

Running head: Multiple systems or task complexity

Procedural memory effects in categorization:
evidence for multiple systems or task complexity?

Safa R. Zaki and Dave F. Kleinschmidt

Williams College

Safa Zaki

Department of Psychology

Bronfman Science Center

Williams College

Williamstown, MA 10267

413-597-4594

Email: szaki@williams.edu

Abstract

According to an influential multiple-systems model of category learning, an implicit procedural system governs learning of information-integration category structures, whereas a rule-based system governs learning of explicit rule-based categories. Support for this idea has come in part from demonstrations that motor interference, in the form of inconsistent mapping between response location and category labels, results in observed deficits but only for learning information-integration category structures. In this article, we argue that this response location manipulation resulted in a potentially more cognitively complex task in which the feedback was difficult to interpret. In one experiment, we attempt to attenuate the cognitive complexity by providing more information in the feedback and demonstrate that this eliminates the observed performance deficit for information-integration category structures. In a second experiment, we demonstrate similar interference of the inconsistent mapping manipulation in a rule-based category structure. We claim that task complexity, and not separate systems, might be the source of the original dissociation between performance on rule-based and information-integration tasks.

In recent years, there has been considerable debate regarding whether or not multiple qualitatively different systems are engaged in categorization tasks. Multiple-systems theorists argue that there are at least two systems: an explicit system that has access to executive functions, working memory, and conscious awareness, and an implicit system that operates unconsciously. There is some debate among such theorists as to the nature of the implicit system. One influential multiple-systems model is COVIS (COmpetition between Verbal and Implicit Systems: Ashby, Alfonso-Reese, Turken, & Waldron, 1998) which posits two systems that run in parallel and compete to provide the final categorization response. According to this model, the verbal explicit system is a hypothesis-forming and testing one, and the implicit system is procedurally based.

This model makes a distinction between two kinds of categorization tasks: rule-based tasks, which are learned primarily by the explicit, rule-based system, and information-integration tasks, which are learned primarily by the implicit, procedural system. Rule-based categorization tasks are often easily described by explicit, verbal rules (“If the line is long, it’s an A; otherwise it’s a B”). These category structures are ones in which optimal performance is possible by making independent decisions about one or more attributes of the stimulus, and responding based on those decisions. The paradigmatic example of a rule-based task is the unidimensional task, where optimal accuracy is achieved by selectively attending to a single stimulus dimension (such as length) and ignoring the others completely. Figure 1A gives an example of the abstract structure of a unidimensional category learning task. In contrast, information-integration tasks, which according to COVIS are learned by the implicit procedural system, are

tasks in which optimal performance is possible only by predecisionally integrating information from a number of different stimulus attributes or dimensions (Ashby & Gott, 1988). Consequently, they are not easily described by verbal rules. The most common type of information-integration task and the one used in the current study requires subjects to compare quantities expressed in different units (such as length and angle). See Figure 1B for an abstract structure of this type of task.

Although the idea of multiple systems in categorization has existed in the literature for quite some time (e.g., Erickson & Kruschke, 1998; Homa, Sterling, & Trepel, 1981; Nosofsky, Palmeri, & McKinley, 1994), one of the intriguing aspects of COVIS is that it specifically hypothesizes that the implicit learning system is a procedural learning system. In proposing COVIS, Ashby et al. (1998) argued that implicit learning is mediated by the striatum, which has been strongly implicated in procedural learning (Willingham, 1998, 1999). Consequently, COVIS predicts a close relationship between implicit category learning and motor learning, and conversely that disruption of motor processing should affect information-integration categorization, but not rule-based categorization. Ashby, Maddox, and colleagues (e.g., Ashby, Ell, & Waldron, 2003; Maddox, Bohil, & Ing, 2004; Maddox, Glass, O'Brien, Filoteo, & Ashby, 2010) have conducted a number of studies aimed specifically at testing this prediction and generally at providing evidence in favor of this multiple-systems theory of categorization.

In one such study, Ashby et al. (2003) trained subjects on either a rule-based or an information-integration category structure. In the transfer trials, participants once again saw these same category structures. However, whereas one group followed

precisely the same procedure as in the training trials, the other group was instructed to switch the buttons they used to indicate their categorization response. Ashby and colleagues (2003) found that this button-switch caused a significant detriment in performance in the information-integration task, but not in the rule-based task. They argued that this result is predicted by a multiple-systems approach to category learning. That is, according to COVIS the implicit system that learns information-integration tasks relies on procedural learning of mappings between category response and motor response location, which is interrupted in the button-switch condition. In contrast, there is no reason to expect interference for the button switching for rule-based tasks.

However, in a subsequent study Nosofsky, Stanton, and Zaki (2005) suggested that this dissociation was due to a difference in cognitive complexity of the different categorization tasks. Nosofsky et al. (2005) hypothesized that all categorization tasks involve a certain amount of procedural learning, and thus should suffer a performance deficit due to the button switch in a sufficiently sensitive experiment. They argued that, given long enough to respond on a sufficiently simple task, subjects could correct for the cost of the button switch and a performance cost would not be observed. Indeed, in their experiments, by reducing the amount of time given to subjects to respond, they observed a button switch cost in both the information-integration and rule-based category structures, implying that cognitive complexity plays a main role in the observed effect (although see Maddox, Lauritzen, and Ing (2007) for a rebuttal).

However, other evidence has been proposed as support for the procedural nature of the implicit system in COVIS. For example, Maddox, Bohil, et al. (2004) observed a dissociation between rule-based and information-integration learning tasks

when the mapping of categorization response to motor-response location was disrupted during learning. Specifically, on each trial the control group was asked "Is this an A or a B" and pressed button A to indicate a choice of category A and likewise for B. In the Yes-No condition, on each trial subjects were randomly asked one of two Yes-No questions (either "Is this an A?" or "Is this a B?") and responded by either pressing the Yes button or the No button. So, to indicate a choice of category A, the subject would sometimes press "Yes" and sometimes "No", and likewise with a choice of B. Thus, in the Yes-No condition there was a consistent mapping from categorization choice to response location whereas there was a clear mapping of category to response location in the A-B condition. Under the assumptions of the COVIS model that information-integration categories are mainly learned using the implicit procedural system, the lack of consistent mapping in the Yes-No condition should result in a decrease in performance on the information-integration task relative to the control group. For the rule-based task, which is primarily learned using an explicit, hypothesis forming and testing system, there should be no difference in performance between the control and Yes-No conditions. In fact, consistent with the COVIS model, this pattern of results was observed in their experiment.

