The role of visuospatial and verbal working memory in perceptual category learning

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

The role of verbal and visuospatial working memory in rule-based and information-integration category learning was examined. Previously, Maddox, Ashby, Ing, and Pickering (2004) found that a sequentially presented verbal working memory task did not affect information-integration learning, but disrupted rule-based learning when the rule was on the spatial frequency of a Gabor stimulus. This pattern was replicated in Experiment 1, in which the same category structures were used, but in which the verbal working memory task was replaced with a visuospatial analog. Experiment 2A examined rule-based learning on an oblique orientation and also found both verbal and visuospatial working memory tasks disrupting learning. Experiment 2B examined rule-based learning on a cardinal orientation and found a minimal effect of the verbal working memory task, but a large effect of the visuospatial working memory task. The conceptual significance of cardinal orientations and the role of visuospatial and verbal working memory in category learning are discussed.

Perceptual categorization is a systematic differentiation among classes (or categories) of objects based on their perceptual features. A large body of evidence suggests that humans have available several category learning systems (see Ashby & Maddox, 2005, or Keri, 2003, for a review). The classical approach to categorization (Bruner, Goodnow, & Austin, 1956), and most multiple-systems approaches, postulate one categorization system based on an extraction of a categorization rule via hypothesis testing during an explicit reasoning process (see Anderson & Betz, 2001; Ashby & Ell, 2001; Erickson & Kruschke, 1998; Nosofsky, Palmeri, & McKinley, 1994; Patalano, Smith, Jonides, & Koeppe, 2001; E. E. Smith, Patalano, & Jonides, 1998). This hypothesis-testing system generates and tests hypotheses (usually verbalizable) and involves working memory and attentional processes (Maddox, Ashby, Ing, & Pickering, 2004; Waldron & Ashby, 2001; Zeithamova & Maddox, 2006). The multiple-system models differ in the nature of the second proposed system. These include exemplar-similarity (Erickson & Kruschke, 1998; Patalano et al., 2001), holistic family-resemblance (J. D. Smith & Shapiro, 1989), or perceptual-representation–based systems (Reber, Gitelman, Parrish, & Mesulam, 2003). One multiple-system model that also offers neurobiological bases for the two systems is the COmpetition between Verbal and Implicit Systems model (COVIS; Ashby, Alfonso-Reese, Turken, & Waldron, 1998), which proposes a procedural-learning–based system.

In the COVIS model, both systems attempt to learn each categorization task encountered, but each system dominates the acquisition of different category structures.

The hypothesis-testing system dominates the acquisition of *rule-based category structures*. The procedural system dominates the acquisition of *information-integration category structures* (Ashby & Ell, 2001). Rule-based category structures are those in which category membership may be determined from some explicit rule. Suppose each stimulus is a Gabor patch. (A Gabor patch is a sine wave grating enclosed in a Gaussian envelope that can vary in spatial frequency and orientation.) An example of a rule-based (RB) category structure is presented in Figure 1 (left panel). Each point in this scatterplot represents a unique Gabor patch stimulus. The dashed vertical line denotes the optimal categorization rule and can be verbalized as "Respond A when the stimulus consists of wide bars; respond B when the stimulus consists of narrow bars." Information-integration category structures include those in which at least two dimensions expressed in different units need to be combined (integrated). An example of an information-integration (II) category structure using Gabor stimuli is presented in Figure 1 (right panel). Although the optimal categorization rule could be verbalized as "Respond A when the orientation is greater than the spatial frequency; respond B when the spatial frequency is greater than the orientation," such a rule is difficult to comprehend because it compares dimensions expressed in incommensurable units.

The assumption is that the two categorization systems employ different cognitive learning mechanisms, each subserved by distinct neural pathways. The hypothesis-testing system seeks explicit rules that discriminate between members of categories based on some criterial dimension, or a limited set of criterial dimensions. Hypothesis testing is an

---
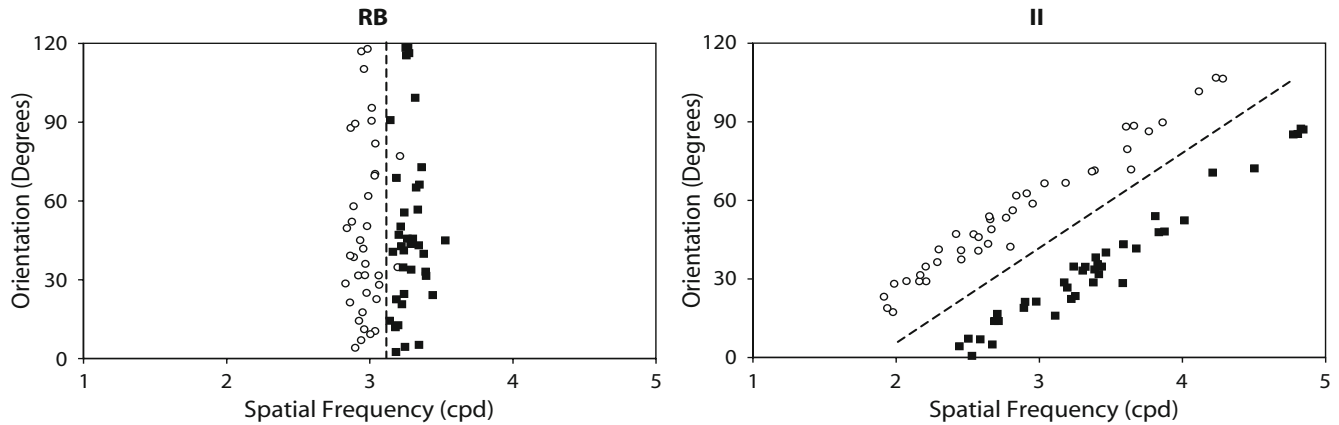
D. Zeithamova, zeithamova@mail.utexas.edu

Figure 1. Rule-based (RB) and information-integration (II) category structures. Open circles denote Category A, filled squares denote Category B. The dashed lines represent the optimal decision bound (cpd = cycles per degree).

effortful process involving selective attention to individual stimulus dimensions and working memory for storing the currently tested rule and for feedback evaluation. The neurobiological pathways involved in working memory, attention and the proposed hypothesis-testing category learning system include dorsolateral prefrontal cortex, anterior cingulate and the head of the caudate nucleus (Fuster, 1989; Goldman-Rakic, 1987; Hikosaka, Sakamoto, & Usui, 1989; Levy, Friedman, Davachi, & Goldman-Rakic, 1997; Posner & Petersen, 1990; Schultz, Apicella, Romo, & Scarnati, 1994). Because the hypothesis-testing system uses limited-capacity resources (e.g., working memory and selective attention), rule-based learning may be disrupted by a concurrent or sequential working memory-demanding task. Indeed, a concurrent Stroop task (Waldron & Ashby, 2001; Zeithamova & Maddox, 2006) and sequential Sternberg working memory task (Maddox et al., 2004) impaired rule-based but not information-integration category learning. One aim of this article is to provide a more detailed examination of the role of a sequential working memory task on rule-based and information-integration learning.

The procedural system is assumed to learn gradually and incrementally to associate specific regions in the perceptual space with category labels. The neurobiological pathway of the proposed procedural category learning system involves a posterior circuit including the inferotemporal cortex and the posterior caudate nucleus. The current thinking is that a low-resolution map of the perceptual space is represented within the tail of the caudate (Ashby et al., 1998; see, however, Bar-Gad, Morris, & Bergman, 2003). This assumption is supported by the fact that there exists a many-to-one convergence of inferotemporal cells onto the neurons in the tail of the caudate nucleus (Brown, Desimone, & Mishkin, 1995; Caan, Perrett, & Rolls, 1984; Wilson, 1995). The involvement of the body and tail of the caudate in information-integration learning is also supported by two recent fMRI studies (Nomura et al., 2007; Seger & Cincotta, 2005). Finally, the fact that the caudate nucleus projects to prefrontal cortex and motor output areas via connections through the thalamus (Alexander, DeLong, & Strick, 1986) further

supports the notion that the caudate might be involved in linking stimuli with classification responses (Packard & Knowlton, 2002). A consistent stimulus–response mapping (Ashby, Ell, & Waldron, 2003; Maddox, Bohil, & Ing, 2004) and immediate feedback (Maddox, Ashby, & Bohil, 2003; Maddox & Ing, 2005) are crucial for the procedural learning system. The reliance on a consistent stimulus–response mapping follows from the reliance of this system on the traditional procedural learning mechanism (Jahanshahi, Brown, & Marsden, 1992; Willingham, 1998; Willingham, Nissen, & Bullemer, 1989). The need for immediate feedback follows from the system's reliance on dopamine-mediated reward learning in the caudate (Aron et al., 2004; Wickens, 1993). The details can be found elsewhere (Ashby & Ennis, 2006; Maddox et al., 2003), but for now it suffices to say that learning involves strengthening of corticostriatal synapses in the caudate via dopamine release from the substantia nigra (pars compacta). The dopamine release must occur while the relevant synapses are still active, thus putting an upper bound on the timing of the feedback.

Working Memory Role in the Hypothesis-Testing and Procedural Learning Systems

Most multiple systems models of category learning postulate a hypothesis-testing (or rule-based) component. These include models that make no claims about the neurobiological underpinnings, such as RULEX (Nosofsky et al., 1994) or ATRIUM (Erickson & Kruschke, 1998), as well as neurobiologically inspired models like COVIS (Ashby et al., 1998; see also Patalano et al., 2001). As stated explicitly in COVIS (and perhaps implicitly assumed in RULEX and ATRIUM), processing of the feedback in the hypothesis-testing system is effortful and requires working memory and attentional capacity. Following feedback on an incorrect trial, a number of events occur in the hypothesis-testing system: (1) The salience of the current rule decreases; (2) a decision is made about whether to reuse the current rule or to generate and select a new one; (3) if applicable, attention is switched from the old rule to the new rule. These events in the hypothesis-testing

system require both time and availability of the limited resources (working memory and attention: Maddox et al., 2004; Waldron & Ashby, 2001; Zeithamova & Maddox, 2006). On the other hand, the procedural system requires the feedback to follow the categorization response immediately (Maddox et al., 2003; Maddox & Ing, 2005), but feedback is then processed automatically and does not require working memory or attention (Maddox et al., 2004).

Maddox et al. (2004) tested the prediction that feedback processing is effortful and time consuming for the hypothesis-testing system, but not for the procedural system. They contrasted rule-based and information-integration category learning using the category structures depicted in Figure 1 across the three experimental conditions displayed in Figure 2A. In the control condition, participants viewed a stimulus, generated a categorization response and received 500 msec of corrective feedback followed by a 2,000-msec (blank screen) intertrial interval (ITI). In the "long" feedback-processing condition, the categorization response feedback was followed by 2,500-msec blank screen display to allow feedback processing, after which a trial of a *verbal* working memory task was presented. The trial was concluded with a 2,000-msec ITI. As the verbal working memory task, the Sternberg (1966) memory scanning task was used (Figure 2B). In the Sternberg memory scanning task, four digits between 1 and 9 were presented for 500 msec (memory

set). Next, a 1,000-msec delay (blank screen) was presented followed by a single digit (probe). The participants had to indicate whether this digit was a part of the memory set or not. In the "short" condition, the categorization response feedback was followed immediately by the working memory task.

Maddox et al. found a larger disruption of rule-based category learning in the "short" condition than in the control condition, and a small (and not statistically significant) disruption of rule-based category learning in the "long" condition. The presence of the Sternberg task had no effect on information-integration category learning.

Overview of the Present Study

Maddox et al. (2004) provided evidence that working memory and attention are necessary for accurate feedback processing in rule-based category learning but not in information-integration category learning. They used a version of Sternberg's memory-scanning task to tax working memory and attentional processes. However, several questions remain.

