Perception & Psychophysics
2006, 68 (7), 1176-1190

Stimulus modality interacts with category structure in perceptual category learning

W. TODD MADDOX, A. DAVID ING, and J. SCOTT LAURITZEN
University of Texas, Austin, Texas

Two experiments were conducted that examined information integration and rule-based category learning, using stimuli that contained auditory and visual information. The results suggest that it is easier to perceptually integrate information within these sensory modalities than across modalities. Conversely, it is easier to perform a disjunctive rule-based task when information comes from different sensory modalities, rather than from the same modality. Quantitative model-based analyses suggested that the information integration deficit for across-modality stimulus dimensions was due to an increase in the use of hypothesis-testing strategies to solve the task and to an increase in random responding. The modeling also suggested that the across-modality advantage for disjunctive, rule-based category learning was due to a greater reliance on disjunctive hypothesis-testing strategies, as opposed to unidimensional hypothesis-testing strategies and random responding.

Human beings constantly combine information from different sensory modalities to guide their behavior. For example, when searching for a crying baby, a mother must rely on both audio and visual information and combine this information to locate the child. We refer to these types of problems as *perceptual categorization* problems, because humans must categorize their (noisy) percepts onto unique states of the environment to guide their behavior.

A large body of research has been devoted to the study of perceptual categorization across several modalities. A significant amount of work has focused on highly learned categories in visual object categorization (Goldstone, 1998; Lamberts, 2002; Newell & Bülthoff, 2002; Stankiewicz, 2002), auditory speech perception (Diehl & Kluender, 1987; Maddox, Molis, & Diehl, 2002; Nearey, 1997), olfaction (Ennis & Mullen, 1985), and tactile sensation (Klatzky & Lederman, 1995). Some work has also been devoted to understanding how information is integrated across modalities. One area that has been of special interest is speech perception by eye and ear (Massaro, 1987, 1998).

All of this work has focused on the nature of highly learned categories. In contrast, this report is about the perceptual categorization of novel, artificial categories and focuses on how these categories are *learned*. Category learning has been studied extensively, but in nearly all cases, the stimulus dimensions are processed within the same sensory modality (e.g., vision, Medin & Schaffer, 1978, and Estes, 1994; audition, Grau & Kemler Nelson, 1988; tactile, Klatzky, Lederman, & Reed, 1987, and Reed, Lederman, & Klatzky, 1990).

This study has four goals. The first goal is to extend research on category learning to situations in which stimuli are defined by information from different sensory modalities. Here, we focus on information from visual and auditory modalities and refer to these types of categorization conditions as *across-modality* conditions. The second goal is to compare across-modality category learning with *within-modality* category learning for both auditory and visual modalities. The third goal is to explore how within- and across-modality category learning interacts with the type of categorization problem: *information integration* or *rule based* (defined shortly). Finally, we relate the results to the predictions made by a neurobiological model of visual category learning called the competition between verbal and implicit systems model (COVIS; Ashby, Alfonso-Reese, Turken, & Waldron, 1998), which we have extended to the auditory and audiovisual domains.

In the next (second) section, we will introduce the information integration and rule-based category-learning tasks used in the present report. In addition, we will describe COVIS and briefly review the empirical evidence in support of COVIS. In the third section, we will extend COVIS to the auditory and across-modality domains and derive predictions from the extended model regarding information integration and rule-based category learning. We then will turn to the results from two experiments and conclude with some general remarks.

Category Structures and COVIS

In an information integration categorization problem, the optimal decision rule requires a *predecisional* com-

This research was supported in part by National Institutes of Health Grant R01 MH59196 to W.T.M. We thank Randy Diehl and Sarah Sullivan for help with auditory stimulus generation. We also thank Greg Ashby, Shawn Ell, and Vince Filoteo for their expertise regarding the neuroanatomy of the caudate nucleus. Finally, we thank Jeffrey Rouder and two anonymous reviewers for helpful comments on an earlier version of the manuscript. Correspondence concerning this article should be addressed to W. T. Maddox, Department of Psychology, University of Texas, 1 University Station A8000, Austin, TX 78712 (e-mail: maddox@psy.utexas.edu).

bination of the perceptual information (Ashby & Gott, 1988; Maddox & Ashby, 2004). For example, consider the scatterplot displayed in Figure 1A. Each point in the scatterplot denotes a stimulus with a unique value along dimensions x and y. The filled squares denote stimuli from Category A, and the open circles denote stimuli from Category B. Suppose that dimension x is the spatial frequency of a sine wave grating and dimension y is the spatial orientation of the grating. The broken line denotes the optimal decision bound. It has no verbal or rule-based analogue, because frequency and orientation are measured in very different units. Although one can certainly state the rule as "respond A if the spatial orientation is greater than the spatial frequency; otherwise, respond B," it is unclear how to interpret the term "greater than," since the dimensional values are measured in different units; so this type of decision rule makes no sense to naive participants.

The second type of category-learning problem is referred to as a *rule-based* category-learning problem. Because our interest is in participants' ability to combine information across dimensions during category learning, we focus on a two-dimensional rule-based category-learning problem that involves a disjunctive rule. In this problem, the participant must perform a *postdecisional* combination of information. A postdecisional combination requires that a separate decision first be made about

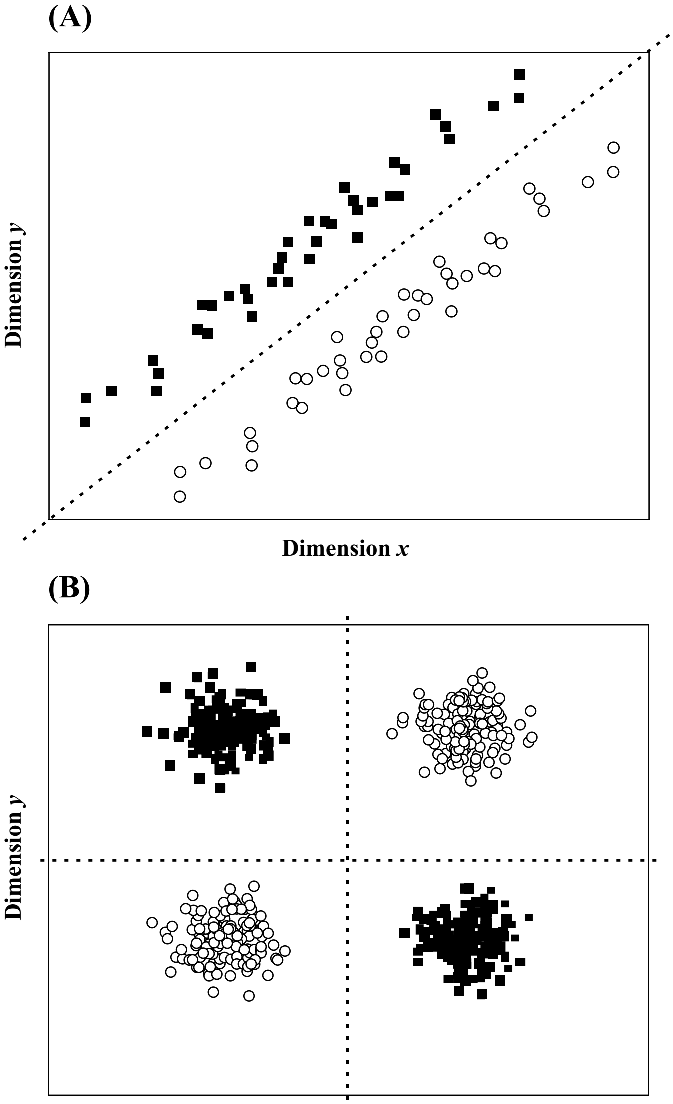

Figure 1. Scatterplots of the stimuli, along with the optimal decision bound from the (A) information integration and (B) disjunctive rule-based conditions. Filled squares denote stimuli from Category A. Open circles denote stimuli from Category B.

the level along each dimension and that these decisions be combined to determine a response. For example, in Figure 1B, the optimal decision rule requires the participant to determine whether dimension x is *low* or *high*, to determine whether dimension y is *low* or *high*, and then to combine those decisions to generate a response. Assuming, again, that the dimensions are spatial frequency and spatial orientation, the optimal bounds require the participant to use the following disjunctive rule: "Respond A if the spatial frequency is low *and* the spatial orientation is high or if the spatial frequency is high *and* the spatial orientation is low; otherwise, respond B." Note that this decision requires an overt, verbalizable rule that relies on separate categorization decisions for each dimension. Such a rule-based decision stands in stark contrast to the type of decision bound shown in Figure 1A.