Although a dissociation between performance in rule-based and information-integration tasks could indeed be evidence for multiple representational systems, alternative possibilities need to be ruled out. In particular, there is a possibility that the Yes-No condition is more cognitively demanding, due to the complicated reasoning required to deduce the correct category from the feedback. For example, during a typical trial in the Yes-No condition, the subject might be asked "Is this an A?", respond

in the negative, and be told that that response was incorrect. To determine to which category the previous stimulus actually belonged, the subject would then have to reason that, since their response of “no” was incorrect, the correct response was “yes,” which, given that prompt of “Is this an A?” indicates that the stimulus was in fact a member of category A. Contrast this with the A-B condition, in which a subject, having responded “A” and finding that they are incorrect, would know immediately that the previous stimulus was a “B”.

Maddox, Bohil, et al. (2004) acknowledged that difficulty might play a role in the observed effect, and indeed took a number of steps to control various types of difficulty. In order to equate performance on the control conditions for the rule-based and information-integration tasks, they reduced the between-category discriminability in the rule-based condition. They also pointed out that the Yes-No condition might place greater demand on working memory as the subject must keep track of the stimulus itself, the prompt, and their response in order to correctly process feedback. To control for this, they performed an additional study in which both the Yes-No prompt and the subject’s response were displayed during feedback, in addition to the correct/incorrect information from the original study. Even though the demands on working memory were by all measures less than in the original A-B condition, the same pattern of results was observed as in the original Yes-No condition, suggesting that the selective impairment of information-integration learning under the Yes-No condition was not due to additional working memory demands, but the lack of consistent response locations for each category.

However, Spiering and Ashby (2008) found support for the idea that a consistent response location was not in fact required for the learning of information-integration category structures. In their experiments, they compared three conditions in which participants completed an information-integration classification task. In the control condition, participants pressed the button beneath a blue circle on one side of the screen if the stimulus was associated with one category and a red circle on the other side of the screen if it was associated with the alternative category. In the random-location condition, those blue and red circles alternated randomly between the right and left location. Finally, a control condition tested a Yes-No version of the task in which participants were asked to decide if the stimulus was a member of either the blue or red category by pressing the appropriate yes or no key. They found that although the early learning of these information-integration categories was slowed in the random-location condition relative to the control condition, later learning was indistinguishable across the two conditions.

Nonetheless, Spiering and Ashby (2008) interpret their findings as evidence for COVIS. Instead of being limited to simple motor responses to particular spatial locations, they propose that the procedural learning system can also learn tasks in which there is a consistent mapping between stimulus features and visual features labeling the response locations, even if the locations themselves are random.

This still leaves unexplained why exactly there is such a large impairment in the Yes-No, information-integration task. In answer to this question, Spiering and Ashby (2008) claim that "the only difference between the Yes–No experimental conditions and the Random-Location condition was that the Yes–No response required an extra logical

decision" (p. 336), which requires some kind of executive functioning to implement and is thus difficult for the procedural learning system to accommodate. On the face of it, this finding that additional executive-function load selectively interferes with an information-integration task seems to contradict previous reports that increasing working memory demands (another executive task) selectively affects rule-based and *not* information-integration tasks (e.g., Waldron & Ashby, 2001; Zeithamova & Maddox, 2006).

Spiering and Ashby (2008) provide the following rationale. They argue that because rule-based tasks, according to COVIS, already tap executive functioning, they should be able to more easily accommodate the extra demands of generating a Yes-No response. That is, they propose that additional low levels of working memory demands that do not reach capacity limits might be more readily integrated into learning rule-based category structures because working memory is already in play. On the other hand, higher levels of working memory demands might begin to hurt the learning of rule-based tasks. As they put it, the "extra logical step loads on PFC processes (e.g., working memory) but not enough to challenge capacity limits" (p.335). One important question, therefore, is whether such a complex explanation is necessary.

To preview our plan for this article, we take a dual-pronged approach to providing an alternative account of the dissociation observed by Maddox, Bohil, et al. (2004) and the findings and conclusions of Spiering and Ashby (2008). In the first experiment, we replicate the original dissociation demonstrated by Maddox, Bohil, et al. (2004). Importantly, we further test the idea that information-integration category structures might be learned if the cognitive complexity of the Yes-No condition is alleviated. To test

this, we include a condition in which the cognitive complexity of the Yes-No condition is reduced by simplifying the processing of feedback. In the second experiment, we test the hypothesis that the interference of the Yes-No manipulation is not limited to information-integration category structures, but rather is mediated by the complexity of the task.

Experiment 1

In the present study, we further examined the potential role of cognitive complexity of the Yes-No condition in the Maddox, Bohil, et al. (2004) results. Given the Spiering and Ashby (2008) findings, there is good reason to believe that the inconsistent response locations were not the source of the observed dissociations. In this experiment, we examined the hypothesis that it was additional complexity of the Yes-No condition which differentially affects the already more difficult information-integration task. Specifically, we focus on the added difficulty of processing the correct/incorrect feedback in the Yes-No condition. Anything which disrupts feedback processing might make it harder for subjects to accurately infer the correct category of the stimulus they just categorized, which would impair their ability to recover the underlying category structure. This, we hypothesize, would be especially problematic in situations where the performance is limited by the complexity of the category structure, as in the information-integration task.