In the working memory literature, a distinction is made between at least two kinds of working memory: verbal and visuospatial (e.g., Baddeley & Logie, 1999; Jonides et al., 1996). This distinction holds behaviorally (Cocchini, Logie, Della Sala, MacPherson, & Baddeley, 2002; Logie, Zucco, & Baddeley, 1990; Shah & Miyake, 1996), and there is also good evidence that verbal and visuo-
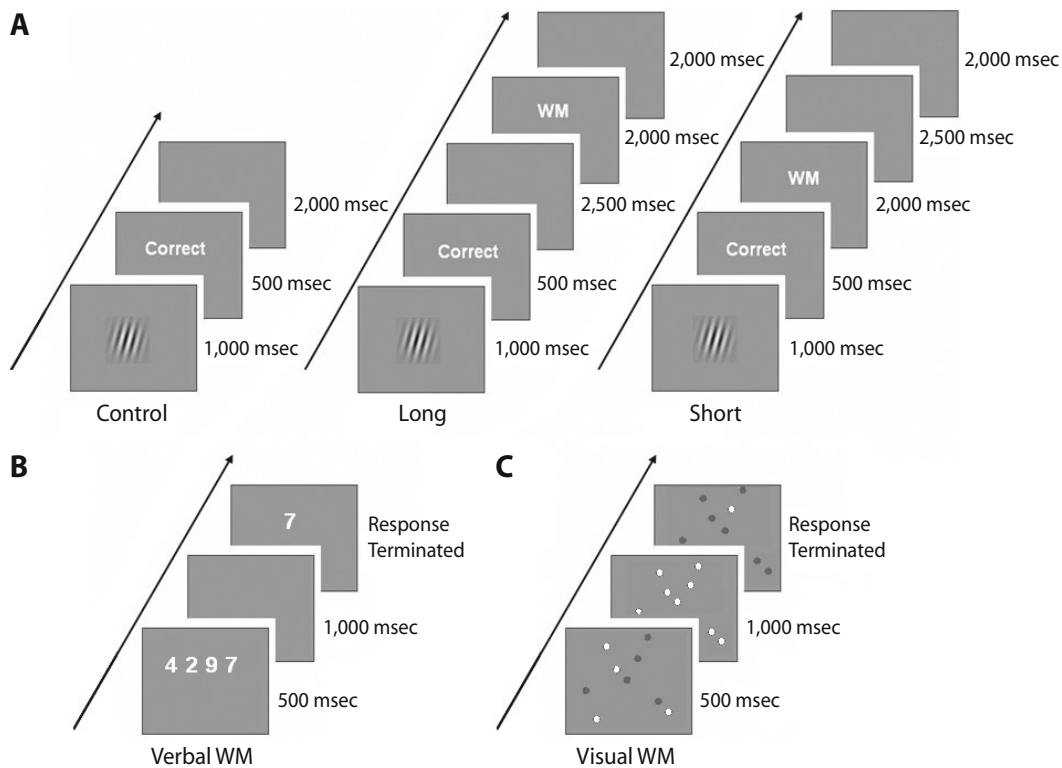

Figure 2. In (A) trial design of three conditions in Maddox et al. (2004) and Experiment 1: Control = control condition; long = long feedback processing time condition; short = short feedback processing time condition. (B) Verbal working memory (WM) task design. (C) Visuospatial WM task design.

spatial working memory rely on different neural systems (Goldman-Rakic, 1998; E. E. Smith, Jonides, & Koeppe, 1996). However, the existing models of the hypothesis-testing system do not address the distinction and do not make any a priori prediction regarding the effects of a secondary verbal versus visuospatial working memory task on category learning. The Sternberg task is a standard verbal working memory task (see Raghavachari et al., 2001); the question remains how a visuospatial working memory task affects rule-based and information-integration category learning. In this section, we discuss a series of reasonable predictions.

Maddox et al. (2004) showed that a sequential verbal working memory task did not affect information-integration category learning. One hypothesis is that this result would replicate with a visuospatial working memory task. The logic is as follows: If the information-integration task is learned primarily via the procedural learning system, and if the procedural system processes feedback automatically without the need for any kind of working memory or attentional resources, there should be no effect of visuospatial working memory on information-integration category learning. A second hypothesis is that the presence of a visuospatial working memory task will affect information-integration category learning. The logic is as follows: Processing in the procedural system depends critically on the visual stimulus representation in the inferotemporal cortex (Freedman, Riesenhuber, Poggio, & Miller, 2003). This representation may be disrupted by the presence of a visuospatial task, because the stimuli are visuospatial and difficult to encode verbally. Additionally, the procedural system is assumed to rely on the basal ganglia (e.g., Filoteo, Maddox, Salmon, & Song, 2005; Poldrack, 2002), in much the same way as visuospatial working memory is assumed to depend on the basal ganglia, perhaps even to a larger extent than do other types of working memory (Lawrence, Watkins, Sahakian, Hodges, & Robbins, 2000; Postle, Jonides, Smith, & Corkin, 1997).

In the hypothesis-testing system, working memory is needed for holding the currently active rule, comparing that rule with the current feedback, and selecting and switching to a new rule if necessary. Because rules learned by the hypothesis-testing system are usually verbalizable (Ashby et al., 1998), one reasonable prediction is that a sequential verbal working memory task will adversely affect rule-based category learning, but that a sequential visuospatial working memory task will not. This prediction assumes that rule-based learning involves generating a verbal representation of the stimulus, response and feedback. Thus, placing a load on a separate visuospatial working memory store will not affect these verbal processes. Although verbal and visuospatial working memory stores are separable, a second possibility is that a sequential visuospatial working memory task will adversely affect rule-based category learning. Two general mechanisms may underlie such a disruption. First, a visuospatial working memory task may affect rule-based learning indirectly because it, like the verbal working memory task, relies on the central executive as a common, limited-capacity resource (Baddeley, 1995; Baddeley & Logie, 1999). A

load on the visuospatial working memory store would thus influence rule-based category learning via the central executive or general attention demand. Second, visuospatial working memory may be involved in some aspect of rule-based category learning that does not require verbal working memory.

To conceive which aspects of rule-based learning may be differentially influenced by the secondary verbal and visuospatial working memory tasks, we need to consider what steps or processes may possibly be parts of rule learning and discovery. With unidimensional rules on a small number of continuous-valued dimension stimuli (as used in our experiments), the process may include the following intertwined steps: (1) selection and focused attention on one *stimulus dimension* (e.g., "spatial frequency"); (2) generation, representation, and testing of a *categorization rule* (in the narrow sense of the meaning) along that dimension (e.g., "narrow stripes are Category A, wide stripes are Category B"); and (3) learning, storing and application of a *categorization criterion* (e.g., the optimal spatial frequency distinguishing between narrow and wide stripes). Verbal working memory would then seem critical primarily for rule generation, maintenance and testing (Step 2), while visuospatial working memory may then be critical for learning and representation of the actual categorization criterion (image of a particular spatial frequency: Step 3) and/or for identification of individual stimulus dimensions (analytical decomposition of the stimulus) that is a basis for Step 1. A second reasonable prediction is thus that the visuospatial working memory may disrupt rule-based learning, either in a similar fashion to a verbal working memory task, or differently.

The goal of Experiment 1 was to test these hypotheses using the same procedure as Maddox et al. (2004) but replacing the verbal Sternberg working memory task with a visuospatial working memory analog. The basic experimental design is depicted in Figure 2A; the design of the verbal Sternberg task and our visuospatial analog are presented in Figure 2B and 2C. To anticipate, we found no effect of the sequential visuospatial working memory task on information-integration category learning, but we did find an effect on rule-based category learning.

Experiment 2 explored the generality of the working memory effects in perceptual category learning. Whereas Maddox et al. (2004) and Experiment 1 examined a rule on the spatial frequency of a Gabor patch stimulus, Experiment 2 examined rules on the orientation of a Gabor patch stimulus. Gabor patch stimuli have several desirable properties for perceptual category learning researchers. For example, they have a known dimensional structure with two separable dimensions. The two dimensions have simple verbal labels, are measured in different units, and have no emergent properties. However, orientation has two special properties that spatial frequency does not have. First, it is periodic, with 0º being equivalent to 360º. Second, it contains special values called *cardinal orientations*. Cardinal orientations, both vertical and horizontal, are processed differently from other values at both the neural and behavioral level. People are more sensitive and more accurate when asked to judge orientations around cardinal orienta-

tions (Campbell & Kulikowski, 1966; Heeley & Timney, 1988; Orban, Vandenbussche, & Vogels, 1984). This is probably because more neurons in primary visual cortex are tuned to cardinal orientations (Furmanski & Engel, 2000).

In Experiment 2, we examined the effect of both visuospatial and verbal working memory tasks on rule-based learning when the optimal categorization rule required the participant to separate orientations above and below 70º from horizontal (oblique, Experiment 2A), or orientations above and below 90º (cardinal, Experiment 2B). When the criterion is set at 70º, we expect the findings to replicate those observed when spatial frequency was relevant, because learning a rule on the (arbitrary) orientation of 70º likely requires the same processes needed to learn a rule on spatial frequency. In both cases, there are two obvious dimensions (spatial frequency and orientation) along which to generate explicit rules; these rules need to be tested, the irrelevant dimension ignored, and the optimal criterion on the relevant dimension needs to be learned. We will test this prediction in Experiment 2A.

The processes involved in cardinal orientation-based category learning may be different. As noted above, people exhibit greater sensitivity to orientation changes near cardinal orientations. We argue that cardinal orientations are *perceptually special*. This well-established higher perceptual sensitivity may lead to more precise categorization (leading to higher asymptotic accuracy for learners), once a correct rule (Step 2) is discovered. In other words, the perceptual advantage for cardinal orientations should improve *categorization criterion* learning (Step 3).

Furthermore, cardinal orientations may not simply constitute an easier *categorization criterion* value on the general *categorization rule* "Respond A if orientation is greater than a criterion; respond B if orientation is smaller than a criterion." Rather, as soon as the participant notices that stimuli vary in orientation, the cardinal orientations may constitute salient, spontaneously used categorization rules that create intuitive categories, or concepts (e.g., right tilted, left tilted). We argue that cardinal orientations may be *conceptually special*. In other words, cardinal orientations may constitute a highly salient *categorization rule* per se, with Step 2 and Step 3 being merged together. To our knowledge, no categorization studies have explicitly tested this assumption. However, Huttenlocher, Hedges, and Duncan (1991) reported that their participants used cardinal orientations as reference points in location estimation. If cardinal orientations are conceptually special, people may tend to use these early in learning as a rule of first choice and most of the categorization rule discovery stage of learning may be skipped. This possible conceptual significance of cardinal orientations may lead to an increase in the proportion of learners, and much more rapid learning with minimal working-memory load (and thus minimal effect of a secondary working-memory task), as most participants will simply select the correct categorization rule as their first choice.

Now let us consider the possible effects of a working memory task when one is learning a rule on a cardinal ori-

entation. First, consider verbal working memory effects. If cardinal orientations are special perceptually but not conceptually, all the effortful hypothesis-testing processes still need to take place to find the correct rule, and we would expect to observe a verbal working memory task effect. If cardinal orientations are special perceptually *and* conceptually, the highly salient cardinal rule will be chosen early and much of the hypothesis-testing process will be bypassed. Under these conditions, we would expect to see no effect, or a minimal one, of the verbal working memory task.

Second, consider visuospatial working memory effects. If the mechanism of the visuospatial working memory task effect is the same as that for the verbal working memory task (i.e., through the central executive), we predict the same effect, or lack thereof, for both types of working memory task. If the mechanisms of visuospatial and verbal working memory task effects differ, the predictions will differ depending on the role of the visuospatial working memory in rule-based category learning.

Let us return to the three steps that take part in unidimensional rule-based learning: (1) selection and focused attention on one stimulus dimension; (2) categorization rule generation, representation and testing; and (3) criterion learning, representation, and application. In order for the first step to occur, the participant needs to notice how the stimuli vary across trials and to decompose them into their individual constituent dimensions. If visuospatial working memory is necessary for such analytic perception of individual stimulus dimensions, the participant may have difficulty identifying the dimensions along which the stimuli vary and how they vary, interfering with the first step of rule-based learning and leading to a learning deficit in the visual condition (even if the verbal working memory task had no effect). If visuospatial working memory is crucial for learning and representing the categorization criterion (Step 3), a cardinal orientation criterion may lead to one of two opposing results. First, learning of a criterion on the cardinal orientation may lead to little or no visuospatial working memory effect. This may follow either from (1) the existing higher perceptual sensitivity around cardinal orientation, which should make learning of cardinal orientation criterion easier and less working memory demanding; or (2) from possible conceptual significance of the cardinal orientation, that is, whether sorting on the basis of a cardinal orientation criterion constitutes an intuitive, highly salient categorization rule per se and does not need to be actually learned the same ways as an oblique criterion. Second, and making an opposite prediction, a criterion on a cardinal orientation may be more working memory demanding, because the increased perceptual sensitivity to orientations around the cardinal orientation would lead participants to consider and test a larger number of possible criteria. This argument assumes that cardinal orientations are not conceptually special and that a criterion on a cardinal orientation needs to be learned in very much the same way as it does on an oblique orientation. In other words, it assumes that a participant would be equally likely to consider a criterion on 88º as a criterion on 90º. In Experiment 2B, we will test

the working memory effects on categorization on the basis of cardinal orientation.