Ashby et al.'s (1998) neurobiologically based multiple-system model of *visual* category learning, COVIS, postulates two systems that compete throughout learning. COVIS assumes that information integration category learning is dominated by an implicit, procedural-learning-based system that depends on a reward signal to strengthen the appropriate (stimulus–category) associations (Ashby et al., 1998; Ashby & Ell, 2001). With visual stimuli, the procedural-learning system is mediated largely within the *tail* of the caudate nucleus. High-level visual areas, such as the inferotemporal cortex, project directly to the tail of the caudate nucleus, with about 10,000 visual cortical synapses converging onto each medium spiny cell in the caudate (Van Hoesen, Yeterian, & Lavizzo-Mourey, 1981; Webster, Bachevalier, & Ungerleider, 1993; Wilson, 1995). These medium spiny cells then project to the prefrontal and premotor cortex (via the globus pallidus and thalamus; see, e.g., Alexander, DeLong, & Strick, 1986). Ashby and Waldron (1999) hypothesized that through a procedural-learning process, each unit in the tail of the caudate learns to associate a category label or behavioral response to activity in a large group of visual cortical cells that project onto that unit. It is important to understand that the projections from the visual cortex into the caudate nucleus are critical for learning in the procedural-based system and are localized in the tail. Recent fMRI results in humans support this claim (Aron et al., 2004; Seger & Cincotta, 2005).

COVIS also assumes that a different subsystem is involved in rule-based learning: an explicit hypothesis-testing system that involves working memory and executive attention. This system is mediated by a circuit that includes the anterior cingulate, the prefrontal cortex, and the head of the caudate nucleus. High-level visual representations (such as verbal–perceptual codes formed in the inferotemporal cortex) project into the prefrontal cortex and the head of the caudate nucleus and are thought to involve excitatory loops that travel through the thalamus, high-level perceptual cortices such as the inferotemporal cortex, and the prefrontal cortex (Alexander et al., 1986; Goldman-Rakic, 1995). COVIS claims that the anterior cingulate and the head of the caudate nucleus are instrumental in gating the inhibition of these loops.

Focusing exclusively on visually presented stimulus dimensions (e.g., the spatial frequency and spatial orientation of a sine wave grating), Ashby, Maddox, and colleagues have examined information integration and rule-based category learning to test a number of specific predictions made by COVIS (for a review, see Ashby & Maddox, 2005; Maddox & Ashby, 2004). One prediction is that the rule-based category-learning system is under conscious control and uses the resources of working memory and executive attention. As a result, the placement and timing of a feedback signal is unimportant in rule-based tasks, because relevant information can be kept alive by working memory, but only when working memory resources are available. In contrast, the procedural-learning-based system does not have access to working memory, and although it may be strongly affected by the timing of a feedback signal, it is relatively unaffected by the availability of working memory resources.

To test the prediction that only information integration, but not rule-based, learning should show specific deficits, Ashby, Maddox, and Bohil (2002) compared rule-based and information integration category learning across an observational training condition (in which the participants were informed before stimulus presentation of what category the ensuing stimulus was from) and a traditional feedback training condition (in which the category label followed the response). In a related study, Maddox, Ashby, and Bohil (2003; see also Maddox & Ing, 2005) compared rule-based and information integration category learning across an immediate feedback condition (in which corrective feedback was provided immediately following the response) and a delayed feedback condition (in which corrective feedback was delayed by 2.5, 5, or 10 sec following the response). In line with predictions from COVIS, observational training and delayed feedback negatively impacted information integration category learning but had little effect on rule-based category learning. Experimental manipulations known to affect performance in classic procedural-learning paradigms, such as the serial reaction time task, have also been found to affect information integration, but not rule-based, category learning. For example, Ashby, Ell, and Waldron (2003; see also Maddox, Bohil, & Ing, 2004) found that a switch of the category button response locations affected information integration, but not rule-based, category learning.

Analogously, experimental manipulations that increase working memory load affected rule-based, but not information integration, category learning. For example, when a working-memory-demanding task was completed simultaneously with categorization or immediately following presentation of the categorization feedback, only rule-based category learning was affected (Maddox, Ashby, Ing, & Pickering, 2004; Waldron & Ashby, 2001; Zeithamova & Maddox, 2006). This collection of studies, along with a number of studies in which brain-damaged patients were used (for a review, see Ashby & Maddox, 2005; Maddox & Filoteo, 2005), provides strong support for the neurobiological underpinnings proposed in COVIS.

Extending COVIS to Auditory and Cross-Modal Perceptual Category Learning

COVIS was originally proposed as a neurobiological model of *visual* category learning. In this section, we will extend COVIS into the domains of auditory and audiovisual information.

Extension of the procedural-learning-based system. COVIS assumes that efferent cortical projections from the visual cortex to the tail of the caudate nucleus (Van Hoesen et al., 1981; Webster et al., 1993; Wilson, 1995) onward to the prefrontal and premotor cortex via the globus pallidus and thalamus (Alexander et al., 1986) form the basis of procedural category learning of visual stimuli.

To extend COVIS's procedural-learning system, we need to examine the nature of the projections from the auditory cortex to the caudate nucleus and the nature of any audiovisual overlap. There have been more examinations of visual projections (e.g., Van Hoesen et al., 1981; Webster et al., 1993) than of auditory ones, but a general picture has been painted. First, the primary auditory cortex has limited direct connectivity with the caudate nucleus (Yeterian & Pandya, 1998). A similar paucity of connectivity holds between the primary visual cortex and the caudate (Kemp & Powell, 1970). Second, secondary auditory cortical areas do project to the caudate nucleus (Arnauld, Jeantet, Arsaut, & Demotes-Mainard, 1996; Hikosaka, Sakamoto, & Usui, 1989; Selemon & Goldman-Rakic, 1985; Yeterian & Pandya, 1998). However, unlike the localized nature of secondary visual cortical projections to the tail of the caudate nucleus, projections from secondary auditory areas to the caudate are more diffuse, projecting to the head, body, and tail (Arnauld et al., 1996; Yeterian & Pandya, 1998). For example, secondary auditory areas in the supratemporal plane project to the body but also to the tail of the caudate nucleus (Yeterian & Pandya, 1998). In addition, secondary auditory areas in the superior temporal gyrus also project to the body of the caudate but have projections to the head of the caudate as well (Selemon & Goldman-Rakic, 1985; Yeterian & Pandya, 1998). Third, there is some evidence for multisensory processing areas within the caudate (Chudler, Sugiyama, & Dong, 1995). In addition, the fact that at least some auditory projection areas project to the tail of the caudate suggests that there is likely some audiovisual overlap there.

It is important to recognize that not enough research has been conducted to describe fully the nature of auditory projections to the caudate and that more work on this topic is desirable. Even so, two points can be made. First, it seems clear that auditory projections to the caudate are more diffuse than are visual ones. Auditory projections to the body of the caudate appear to dominate, but projections to the head and tail also exist. On the other hand, visual projections are more focal and are confined to the tail of the caudate. Second, although some units in the caudate process audio and visual information, it is more common for the information to be segregated.

These two facts, along with the fact that only about 5% of the cells in the caudate are interneurons (Wilson, 1995), lead to one strong and one weaker prediction with respect to the procedural category-learning system. First, these facts lead to the strong prediction that within-modality information integration category learning should be easier than across-modality information integration category learning, because information is usually segregated by modality. If many cells in the caudate were interneurons, it would be possible that across-modality information integration category learning could be as good as (or even better than) within-modality information integration category learning. However, this would require there to be more interneurons connecting across-modality than within-modality caudate cells. With only 5% of the caudate being devoted to interneurons, this is simply not the case. Second, these facts lead to the weaker prediction that visual information integration should be better than auditory information integration. This prediction is weaker, because it is possible that the auditory features relevant to a specific information integration category-learning problem might be fairly segregated (say, within the body of the caudate). In addition, even if we were to find poorer auditory information integration category learning, it could be due to the fact that more of the sensory cortex is devoted to visual than to auditory processing. Regardless, the critical prediction is that within-modality auditory and visual information integration category learning should be superior to across-modality information integration category learning, and this is predicted directly from the architecture of the caudate nucleus.