Therefore, Experiment 1 was designed to simultaneously replicate the results from Maddox, Bohil, et al. (2004) and test the effects of providing additional, simplifying feedback in the Yes-No condition. In addition to the A-B and Yes-No conditions of Maddox, Bohil, et al. (2004), we added a third response-type condition in which some

subjects were prompted exactly as in the Yes-No condition but during feedback they were also told the category to which the stimulus belonged. We hypothesize that, by providing this extra information during feedback, the additional complexity of the Yes-No task will be at least partially mitigated. Note that this extra feedback does not provide any information that was unavailable before, but merely simplifies the interpretation of the feedback information. It also does not change the structure of the category-response mapping: in the Yes-No (extra feedback) condition, there is still no consistent mapping between category and response location. Thus, if the COVIS-based interpretation put forth by Maddox, Bohil, et al. (2004) is correct, then this manipulation should not eliminate the performance deficit observed in the information-integration Yes-No condition, since it preserves the lack of consistent response-location mapping. Similarly, while this condition simplifies the processing of feedback, it preserves the extra logical step in formulating the Yes-No *response*, the factor to which Spiering and Ashby (2008) attribute the selective deficit in Yes-No information-integration learning. Conversely, if this manipulation eliminates the observed performance deficit, then task complexity cannot be overlooked as a source of the observed dissociation between performance on rule-based and information-integration tasks.

Method

Participants. One hundred and seventeen participants were recruited from the population of undergraduates at Williams College and participated for either course credit or a ten dollar payment. For the information-integration category-structure conditions, there were 21, 19, and 19 subjects who participated in the A-B, Yes-No, and

Yes-No with extra feedback conditions, respectively. For the rule-based category-structure conditions, there were 20, 20, and 18 participants in the A-B, Yes-No, and Yes-No with extra feedback conditions, respectively.

Stimuli. The stimuli used in this experiment were similar to those used in Maddox, Bohil, et al. (2004). Stimuli were Gabor patches that varied in orientation and spatial frequency, and were generated for each condition by sampling randomly from two bivariate normal distributions (see Table 1 for means and variances). Each patch was 819 pixels square and displayed on a 20-inch monitor with a resolution of 1280 by 1024. The specific parameter values used in generating each patch were identical to those used in the original experiment. For each category, a total of 160 patches were generated. For each of the 4 blocks, 40 A and 40 B were randomly selected from the stimuli that had not previously been displayed, and displayed in a different random order for each participant.

Procedure. Participants were randomly assigned to one of the six conditions. These conditions consisted of either of the two category structures (information-integration or rule-based) factorially combined with one of the three query/feedback types (A-B, Yes-No, Yes-No with extra feedback). Participants were tested individually in a dimly lit room. At the start of the experiment, participants were told each patch they would see would belong to one of two categories, and that their task was to decide whether the stimulus belonged to category A or B. Participants were instructed that at first they would have no idea which stimuli belonged to which category, but by paying attention to the feedback, they could learn to categorize the stimuli almost perfectly. Participants were asked to focus on accuracy and not worry about time. In addition, a

prize of \$25 was offered for participants who achieved the highest accuracy in the experiment.

On each trial, a single patch was displayed on the monitor, and the participant was either asked "Is this an A or a B?" or a Yes-No question of the form "Is this an A?" or "Is this a B?". In the Yes-No condition, the choice of the two potential questions was randomly determined on each trial. The patch remained on the screen until the participant provided a response. Upon indicating their response, the stimulus was removed from the screen, and feedback was given. In the A-B and Yes-No conditions, feedback consisted simply of the word "Correct" or "Incorrect". In the extra feedback conditions, feedback consisted of the word "Correct" or "Incorrect", followed by the correct category for the stimulus. The feedback remained on the screen for 750ms. There was a 350ms intertrial interval.

Results

In analyzing the data, we take three different approaches. In the first section, we report a set of ANOVA analyses that are similar to those conducted by Maddox, Bohil, et al. (2004). In the second section we report a supplemental Bayesian analysis which accounts for the presence of non-learners and delayed learners in the data. Finally, we report an analysis of the strategies that subjects used to solve the task by fitting various decision-bound models to the data.

ANOVA analyses. The probability of correct classification as a function of block for the different conditions is shown in Figure 2. Visual inspection of the figure (left panel) reveals a pattern of results for the Yes-No and A-B conditions that replicates the pattern reported in Maddox, Bohil, et al. (2004). That is, the Yes-No response condition

interferes only with performance in the information-integration category structure. To confirm this pattern, we conducted a set of analyses analogous to the analyses reported by Maddox, Bohil, et al. (2004). That is, we first restricted our analysis to the A-B conditions to verify that performance was equivalent for the two category structures. A 2(category structure: information-integration vs. rule-based) X 4(block) mixed model ANOVA showed no category structure effect for A-B responses, $F(1,39)=.32$, $MSE=.015$, $p=.58$. This replicates the important result of equivalent performance across category structures for A-B responses observed by Maddox, Bohil, et al. (2004). In addition, the main effect of block was significant, $F(3,117)=10.18$, $MSE=.056$, $p<.001$ indicating that subjects were learning over time. The interaction between block and category structure was not significant, $F(3,117)=1.04$, $MSE=.006$, $p=.38$.

Next we restricted the analysis to the Yes-No conditions across the two category structures, and conducted another 2 (category structure: information-integration vs. rule-based) X 4 (block) ANOVA. Importantly, we replicated the main effect of category structure with the rule-based condition outperforming the information-integration condition for the Yes-No training procedure, $F(1,37)=4.79$, $MSE=.255$, $p=.04$. Once again, a block main effect indicated learning occurred with experience, $F(3,111)=9.35$, $MSE=.061$, $p<.001$. The interaction was not significant, $F(3,111)=1.43$, $MSE=.009$, $p=.24$.

Next, following Maddox, Bohil, et al. (2004), we conducted a series of analyses aimed at testing the prediction that Yes-No training would impair performance in the information-integration condition but not the rule-based condition. For the following analyses, we restricted the ANOVA to each of the category structures and compared

the Yes-No and A-B training procedures. For the rule-based category structures, the type of training had no effect $F(1,38)=.25$, $MSE=.018$, $p=.62$. However, as would be expected, there was a main effect of block, $F(3,114)=17.35$, $MSE=.099$, $p<.001$ and there was no significant interaction between block and training procedure, $F(3,114)=1.27$, $MSE=.007$, $p=.29$.