EXPERIMENT 1

The goal of Experiment 1 was to test a number of hypotheses regarding the effect of a visuospatial working memory task on rule-based and information-integration category learning. To achieve this goal, we replicated Maddox et al.'s (2004) procedure, but replaced the Sternberg (verbal) working-memory task with a visuospatial working memory task (Figure 2). We were interested in determining whether the effects on rule-based and information-integration category learning observed for a verbal working memory task were replicated when the verbal task was replaced with a visuospatial working memory task. If the pattern was not replicated, we wanted to determine how the pattern changed.

Method

Participants and Design
Two hundred ninety-three students at the University of Texas at Austin participated in the experiment in partial fulfillment of a class requirement or for pay. All participants were tested for 20/20 vision. The experimental design was 2 category structures (rule-based vs. information-integration) × 3 sequential working memory conditions (control, long feedback-processing time, or short feedback-processing time). Each participant completed one of the six experimental conditions: rule-based control (RB control, 46 participants), rule-based long feedback-processing time (RB long, 51 participants), rule-based short feedback-processing time (RB short, 53 participants), information-integration control (II control, 38 participants), information-integration long feedback-processing time (II long, 52 participants), or information-integration short feedback-processing time (II short, 53 participants).

Stimuli and Apparatus
Category learning. The stimulus dimensions and category structures were identical to those from Maddox et al. (2004). The categorization stimuli were Gabor patches that varied across trials in spatial frequency and orientation. For the rule-based and information-integration category structures, 40 Category A and 40 Category B stimuli were obtained by randomly sampling from two bivariate normal distributions. The rule-based task was unidimensional, with spatial frequency being the relevant dimension and orientation being the irrelevant dimension. The optimal rule was "Respond A if the frequency of the Gabor is small, below 3.13 cycles per degree (cpd); respond B if the frequency of the Gabor is large, above 3.13 cpd." Both dimensions were relevant in the information-integration task and no simple verbal rule discriminated between the two categories. A schematic representation of the two category structures is depicted in Figure 1; the category distribution parameters for both category structures are listed in Table 1.

Each Gabor stimulus was generated and presented using MATLAB (MathWorks, Natick, MA) running Psychophysics Toolbox (Brain-ard, 1997; Pelli, 1997). The stimuli were 200 × 200 pixels, centered on a computer screen, and subtended about 4º of visual angle.

Visuospatial working memory. A visuospatial working memory task that was analogous to the Sternberg working memory task used in Maddox et al. (2004) was created (Figure 2C). The participant was asked to remember four locations out of nine possible (analogous to remembering four numerical digits sampled from nine possible digits). First, nine locations were randomly placed in an imaginary 9 × 9 grid, with the restriction that there be one location in each imaginary row and one location in each imaginary column. The nine locations were marked by dark gray circles each with a radius of 48 pixels and remained visible throughout the visuospatial task trial. After 500 msec, four out of the nine locations were highlighted by a white circle with radius of 40 pixels for 500 msec (memory set). The participant needed to remember those four locations. Next, all nine locations were highlighted for 1,000 msec (delay period). Finally, only one location (probe) was highlighted, and the participant's task was to indicate whether that location was one of the four initially highlighted locations. The probability of a probe being one of the memory set was .5.

Procedure
The procedure was identical to that of Maddox et al. (2004). There were three conditions: control, long, and short (Figure 2A). Each condition consisted of four randomly ordered 80-trial blocks. The participants were informed that there were two equally likely categories and that their task was to learn which patterns went into which category via corrective feedback. In the control condition, a categorization stimulus was presented on each trial and remained on the screen for 1,000 msec or until the participant categorized it in either Category A or Category B. If the participant did not respond during the 1,000 msec, the Gabor stimulus disappeared and only the response prompt ("Categorize the pattern as A or B") remained on the screen. The participant had as much time as needed to make a response. Corrective feedback was then provided for 500 msec followed by a 2,000-msec ITI (blank screen).

In the long feedback-processing time condition, the categorization response feedback was followed by a 2,500-msec blank screen to allow feedback processing, after which a trial of visuospatial working memory task was presented. The visuospatial task response was followed by a 2,000-msec blank screen ITI and no feedback was provided. The short feedback-processing time condition was similar to the long feedback-processing time condition; however, the categorization response feedback was followed immediately by the visuospatial working memory task and the 2,500-msec delay was placed after the working memory task response. After each block of 80 trials, participants were given a short self-paced break, during which they were informed how many trials had passed and were urged to keep their visuospatial task accuracy high.

Results and Discussion

Working Memory Task Performance
The mean percent correct in the working memory task was high, 96.0% for RB long ($SD = 7.0\%$); 94.1% for RB short ($SD = 7.6\%$); 96.0% for II long ($SD = 9.8\%$); 94.8% for II short groups. There were no differences between rule-based ($M = 95.0\%$, $SD = 7.3\%$) and

Table 1
Category Distribution Parameters for the Rule-Based and Information-Integration Category Structures Used in Experiment 1

Category Structure	μ_{fA}	μ_{oA}	μ_{fB}	μ_{oB}	σ_f	σ_o	cov_{fo}
Rule-based	2.97	45	3.28	45	0.087	34	0
Information-integration	2.84	55	3.41	35	0.674	24	16

Note—μ, mean; σ, standard deviation; cov, covariance; f, spatial frequency (cycles per degree); o, orientation (degrees).

information-integration (M = 94.5%, SD = 8.6%) category structure groups [$t(207)$ = 0.413, p = .680], suggesting that the resources allocated to the working memory task were distributed about equally. There was a difference between participants in long (M = 96.0%, SD = 7.0%) and short (M = 93.6%, SD = 8.7%) feedback-processing time conditions collapsed over the two category structures [$t(207)$ = 2.125, p = .035].

Category Learning Performance

Distribution of accuracy scores. In a recent article (Zeithamova & Maddox, 2006), we identified substantial differences between the distributions of accuracy scores in rule-based and information-integration category learning. We found a bimodal distribution of scores in rule-based learning and a normal distribution of scores in information-integration learning. Thus, we began our analysis of category learning performance by examining the distribution of accuracy scores. The distribution of scores for the rule-based groups and information-integration groups (collapsed over the three working memory conditions) from the final block of trials are presented in Figure 3. As is apparent from Figure 3, the score distribution for the rule-based category structure participants deviates from normality [Kolmogorov–Smirnov $D(150)$ = .156, p = .001], whereas the score distribution for the information-integration category structure participants does not deviate from normality [K–S $D(143)$ = .056, p = .510]. The same pattern holds for each of the working memory groups (control, long, and short) within each category structure, with the rule-based groups appearing bimodal and the information-integration groups appearing normal. These results suggest that different processes underlie rule-based and information-integration category learning that lead to very different performance profiles.

To gain further insight into the nature of the bimodality in the rule-based score distributions, we applied a series of models separately to the control, long, and short score distributions. Each model assumes that the observed score distribution results from a mixture of two underlying distributions. A detailed description of the models

and the results are presented in the Appendix. To summarize, the score distribution for each of the three rule-based category learning groups was best fit by a mixture of two underlying distributions, one with mean at .5 (chance) and one with a mean around .8. As suggested by this analysis, there were two types of participants in each condition: learners, who discovered the appropriate categorization rule; and chance performers, who failed to discover the rule. The difference between the three rule-based groups was in the relative weight of the two distributions, that is, the proportion of participants that fell under the .5 modus (nonlearners) and under the .8 modus (learners). More specifically, the effect of the secondary visuospatial working memory task was to decrease the proportion of learners (participants who discovered the rule and constituted the .8 accuracy modus) from 69.2% in the RB control group to 44.0% in the RB long and 48.8% in the RB short group.

Proportion of learners. The distributional analyses suggest that the score distributions in the rule-based conditions are composed of a mixture of two populations of participants (chance performers and learners), and that the relative ratio of learners to chance performers decreases when a sequential working memory task is included. As a further test, we compared the proportion of learners in each condition by defining "learners" as participants who reached .65 correct and higher in the last block of trials and "nonlearners" as participants who failed to reach .65 correct in the last block of trials. We choose .65 correct as it appears to be a natural cut-off in the score distributions (see Figure 3, left panel) and constitutes an average between the means of the chance distribution (.5) and the high performance distribution (.8) for the rule-based groups. The results are depicted in Figure 4 (left panel).

In the RB control group, 32 out of 46 participants learned. There were significantly fewer learners in both the RB long [23 out of 51, $\chi^2(1)$ = 5.897, p = .015] and RB short conditions [26 out of 53, $\chi^2(1)$ = 4.26, p = .039], compared with the RB control condition. In the II control group, 30 out of 38 participants learned. The proportions of learners in II long (36 of 52) and II short
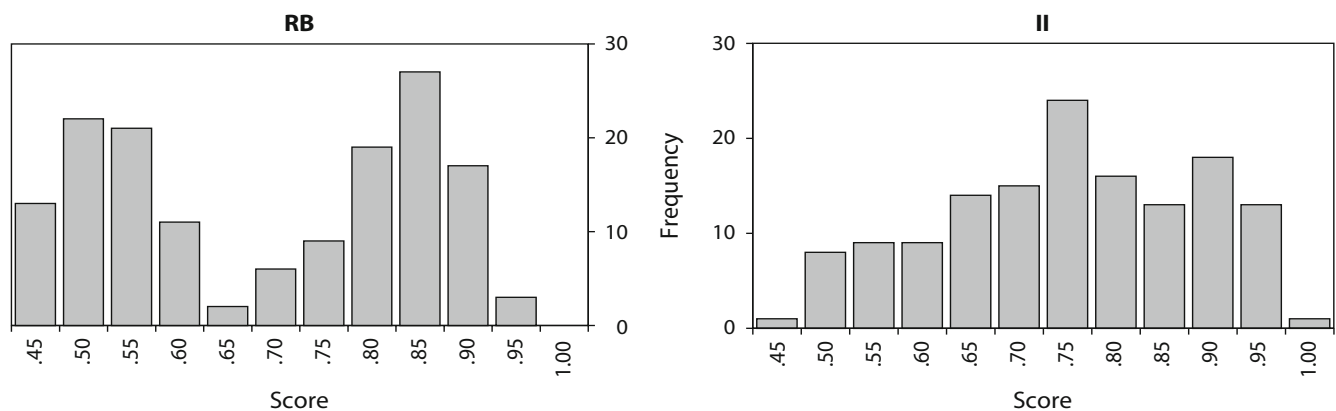

Figure 3. Distribution of the final block accuracy scores of rule-based (RB) and information-integration (II) category learning groups collapsed over the three feedback processing conditions (control, short, and long).
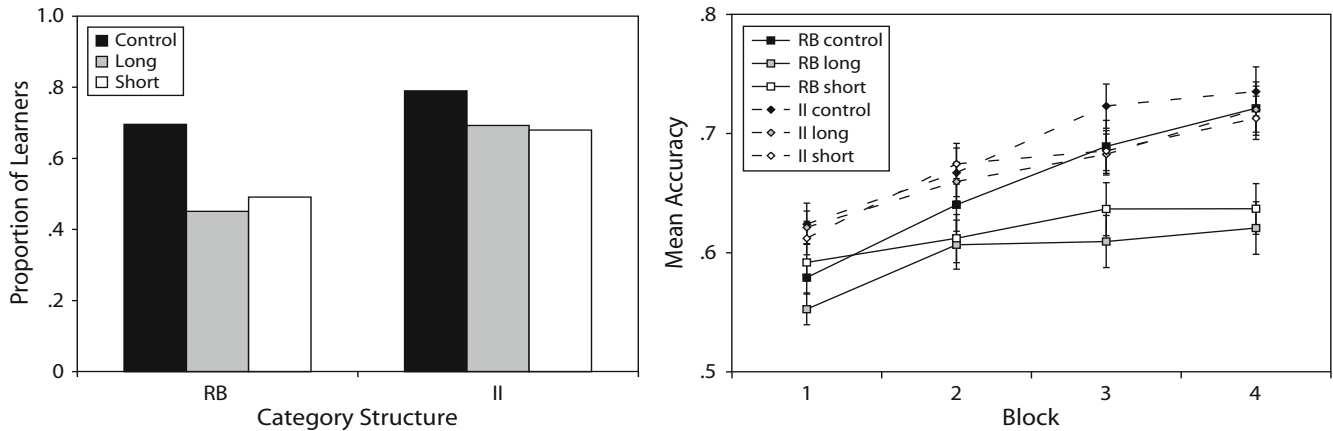
Figure 4. Left panel: Proportion of learners in Experiment 1. RB, rule-based groups; II, information-integration groups. Right panel: Mean accuracy (proportion correct) for each group in Experiment 1. Unidimensional rule-based (RB) groups are denoted with square symbols and solid lines, information-integration (II) groups with diamond symbols and broken lines. Error bars denote bootstrapped 68% confidence intervals (equivalent to a standard error of mean).