Extension of the hypothesis-testing system. For visual stimuli, COVIS assumes that projections from higher perceptual cortices (e.g., the inferotemporal cortex) resonate with the prefrontal–thalamic loops to which they are connected (Alexander et al., 1986; Goldman-Rakic, 1995). This resonation is gated by the anterior cingulate and the head of the caudate to form the basis of category learning by hypothesis testing. This process is mediated by a visually deictic code that arises because cells in the inferotemporal cortex stand for visual meanings, such as *narrow*, *wide*, *steep*, and *shallow*. The hypothesis-testing system simply learns the relations between deictic symbols that predict category membership.

Extending COVIS's hypothesis-testing system into the domain of auditory processing is straightforward. Auditory information is sent from the auditory cortex to the prefrontal cortex (Goldman-Rakic, 1995) and the caudate nucleus (e.g., Nenadic et al., 2003), and these connections resonate to form the basis of an auditory deictic code. The auditory code contains symbols, such as *high pitch* and *low pitch*, and just like the visual code, it interfaces with the hypothesis-testing system via resonating loops. Thus, the simplest prediction is that there will be no difference in rule-based processing for within- and across-modality stimulus dimensions, since verbal codes are operated upon. On the other hand, it is possible that hypothesis testing will be easier for across-modality stimuli than for within-modality stimuli, because the across-modality stimulus dimensions are highly separable (Garner, 1974; Maddox, 1992; Shepard, 1957). Briefly, separable dimensions are thought to be easy to process separately, whereas integral dimensions are thought to be difficult to process separately. Although integrality–separability is currently thought to represent a continuum and not a dichotomy, in general, stimulus dimensions processed by different modalities fall more toward the separable than toward the integral end of the continuum (however, see work on synesthetically related cross-modal dimensions; Melara, 1989).

EXPERIMENT 1
Within- and Across-Modality Information Integration Category Learning

The goal of Experiment 1 was to examine information integration category learning when the stimulus dimensions were from the same or different modalities. Three experimental conditions were examined. These included a within-modality visual condition, a within-modality auditory condition, and an across-modality visual/auditory condition. In the within-modality visual condition, the stimuli were Gabor patches that varied across trials in spatial frequency and spatial orientation. In the within-modality auditory condition, the stimuli were auditory tones that varied in auditory frequency and auditory duration. In the across-modality condition, the stimuli were Gabor patches that varied across trials in spatial frequency (with spatial orientation fixed) and auditory frequency.

We used the randomization technique originally proposed by Ashby and Gott (1988). In the randomization technique, each category is represented by a normal distribution. In the present study, there are two categories, with each distribution being bivariate normally distributed because the stimuli are 2-D. The optimal decision bound can also be defined and, importantly, can be constructed in such a way that an information integration strategy is optimal (i.e., maximizes accuracy). Each category exemplar is generated by taking a random sample from the relevant bivariate normal distribution and constructing the associated stimulus. Because of the 2-D nature of the stimuli, each stimulus can be represented by a point in a 2-D stimulus space. The collection of stimuli used in a particular experimental condition can be displayed in a 2-D scatterplot. Scatterplots of the stimuli and optimal decision bound in arbitrary units can be found in Figure 1A. The category distribution parameters can be found in Table 1. The transformations from arbitrary units to visual and auditory dimension units will be provided in the Method section below.

As was outlined above, we predicted that information integration would be harder across modalities than within modality, because information is largely segregated by modality in the caudate nucleus. Using a model-based analysis technique (Ashby & Maddox, 1993; Maddox & Ashby, 1993), we also tested a hypothesis, suggested by other work (see Maddox & Ashby, 2004, for details), that the hypothesis-testing system would dominate behavior when the procedural-learning-based system performed poorly. To anticipate, both predictions were supported by the data.

Table 1
Category Distribution Parameters From the Information Integration Experiment

Category	μ_x	μ_y	σ_x	σ_y	cov_{xy}
A	.418	.578	.188	.188	.035
B	.578	.418	.188	.188	.035

Method

Participants and Design. Eighty-four participants were solicited from the University of Texas community and received course credit or $6 payment for participation. Each participant received a $1 bonus for accuracy exceeding 65% during the first half of the study (320 trials) and another $1 bonus for accuracy exceeding 70% during the last half of the study. Thirty participated in the within-modality-visual condition, 26 participated in the within-modality-auditory condition, and 28 participated in the across-modality condition.

Stimuli and Stimulus Generation. In the within-modality-visual condition, the relevant dimensions were the spatial frequency and spatial orientation of a 200 × 200 square-pixel Gabor patch viewed from a distance of approximately one arm's length on a 21-in. monitor with 1,360 × 1,024 resolution. The stimulus was presented for 1,000 msec. The arbitrary stimulus coordinates were converted to physical units, using the following transformations. Spatial frequency (in cycles/pixel) was converted using $f(x) = 0.08x + 0.02$. Spatial orientation (in radians) was converted using $f(x) = \pi x/2$. In the within-modality auditory condition, the relevant dimensions were the auditory frequency and duration of a narrowband white noise stimulus consisting of 50 independent sine waves that were randomly selected from a Gaussian distribution with a mean indicated by the arbitrary stimulus coordinate, x, and a standard deviation of 25 mel units. The *mel* (Stevens & Volkman, 1940) is a unit of pitch that is a function of frequency (in hertz) as defined by $f(x) = 700\,[\exp(x/1{,}127) - 1]$. The durations were converted to seconds by the transform $f(x) = 2\,(0.1 + 10^{x-1})$. In the across-modality condition, the relevant dimensions were the spatial frequency of a Gabor patch (with orientation fixed at 45º) and the auditory frequency of the narrowband white noise stimulus described above (with duration fixed at 1,000 msec). When an auditory dimension was present, the participant listened to the stimuli, using high-quality commercial headphones that surround the pinna of each ear. In all the conditions, each experimental session consisted of the presentation of eight, 80-trial blocks of trials. Each block contained 40 stimuli from Category A and 40 stimuli from Category B in random order.

Procedure. The participants were informed that there were two equally likely categories. They were told to emphasize accuracy over speed of responding. After viewing or hearing the stimulus on each trial, the participant pressed either an A or a B button to denote their categorization response; 750 msec of corrective feedback followed, consisting of the word "correct" or "wrong" alongside information about the correct category label.

Accuracy Results

A modality condition (within-modality visual, within-modality auditory, or across modalities) × block (eight 80-trial blocks) mixed design ANOVA was conducted on the accuracy rates. The accuracy rates, averaged across participants, are presented in Figure 2. The main effects of modality condition [$F(2,81) = 11.64$, $MS_e = 0.047$, $p < .001$] and block [$F(7,756) = 24.89$, $MS_e = 0.004$, $p < .001$] and the interaction [$F(14,567) = 2.06$, $MS_e = 0.004$, $p < .05$] were significant. To examine the locus of the interaction, we examined the effect of modality condition separately for each block. The effect of modality condition was significant in all but the first block of trials

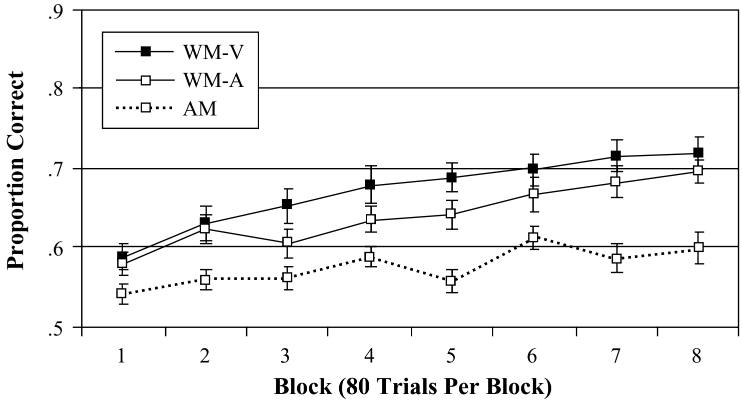

Figure 2. Proportion correct (averaged across participants) from the information integration experiment, along with standard error bars. Solid lines denote the within-modality visual (WM-V) and within-modality auditory (WM-A) conditions. The broken line denotes the across-modality (AM) condition.