For the information-integration category structure, the critical main effect of training procedure was significant, $F(1,38)=8.54$, $MSE=.272$, $p=.006$. That is, the result of selective interference of the Yes-No procedure for the information-integration category structure was replicated. Once again, there was a main effect of block, $F(3,114)=4.06$, $MSE=.025$, $p=.009$, but no interaction between training procedure and block, $F(3,114)=.03$, $p=.99$. Collectively, the analyses replicate the key findings of Maddox, Bohil, et al. (2004).

We now move onto a series of analyses aimed at testing the hypothesis that this selective impairment may be due to the difficulty in interpreting the feedback in the more complex information-integration group. First, we tested whether the addition of the extra feedback benefited participants in the rule-based structure. For this category structure, the Yes-No and the Yes-No with extra feedback conditions were not significantly different, $F(1,36)=.08$, $MSE=.006$, $p=.78$. Next, we conducted the same test for the information-integration category structure, and found that difference in performance of the Yes-No condition and the Yes-No condition with extra feedback were marginally different across all four blocks, $F(1,36)=3.97$, $MSE=.137$, $p=.05$, and were significantly different in the last two blocks, $F(1,36)=4.71$, $MSE=.130$, $p=.04$. In fact, for the information-integration category structure, the Yes-No with extra feedback condition did

not differ from the A-B group, when the two feedback conditions were made more comparable, $F(1,38)=.47$, $MSE=.020$, $p=.47$. These results together indicate that the effects of the Yes-No manipulation can be mitigated by more easily digestible feedback suggesting that the locus of the effect reported in Maddox, Bohil, et al. (2004) might not be procedural in nature.

Bayesian analysis of accuracy and learner-rates. In the above analyses, following Maddox, Bohil, et al. (2004), we included all subjects, regardless of whether they showed any learning at all. This is potentially problematic, in part because the presence of non-learners both reduces the average accuracy in a condition and increases the between-subject variance, leading to weaker statistical power and potentially obscuring real differences in the accuracy of subjects who do manage to learn the task. More problematically, this also conflates different aspects of category learning behavior, namely, how quickly and well subjects learn once they do start learning and how likely subjects are to be able to learn the task at all after some amount of training. Some authors have advocated simply removing data from subjects whose final-block accuracy falls below a certain threshold, but this ignores the fact that non-learners may be more likely in one condition than another, meaning that lower accuracy in that condition reflects a real increase in difficulty. Ideally, we would like to evaluate the effect of our experimental manipulations on non-learners and accuracy simultaneously.

Bayesian data analysis techniques offer the ability to do just that. In Bayesian data analysis, one first defines a model of how observed data are generated, as a function of underlying parameters (Kruschke, 2012). In our case here, this model allows

for a mix of non-learners (who perform at chance) and learners who improve over the course of training, with the learning curve depending on their task/response condition (and possibly a delay in learning; see the Appendix A for details). This model is fit to the available data by Bayesian inference, which assigns probability mass to parameter values to the degree that they are consistent with the data. For our purposes, the main benefit of such an approach is that there is more flexibility in defining the data model, which allows us to take into account the inherent uncertainty in whether a subject with very bad accuracy is simply guessing or is learning but struggling to do so. A secondary but more general and more powerful benefit of this approach is that the resulting posterior distribution over model parameters reflects the degree of uncertainty that remains about these explanatory parameters given the limited and possibly ambiguous data. This allows us to ask questions about the explanatory parameters, such as "was learning faster in condition A vs. condition B?", in addition to "was there some difference?" (see Krushcke, 2012 for an excellent overview of the relevant literature).

We summarize, very briefly, the results of this analysis here; details can be found in Appendix A. We report both the 95% highest posterior interval (HPI) of each effect, which quantifies the range of effect sizes and directions which are most credible given the data, and the probability that the effect is different from zero (pMCMC)¹. The HPI provides information about the range of credible effect sizes, and critically whether this range includes zero, while the pMCMC gives a sense of what proportion of the credible parameter values are compatible with the observed effect direction. It's important to note that neither of these measures is primary. Both are derived from the full, joint

distribution over parameter values (condition learning-curve slopes and intercepts, and subject learning delays) given the data.

First, remarkably, only about half of the subjects in the Yes-No, information-integration condition were classified as learners (HPI = [47%, 58%]), less than any other condition (HPIs all > 60%, pMCMC < 0.01). However, the lower proportion of learners is not enough to account for the decreased accuracy in the Yes-No, information-integration condition. Comparing the slope and intercept of the learning curve in the Yes-No information-integration condition with each other condition (Figure A1), we find that these subjects both had poorer performance overall (all intercepts difference pMCMC < 0.0004, HPI < -0.15) and learned more slowly (all slopes difference HPI < -0.03, pMCMC < 0.01, except for Yes-No, unidimensional task, HPI = [-0.30,0.03], pMCMC = 0.059). There is also a significant interaction between task and response/feedback condition: the effect of the information-integration task vs. the unidimensional task is greater under Yes-No responses than either AB or Yes-No-Extra response conditions (slope interaction between categorization task and condition AB vs. Yes-No, pMCMC = 0.02, HPI = [-0.20, -0.005] and Yes-No vs. Yes-No-Extra, pMCMC = 0.005, HPI = [-0.22, -0.03]; intercept interactions both pMCMC = 0.0003, HPI = [-0.31, -0.07] and [-0.35, -0.11], respectively), while there is no substantial difference in the effect of information-integration vs. unidimensional on AB vs. Yes-No-Extra (intercept HPI=[-0.07, 0.15], pMCMC=0.22, slope HPI=[-0.06, 0.10], pMCMC=0.30).

This analysis picks up on other effects that are missed when ignoring the effect of delayed learning and non-learners. First, even though the effect of information-integration vs. unidimensional structures was largest in the Yes-No response condition,

overall performance (measured by intercept) on the information-integration task was also worse than the unidimensional task for AB and Yes-No-Extra as well (for AB, HPI = $[-0.49, -0.17]$, $p\text{MCMC} < 0.0004$, and for Yes-No-Extra, $p\text{MCMC} = 0.003$, HPI = $[-0.41, -0.08]$). This was balanced out by slightly faster learning in the information-integration task (steeper slope for the information-integration task; for AB, $p\text{MCMC} = 0.066$, HPI = $[-0.02, 0.19]$, and for Yes-No-extra, $p\text{MCMC} = 0.017$, HPI = $[0.08, 0.41]$), resulting in equivalent accuracies on the information-integration and unidimensional task by the final block in the AB and Yes-No-Extra conditions (see Appendix Figure A2).