(36 of 53) conditions were not significantly different from those observed in the II control condition [long vs. control, $\chi^2(1) = 1.060$, $p = .303$; short vs. control, $\chi^2(1) = 1.350$, $p = .245$]. It is worth noting that these proportions of learner analyses for the information-integration groups were included for completeness and for comparison with the analyses for the rule-based conditions. This learning criterion is less meaningful for the information-integration conditions, because it cuts the distribution of scores at an arbitrary value.

Mean proportion correct. Unlike proportion of learners, mean accuracy, or proportion of correct responses, is a more suitable performance measure for the information-integration groups. For the rule-based groups, mean accuracy is also applicable, but reflects the relative proportion of two populations of participants rather than reflecting performance of a typical participant. We decided to use bootstrapping procedures (Efron & Tibshirani, 1993) that are more appropriate than traditional parametric statistics when the distribution shape is nonnormal or unknown.[1] Categorization accuracy for each group in each block of trials is presented in Figure 4 (right panel).

We found a significant effect of the secondary visuospatial task on rule-based category learning but not on information-integration category learning. Average final block performance in the RB long group dropped by .10 relative to the RB control group [bootstrapped p(control = long) = .002]. Performance in the RB short group dropped by .084 relative to the RB control [p(control = short) = .009].[2]

There were no significant differences among the three information-integration category learning groups in any block of trials or in overall accuracy. Average performance in the final block from the II control group was only .015 higher than that observed in the long group [bootstrapped p(control = long) = .602], and was only .022 higher than that observed in the II short group [p(control = short) = .415].

Brief Summary

To summarize, we found an adverse effect of the visuospatial working memory task on rule-based, but not information-integration category learning. The lack of interference between the visuospatial working memory and the information-integration category learning task is interesting, as both tasks use visual stimuli and are thought to rely on the basal ganglia (visuospatial working memory: see Lawrence, Sahakian, Hodges, & Rosser, 1996; Lawrence et al., 2000; discrimination and category learning: see Packard & McGaugh, 1992; Poldrack, 2002). This result, together with that observed in Maddox et al. (2004) with a verbal working memory task, provides support for the assumption that learning of information-integration categories is mediated by a procedural system that processes feedback automatically, without relying on attention or working memory.

The significant effect of visuospatial working memory on rule-based category learning replicates the effect observed in Maddox et al. with a verbal working memory task, and extends it to a visuospatial working memory task. The effect of the verbal working memory task on rule-based learning found in Maddox et al. (2004) was expected, because attention and (verbal) working memory have been implicated in rule generation, rule maintenance, rule selection and rule switching (Ashby et al., 1998; Dougherty & Hunter, 2003). On the surface, none of these processes appears to require visuospatial working memory, yet visuospatial working memory affected rule-based category learning.

As outlined in the introduction, the visuospatial working memory task may affect rule-based learning, indirectly via cognitive resources shared with verbal working memory, or directly through mechanisms other than those impacted by the verbal working memory task. We will discuss this issue in more detail in the General Discussion. Let us now turn to Experiment 2, which examines the gen-

eralizability of the working memory effects in rule-based learning, which may provide evidence for one or the other explanations of the visuospatial working memory effects observed in Experiment 1.

EXPERIMENT 2

The aim of Experiment 2 was threefold. First, we tested the generality of the working memory effect on rule-based category learning by examining a rule on spatial orientation, instead of spatial frequency; second, we explicitly tested the effect of cardinality of a categorization criterion on category learning; third, we began to examine possible differences in the mechanisms of the visuospatial and verbal working memory effects. Experiment 1 and previous studies (Maddox et al., 2004; see also Zeithamova & Maddox, 2006) suggested that working memory is essential for rule-based but not for information-integration category learning. An effect of visuospatial and verbal working memory tasks was observed when a unidimensional rule on spatial frequency of the Gabor stimulus was relevant. To assess the generality of these effects in rule-based learning, we examined the role of visuospatial and verbal working memory tasks on unidimensional rule-based learning when the orientation of the Gabor stimulus was relevant. To maximize comparison with previous work, we used the same type of stimulus (a Gabor patch), but rotated the rule-based category structure from Experiment 1 by 90º, so that orientation became the relevant dimension and spatial frequency the irrelevant dimension (Figure 5).

Importantly, the orientation dimension contains a set of values—cardinal orientations—that are perceptually special (e.g., Campbell & Kulikowski, 1966). We hypothesized that the cardinal orientations may be also conceptually special and that processes and working memory demands involved in cardinal orientation learning may be different from those involved in oblique orientation learning. To examine this possibility empirically, and to test a number of hypotheses stated in the introduction, we separated Experiment 2 into two parts: Experiment 2A,

in which the optimal criterion was the oblique orientation of 70º from horizontal with all presented orientations being between 50º and 90º, and Experiment 2B, in which the optimal criterion was the cardinal orientation of 90º with all presented orientations between 70º and 110º. The category structures used in Experiment 2A and Experiment 2B are depicted on Figure 5.

EXPERIMENT 2A

In Experiment 2A we tested the effects of a sequentially presented visuospatial or verbal working memory task on rule-based learning when the optimal categorization rule was: "Respond A when the orientation of the stimulus is larger than 70º from horizontal; respond B when the orientation of the stimulus is smaller than 70º from horizontal." We compared a control condition with a short feedback-processing visuospatial working memory condition (as in Experiment 1), and with a short feedback processing verbal working memory condition (as in Maddox et al., 2004). We dropped the long feedback processing condition because its effects were modest in Maddox et al. (2004).

Method

Participants and Design

Seventy-two students at the University of Texas at Austin participated in the experiment in partial fulfillment of a class requirement or for pay. All participants were tested for 20/20 vision. Each participant completed one of the following experimental conditions: control (no secondary task), visual (each category learning trial was immediately followed by a visuospatial working memory task trial), or verbal (each category learning trial was immediately followed by a Sternberg verbal working memory trial). There were 24 participants in each condition.

Stimuli and Apparatus

Category learning task. The stimuli were Gabor patches that varied across trials in spatial frequency and orientation. A rule-based category structure was used with the optimal rule being: "Respond A if the orientation of the Gabor is larger than 70º; respond B if the orientation of the Gabor is smaller than 70º." The stimuli were randomly sampled from two bivariate normal distributions both with a
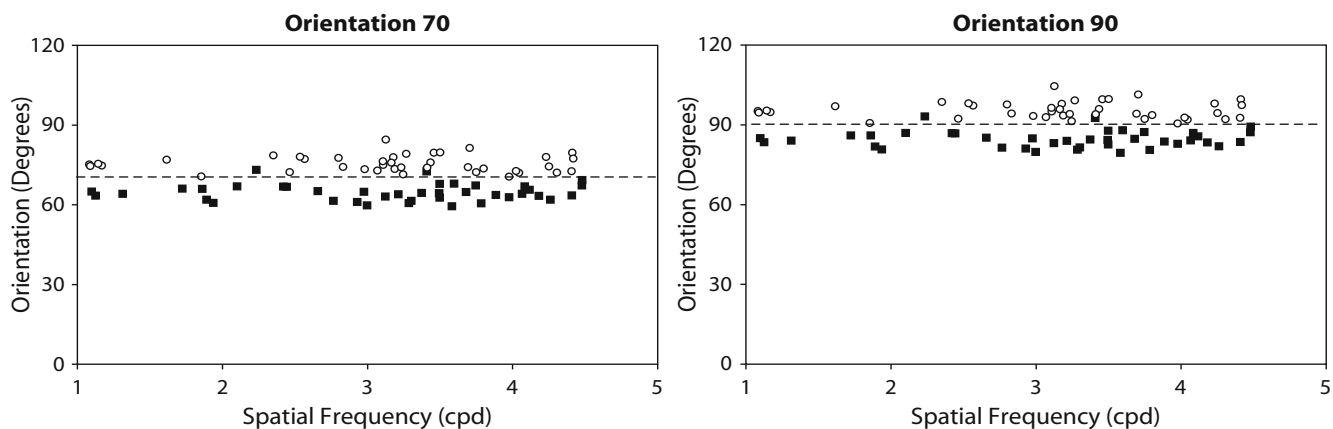

Figure 5. Category structures used in Experiment 2A (Orientation 70) and Experiment 2B (Orientation 90). Open circles denote Category A stimuli, filled squares denote Category B stimuli (cpd = cycles per degree).

mean spatial frequency of 3.13 cpd and a standard deviation (*SD*) of 0.95 cpd. Category A stimuli had a mean orientation of 75.8º with an *SD* of 3.1º. Category B stimuli had a mean orientation of 64.2º with an *SD* of 3.1º. Orientation and spatial frequency of the Gabors were uncorrelated. Forty Category A and 40 Category B stimuli were generated. A schematic representation of the category structure is depicted in Figure 5, left panel. The apparatus was identical to that from Experiment 1.

Working memory tasks. The visuospatial task was identical to that used in Experiment 1 (Figure 2C). The verbal working memory task (Figure 2B) was taken from Maddox et al. (2004). On each trial, four digits (memory set) were randomly selected without replacement from the set of digits from 1 to 9. The memory set was displayed for 500 msec in a horizontal array centered on the screen and subtended approximately 8º of visual angle horizontally and 4º of visual angle vertically. A blank screen followed for 1,000 msec. Finally, a single digit (probe) was presented in the center of the screen and the participant was asked to indicate whether the probe was a part of the memory set. The probability that the probe was a member of the memory set was .5.

Procedure

The procedure was similar to that used in Experiment 1, except that we equated the duration of the ITI across conditions and added a fixation cross to prepare participants for the next trial. Each condition consisted of four blocks of 80 randomly ordered trials. The participants were informed that there were two equally likely categories and that their task was to learn which pattern goes into which category via corrective feedback. In the control condition, each trial started with a 500-msec fixation cross (a plus sign) to prepare the participant for the upcoming trial. A categorization stimulus was then presented and remained on the screen for 1,000 msec, or until the participant categorized it in either Category A or Category B. Corrective feedback was provided for 500 msec followed by 2,500-msec ITI (blank screen).

The two working memory (visual and verbal) conditions were similar to the "short" condition from Experiment 1. Each trial also started with a 500-msec fixation cross, followed by a categorization stimulus presented for 1,000 msec, or until the participant responded. Corrective feedback was presented for 500 msec after a response was made and a working memory task trial immediately followed (verbal, Figure 2B; or visuospatial, Figure 2C). After the participant responded, a 2,500-msec ITI (blank screen) concluded the trial, and no working memory task feedback was provided. As in Experiment 1, after each block of 80 trials participants were given a short self-paced break, during which they were informed how many trials had passed and were urged to keep their working memory task accuracy high.

Results

Working Memory Task Performance

The mean percent correct was 95.5% in the visuospatial task (*SD* = 4.5%) and 96.6% in the verbal task (*SD* = 3.4%). This difference was not statistically significant [$t(46) = .940, p = .352$].