[$F(2,81) = 0.267$, $MS_e = 0.006$, $p = .08$; $F(2,81) = 4.41$, $MS_e = 0.009$, $p < .05$; $F(2,81) = 5.93$, $MS_e = 0.010$, $p < .01$; $F(2,81) = 6.17$, $MS_e = 0.010$, $p < .01$; $F(2,81) = 14.23$, $MS_e = 0.009$, $p < .001$; $F(2,81) = 5.72$, $MS_e = 0.010$, $p < .01$; $F(2,81) = 11.99$, $MS_e = 0.011$, $p < .001$; and $F(2,81) = 11.04$, $MS_e = 0.010$, $p < .001$, for Blocks 1–8, respectively]. The locus of the modality condition effect was different for the early blocks (2–4) than for the later blocks (5–8). During Blocks 2–4, performance was significantly better in the within-modality visual condition than in the across-modality condition (all ps $< .05$) and did not differ from that in the within-modality auditory condition (all ps $> .10$). However, performance in the within-modality auditory condition was not statistically superior to that in the across-modality condition (all ps $> .05$). In each of Blocks 5–8, performance was significantly better in the within-modality visual condition than in the across-modality condition (all ps $< .05$), performance was significantly better in the within-modality auditory condition than in the across-modality condition (all ps $< .05$), and performance did not differ between the two within-modality conditions (all ps $> .10$). Averaged across Blocks 5–8, the observed proportions correct were .70, .67 and .59 in the within-modality visual, within-modality auditory, and across-modality conditions, respectively.[1]

These results support the prediction generated from the extended version of COVIS that information integration is easier for within-modality stimuli than for across-modality stimuli. Interestingly, the prediction held across early and late training when the within-modality visual condition was compared with the across-modality condition but held only in the latter half of learning when the within-modality auditory condition was compared with the across-modality condition. In addition, although never statistically significant, within-modality auditory performance was consistently worse than within-modality visual performance.

Theoretical Analysis and Modeling Results

Accuracy-based analyses provide important information regarding overall performance, but they tell us little about the types of strategies participants might use to solve these tasks. Model-based analyses can reveal patterns obscured by averaging (Ashby, Maddox, & Lee, 1994; Estes, 1956; Estes & Maddox, 2005; Maddox, 1999; J. D. Smith & Minda, 1998) and shed light on the strategies that participants use. Therefore, a number of different decision bound models (Ashby, 1992a; Maddox & Ashby, 1993) were fit to the data separately by condition and participant. Because the effects of modality increased with training, we focused the model-based analyses on the final block of trials. The model parameters were estimated using maximum likelihood (Ashby, 1992b; Wickens, 1982), and the goodness-of-fit statistic was

$$\mathrm{AIC} = 2r - 2\ln L,$$

where r is the number of free parameters and L is the likelihood of the model, given the data (Akaike, 1974; Takane & Shibayama, 1992). The AIC statistic penalizes a model for extra free parameters in such a way that the smaller the AIC, the closer a model is to the "true model," regardless of the number of free parameters (for a discussion of the complexities of model comparisons, see Myung, 2000; Pitt, Myung, & Zhang, 2002).

Two different classes of decision bound models were fit to each participant's responses (see Ashby, 1992a, and Maddox & Ashby, 1993, for a more formal treatment of these models, and Maddox & Ashby, 2004, for a review of similar applications). One type is compatible with the assumption that the participants used an implicit information integration (procedural-learning-based) strategy, and one type assumes an explicit hypothesis-testing strategy.

Procedural-learning-based models. Three procedural-learning-based models were applied to the data. The *optimal model* assumes that the participant used the optimal decision bound (see Figure 1A) and contains the single noise parameter σ^2. The *general linear classifier* (GLC) assumes that the participant used a linear decision bound but allows the linear bound to differ from the optimal. The GLC has three parameters (i.e., the slope and intercept of the linear bound and σ^2). The *piecewise bilinear classifier* assumes that the participant constructs two linear decision bounds to separate the A and B categories. The model has five free parameters (two slopes, two intercepts, and one noise variance). This model has been found to provide a good computational model of the participants' response regions in previous information integration category-learning studies (e.g., Ashby & Waldron, 1999; Waldron & Ashby, 2001; for applications to stimulus identification, see Ashby, Waldron, Lee, & Berkman, 2001, and Maddox, 2001, 2002).

Hypothesis-testing models. Four models compatible with the assumption that the participants used an explicit hypothesis-testing strategy were applied to the data. Two *unidimensional models* were applied to the data. One assumes that the participant attends only to dimension x, and the other assumes that the participant attends only to dimension y. Both of these models have two parameters (one decision criterion and one noise parameter). Two *conjunctive models* were applied to the data. Both assume that the participant sets a criterion on dimension x and one on dimension y, and then integrates that information post-decisionally. One version assumes that the participant uses the decision rule, "respond A if dimension x is *low* and dimension y is *high*; otherwise, respond B." The other assumes that the participant uses the decision rule "respond B if dimension x is *high* and dimension y is *low*; otherwise, respond A." Both of these models have three parameters (two decision criteria and one noise parameter).

It is worth mentioning that each of these models can be thought of as a special case of the piecewise bilinear model outlined above, where the decision bound slopes would be zero or infinite. Under these conditions, the notion of a decision criterion is more parsimonious, and more important, the decision criterion can be interpreted verbally. For example, the spatial frequency is *low* because it is less than one's decision criterion. On the other hand, the bilinear model described above allows the slope to be nonzero and noninfinite. Under these conditions, the notion of a decision bound is useful, and verbalizability of the decision rule does not make sense, because the two dimensions are measured in different units.

Random-responding models. The procedural-learning-based and hypothesis-testing models assume that the participant chooses (perhaps subconsciously) a specific strategy to apply when solving the task. The models make this assumption, and the associated parameters are estimated. Sometimes, though, participants choose not to attempt to solve the task and, instead, respond in a nearly random fashion, or sometimes a participant simply cannot solve the task. It is important to separate these participants from those using procedural-learning-based or hypothesis-testing strategies. To achieve this goal, we applied two models to the data. The *equal-response-frequency random-responding model* assumes that the participant is equally likely to respond A or B to every stimulus. This model assumes that the predicted probability of responding A is .5, and no parameters are necessary. The *biased-response-frequency random-responding model* assumes that the participant has a bias in his or her responding and, thus, will respond A with some fixed probability for all stimuli. This model assumes that the predicted probability of responding A is some value that is estimated from the data, yielding one free parameter.

Model fits. For each participant, we determined which model provided the best account of the data. For ease of exposition, we then summarized the data by collapsing across the different procedural-learning-based, hypothesis-testing, and random-responding models. The percentages of the participants' data sets for which a procedural-learning-based, hypothesis-testing, or random-responding model provided the best account of the data by experimental condition are presented in Table 2. To provide insight into the categorization performance achieved with the different types of response strategies, we computed the average accuracy rate obtained by the participants whose data were best fit by each type of model. For example, in the within-modality visual condition, 70% of the participants' data from the final block were best fit by a procedural-learning-based model, and the average proportion correct for these participants was .76 (see Table 2).

An examination of Table 2 reveals four things about the strategies used by the participants. First, in the two within-modality conditions, the percentage of participants whose

Table 2
Percentages of Participants for When a Procedural-Learning-Based, Hypothesis-Testing, or Random Responding Model Provided the Most Parsimonious Account of the Final Block of Data and the Average Accuracy Rates Associated With These Data Sets for the Three Experimental Conditions in the Information Integration Experiment

	Model Class					
	Procedural Learning Based		Hypothesis Testing		Random Responding	
Condition	Model Percent	Proportion Correct	Model Percent	Proportion Correct	Model Percent	Proportion Correct
Within-modality visual	70	.76	23	.65	7	.49
Within-modality auditory	69	.74	27	.61	4	.55
Across modality	36	.71	50	.55	14	.49

data were best fit by a procedural-learning-based model was high (approximately 70%), whereas the same percentage was nearly half that for the across-modality condition (36%). As a statistical test of these differences, we conducted χ^2 tests on the procedural-learning-based versus non–procedural-learning-based (hypothesis testing plus random responding) condition frequencies for all pairs of conditions. As was expected, the model frequencies differed between the within-modality visual and the across-modality conditions [$\chi^2(1) = 6.84, p < .01$] and between the within-modality auditory and the across-modality conditions [$\chi^2(1) = 6.07, p < .05$], but not between the two within-modality conditions [$\chi^2(1) = 0.004$, n.s.]. Second, the pattern was reversed with respect to hypothesis-testing models. Specifically, the percentage of hypothesis-testing participants was low in the two within-modality conditions (approximately 25%) but was nearly twice that large in the across-modality condition (50%). As was expected, the model frequencies differed between the within-modality visual and the across-modality conditions [$\chi^2(1) = 4.46$, $p < .05$] and (marginally) between the within-modality auditory and the across-modality conditions [$\chi^2(1) = 3.02$, $p = .08$], but not between the two within-modality conditions [$\chi^2(1) = 0.10$, n.s.]. Third, only about 5% of the participants responded randomly in the within-modality conditions, whereas 14% were random responders in the across-modality condition. Finally, across all three conditions, accuracy rates were highest for procedural-learning-based participants and were worst for random responders.