In summary, Bayesian analysis confirms the main results of the ANOVA analysis: performance is selectively impaired in the Yes-No information-integration condition, and this impairment goes away when additional feedback is provided in the Yes-No-Extra condition. This analysis also suggests that the equivalent overall accuracy in the information-integration and unidimensional tasks can mask differences between these tasks in the delay of learning onset (including the number of non-learners) and the speed/quality of learning once it begins.

Decision-bound model analyses. In their article, Maddox, Bohil, et al. (2004) investigated the strategies that subjects used to complete the task by fitting various decision bound models (e.g., Ashby & Gott, 1988). Specifically, they fit two models corresponding to rule-based strategies (a unidimensional decision boundary rule or a conjunction of two such rules) and a model corresponding to information-integration strategies (a general linear classifier, or GLC, allowing diagonal decision boundaries). They found that while only about 33% of subjects used a (sub-optimal) rule-based strategy in the information-integration A-B condition, 70% of subjects in the information-

integration Yes-No condition used a rule-based strategy. This finding is intriguing in that it supported their claim that the Yes-No condition selectively interferes with the use of information-integration strategies and drives subjects to use rule-based strategies.

We conducted a similar analysis, fitting various decision-bound models to our data. However, we added one important model to the mix: a random-responder model (e.g., Maddox, Ing, & Lauritzen, 2006). Note that for subjects who do not learn the task and guessed randomly on every trial, both the unidimensional model and the GLC will fit the data in the same way, but the unidimensional model will be preferred because it has fewer free parameters². Thus, without including a random-guessing model, subjects who are in fact randomly guessing would be classified as rule-users. Table B1 shows the percentage of subjects in each condition best fit by each of the models.

If we conduct the same modeling analyses as Maddox et al., essentially treating guessers as rule-users by omitting the random-guessing model, the proportion of rule-users in the information integration task jumps from 43% in the A-B condition to 68% in the Yes-No condition. This pattern of data is very similar to the one reported by Maddox et al. However, with the addition of the critical random-guessing model, contrary to the findings of Maddox, Bohil, et al. (2004), we found that the percentage of rule-use in the information-integration task remained essentially unchanged in the A-B condition (43%) versus the Yes-No condition (37%). Notably, the percentage of guessing was dramatically higher in the Yes-No (32%) versus the AB (5%) information-integration conditions. A subsequent analysis revealed that in every case where the guessing model was selected as the best fitting model, the unidimensional model was the next-best fitting model.

Finally, we considered whether the inconsistent response locations in the Yes-No information integration task interfered with participants' ability to learn the categories using information-integration strategies, as might be predicted by the COVIS model. A comparison of the use of these strategies by the Yes-No condition with extra feedback and the AB condition showed no differences (52% for A-B information-integration, vs. 53% for Yes-No-Extra, information-integration; see Table B1).

Discussion

Spiering and Ashby did not test rule-based tasks in their experiment, but rather relied on the data from Maddox, Bohil, et al. (2004) in order to interpret their results as selective interference of information-integration category structures. Therefore, Experiment 1 provides an important replication of this selective interference of the Yes-No condition on learning of information-integration category structures. In addition, we demonstrated that this selective interference could be alleviated by adding additional feedback. Taken together with the findings of Spiering and Ashby (2008), our results suggest that it was not the response location in the Yes-No condition that resulted in the selective interference with information-integration category structures observed by Maddox, Bohil, et al. (2004). The results of our Experiment 1 indicate that, in particular, it was the demands related to processing of the feedback in the Yes-No condition that were responsible for deficits observed by Maddox, Bohil, et al. (2004) and Spiering and Ashby (2008). In addition, the Bayesian analyses hint at important differences in the learnability of the information-integration and rule-based category tasks that were missed by the traditional null hypothesis testing approach. Separating the effect on accuracy of subjects who learn the task and those that don't (or are delayed), we find a

decrease in overall accuracy in the information-integration task relative to the unidimensional task, even in the A-B response condition (Appendix Figures A1 and A2).

Importantly, the strategy-use modeling analyses reported in Experiment 1 suggests that, at least in our own data set, the pattern of increased rule-use reported by Maddox, Bohil, et al. (2004) in the Yes-No condition may in fact be due to participants using a guessing strategy rather than a rule-based one. Moreover, participants in the Yes-No extra-feedback condition show a similar level of information-integration strategy use to those in the A-B version of the task.

Experiment 2

One framework for understanding the results of Experiment 1 is within the context of the COVIS model, as proposed in Spiering and Ashby (2008). That is, perhaps executive functions needed in order to process the executive demands in the Yes-No task are already operational in the rule-based condition. However, given that these same types of working memory demands have been shown to have a detrimental effect on rule-based tasks (Waldron & Ashby, 2001; Zeithamova & Maddox, 2006), the explanation lacks parsimony.

Alternatively, differential effects of the Yes-No task on the two different category structures might be understood by considering differences in the baseline difficulty of the two tasks. Remember that in order to equate performance in the two category structures, Maddox, Bohil, et al. (2004) needed to increase the separation between the two category distributions in the information-integration task (see Figure 1). That is, when all else is equal, when the optimal decision bound is on the diagonal in this task, and subjects do not perform as well as when the optimal decision bound is orthogonal

to one of the stimulus dimensions (e.g., Ell & Ashby, 2006). Therefore, to equate accuracy of performance in the two tasks, Maddox, Bohil, et al. (2004) increased the distances between the two distributions in the information-integration condition. The authors make the assumption that statistically equivalent accuracy means that the tasks are equal in their difficulty level. However, as our Bayesian analysis of Experiment 1 suggests, there may be differences between the unidimensional rule-based and information-integration tasks used in our Experiment 1 and by Maddox, Bohil, et al. (2004) in terms of learning delay and speed and quality of learning once it begins. This is consistent with our reasoning (based on the difference in perceptual discriminability required to equate accuracy), that these tasks are difficult for different reasons.