Category Learning Performance

Distribution of scores and proportion of learners. As expected from Experiment 1, the distribution of scores in the final block of trials violated normality [K–S $D(72) = .221, p = .002$], and instead was bimodal (Figure 6, left panel). Using the same procedure applied in Experiment 1 (see the Appendix), we found that the score distribution for each of the three conditions was best fit by a mixture of two underlying distributions, one with mean at .5 (chance) and one with mean around .9. The effect of both secondary working memory tasks was again to decrease of the proportion of participants who discovered the rule (learners) and constituted the .9 accuracy modus, from 83.3% in the control group to 50.0% in the visual and 50.0% in the verbal condition. Both the mean performance level for the learners and relative proportion of learners in the observed score distribution were higher than in Experiment 1. Specifically, whereas the mean performance level for learners in Experiment 1 was about .8, it was .9 in Experiment 2A. Similarly, whereas the relative proportion of learners was 69.6% in Experiment 1 control condition, it was 83.3% in Experiment 2A control condition. Taken together, these data suggest that the rule on the orientation of the Gabor stimuli was somewhat easier to learn.

We also analyzed the proportion of learners in each group using the same criterion as in Experiment 1 (at least .65 correct in the last block). The results are depicted in Figure 7 (left panel). There were 20 learners (out of 24) in the control condition, which is significantly more than 12 learners (out of 24) in either the visual or the verbal condition [$\chi^2(1) = 6.0, p = .014$].[3]

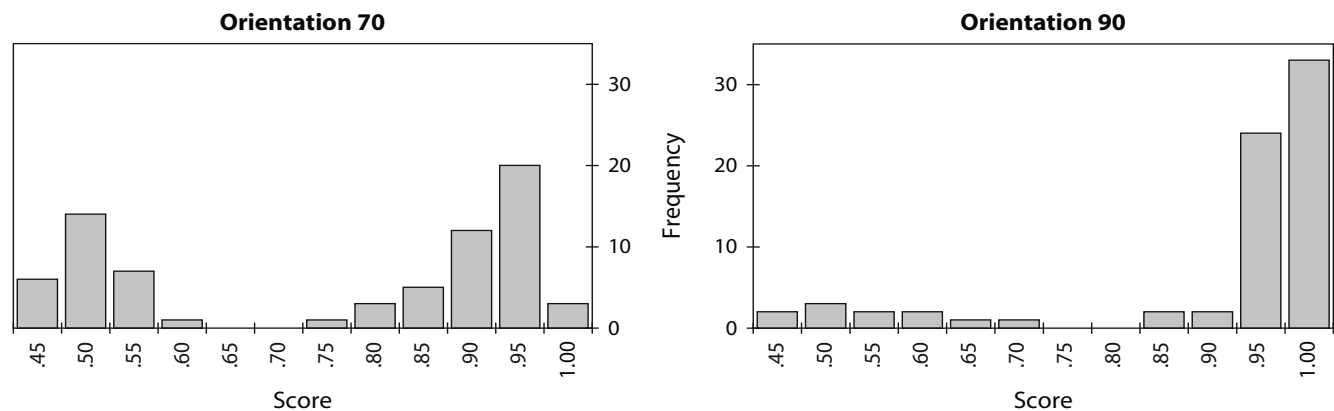**Mean proportion correct**. Block accuracies in Experiment 2A are presented in Figure 7 (right panel).

Figure 6. Distribution of the final block accuracy scores in Experiment 2A (left panel; Orientation 70) and Experiment 2B (right panel; Orientation 90), collapsed over the three working memory conditions (control, visual, verbal).

We found that both the visuospatial and verbal working memory tasks significantly disrupted category learning. Specifically, the average final block performance dropped by .134 in the visual condition relative to the control condition [bootstrapped p(control = visual) = .017], and by .139 in the verbal condition relative to the control condition [bootstrapped p(control = verbal) = .009].

To summarize, we found an adverse effect of both visuospatial and verbal working memory tasks on rule-based category learning when the criterion was on an (arbitrary) oblique orientation. As expected, these results are similar to those found when the criterion was on spatial frequency of a Gabor, suggesting that these working memory effects generalize. We now turn to Experiment 2B, in which the generality of these effects is test more thoroughly.
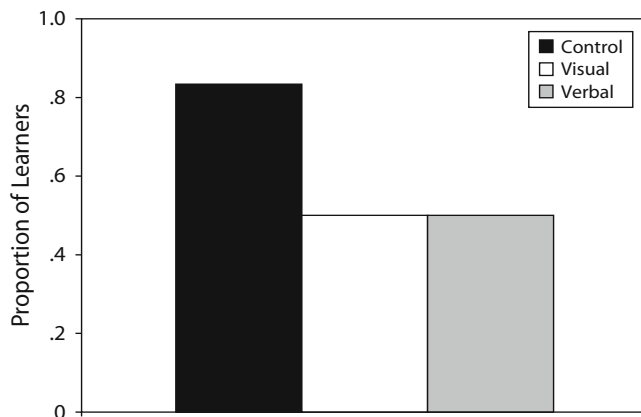
EXPERIMENT 2B

In Experiment 2B, we used a category structure formally identical to that in Experiment 2A, except that the criterion was shifted from 70º to 90º, a cardinal orientation (Figure 5, right panel). As stated in the introduction to Experiment 2, we suspect that this slight manipulation might have a substantial effect on rule-based category learning under both the control condition and when a sequentially presented working memory task is included. The two main aims of Experiment 2B were to investigate whether cardinal orientations were conceptually special and whether we would find evidence for dissociability of the visuospatial and verbal working memory effects.

Method

Participants and Design

Seventy-two students at the University of Texas at Austin participated in the experiment in partial fulfillment of a class requirement or for pay. All participants were tested for 20/20 vision; all completed one of the following experimental conditions: control, visual, or verbal. There were 24 participants assigned into each condition. No student participated in both Experiment 2A and Experiment 2B.

Stimuli, Apparatus, and Procedure

Category learning task. The stimuli, apparatus, and category structure were identical to that from Experiment 2A, except that the optimal categorization rule was "Respond A if the orientation of the Gabor is larger than 90º (is left tilted); respond B if the orientation of the Gabor is smaller than 90º (is right tilted)." The stimuli were randomly sampled from two bivariate normal distributions that both had a mean spatial frequency of 6.25 cpd with an SD of 2.14 cpd. Category A stimuli had a mean orientation of 95.8º with an SD of 3.1º. Category B stimuli had a mean orientation of 84.2º with an SD of 3.1º. The orientation and spatial frequency of the Gabors were uncorrelated. Forty Category A and 40 Category B stimuli were generated. A schematic representation of the category structure is depicted in Figure 5 (right panel). The secondary working memory tasks and experimental procedure were identical to those from Experiment 2A.

Results

Working Memory Task Performance

The mean percent correct was 95.8% in the visuospatial task (SD = 3.2%) and was 96.6% for the verbal task (SD = 2.9%). This difference was not statistically significant [t(46) = .898, p = .374]. These accuracies are very similar to those obtained in Experiment 2A.

Category Learning Performance

Distribution of scores, proportion of learners and mean proportion correct. The distribution of scores in the final block of trials deviated from normality [K–S D(72) = .373, p < .001], but unlike in Experiments 1 and 2A, it did not appear bimodal (see Figure 6). Rather, the vast majority of participants learned the correct categorization rule and performed at a high rate of accuracy (above .90 correct). In addition, there were no differences among groups in the proportion of learners (Figure 8, left panel). Only 2 participants in the control condition, 3 participants in the visual condition, and 4 participants in the verbal condition did not reach .65 proportion correct in the last block of trials. Block accuracies in Experiment 2B are presented in Figure 8 (right panel). As is apparent from the figure, there were small and non-
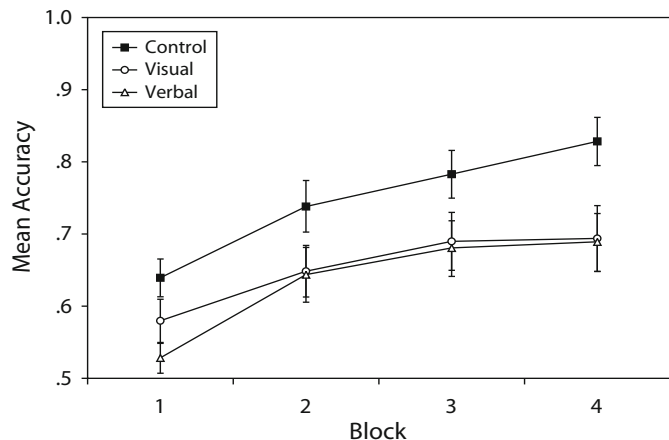

Figure 7. Left panel: Proportion of learners in Experiment 2A. Right panel: Mean accuracy for each condition in Experiment 2A. Error bars denote bootstrapped 68% confidence intervals (equivalent to a standard error of mean).

significant differences among the groups during the final block of trials [bootstrapped p(control = visual) = .270, p(control = verbal) = .232] and performance was high. There were performance differences in the initial block of trials. In the first block, performance in the visual condition dropped by .148 in comparison with the control condition [p(control = visual) = .004] and performance in the verbal condition dropped by .108 in comparison with the control condition [p(control = verbal) = .008].

A comparison of the control conditions from Experiment 2A and Experiment 2B suggests that cardinal orientations are special both perceptually *and* conceptually. The higher asymptotic accuracy for learners in Experiment 2B than for those in Experiment 2A (e.g., compare left and right panels in Figure 6) supports the notion that cardinal orientations are perceptually special. A comparison of early control performance in Experiment 2A and Experiment 2B (e.g., compare Block 1 in the right panel of Figure 7 with Block 1 in the right panel of Figure 8) supports the notion that cardinal orientations are also conceptually special and constitute a highly salient categorization rule. High accuracy in the control condition starting in Block 1 suggests that a majority of the participants selected the correct categorization rule very early on.

In Experiments 1 and 2A, by the end of training the secondary working memory task led primarily to a decrease in the proportion of learners, compared with the proportion in the control condition. This pattern did not replicate in Experiment 2B, because the proportion of learners was high and approximately equal in all conditions by the end of training. Because there were no differences in the proportion of learners among the conditions and almost every participant learned the rule, we turned to an alternative performance measure that is more sensitive under these conditions and is often used in categorization research: the number of trials needed to reach an accuracy criterion (*trials to criterion*). Because the vast majority of participants in Experiment 2B succeeded in learning the task, using trials to criterion is suitable to characterize the speed of learning in all conditions. The trials-to-criterion measure is not suitable in Experiment 1 and Experiment 2A, because half or more of the participants in the visual and verbal working memory conditions failed to reach any reasonable learning criterion, whereas the majority of participants in the control conditions did.

Trials to criterion. As one trials-to-criterion measure, we recorded the trial on which each participant reached or exceeded .65 correct (the learning criterion used in the previous experiments) over the last 80 trials. These analyses are presented in Figure 9 (C6580; i.e., criterion of .65 correct over previous 80 trials). Participants in the visuospatial working memory condition needed significantly more trials to reach the criterion [bootstrapped p(control = visual) = .013] than did participants in the control condition, whereas there was no difference in the trials-to-criterion between the verbal working memory group and the control group [bootstrapped p(control = verbal) = .859]. Because a large proportion of participants reached the .65 correct criterion right at trial number 80 (i.e., discovered the rule during the first 80 trials of the experiment), we examined a number of other performance criteria and smaller window sizes to ensure that the results were robust. We typically found a large and significant effect of the visuospatial working memory task and a small and nonsignificant effect of the verbal working memory task relative to the control condition. For example, for the criterion of .75 correct over the last 40 trials (Figure 9, C7540), we found that the visuospatial working memory group needed on average 48 more trials than the control group to reach the criterion [bootstrapped p(control = visual) = .002], whereas the verbal working memory group needed on average only 12 more trials did than the control group [bootstrapped p(control = verbal) = .127].[4]