As a follow-up to these model-based analyses, we decided to compare the performance of only those participants who were procedural-learning-based participants. The idea was to determine whether the modality manipulation affected performance even when we restricted attention to participants who were using a strategy of the same form as the optimal (i.e., a procedural-learning-based strategy). Specifically, we included data only from those participants whose final block data were best fit by a procedural-learning-based model. We conducted a 3 modality condition × 8 block ANOVA on these accuracy rates. The main effect of block was significant [$F(7,322) = 26.70$, $MS_e = 0.004$, $p < .001$], whereas the interaction was not ($F = 1$). More important, the main effect of modality condition was significant [$F(2,46) = 3.22$, $p < .05$] and suggested better performance for the within-modality visual (70%) condition than for the across-modality (63%) condition, equivalent performance for the two within-modality conditions, but no difference across the within-modality auditory (67%) and across-modality conditions ($p = .12$). Thus, when we focused exclusively on procedural-learning-based participants, the advantage over the across-modality condition remained for the within-modality visual condition, but failed to reach significance when the within-modality condition was auditory. To determine the locus of the effect, we examined the slope and noise parameter values for these participants during the final block of trials. The slope values were further from optimal in the across-modality condition (average absolute slope deviation = .90) than in the within-modality visual condition (average absolute slope deviation = .43), with the within-modality auditory condition yielding an intermediate result (average absolute slope deviation = .58). The noise values did not differ across conditions (within-modality visual = .74, within-modality auditory = .70, and across modality = .73).

Brief Summary

The accuracy-based analyses support the predictions derived from the extended version of COVIS, in that performance was better when the stimulus dimensions came from a single modality (whether visual or auditory) than when they came from separate modalities. The modeling analyses suggest that this effect on performance was due to greater use of hypothesis-testing strategies in the across-modality condition than in the within-modality conditions and, thus, a smaller likelihood that a procedural-learning-based strategy was used. Even so, when we restricted attention to the participants whose final block of data was best fit by a procedural-learning-based model, we continued to find an accuracy advantage of 7% for the within-modality visual participants over the across-modality participants; but in this case, the accuracy advantage fell to only 4% for the within-modality auditory participants over the across-modality participants, and this difference was not statistically significant.

EXPERIMENT 2
Within- and Across-Modality Rule-Based Category Learning

The goal of Experiment 2 was to examine rule-based category learning when the stimulus dimensions were from the same or different modalities. The three modality conditions in Experiment 1 were replicated in Experiment 2, with the only difference being in the nature of the categorization rule. Because it was important to select a rule-based category structure for which both stimulus dimensions were relevant, we used a disjunctive rule-based category structure modeled after one used successfully by Maddox, Filoteo, Hejl, and Ing (2004). A scatterplot of the stimuli and optimal disjunctive category rule is displayed in Figure 1B, and the distribution parameters are presented in Table 3.

As was outlined above, the extended version of COVIS makes only a weak prediction regarding the effects of stimulus modality on disjunctive, rule-based category learning. Because rule-based category learning, by definition, requires that separate decisions be made for each stimulus dimension and since, in general, across-modality dimensions tend to be more separable, we predicted enhanced performance when the stimulus dimensions were processed by different modalities. To anticipate, this prediction was supported by the data.

Method

Participants and Design. Eighty-six participants were solicited from the University of Texas community and received course credit or $6 payment for participation. Each participant received

Table 3
Category Distribution Parameters From the Disjunctive, Rule-Based Experiment

Category	μ_x	μ_y	σ_x	σ_y	cov_{xy}
A_1	.300	.700	.040	.040	0
A_2	.700	.300	.040	.040	0
B_1	.300	.300	.040	.040	0
B_2	.700	.700	.040	.040	0

a $1 bonus for accuracy exceeding 65% during the first half of the study (320 trials) and another $1 bonus for accuracy exceeding 70% during the last half of the study. Twenty-seven participants were randomly assigned to the within-modality visual condition, 28 were randomly assigned to the within-modality auditory condition, and 31 were randomly assigned to the across-modality condition.

Stimuli, Stimulus Generation, and Procedure. The stimuli, stimulus generation, and procedure were identical to those in Experiment 1.

Accuracy Results

As an initial examination of the data, we constructed frequency histograms that plotted the proportions of participants whose performance level fell within a range of accuracy rates (e.g., proportion correct < .5, proportion correct between .5 and .6, .6 and .7, .7 and .8, or .8 and .9, or proportion correct > .9). We generated histograms separately for each of the three modality conditions and blocks of trials. The pattern was similar across all blocks, so for illustrative purposes, we will present data from the final block of trials in Figure 3A. For completeness, we present frequency histograms for the final block of trials from the information integration experiment in Figure 3B. Two aspects of the data are of particular interest. First and foremost, performance appears to be bimodal in the disjunctive rule-based conditions (panel A), with the majority of the participants performing very well (i.e., above .9) or very poorly (i.e., below .5 or .6) and far fewer par-

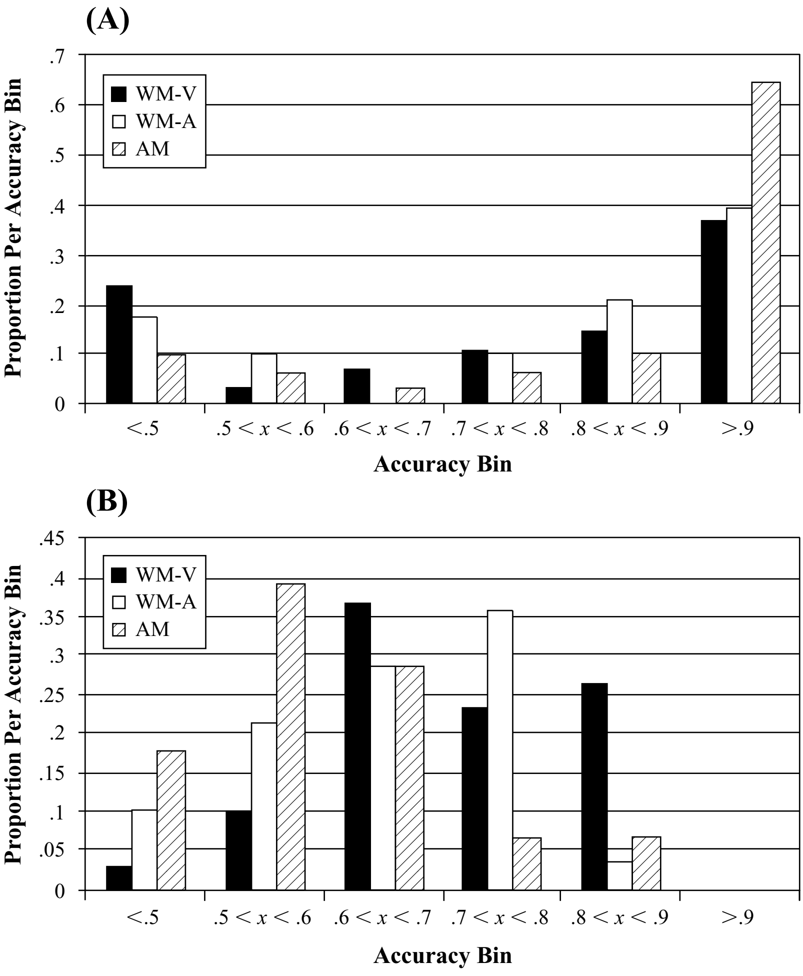

Figure 3. (A) Proportion of the participants from the rule-based experiment in each condition whose final block performance level fell within a range of accuracy rates (<.5, between .5 and .6, .6 and .7, .7 and .8, or .8 and .9, or >.9). (B) Same proportions from the information integration experiment. WM-V, within-modality visual; WM-A, within-modality auditory; AM, across modalities.

ticipants performing at intermediate levels. This pattern is indicative of an all-or-none type learning process that converges nicely with the performance pattern predicted by a hypothesis-testing system. Second, performance is more normally distributed in the information integration conditions (panel B), with the majority of participants falling in an intermediate performance range (.5–.7 for the across-modality condition and .5–.8 for the within-modality conditions) and fewer in the extreme range. This pattern is indicative of a gradual, incremental learning process that converges nicely with the performance pattern predicted by a procedural-learning-based system.