The difficulty hypothesis makes an interesting prediction. If in fact it is the overall difficulty of the task structure (vs. the perceptual difficulty) that interacts with the difficulty of the Yes-No task, then it should be possible to observe evidence of similar interference with a difficult rule-based task. In fact, Maddox, Bohil, et al. (2004) pursued exactly this possibility in their original study by testing a conjunctive rule-based categorization structure in the Yes-No condition. This task uses two unidimensional decision bounds, one on each dimension, to divide the stimulus space into four quadrants, one of which is assigned to one category and the other three to the other category. However, they did not find evidence of interference of the Yes-No condition in that task.

In the following experiment, we tested participants on the potentially more complex biconditional rule-based task shown in Figure 3. The stimuli were lines that varied in their orientation and length. In this task, a stimulus is a member of Category A

if its value on Dimension 1 is small and its value on Dimension 2 is large or conversely if its value on Dimension 1 is large and its value of Dimension 2 is small. The biconditional task is considered a rule-based task because the category boundaries are orthogonal to the axes of the stimulus space. In addition, participants might verbalize a rule such as “if it is long and flat or short and steep it is an A, otherwise it is a B.”

This design provides a nice test of the difficulty hypothesis. If the locus of the Yes-No interference observed in Maddox, Bohil, et al. (2004) and Experiment 1 is in the difficulty of the information-integration task, then it is reasonable to expect interference in learning a rule-based category structure that is sufficiently complex.

Method

Participants. Once again, we recruited 66 participants from the population of students at Williams College. Half of the participants were randomly assigned to the Yes-No condition, and the other half were assigned to the A-B condition. Participants either received course credit or \$10.

Stimuli. The stimuli were lines that varied along two dimensions: length and orientation that have been used in a number of studies (e.g., Nosofsky et al., 2005). The stimuli were centered on a screen with 1920 by 1080 resolution. A total of 320 stimuli were sampled randomly from the four bivariate normal distributions represented by the parameters listed in Table 2. There were therefore 160 stimuli assigned to each of two categories. The category structures are shown in Figure 3.

Procedure. Participants were randomly assigned to either a A-B or Yes-No version of the task. All other aspects of the procedure were identical to Experiment 1.

Results

The probability of correct classification as a function of block for the two different response conditions is shown in Figure 4. The main question of interest lies in the comparison of learning for the two training procedures. We therefore conducted a 2(training procedure: Yes-No or AB) X 4(block) mixed model ANOVA to address this question. There was indeed a main effect of training procedure, with higher accuracy for participants in the A-B condition than the Yes-No condition, $F(1,64)=4.83$, $MSE=.311$, $p=.03$. As would be expected, there was also a main effect of block, with increasing accuracy as a function of training block, $F(3,64)=17.62$, $MSE=.123$, $p<.001$. There was no significant interaction between block and training procedure, $F(3,64)=2.05$, $MSE=.014$, $p=.11$.

Bayesian analysis of accuracy and learner-rates. We also analyzed the results of Experiment 2 using the same Bayesian model as Experiment 1 (see Appendix A for details). The results confirm the ANOVA analysis: accuracy was worse overall in the Yes-No condition (intercept difference HPI = [-0.48, -0.04], $pMCMC = 0.010$), but the learning curve slopes were comparable (difference HPI = [-0.14, 0.11], $pMCMC = 0.46$; Figure A3). As was the case in Experiment 1 for the information-integration task, there were fewer learners in the Yes-No condition (HPI = [42%, 67%]) than in the AB condition (HPI = [64%, 70%]; difference HPI = [-21%, 3%], $pMCMC = 0.027$).

Decision-bound model analyses. Finally, we performed a decision bound strategy analyses as in Experiment 1. Because the biconditional task used in this experiment was different than the tasks in the first experiment, we included additional candidate models. In addition to the guessing, unidimensional, diagonal (GLC), and

conjunctive models, we fit two other models. The first of these was a rule-based strategy with two orthogonal decision boundaries, combined via a disjunctive rule, and the second was an information-integration strategy which has two parallel diagonal decision boundaries (see Appendix B for details).

Figure B2 and Table B2 show the proportions of participants best fit by each model across blocks in Experiment 2. In the AB condition, the majority of participants are identified as using the optimal, rule-based disjunctive strategy by the third and fourth blocks (58% and 55%, respectively). Second, in the Yes-No condition, the majority of participants are identified as using suboptimal strategies (only 24% and 21% classified as using the optimal disjunctive strategy in blocks three and four, respectively).

Intriguingly, contrary to the COVIS's prediction that the Yes-No response condition should lead to less information-integration strategy use, there are in fact more participants classified as using the (slightly) sub-optimal information-integration strategy of two parallel diagonal decision bounds (42% in the final block of Yes-No vs. only 18% in the final AB block), although this may be a result of increased response variability due to overall poor performance, which can lead to unstable AIC values (see the Appendix). Also, as in Experiment 1, there were more participants who were classified as simply guessing randomly in the Yes-No conditions (42% and 27% in the third and fourth Yes-No blocks, vs. 9% and 18% for AB).

In order to obtain a more stable estimate of each participant's strategy use in the later phase of the experiment, we also fit the same models to the trials in the second half of the experiment (blocks three and four together). The results are very similar to the fits to these two blocks individually: more subjects in the Yes-No condition use

suboptimal strategies (only 36% using the disjunctive rule) than in the AB condition (58% using the disjunctive rule). One noteworthy aspect of these combined fits is that there are fewer Yes-No participants classified as simply guessing, and more classified as using the disjunctive strategy. This suggests that some of what appears to be random guessing when either the third or fourth block is fit independently is better explained as noisy or inconsistent application of a strategy. In model comparison which penalizes free parameters it can be hard to distinguish such noisy but systematic behavior from truly random behavior when there is limited data available. The mean and median AIC (goodness of fit) values for the best model are substantially better (lower) in the AB condition (mean AIC=89, median AIC=98) than in the Yes-No condition (mean AIC=99, median AIC=107), indicating participants in the AB condition generally responded in a more systematic way³. In fact, in the Yes-No condition, the median AIC value is only four points better than the upper bound AIC value associated with the random-guessing model.