Brief Summary

To summarize, we examined visuospatial and verbal working memory effects in rule-based category learning when the criterion to be learned was on a cardinal orientation. The results differed from those obtained for a formally identical task that used an oblique orientation as the
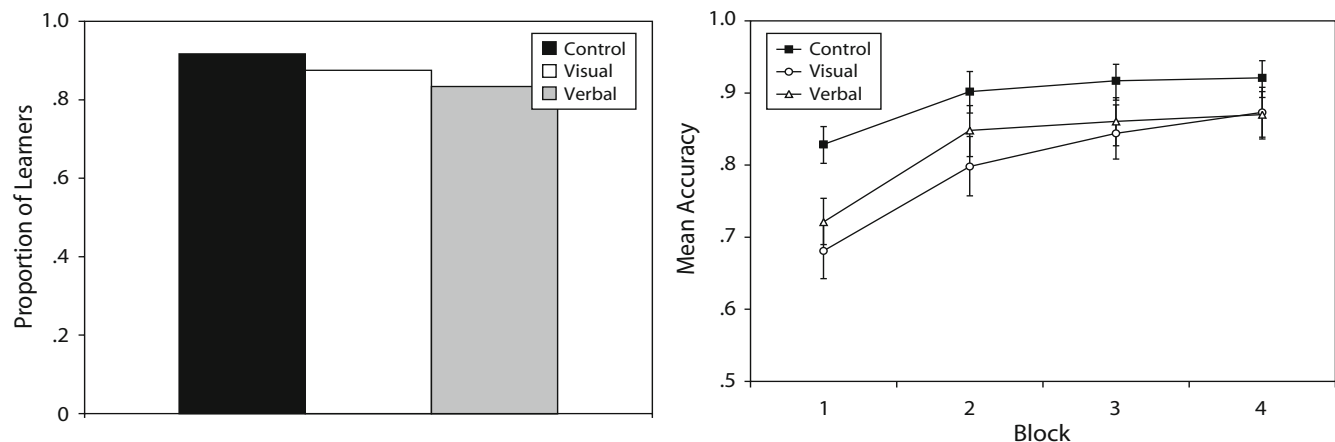

Figure 8. Left panel: Proportion of learners in Experiment 2B. Right panel: Mean accuracy in Experiment 2B. Error bars denote bootstrapped 68% confidence intervals (equivalent to a standard error of mean).
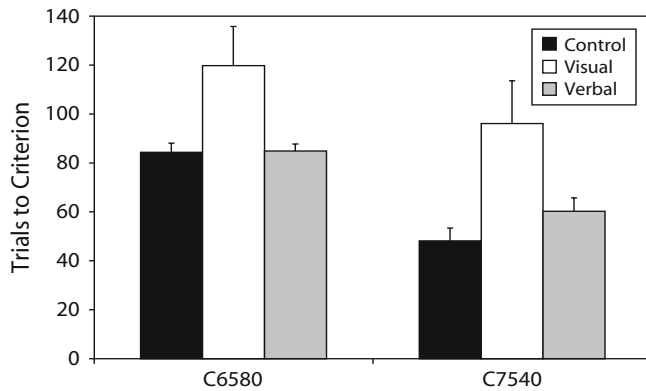
Figure 9. Mean number of trials to reach the criterion of .65 correct over last 80 trials (C6580; left half) and the criterion of .75 correct over last 40 trials (C7540; right half) in Experiment 2B. Error bars denote 68% confidence interval (equivalent to a standard error of mean).

criterion. We found that the task was much easier to learn and that neither the visuospatial nor the verbal working memory task had a significant effect on the proportion of learners or final (asymptotic) accuracy. We found adverse effects of both the visuospatial and the verbal working memory tasks on mean accuracy during the early stages of learning; but we also found a significant effect of the visuospatial but not the verbal working memory task on the speed of learning, as measured by the mean number of trials to reach an accuracy criterion.

How do these results address the hypotheses outlined in the introduction? First, the control data support the hypothesis that cardinal orientations are special both perceptually (leading to a higher asymptotic accuracy for learners) and conceptually (leading to high proportion of learners, even early in the experiment). Second, the difference between the visuospatial and verbal working memory effects suggests that they may affect different processes associated with rule-based learning. The minimal effect of the verbal working memory task fits with the hypothesis that cardinal orientations are conceptually special, as supported by the control data. The high salience of the cardinal orientation rule leads participants to select the rule early in learning, bypassing much of the working memory demanding hypothesis-testing process. The effect of the visuospatial working memory task was larger and significantly affected the speed of learning. These results would not be likely if the visuospatial working memory task were primarily affecting domain nonspecific resources (central executive), because the results would be similar to those from the verbal condition. In addition, these results would not be likely if visuospatial working memory were only involved in learning and representation of the optimal categorization criterion. Because the cardinal orientation seems conceptually special (from the control and verbal-condition data), the cardinal orientation criterion does not need to be learned in the same way as does an oblique orientation, and thus should require minimal visuospatial working memory resources. Rather, the results seem in accordance with the hypothesis

that the visuospatial working memory task disrupts the analytic perception of stimuli and causes the participant to take longer to notice the variation of the stimuli around the cardinal orientation. However, as the first attempt to address the visuospatial working memory role in category learning, this notion needs to be taken as a working hypothesis only. We elaborate on this possibility below.

GENERAL DISCUSSION

Many categorization theories assume that an effortful, working memory demanding process of hypothesis testing is involved in at least some types of category learning (e.g., Ashby et al., 1998; Bruner et al., 1956; Erickson & Kruschke, 1998; Feldman, 2000; Nosofsky et al., 1994). Previous research (Maddox et al., 2004) and our Experiment 1 showed that a sequentially presented verbal or visuospatial working memory task disrupt rule-based learning, confirming the existence of the effortful hypothesis-testing–based category learning. Importantly, Maddox et al. (2004) and Experiment 1 also demonstrated that effortful, working memory–hypothesis testing is not the only existing process of category learning, because information-integration category learning was not affected at all by a secondary verbal (Maddox et al., 2004) or visuospatial task (present Experiment 1). The lack of the visuospatial working memory task effect on information-integration category learning is nontrivial, as both tasks use visual stimuli and are thought to rely on the basal ganglia (information-integration/visual discrimination, Filoteo et al., 2005; Maddox & Filoteo, 2001; Packard & McGaugh, 1992; Poldrack, 2002; visuospatial working memory, Lawrence et al., 1996; Lawrence et al., 2000; Postle et al., 1997). On the basis, for example, of the COVIS theory of category learning (Ashby et al., 1998), we propose that visuospatial working memory relies on the head of the caudate nucleus (Levy et al., 1997), whereas information-integration category learning relies on the body and/or the tail of caudate nucleus (Nomura et al., 2007; Seger & Cincotta, 2005). The data from Experiment 1 indeed indicate that there is no interference between the two (at least with respect to the present tasks). Taken together, these results support the notion that both visuospatial and verbal working memory tasks impact the processes involved in the hypothesis-testing system that mediates rule-based category learning, but that neither visuospatial nor verbal working memory is crucial for the procedural system that mediates information-integration category learning.

Experiment 2 tested the generality of working memory effects in rule-based learning by investigating the effects of a visuospatial and a verbal working memory task on rule-based category learning when the criterion was on an oblique orientation (70°) or on the cardinal orientation (90°) of a Gabor stimulus, instead of on spatial frequency. Experiment 2 demonstrated that not all rule-based categories are treated equally. When the criterion was on an oblique orientation, the results replicated those from Maddox et al. and Experiment 1, for which the criterion was

on spatial frequency. When the criterion was on a cardinal orientation, we found faster learning with a higher asymptotic accuracy in the control condition. This result confirmed that cardinal orientations are perceptually special (as has been previously established; see, e.g., Campbell & Kulikowski, 1966; Furmanski & Engel, 2000). Furthermore, this result suggested that cardinal orientations are also conceptually special; that is, cardinal orientations constitute salient, spontaneously used categorization boundaries that create intuitive concepts (e.g., right-tilted vs. left-tilted in the present experiment). Additionally, when the criterion was on a cardinal orientation, we found significantly slower learning when the visuospatial task was present, but minimally slower learning when the verbal working memory task was present. These results suggest that visuospatial and verbal working memory effects on rule-based learning may be due to dissociable mechanisms. In the remainder of this discussion, we will discuss these issues in more detail.

Category Structure Effects on the Distribution of Accuracy Scores

One important result observed in the present studies was the existence of qualitatively different distributions of accuracy scores for the rule-based and the information-integration category structures (see Figure 3). Whereas the accuracy scores in the information-integration conditions were normally distributed, a bimodal distribution was observed in the rule-based conditions, with one modus at chance and a second modus at a high level of accuracy. This pattern held whether a secondary working memory demanding task was present or not, and it replicates the pattern observed in Zeithamova and Maddox (2006) that used a dual-task procedure.

These findings provide evidence that rule-based and information-integration category learning is mediated by separate systems. We argue that rule-based category learning is mediated by a hypothesis-testing system whose processing is effortful and attention demanding. Hypothesis-testing systems of this sort are known to have an all-or-none characteristic to their learning that has been studied since the 1960s (Bower & Trabasso, 1963; Trabasso & Bower, 1964). The bimodality observed in our rule-based score distributions is consistent with this hypothesis. We also argue that information-integration category learning is not mediated by a hypothesis-testing system, but is mediated instead by a procedural-based system, whose processing is automatic and does not require attention. The fact that the information-integration score distributions were normally distributed and were not affected by a secondary working memory task follows from this hypothesis.

Working Memory Task Effects on Rule-Based Accuracy Score Distributions

As outlined earlier, the results from Experiment 1, along with those from Maddox et al. (2004), suggest that a sequential verbal or visuospatial working memory task has no effect on the distribution of information-integration scores, but has a large effect on the distribution of rule-based scores. The same effect on rule-based learning holds in Experiment 2A that focused on a rule with an oblique orientation criterion. In this section, we elaborate on the nature of the working memory effect on rule-based learning, leaving a discussion of the "cardinal orientation" results from Experiment 2B for later.

As outlined in the previous section, the distributions of accuracy scores in the rule-based conditions were bimodal, with one modus at chance (nonlearners, .50 correct) and a second modus at a high level of accuracy (learners, .80 correct in Experiment 1 and .90 correct in Experiment 2A). One of the most interesting findings from the present study was the fact that the accuracy achieved by learners remained constant across the control and working memory conditions. Rather, the effect of the working memory task was to increase the proportion of participants who fell under the chance (nonlearner) modus relative to the proportion of participants who fell under the higher accuracy (learner) modus. Thus, both the visuospatial and verbal working memory tasks seemed to disrupt primarily the process of rule discovery, leading fewer participants to discover the correct rule. However, once the correct rule was discovered, participants were just as accurate in applying the rule as were learners in the control condition. This finding is important, especially given the general focus on learning curves in category learning research. The typical interpretation of a performance deficit is to assume a "shallowing" of the learning curve and thus a general effect on "average" performance. What the present data suggest is that the effect is not general, but rather increases the probability that a participant will fail to discover the correct rule.

Interaction Between the Hypothesis Testing and Procedural System Under Secondary Task

One reasonable prediction from the COVIS model (Ashby et al., 1998), and perhaps also an implicit expectation in other multiple-system models of learning, is that the procedural system would take over and dominate rule-based category learning when the hypothesis-testing system is disrupted by a working memory task. COVIS assumes that both category learning systems (the hypothesis-testing system and the procedural system) aim to learn every categorization task and that the two systems compete to generate the response on each trial. If the procedural system is unaffected by the presence of a working memory task, it should dominate the hypothesis-testing system. This did not seem to be the case in the present studies, because we found no compensation in the rule-based task by the procedural system—the accuracy scores remained low and were not normally distributed—when the working memory task was present. In our previous study (Zeithamova & Maddox, 2006), we similarly found that participants learning a more complex conjunctive rule-based categorization task under the dual Stroop-task interference tended to use (rule-based) unidimensional strategies, or guessing, rather than converting to information-integration strategies.

There are at least two possible explanations for this finding. First, the rule-based category structures used in the present experiments may be unfavorable for learning in the procedural system; for instance, because of their relatively high intracategory variability and low intercategory variability. Second, the procedural system may be learning the task, but the participant may be highly biased towards the (unsuccessful) hypothesis-testing system by the secondary working memory task. COVIS assumes that there is an initial bias toward the hypothesis-testing system. Participants may never abandon this bias under the secondary working memory task conditions, perhaps also because processing of the feedback regarding each system's performance is compromised. Given the lack of research that directly examines system level interactions, these hypotheses should be considered speculative at this time.