The clear bimodality in the data precludes the use of an ANOVA. In light of this fact, we will compare the proportions of learners across conditions, using different criteria for learning. Because the category distributions are widely displaced, there is no overlap across categories, and a participant who learns the disjunctive rule can easily achieve nearly perfect performance (i.e., near 100%). Thus, we used a stringent learning criterion. Specifically, a learner is a participant who exceeds 95% correct during the final block of trials, and a nonlearner is a participant who fails to exceed 95% correct during the final block of trials. To determine whether the proportions of learners differed across modality conditions, we conducted χ^2 tests for each pair of conditions. As was expected, the proportion of learners was higher in the across-modality condition than in the within-modality visual condition [$\chi^2(1) = 4.28, p < .05$] and was higher than in the within-modality auditory condition [$\chi^2(1) = 7.84, p < .01$]. The proportion of learners did not differ across the two within-modality conditions [$\chi^2(1) = 0.58$, n.s.]. The same pattern of results held if we lowered the learning criterion to 90%. Interestingly, though, if we lowered the criterion to 80%, the effect weakened. The proportion of learners continued to be higher for the across-modality than for the within-modality visual condition [$\chi^2(1) = 3.12, p = .08$], and the proportions continued to be equivalent across the two within-modality conditions [$\chi^2(1) = 0.44$, n.s.], but the proportion of learners did not differ between the across-modality and the within-modality auditory conditions [$\chi^2(1) = 1.22, p = .27$].[2]

These results suggest that rule-based learning is superior when the stimulus dimensions come from different modalities, rather than from the same modality. This finding is robust across learning criteria when both the dimensions are visual (spatial frequency and spatial orientation) but holds only for the most stringent of learning criteria when both the dimensions are auditory (auditory frequency and auditory duration). Although speculative, it might be the case that the two auditory dimensions chosen are more separable than the two visual dimensions, leading to better within-modality auditory rule-based category learning. Another interesting possibility is that within-modality auditory rule-based category learning is superior because there are some auditory projections directly to the head of the caudate, which is known to be involved in hypothesis testing.

Theoretical Analysis and Modeling Results

As in Experiment 1, procedural-learning-based and hypothesis-testing models were applied to the data from the final block of trials at the level of the individual participant.

Procedural-learning-based models. The *bilinear classifier* was applied to these data, which assumed that the participant constructs two decision bounds to separate the A and B categories. The model had five free parameters (two slopes, two intercepts, and one noise variance).

Disjunctive hypothesis-testing models. Four rule-based models that assumed that the participant used a disjunctive rule were applied to the data. The *optimal model* assumed that the participant used the optimal decision bound (see Figure 1B) and contained the single noise parameter. The *suboptimal-x model* assumed that the participant used the optimal decision criterion along the y dimension but a suboptimal criterion along the x dimension. The *suboptimal-y model* assumed that the participant used the optimal decision criterion along the x dimension but a suboptimal criterion along the y dimension. Both models contained two free parameters (one decision criterion and the noise variance). The *suboptimal-xy model* assumed that the participant used a suboptimal criterion along both the x and the y dimensions and contained three free parameters (two decision criteria and the noise variance).

Unidimensional hypothesis-testing models. Four rule-based models that assumed that the participant ignored one stimulus dimension and set a criterion along the other stimulus dimension were applied to the data. Because of the disjunctive nature of the optimal decision bound, unidimension decision rules will yield near-chance performance. Even so, these are qualitatively different, psychologically, from the random-responding models (despite the fact that they yield similar poor accuracy rates) and need to be applied to the data. Two of the unidimensional models assume that the participant sets a criterion along the x dimension and ignores the y dimension. One assumes that *low* values on x are associated with Category A, and the other assumes that *low* values on x are associated with Category B. The other two unidimensional models assume that the participant sets a criterion along the y dimension and ignores the x dimension. One assumes that *low* values on y are associated with Category A, and the other assumes that *low* values on y are associated with Category B.

Random-responding models. The two random-responding models described above were also applied to the disjunctive, rule-based category-learning data.

Model fits. The percentages of the participants' data sets for which a procedural-learning-based, hypothesis-testing, or random-responding model provided the best account of the data, along with the associated accuracy rates, are presented in Table 4. As with the information integration data, the modeling results are very suggestive. First, the use of hypothesis-testing strategies was high in all the conditions but was somewhat larger in both the within-modality auditory and the across-modality

Table 4
Percentages of Participants for Whom a Procedural-Learning-Based, Hypothesis-Testing, or Random Responding Model Provided the Most Parsimonious Account of the Final Block of Data and the Average Accuracy Rates Associated With These Data Sets for the Three Experimental Conditions in the Rule-Based Experiment

	Model Class					
	Procedural Learning Based		Hypothesis Testing		Random Responding	
Condition	Model Percent	Proportion Correct	Model Percent	Proportion Correct	Model Percent	Proportion Correct
Within-modality visual	19	.88	63	.79	19	.46
Within-modality auditory	14	.90	79	.79	7	.43
Across modality	10	.81	77	.92	13	.48

conditions. Interestingly though, the accuracy rates for hypothesis testers in the across-modality condition was much higher (.92) than that for hypothesis testers in the two within-modality conditions (visual and auditory = .79). In Table 4, the participants classified as hypothesis testers were the participants who used either a disjunctive *or* a unidimensional strategy to solve the task. A disjunctive rule can yield very high levels of performance, whereas a unidimensional rule yields essentially chance performance. To investigate this further, we isolated only those participants who used a disjunctive rule-based strategy. In the within-modality visual condition, only 52% of the participants used a disjunctive rule. In the within-modality auditory condition, only 64% used a disjunctive rule. In contrast, all 77% classified as hypothesis testers in the across-modality condition used disjunctive rules. As a statistical test of these differences, we conducted χ^2 tests on the disjunctive versus nondisjunctive (unidimensional plus procedural-learning-based plus random-responding) condition frequencies for all three pairs of conditions. The only comparison that reached significance was that between the within-modality visual and the across-modality condition, suggesting significantly greater use of a disjunctive rule in the across-modality condition [$\chi^2(1) = 4.18, p < .05$].

Following the approach taken for the information integration conditions, we followed up these model-based analyses with a comparison of performance for only those participants whose data were best fit by a disjunctive hypothesis-testing model. The idea was to determine whether the modality manipulation would affect performance even when we restricted attention to participants who were using a strategy of the same form as the optimal (i.e., a disjunctive rule strategy). Specifically, we included data only from those participants whose final block of data was best fit by a disjunctive hypothesis-testing model. We used the same learning criteria (95%, 90%, and 80%) and χ^2 tests. For all three learning criteria, we never found a difference across the two within-modality conditions [95%, $\chi^2(1) = 2.67$, n.s.; 90%, $\chi^2(1) = 0.51$, n.s.; 80%, $\chi^2(1) = 0.79$, n.s.]. The proportion of learners was higher in the across-modality condition than in the within-modality visual condition for the 90% and 80% criteria, but not for the most stringent 95% criterion [95%, $\chi^2(1) = 1.38$, n.s.; 90%, $\chi^2(1) = 3.43, p < .05$; 80%, $\chi^2(1) = 4.51, p < .05$]. The proportion of learners was higher in the across-modality condition than in the within-modality auditory condition, but only for the most stringent 95% criterion [95%, $\chi^2(1) = 8.82, p < .01$; 90%, $\chi^2(1) = 1.75$, n.s.; 80%, $\chi^2(1) = 1.56$, n.s.]. To determine the locus of the effect, we examined the decision criterion and noise values. The decision criterion values were closer to optimal in the across-modality than in the two within-modality conditions (average absolute deviation from optimal: within-modality visual = .013; within-modality auditory = .013; across modality = .002), and there was less noise associated with the memory and placement of these criteria in the across-modality than in the two within-modality conditions (within-modality visual = .119; within-modality auditory = .114; across modality = .101).

Brief Summary

The accuracy-based analyses support the weak prediction derived from the extended version of COVIS, in that performance was better when the stimulus dimensions came from separate modalities than when they came from the same modality. Even so, the effect was larger and more robust when the across-modality condition was compared with the within-modality visual condition. The modeling analyses suggest that this effect on performance was due to greater use of disjunctive hypothesis-testing strategies in the across-modality than in the within-modality conditions, especially the within-modality visual condition. Interestingly, when we restricted attention to the participants whose final block of data was best fit by a disjunctive, hypothesis-testing model, the same pattern in the accuracy data emerged.