Discussion

We hypothesized that in the case of a more difficult rule-based category structure, the complexity involved in deciphering the Yes-No condition feedback could interfere with learning. The results of Experiment 2 support this idea, in that subjects performed more accurately in the A-B condition relative to the Yes-No condition when learning a bi-conditional rule-based category structure. Thus, the pattern of interference observed in Maddox, Bohil, et al. (2004) does not seem to be associated with whether the task involves explicit verbal rules or implicit information integration but, rather, seems to be related to the overall difficulty of the task.

General Discussion

COVIS (Ashby et al., 1998), a multiple-systems model of categorization, makes specific predictions about the nature of the component systems. In particular, COVIS posits two systems: a working memory-mediated hypothesis-testing system which learns explicit rule-based categories, and a procedural learning system which learns implicit categories.

COVIS has been gaining popularity in recent years due to accumulating evidence in its favor (for reviews see: Ashby & Maddox, 2005; Ashby & O'Brien, 2005). Much of the evidence advanced in support of COVIS comes from dissociation studies, in which an experimental manipulation differentially affects the learning of rule-based and information-integration category structures. That is, a manipulation that affects one type of task but not the other is taken as evidence of multiple dissociable systems, and the characteristics of the specific systems posited by COVIS predict which manipulations should have a dissociable effect.

The particular manipulation that is the focus of this current article is the one in Maddox, Bohil, et al. (2004). In their article, the authors examined subjects' performance on either an implicit or explicit categorization task, where some subjects received A-B prompts (of the form "Is this an A or a B?"), while others received Yes-No prompts (of the form "Is this an A?"). Maddox, Bohil, et al. (2004) hypothesized that the Yes-No response condition would impair performance on the implicit task but not the explicit task, since COVIS posits that implicit categories are learned by a procedural system that relies on a consistent mapping between categorization response and motor-

response location. This prediction was played out in their results, where Yes-No performance was impaired relative to A-B performance, but only in the information-integration conditions. However, the present results suggest that Maddox, Bohil, et al.'s (2004) design failed to rule out a single-system, complexity-based explanation. Maddox, Bohil, et al. (2004) tacitly assumed that by controlling the baseline accuracy by making the explicit task perceptually more difficult, there were no essential differences in terms of complexity between these tasks.⁴ Our results suggest that this assumption might not be warranted. In our delayed-learning Bayesian analysis of our replication of the original Yes-No and A-B conditions, we found differential effects of these conditions in the rule-based and information-integration tasks on both the delay of learning and the speed and quality of learning once it begins. These effects suggest that the equivalent overall accuracy in the information-integration and unidimensional tasks can at least sometimes mask potentially important behavioral differences in these groups.

In the present study, we took a different approach to controlling for difficulty, and instead attempted to correct potential differences in difficulty between the A-B and Yes-No response conditions. We reasoned that, given the greater number of logical steps that must be performed in order to glean the correct categorization based on correct/incorrect feedback to a Yes-No question, providing subjects in the Yes-No condition with the correct category at feedback might eliminate any differences in performance between the response conditions.

This is precisely what was obtained in Experiment 1. Note that adding additional feedback to the Yes-No condition does not in any way change the lack of a consistent response location for each category, which is the feature to which Maddox, Bohil, et al.

(2004) attributed the drop in performance in the implicit/Yes-No condition. Rather, the present results best support an explanation in terms of cognitive difficulty or complexity of the implicit learning task and the Yes-No response condition. In Experiment 2, we test a specific prediction of this difficulty hypothesis by showing that the same interference could be demonstrated in a challenging rule-based task.

The present study is just one piece of a growing body of work (Gureckis, James, & Nosofsky, 2011; Newell & Dunn, 2008; Newell, Dunn, & Kalish, 2010; Nosofsky et al., 2005; Zaki & Nosofsky, 2001) that raises issues with the common interpretation of dissociation studies as supporting multiple-system theories and highlights some various methodological issues. In particular, the common practice of controlling averaged baseline accuracy between unidimensional and information-integration tasks by reducing the perceptual discriminability of the unidimensional category introduces a potential confound, where the structurally-simpler rule-based categories are perceptually more difficult. That is, there are at least two ways in which these tasks are difficult, all else being equal. Tasks where the categories are less perceptually discriminable are harder, which we might call perceptual difficulty. Likewise, when discriminability is controlled, subjects find information-integration tasks harder than unidimensional tasks (Ell & Ashby, 2006), which we attribute to cognitive complexity.

These different types of difficulty presumably reflect different functional parts of the category learning system, broadly construed, and poor performance can result from a bottleneck at either. Manipulations which increase the perceptual difficulty will selectively impair performance on tasks where the primary constraint on performance is perceptual discriminability, while manipulations which increase the cognitive complexity

of a task will impair performance on tasks where cognitive complexity is the main constraint.

Our view that difficulty in a task might arrive from both "cognitive complexity" and perceptual factors is not incompatible with a single system approach, as even so-called single-system models rely on distinctions between various components such as perception, memory, and decisional processes (e.g., see Nosofsky & Palmeri, 1997 for a process-oriented instantiation of a single-system model). Some might object that this is not a single-system approach at all, but rather a multiple-systems one. However, when defined in this way, a multiple-systems theory of categorization is tautologically true: in order to make a categorization response, information must pass through a wide variety of different systems, ranging from perceptual to cognitive to motor. The relevant question is not whether multiple systems are engaged in categorization, but rather how those systems are functionally arranged, and whether the evidence presented by Maddox, Bohil, et al. (2004) justifies positing multiple *parallel* categorization systems.

Note that what we call cognitive complexity is likely itself not a unitary construct. In fact, many variables affect whether or not a particular categorization task is cognitively difficult to learn (e.g., Shepard, Hovland, and Jenkins, 1961). For example, the complexity of optimal bound, the number of different categories or clusters, the number of candidate dimensions and the number of relevant dimensions are all likely contenders, and they might affect processing at different stages and thus be dissociable. The question of what makes a categorization task difficult is a very interesting one (see Alfonso-Reese, Ashby, & Brainard, 2002) and deserves continued theoretical attention.