Dissociating Visuospatial and Verbal Working Memory Effects on Rule-Based Category Learning

The phenomena of visuospatial and verbal working memory are behaviorally and neurally dissociable (e.g., Baddeley, 1995; Goldman-Rakic, 1998; Shah & Miyake, 1996). It is thus reasonable to consider the possibility that, although both visuospatial and verbal working memory tasks adversely affect rule-based learning, the locus of their effect may differ. The effect of the verbal working memory task on rule-based category learning reported in Maddox et al. (2004) and replicated in Experiment 2A would be expected by any hypothesis-testing model. The effect of the visuospatial working memory task on rule-based category learning observed in the present studies is less straightforward, because no existing category learning theory addresses the possible role of visuospatial working memory. We speculated that the observed visuospatial task effect in rule-based category learning may act indirectly, via some kind of general attention or control mechanism common to both visuospatial and verbal working memory

(e.g., central executive in Baddeley & Logie, 1999); or it may be mediated by a different, independent mechanism. Two pieces of evidence argue against the notion of an indirect effect via commonly shared resources and instead support the notion of two different mechanisms for visuospatial and verbal working memory task effects. We will now discuss these two pieces of evidence in more detail.

Comparison of Maddox et al. (2004) and Experiment 1. Although qualitatively similar, the verbal working memory task effect reported in Maddox et al. (2004) seem to differ in magnitude from the visuospatial working memory task effect found in Experiment 1. To compare the working memory effects on rule-based learning across the two experiments, we computed effect sizes (Cohen, 1988)[5] for the final block performance drop in each experimental condition compared with the associated control condition (Figure 10, left panel). These analyses should be interpreted with caution bearing in mind that Cohen's effect size measure is derived from means, and that the means in these experiments represent a relative mixture of two populations of participants rather than an average participant. Nevertheless, these analyses are suggestive and seem to shed some light on the nature of the working memory effects. As is apparent from Figure 10, left panel, the verbal working memory task had a large effect immediately following categorization feedback, but only a small effect when participants were first allowed to process the categorization feedback for 2,500 msec. The visuospatial working memory task had an intermediate effect immediately following categorization feedback, but continued to disrupt performance even when the participant was first allowed to process the categorization feedback for 2,500 msec. Although by itself this finding is inconclusive, it favors the notion that verbal and visuospatial working memory tasks have at least partially dissociable effects on rule-based category learning.

The effect of the cardinality of a categorization criterion. One goal of this article was to investigate the
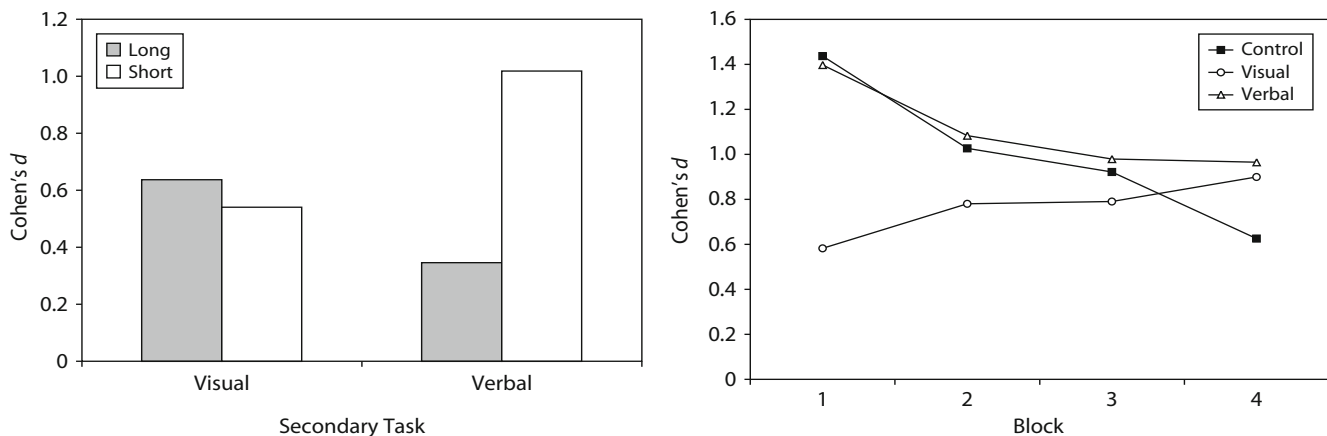

Figure 10. Left panel: Effect size (rule-based category learning performance decrement in Cohen's *d*) of a secondary visuospatial or verbal working memory task following categorization task feedback after 2,500-msec delay ("long," black bars) or immediately ("short," gray bars). Data for visuospatial working memory effects are from Experiment 1, data for verbal working memory effects were computed from Maddox et al. (2004). Right panel: Effect size of moving criterion from 70° to 90°. Data were obtained by comparing Experiment 2A and Experiment 2B.

effect of cardinality of a criterion on rule-based category learning. We asked whether cardinal orientations were conceptually special and whether working memory effects, observed with a general (arbitrary) criterion rule-based learning, replicated for a cardinal criterion. The pattern of results for the cardinal criterion exhibited several differences from those observed when the criterion was on an oblique orientation or on spatial frequency. First, in the control condition, performance reached asymptote much earlier and at a higher proportion correct for learners than for the formally identical structure using a criterion on oblique orientation. The results suggest that a criterion on cardinal orientation is easier to learn than a criterion on an oblique orientation—not only perceptually, but also conceptually. It seems that cardinal orientation plays a role in cognition as an intuitive, salient categorization rule. Second, the speed of learning was adversely affected only by the visuospatial working memory task, with minimal effect of the verbal working memory task. This finding implies that the salience of the cardinal criterion may be occluded by a secondary visuospatial working memory task, perhaps because it disrupts memory traces of previously seen stimuli. In Figure 10 (right panel), we accentuate this by examining the effect of the cardinality of a criterion throughout the learning process. For each block and each condition, we computed the effect size (Cohen's d) of changing the categorization criterion from 70º to 90º (i.e., the difference between scores in Experiment 2A and Experiment 2B). The advantage for the cardinal orientation was large, mainly early in the experiment, and of about the same magnitude for the control and the verbal condition. On the other hand, the advantage for the cardinal orientation was much smaller for the visual condition, especially early in the learning, consistent with the trials-to-criterion analysis. This finding further contributes to the notion that visuospatial and verbal working memory effects in rule-based category learning may be due to different mechanisms. Next we discuss the possible role of verbal and visuospatial working memory in category learning separately.

The role of verbal working memory in category learning. Previous research implicated verbal working memory in hypothesis generation, selection and testing (Ashby et al., 1998; Dougherty & Hunter, 2003). The presented results are all consistent with this notion. An effect of the verbal working memory task was observed when participants were presented with two generally equally salient dimensions and set of criteria along those dimensions, but not when a highly salient rule was offered in Experiment 2B. The minimal effect of the verbal working memory task on cardinal orientation rule learning suggests that cardinal orientations are special not only perceptually, but also conceptually (i.e., they serve as highly salient natural categorization rules). Specifically, the values along the orientation dimension seem to naturally fall into two distinct classes (right-tilted stimuli or left-tilted stimuli) when orientation varies around a cardinal direction. Thus, verbal working memory may not be needed when selecting and applying a highly salient rule, because much of the effortful hypothesis-testing process is skipped.

The role of visuospatial working memory in category learning. What may be the possible role of visuospatial working memory in rule-based category learning? One notion outlined in the introduction was that visuospatial working memory may be needed to represent the optimal categorization *criterion*, while the verbal working memory may be needed to represent the optimal *rule*. In other words, visuospatial working memory may be important in representing the criterion of 70º (in Experiment 2A) or 90º (in Experiment 2B), but may not be needed to represent the rule "Respond A if the orientation is greater than the criterion, and respond B if the orientation is less than the criterion." The proportion of learners is a measure of the proportion of participants that discover the correct rule. The specific accuracy of these learners depends on how well they learn and remember the optimal criterion; that is, if they can find a particular value of the orientation that best discriminates between the high and the low orientations. If visuospatial working memory is needed only to hold the optimal categorization criterion but is not crucial in the process of rule discovery, we would expect the proportion of learners in Experiment 1 and Experiment 2A to be about the same in the visual condition as in the control condition, but for their accuracy to be lower because of noisier categorization criterion representation. Contrary to this hypothesis, we actually observed a lower proportion of learners, but noted their accuracy to be about the same as in the control condition. In Experiment 2B, we would expect a minimal effect of the visuospatial working memory task, because representing a criterion on cardinal orientation—a highly learned natural boundary—should require minimal working memory resources. Contrary to this prediction, we found an adverse effect of the visuospatial task on the speed of learning.

Another notion was that visuospatial working memory may be needed for analytic evaluation of a stimulus and its individual dimensions. The present results are quite consistent with this notion. If the visuospatial working memory task disrupts analytical perception of the stimuli, identifying the dimensional structure of the stimulus, and thus generating possible rules in Experiment 1 and Experiment 2A, might be more difficult. Disrupting analytic processing of the stimuli may also disrupt the perception of variation around a cardinal orientation, perhaps by disrupting the visual memory representation of previous stimuli and thus delaying the time when the rule on cardinal orientation may be selected. On the other hand, the secondary visuospatial working memory task had no effect on information-integration category learning in Experiment 1, because information-integration category learning requires holistic rather than analytic perception of a stimulus. Although consistent with the presented results, the proposed role of visuospatial working memory in analytic evaluation of categorization stimuli remains speculative until further research addresses this issue.

Summary

The results presented in this article extend our understanding of the role of working memory in category learning by examining the effects of sequentially presented

visuospatial and verbal working memory tasks on rule-based and information-integration category learning. In line with the results from Maddox et al. (2004) that used a verbal working memory task, we found no effect of a visuospatial working memory task on information-integration category learning, but a significant effect on rule-based category learning when the categorization criterion was on the spatial frequency of a Gabor stimulus. We also replicated the effect of both visuospatial and verbal working memory tasks on rule-based learning with a categorization criterion on an oblique orientation of a Gabor stimulus. These results add to the evidence for the existence of multiple category learning mechanisms. When examining the effect of the secondary tasks on rule-based learning in more detail, we interestingly found that the presence of both visuospatial and verbal working memory tasks affected primarily the proportion of participants who discovered the rule by the end of training; the accuracy of those learners remained the same across conditions. A different pattern of working memory effects was observed when the rule-based categorization criterion was on a cardinal orientation. We found a minimal effect of the verbal working memory task, but a large effect of the visuospatial working memory task on the speed of learning. These results suggest that the cardinal orientation serves as a highly salient, natural categorization boundary, and that visuospatial and verbal working memory effects on rule-based category learning are at least partially dissociable. A plausible role for visuospatial working memory consistent with the presented results is analytic evaluation of individual stimulus dimensions.

AUTHOR NOTE

REFERENCES

ALEXANDER, G. E., DeLONG, M. R., & STRICK, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience*, **9**, 357-381.

ANDERSON, J. R., & BETZ, J. (2001). A hybrid model of categorization. *Psychonomic Bulletin & Review*, **8**, 629-647.

ARON, A. R., SHOHAMY, D., CLARK, J., MYERS, C., GLUCK, M. A., & POLDRACK, R. A. (2004). Human midbrain sensitivity to cognitive feedback and uncertainty during classification learning. *Journal of Neurophysiology*, **92**, 1144-1152.

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., & ENNIS, J. M. (2006). The role of the basal ganglia in category learning. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 46, pp. 1-36). New York: Elsevier.

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

BADDELEY, A. D. (1995). Working memory. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 755-764). Cambridge, MA: MIT Press.

BADDELEY, A. D., & LOGIE, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28-61). New York: Cambridge University Press.

BAR-GAD, I., MORRIS, G., & BERGMAN, H. (2003). Information processing, dimensionality reduction and reinforcement learning in the basal ganglia. *Progress in Neurobiology*, **71**, 439-473.

BOWER, G., & TRABASSO, T. (1963). Reversals prior to solution in concept identification. *Journal of Experimental Psychology*, **66**, 409-418.

BRAINARD, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, **10**, 433-436.

BROWN, V. J., DESIMONE, R., & MISHKIN, M. (1995). Responses of cells in the tail of the caudate nucleus during visual discrimination learning. *Journal of Neurophysiology*, **74**, 1083-1094.

BRUNER, J. S., GOODNOW, J., & AUSTIN, G. (1956). *A study of thinking*. New York: Wiley.

CAAN, W., PERRETT, D. I., & ROLLS, E. T. (1984). Responses of striatal neurons in the behaving monkey: II. Visual processing in the caudal neostriatum. *Brain Research*, **290**, 53-65.