GENERAL DISCUSSION

This article has reported the results from two experiments in which the effects of stimulus dimension modality (within-modality vs. across-modality stimulus dimensions) on information integration and disjunctive rule-based category learning were examined. To our knowledge, this is one of only a few studies to examine category learning using across-modality stimulus dimensions and is the first to compare performance across the qualitatively different information integration and rule-based category-learning tasks. As was predicted from the extended version of

COVIS developed in this article, there was a performance interaction between stimulus dimension modality and the nature of the category structure. Specifically, information integration category learning was better when both stimulus dimensions were processed by the same modality (visual or auditory) than when they were processed by different modalities (visual and auditory), whereas disjunctive rule-based category learning showed the opposite pattern. The modeling results offer important insights into the locus of this performance interaction.

Information Integration Category Learning

The information integration performance advantage for within-modality stimulus dimensions resulted because (1) there were more participants' data sets best fit by a model that instantiated a procedural-learning strategy, (2) there were fewer participants' data sets best fit by a model that instantiated a hypothesis-testing strategy, and (3) there were fewer participants' data sets best fit by a random-responding model. Importantly, though, performance was worse for across-modality stimulus dimensions even when we restricted attention only to those participants who used procedural-learning strategies. In addition, the within-modality advantage was larger for visual than for auditory stimuli.

These findings follow from the underlying neurobiology of information integration category learning proposed in the original and extended versions of COVIS. With visually presented stimuli, the original version of COVIS assumes that information integration category learning involves learning to associate collections of visual cortical cells with specific category labels. These associations are instantiated in the neural connections between visual cortical cells that project to the tail of the caudate nucleus and medium spiny cells within the tail of the caudate. With auditory dimension stimuli, the projections are more diffuse, but the current thinking is that the majority project to the body of the caudate, with some projections to the head and tail (Arnauld et al., 1996; Nenadic et al., 2003; Selemon & Goldman-Rakic, 1985; Yeterian & Pandya, 1998). Thus, with auditory stimuli, the extended version of COVIS assumes that information integration category learning involves learning to associate collections of auditory cortical cells with specific category labels and that these associations are instantiated in the neural connections between auditory cortical cells that project primarily to the body, but also to the head and tail, of the caudate nucleus and medium spiny cells within these regions.

With stimuli composed of one visual and one auditory dimension, the specific nature of the corticocaudate connectivity is less clear. Multisensory association areas have been identified in the caudate (Chudler et al., 1995). In addition, because some auditory cortical projections are to the tail of the caudate, where visual processing occurs, it seems reasonable to assume that across-modality information integration category learning involves learning to associate visual–auditory association cells in the caudate or tail of the caudate cells that receive both visual and auditory information with specific category labels. Because there are fewer of these cells than of cells devoted to single-modality stimuli (Chudler et al., 1995) and because only 5% of the cells in the caudate are interneurons (Wilson, 1995), across-modality information integration category learning should be poor. Since fewer cells in the caudate are devoted to across-modality processing, one would expect participants to be more likely to use another strategy to solve the task, such as a hypothesis-testing strategy, and to be less accurate when using a procedural-learning-based strategy. Both of these predictions were supported in the data.

Rule-Based Category Learning

The disjunctive, rule-based performance advantage for across-modality stimulus dimensions followed because (1) there were more participants' data sets best fit by a disjunctive, rule-based model, (2) no participants used a unidimensional rule-based strategy, and (3) fewer participants' data sets were best fit by a random-responding model. Importantly, when we restricted attention only to those participants using disjunctive, rule-based strategies, the same performance pattern emerged. These findings are important for a number of reasons. First, they suggest that the effect of modality is to increase or decrease the likelihood that the participant will discover the correct rule, with rule discovery increasing for across-modality stimuli. Even so, a participant who identifies the correct disjunctive rule will perform worse if the stimulus dimensions come from the same modality than if they came from separate modalities. Second, this pattern is reversed from that for information integration category learning. With information integration categories, participants in the across-modality condition are less likely to use a (correct) procedural-learning-based strategy, and those who do are still less accurate than those who use procedural-learning-based strategies in the within-modality conditions. Third, as we found in the information integration experiment, the within-modality auditory condition yields a performance that is intermediate between that for the across-modality and the within-modality visual conditions. Specifically, the participants performed better and were more likely to learn the disjunctive rule in the within-modality auditory condition than in the within-modality visual condition. This could be due to the fact that (1) auditory frequency and duration are more separable than spatial frequency and orientation, (2) auditory cortical areas project to the head of the caudate, as well as to the body and tail, or (3) some combination of the two. Clearly, more work is needed.

One interesting finding from the present report was the clear bimodality in the nature of responding in the disjunctive rule-based condition that was not evident in the information integration condition. In the latter case, responding was normally distributed. Importantly, these different patterns held for all three modality conditions. Zeithamova and Maddox (2006; see also Maddox & Ashby, 2004) found a similar pattern across information integration and rule-based tasks for visually presented stimuli. The present report suggests that these distinct performance profiles extend to auditory dimension stimuli, as well as across-modality (visual/auditory) stimulus dimensions.

Alternative Theories of Category Learning

It is important to point out that multiple-system models other than COVIS have been developed. In fact, there is a long and rich tradition in category learning that focuses on the distinction between rule application and similarity-based processing (e.g., Allen & Brooks, 1991; Brooks, 1978; Folstein & Van Petten, 2004; Kemler Nelson, 1984; J. D. Smith & Shapiro, 1989; see also Shanks & St. John, 1994). Several of these other multiple-system models have components similar to those of COVIS. For example, in the model proposed by E. E. Smith, Patalano, and Jonides (1998), rule application involves a high working memory load and requires analytic, serial processing of criterial attributes with differential weighting of attributes, much like the hypothesis-testing system in COVIS. Similarly, one system in Erickson and Kruschke's (2002) ATRIUM model instantiates rules in a manner similar to that of the hypothesis-testing system in COVIS. Where these other multiple-system models differ from COVIS is in their proposal of a second system that is a similarity-based process. In contrast, the second system in COVIS is a procedural-learning-based system that is involved in associating category labels (or responses) to regions of perceptual space.

Another important difference between COVIS and other multiple-system models is the detail with which COVIS proposes the neurobiological underpinnings of the different category-learning systems. Although the neurobiology of category learning has been elaborated in the context of other multiple-system models (e.g., Patalano, Smith, Jonides, & Koeppe, 2001), COVIS has the added advantage of incorporating biological constraints into its architecture that have enabled more detailed predictions of the impact of various experimental manipulations on category learning. For example, COVIS and the extended version offered in this article hypothesize that the procedural-learning system relies on the many-to-one convergence of *visual* cells from the inferior temporal cortex onto cells within the tail of the caudate and of auditory cells from the supratemporal plane and superior temporal gyrus onto the body, but also the tail and head, of the caudate, with some evidence for association areas as well. Such a funneling of information onto the caudate, along with the general separation of visual and auditory processing areas within the caudate, results in the prediction that information integration category learning will be worse for across- than for within-modality stimulus dimensions, which was supported by the findings in the present study. Other biological constraints have also been incorporated into predictions made by COVIS. For example, on the basis of COVIS, the dopamine reward signal associated with feedback has to occur in close temporal proximity to the response in order for efficient information integration category learning to occur. This prediction has been supported by the finding that delayed feedback negatively impacts information integration category learning (Maddox et al., 2003; Maddox & Ing, 2005). Thus, COVIS attempts to provide computationally and biologically plausible accounts of multiple category-learning systems.

Future Directions

Because this work represents only a first step in the examination of information integration and rule-based category learning with across-modality stimulus dimensions, there are a number of avenues for future research. Anecdotal evidence suggests that we frequently integrate information (predecisionally) from multiple sensory dimensions and are quite good at doing so. It is possible, though, that these types of problems take more time to master. One interesting avenue for future research would be to examine the long-run time course of across-modality information integration and to investigate methods for speeding these learning processes. Future research should also examine the generality of the present results across different types of information integration and rule-based category structures and different visual and auditory stimulus dimensions. Although a detailed discussion is beyond the scope of this article, an important distinction in the literature is that between linear information integration decision rules and nonlinear information integration decision rules, with nonlinear rules being generally more difficult to learn (e.g., Ashby & Maddox, 1990, 1992). All of this work has been conducted within modality. It would be interesting to determine whether the same qualitative pattern holds for across-modality stimulus dimensions.

