The c parameter measures the overall discriminability of the stimuli and *should* be high for such highly discriminable stimuli as those used in Waldron and Ashby (2001).

(Manuscript received July 29, 2004;
revision accepted for publication February 28, 2005.)