CAMPBELL, F. W., & KULIKOWSKI, J. J. (1966). Orientational selectivity of the human visual system. *Journal of Physiology*, **187**, 437-445.

COCCHINI, G., LOGIE, R. H., DELLA SALA, S., MacPHERSON, S. E., & BADDELEY, A. D. (2002). Concurrent performance of two memory tasks: Evidence for domain-specific working memory systems. *Memory & Cognition*, **30**, 1086-1095.

COHEN, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

DOUGHERTY, M. R. P., & HUNTER, J. E. (2003). Hypothesis generation, probability judgment, and individual differences in working memory capacity. *Acta Psychologica*, **113**, 263-282.

EFRON, B., & TIBSHIRANI, R. J. (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.

FELDMAN, J. (2000). Minimization of Boolean complexity in human concept learning. *Nature*, **407**, 630-632.

FILOTEO, J. V., MADDOX, W. T., SALMON, D. P., & SONG, D. D. (2005). Information-integration category learning in patients with striatal dysfunction. *Neuropsychology*, **19**, 212-222.

FREEDMAN, D. J., RIESENHUBER, M., POGGIO, T., & MILLER, E. K. (2003). A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *Journal of Neuroscience*, **23**, 5235-5246.

FURMANSKI, C. S., & ENGEL, S. A. (2000). An oblique effect in human primary visual cortex. *Nature Neuroscience*, **3**, 535-536.

FUSTER, J. M. (1989). *The prefrontal cortex: Anatomy, physiology, and neuropsychology of the frontal lobe* (2nd ed.). New York: Raven Press.

GOLDMAN-RAKIC, P. S. (1987). Circuitry of the prefrontal cortex and the regulation of behavior by representational knowledge. In F. Plum (Ed.), *Higher functions of the brain* (Handbook of Physiology: Section 1. Vol. 5, pp. 373-417). Bethesda, MD: American Physiological Society.

GOLDMAN-RAKIC, P. S. (1998). The prefrontal landscape: Implications of functional architecture for understanding human mentation and the central executive. In A. C. Roberts, T. W. Robbins, & L. Weiskranz (Eds.), *Prefrontal cortex: Executive and cognitive functions* (pp. 87-102). New York: Oxford University Press.

HEELEY, D. W., & TIMNEY, B. (1988). Meridional anisotropies of orientation discrimination for sine wave gratings. *Vision Research*, **28**, 337-344.

HIKOSAKA, O., SAKAMOTO, M., & USUI, S. (1989). Functional properties of monkey caudate neurons: III. Activities related to expectation of target and reward. *Journal of Neurophysiology*, **61**, 814-832.

HUTTENLOCHER, J., HEDGES, L. V., & DUNCAN, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review*, **98**, 352-376.

JAHANSHAHI, M., BROWN, R. G., & MARSDEN, C. D. (1992). The effect of withdrawal of dopaminergic medication on simple and choice reaction time and the use of advance information in Parkinson's disease. *Journal of Neurology, Neurosurgery, & Psychiatry*, **55**, 1168-1176.

JONIDES, J., REUTER-LORENZ, P. A., SMITH, E. E., AWH, E., BARNES, L. L., DRAIN, M., ET AL. (1996). Verbal and spatial working memory

in humans. In D. L. Medin (Ed.), *Psychology of learning and motivation: Advances in research and theory* (Vol. 35, pp. 43-88). San Diego: Academic Press.

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

LAWRENCE, A. D., SAHAKIAN, B. J., HODGES, J. R., & ROSSER, A. E. (1996). Executive and mnemonic functions in early Huntington's disease. *Brain: A Journal of Neurology*, **119**, 1633-1645.

LAWRENCE, A. D., WATKINS, L. H. A., SAHAKIAN, B. J., HODGES, J. R., & ROBBINS, T. W. (2000). Visual object and visuospatial cognition in Huntington's disease: Implications for information processing in corticostriatal circuits. *Brain: A Journal of Neurology*, **123**, 1349-1364.

LEVY, R., FRIEDMAN, H. R., DAVACHI, L., & GOLDMAN-RAKIC, P. S. (1997). Differential activation of the caudate nucleus in primates performing spatial and nonspatial working memory tasks. *Journal of Neuroscience*, **17**, 3870-3882.

LOGIE, R. H., ZUCCO, G. M., & BADDELEY, A. D. (1990). Interference with visual short-term memory. *Acta Psychologica*, **75**, 55-74.

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**, 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 perceptual 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., & 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.

NOMURA, E. M., MADDOX, W. T., FILOTEO, J. V., ING, A. D., GITELMAN, D. R., PARRISH, T. B., ET AL. (2007). Neural correlates of rule-based and information-integration visual category learning. *Cerebral Cortex*, **17**, 37-43.

NOSOFSKY, R. M., PALMERI, T. J., & MCKINLEY, S. C. (1994). A rule-plus-exception model of classification learning. *Psychological Review*, **101**, 53-79.

ORBAN, G. A., VANDENBUSSCHE, E., & VOGELS, R. (1984). Human orientation discrimination tested with long stimuli. *Vision Research*, **24**, 121-128.

PACKARD, M. G., & KNOWLTON, B. J. (2002). Learning and memory functions of the basal ganglia. *Annual Review of Neuroscience*, **25**, 563-593.

PACKARD, M. G., & MCGAUGH, J. L. (1992). Double dissociation of fornix and caudate nucleus lesions on two radial maze tasks: 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.

PELLI, D. G. (1997). The Video Toolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, **10**, 437-442.

POLDRACK, R. A. (2002). Neural systems for perceptual skill learning. *Behavioral & Cognitive Neuroscience Reviews*, **1**, 76-83.

POSNER, M. I., & PETERSEN, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, **13**, 25-42.

POSTLE, B. R., JONIDES, J., SMITH, E. E., & CORKIN, S. (1997). Spatial, but not object, delayed response is impaired in early Parkinson's disease. *Neuropsychology*, **11**, 171-179.

RAGHAVACHARI, S., KAHANA, M. J., RIZZUTO, D. S., CAPLAN, J. B.,

KIRSCHEN, M. P., BOURGEOIS, B., ET AL. (2001). Gating of human theta oscillations by a working memory task. *Journal of Neuroscience*, **21**, 3175-3183.

REBER, P. J., GITELMAN, D. R., PARRISH, T. B., & MESULAM, M. M. (2003). Dissociating explicit and implicit category knowledge with fMRI. *Journal of Cognitive Neuroscience*, **15**, 574-583.

SCHULTZ, W., APICELLA, P., ROMO, R., & SCARNATI, E. (1994). Context-dependent activity in primate striatum reflecting past and future behavioral events. In J. C. Houk, J. L. Davis, & D. G. Beiser (Eds.), *Models of information processing in the basal ganglia* (pp. 11-27). Cambridge, MA: MIT Press.

SCHWARZ, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, **6**, 461-464.

SEGER, C. A., & CINCOTTA, C. M. (2005). The roles of the caudate nucleus in human classification learning. *Journal of Neuroscience*, **25**, 2941-2951.

SHAH, P., & MIYAKE, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General*, **125**, 4-27.

SMITH, E. E., JONIDES, J., & KOEPPE, R. A. (1996). Dissociating verbal and spatial working memory using PET. *Cerebral Cortex*, **6**, 11-20.

SMITH, E. E., PATALANO, A. L., & 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.

STERNBERG, S. (1966). High-speed scanning in human memory. *Science*, **153**, 652-654.

TRABASSO, T., & BOWER, G. (1964). Presolution reversal and dimensional shifts in concept identification. *Journal of Experimental Psychology*, **67**, 398-399.

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

WICKENS, J. (1993). *A theory of the striatum*. New York: Pergamon.

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

WILLINGHAM, D. B., NISSEN, M. J., & BULLEMER, P. (1989). On the development of procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 1047-1060.

WILSON, C. J. (1995). *The contribution of cortical neurons to the firing pattern of striatal spiny neurons*. Cambridge, MA: MIT Press.

ZEITHAMOVA, D., & MADDOX, W. T. (2006). Dual-task interference in perceptual category learning. *Memory & Cognition*, **34**, 387-398.

NOTES

1. Bootstrapping is a statistical procedure that uses simulations from sample data (e.g., rather than assuming normality) for obtaining estimates of reliability or error, such as confidence intervals.

2. The tested null hypothesis is mean1 − mean2 = 0. All bootstrapped p values for the difference in means reflect a conservative, two-tailed hypothesis and may be converted to one-tailed values when divided by 2.

3. The learner criterion of .7, which is the average of the .5 chance score distribution and .9 high performing score distribution in Experiment 2A, gives identical results, because no participant had final block accuracy between .65 and .70 (see Figure 6, left panel).

4. The general pattern of results remains similar even if we exclude outliers (so the effect is not driven by extreme values) or include non-learners by assigning 320 as their trials to criterion, which is the maximum possible value.

5. Effect size measures the magnitude of an effect independent of sample size (number of participants in a study). A frequently used measure of effect size is Cohen's d, which is computed as the difference in means divided by the pooled standard deviation (d'). As a rule of thumb, d above .2 is considered a small effect, d above .5 is considered a medium effect, and d above .8 is considered a large effect (Cohen, 1988).

APPENDIX
Description of the Model Fitting Procedure Applied to the
Score Distributions From Experiments 1 and 2A

A visual inspection of the accuracy score distributions from the Experiment 1 and 2A rule-based groups suggested that these distributions are bimodal (see Figure 3, left panel and Figure 6, left panel). In this appendix, we describe a method for characterizing the nature of these bimodal distributions and the effect that the verbal and visuospatial working memory tasks had on these distributions.

We fit a series of models of various degrees of generality to the distribution of accuracy scores from the final block of trials, separately for each condition. After best-fitting parameters for each model were estimated using maximum likelihood method, we used the BIC measure (Schwarz, 1978) to compare the models and determine the best-fitting model.

The following five fitted models were used.

Model 1 is the most general model. It assumes that the score distribution is bimodal, with each modus being best described by a normal distribution. This model has five free parameters: the mean and *SD* of the first normal distribution, the mean and *SD* of the second distribution, and the relative weight of the second distribution (with the relative weight of the first distribution being 1 − relative weight of the second distribution). The relative weight of each distribution represents the proportion of participants whose accuracy scores contributed to that distribution (modus).

Model 2 is a special case of Model 1, for which the mean of one distribution is fixed at chance (.5). This model instantiates the hypothesis that one group of participants did not discover the correct rule and that their mean accuracy is at chance level of .5. This model has four free parameters.

Model 3 is a special case of Model 2, for which the chance distribution is assumed to be binomial (rather than normally distributed), with the probability of success being .5 on each of the 80 trials that constituted the final block. Because the *SD* of a binomial distribution is derived from the mean and the number of trials, there are no free parameters associated with the "chance distribution," leaving the three free parameters associated with the second distribution to be estimated.

Model 4 is a special case of Model 3, for which the nonchance distribution is also assumed to be binomial. This model has two free parameters: the mean and relative weight of the second distribution.

Model 5 was included as an additional check of the Kolmogorov–Smirnov test of normality. This model assumes that the distributions of scores are best characterized by one normal distribution. This model has two free parameters: the mean and *SD* of the normal distribution. This model provided poor fits in all cases.

Results

Best-Fitting Models in Experiment 1
RB control – Model 3:
Distribution 1: binomial (mean = .5, relative weight = .318)
Distribution 2: normal (mean = .814, sd = .065, relative weight = .692)
RB long – Model 4:
Distribution 1: binomial (mean = .5, relative weight = .560)
Distribution 2: binomial (mean = .788, relative weight = .440)
RB short – Model 4:
Distribution 1: binomial (mean = .5, relative weight = .512)
Distribution 2: binomial (mean = .786, relative weight = .488)

Best-Fitting Models in Experiment 2A
Control – Model 3:
Distribution 1: binomial (mean = .5, relative weight = .167)
Distribution 2: normal (mean = .897, *SD* = .046, relative weight = .833)
Visual – Model 3:
Distribution 1: binomial (mean = .5, relative weight = .500)
Distribution 2: normal (mean = .907, *SD* = .058, relative weight = .500)
Verbal – Model 2:
Distribution 1: normal (mean = .5, *SD* = .032, relative weight = .500)
Distribution 2: normal (mean = .879, *SD* = .058, relative weight = .500)