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**, 716-723.

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.

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

Arnauld, E., Jeantet, Y., Arsaut, J., & Demotes-Mainard, J. (1996). Involvement of the caudal striatum in auditory processing: c-fos response to cortical application of picrotoxin and to auditory stimulation. *Molecular Brain Research*, **41**, 27-35.

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. (1992a). Multidimensional models of categorization. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 449-483). Hillsdale, NJ: Erlbaum.

Ashby, F. G. (1992b). Multivariate probability distributions. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 1-34). Hillsdale, NJ: Erlbaum.

Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, **105**, 442-481.

Ashby, F. G., & Ell, S. W. (2001). The neurobiology of human category learning. *Trends in Cognitive Sciences*, **5**, 204-210.

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

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

Ashby, F. G., & Maddox, W. T. (1990). Integrating information from separable psychological dimensions. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 598-612.

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

Ashby, F. G., & Maddox, W. T. (1993). Relations between prototype, exemplar, and decision bound models of categorization. *Journal of Mathematical Psychology*, **37**, 372-400.

Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology*, **56,** 06.1-06.30.

Ashby, F. G., Maddox, W. T., & Bohil, C. J. (2002). Observational versus feedback training in rule-based and information-integration category learning. *Memory & Cognition*, **30**, 666-677.

Ashby, F. G., Maddox, W. T., & Lee, W. W. (1994). On the dangers of averaging across subjects when using multidimensional scaling or the similarity-choice model. *Psychological Science*, **5**, 144-151.

Ashby, F. G., & Waldron, E. M. (1999). On the nature of implicit categorization. *Psychonomic Bulletin & Review*, **6**, 363-378.

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

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

Chudler, E. H., Sugiyama, K., & Dong, W. K. (1995). Multisensory convergence and integration in the neostriatum and globus pallidus of the rat. *Brain Research*, **674**, 33-45.

Diehl, R. L., & Kluender, K. R. (1987). On the categorization of speech sounds. In S. Harnad (Ed.), *Categorical perception* (pp. 226-253). Cambridge: Cambridge University Press.

Ennis, D. M., & Mullen, K. (1985). Theoretical aspects of sensory discrimination. *Chemical Senses*, **11**, 513-522.

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

Estes, W. K. (1956). The problem of inference from curves based on group data. *Psychological Bulletin*, **53**, 134-140.

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

Estes, W. K., & Maddox, W. T. (2005). Risks of drawing inferences about cognitive processes from model fits to individual versus average performance. *Psychonomic Bulletin & Review*, **12**, 403-408.

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

Garner, W. R. (1974). *The processing of information and structure*. Potomac, MD: Erlbaum.

Goldman-Rakic, P. S. (1995). Architecture of the prefrontal cortex and the central executive. In J. Grafman, K. J. Holyoak, & F. Boller (Eds.), *Structure and functions of the human prefrontal cortex* (Annals of the New York Academy of Sciences, Vol. 769, pp. 71-83). New York: New York Academy of Sciences.

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology*, **49**, 585-612.

Grau, J. W. & Kemler Nelson, D. G. (1988). The distinction between integral and separable dimensions: Evidence for the integrality of pitch and loudness. *Journal of Experimental Psychology: General*, **117**, 347-370.

Hikosaka, O., Sakamoto, M., & Usui, S. (1989). Functional properties of monkey caudate neurons. *Journal of Neurophysiology*, **61**, 780-832.

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

Kemp, J. M., & Powell, T. P. (1970). The cortico-striate projection in the monkey. *Brain*, **93**, 525-546.

Klatzky, R. L., & Lederman, S. J. (1995). Identifying objects from a haptic glance. *Perception & Psychophysics*, **57**, 1111-1123.

Klatzky, R. L., Lederman, S. J., & Reed, C. (1987). There's more to touch than meets the eye: The salience of object attributes for haptics with and without vision. *Journal of Experimental Psychology: General*, **116**, 356-369.

Lamberts, K. (2002). Feature sampling in categorization and recognition of objects. *Quarterly Journal of Experimental Psychology*, **55A**, 141-154.

Maddox, W. T. (1992). Perceptual and decisional separability. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 147-180). Hillsdale, NJ: Erlbaum.

Maddox, W. T. (1999). On the dangers of averaging across observers when comparing decision bound models and generalized context models of categorization. *Perception & Psychophysics*, **61**, 354-374.

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

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

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

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

Maddox, W. T., Ashby, F. G., & Bohil, C. J. (2003). Delayed feedback effects on rule-based and information-integration category learning. *Journal of Experimental Psychology: Learning, Memory, & 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. (2005). The neuropsychology of perceptual category learning. In H. Cohen & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science* (pp. 573-596). New York: Elsevier.

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

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

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

Massaro, D. W. (1987). *Speech perception by ear and eye: A paradigm for psychological inquiry*. Hillsdale, NJ: Erlbaum.

Massaro, D. W. (1998). *Perceiving talking faces: From speech perception to a behavioral principle*. Cambridge, MA: MIT Press.

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

Melara, R. D. (1989). Dimensional interaction between color and pitch. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 69-79.

Myung, I. J. (2000). The importance of complexity in model selection. *Journal of Mathematical Psychology*, **44**, 190-204.

Nearey, T. M. (1997). Speech perception as pattern recognition. *Journal of the Acoustical Society of America*, **101**, 3241-3254.

Nenadic, I., Gaser, C., Volz, H. P., Rammsayer, T., Hager, F., & Sauer, H. (2003). Processing of temporal information and the basal ganglia: New evidence from fMRI. *Experimental Brain Research*, **148**, 238-246.

Newell, F. N., & Bülthoff, H. H. (2002). Categorical perception of familiar objects. *Cognition*, **85**, 113-143.

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.

Pitt, M. A., Myung, I. J., & Zhang, S. (2002). Toward a method of selecting among computational models of cognition. *Psychological Review*, **109**, 472-491.

Reed, C. L., Lederman, S. J., & Klatzky, R. L. (1990). Haptic integration of planar size with hardness, texture, and planar contour. *Canadian Journal of Psychology*, **44**, 522-545.

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

Selemon, L. D., & Goldman-Rakic, P. S. (1985). Longitudinal topography and interdigitation of corticostriatal projections in the rhesus monkey. *Journal of Neuroscience*, **5**, 776-794.

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

Shepard, R. N. (1957). Stimulus and response generalization: A stochastic model relating generalization to distance in psychological space. *Psychometrika*, **22**, 325-345.

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

Smith, J. D., & Minda, J. P. (1998). Prototypes in the mist: The early epochs of category learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 1411-1436.

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

Stankiewicz, B. J. (2002). Empirical evidence for independent dimensions in the visual representation of three-dimensional shape. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 913-932.

Stevens, S. S., & Volkman, J. (1940). The relation of pitch to frequency. *American Journal of Psychology*, **53**, 329.

Takane, Y., & Shibayama, T. (1992). Structures in stimulus identification data. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 335-362). Hillsdale, NJ: Erlbaum.

Van Hoesen, G. W., Yeterian, E. H., & Lavizzo-Mourey, R. (1981). Widespread corticostriate projections from temporal cortex of the rhesus monkey. *Journal of Comparative Neurology*, **199**, 205-219.

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.

Webster, M. J., Bachevalier, J., & Ungerleider, L. G. (1993). Subcortical connections of inferior temporal areas TE and TEO in macaque monkeys. *Journal of Comparative Neurology*, **335**, 73-91.

Wickens, T. D. (1982). *Models for behavior: Stochastic processes in psychology*. San Francisco: Freeman.

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

Yeterian, E. H., & Pandya, D. N. (1998). Corticostriatal connections of the superior temporal region in rhesus monkeys. *Journal of Comparative Neurology*, **399**, 384-402.

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

NOTES

1. We also collected data by using across-modality stimuli for which the spatial orientation of the Gabor patch and the auditory frequency of a tone were relevant (spatial frequency was held fixed). The results mirrored those for spatial frequency and auditory frequency.

2. We also collected data by using across-modality stimuli for which the spatial orientation of the Gabor patch and the auditory frequency of a tone were relevant (spatial frequency was held fixed). The results mirrored those for spatial frequency and auditory frequency.

(Manuscript received December 17, 2004;
revision accepted for publication December 28, 2005.)