Whatever the nature of the additional difficulty that human observers have with information-integration category structures, our results show that this additional source of difficulty interacts with the response manipulation used by Maddox, Bohil, et al. (2004) in a way that is inconsistent with COVIS's hypothesis about the source of this difficulty. Namely, COVIS hypothesizes that human observers find information integration tasks more difficult (resulting in slower learning and lower asymptotic accuracy) because they are not amenable to simple, easily verbalizable classification rules which make so-called rule-based tasks easy for human observers, and thus human observers must fall back on a more robust but less efficient procedural learning system. This explanation fails to account for the interaction between category structure and the Yes-No response key manipulation that we continue to explore in this paper, and thus some other notion of why human observers find information-integration tasks more difficult is required.

Specifically, we show that the additional demands of processing correct/incorrect feedback to yes-no prompts, rather than the lack of the consistent response mapping required by COVIS's procedural learning system, is the source of the additional difficulty that participants have with the Yes-No, information-integration condition. Moreover, the Yes-No response condition interfered with learning of a disjunctive task, which is more complex than the unidimensional task in terms of the number of regions that the stimulus space is partitioned into, but still amenable to an explicit rule-based solution and thus predicted by COVIS to be unaffected by the Yes-No condition.

This significantly changes the interpretation of the original results of Maddox, Bohil, et al. (2004) as to why information-integration tasks are harder than rule-based

tasks for human learners. The same logic behind our studies—that controlling for overall poorer performance on an information-integration task by making the equivalent unidimensional task harder in a different way introduces a confound—underlies other recent work challenging dissociation evidence for COVIS. For example, Nosofsky, Stanton, and Zaki (2005) investigated the role of cognitive complexity in the experimental design of Ashby, Ell, and Waldron (2003). By using stricter response deadlines and more complex category structures, they found that response-motor interference (introduced by switching response button labels at test) affected performance not only on implicit categorization tasks (as predicted by COVIS) but on explicit tasks as well.

Likewise, Stanton and Nosofsky (2007) hypothesized that perceptual discriminability, rather than the operation of distinct explicit vs. implicit category learning systems, was the root of another dissociation reported by Maddox, Ashby, Ing, and Pickering (2004). In that paradigm, Maddox, Ashby, et al. (2004) had tested a prediction of COVIS that rule-based tasks are learned by a system that relies on working memory, and thus should be disproportionately impaired by a working-memory scanning task performed during the processing of feedback. However, as in the present paradigm, Maddox, Ashby, et al. (2004) used rule-based categories that had lower perceptual discriminability than the information-integration categories. Critically, the memory scanning task required visual processing, and thus Stanton and Nosofsky (2007) hypothesized that this added burden on visual processing would interfere more with a perceptual category learning task that was more perceptually demanding due to reduced discriminability between categories. Stanton and Nosofsky (2007)

demonstrated that, consistent with this perceptual-difficulty hypothesis, when the relative perceptual difficulty of the rule-based and information-integration tasks was reversed, the opposite result of Maddox, Ashby, et al. (2004) was obtained. The results of Stanton and Nosofsky (2007) therefore show that perceptual difficulty rather than multiple systems account for a previously reported dissociation.

These experimental studies have led to re-evaluation of evidence previously taken in support of the multiple-systems perspective. A different tack in approaching this issue is to analyze extant data in novel ways (c.f. Zaki, 2004). In one such investigation, Newell and Dunn (2008) conducted a state trace analysis of results reported by Maddox, Ashby, and Bohil (2003). Although the behavioral data suggested a multiple-systems model, a state trace analysis failed to reject a single system account of the same data by considering a richer vocabulary of possible functional architectures, which goes beyond the simple dichotomy of single- vs. multiple-systems. In essence, what appear to be behavioral dissociations that imply a multiple systems model can often in fact be produced by single-system models (Newell & Dunn, 2008; Newell, Dunn, & Kalish, 2011; Nosofsky & Zaki, 1999; Shanks & St. John, 1994; Van Orden, Pennington, & Stone, 1990).

Of course, even though the present results suggest an alternative explanation of Maddox, Bohil, et al.'s (2004) observed dissociation and of Spiering and Ashby's (2008) conclusions, there remains a remarkably large body of work that supports COVIS (Maddox & Ashby, 2004; Ashby & Maddox, 2005). However, this large amount of evidence in no way precludes the critical evaluation of each individual piece of evidence. The current study calls into question previous interpretations of behavioral

dissociations between rule-based and information-integration category learning tasks, and is part of a smaller but growing body of work suggesting that at least some of these results may be insufficient to reject a single-system perspective. We note that, given the extant data which has not yet been addressed by single-system theorists, it is entirely possible that multiple-categorization systems operate in human cognition. However, we also believe a high burden of proof should be placed on accepting multiple-system theories that are often complex and hard to falsify and should be subject to investigations of alternative explanations, as we have done in this article. The body of work on both sides of the debate collectively sheds light on the way in which categorization processes interact with other relevant processes, such as perception, memory, and executive function. Even while this complicates the interpretation of dissociation studies, it ultimately advances our understanding of categorization as situated in the context of other cognitive functions.

References

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6): 716–723. doi:10.1109/TAC.1974.110070
- Alfonso-Reese, L. A., Ashby, F. G., & Brainard, D. H. (2002). What makes a categorization task difficult? *Perception & Psychophysics*, 64, 570–83.
- 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., & 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. [Research Support, U.S. Gov't, Non-P.H.S.]. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 33-53. doi: 10.1037/0278-7393.14.1.33
- Ashby, F. G., & Maddox, W. T. (2005). Human Category Learning. *Annual Review of Psychology*, 56, 149-178.
- Ashby, F. G., & O'Brien, J. B. (2005). Category learning and multiple memory systems. *Trends in Cognitive Sciences*, 9, 83-89.
- Ell, S. W., & Ashby, F. G. (2006). The effects of category overlap on information-integration and rule-based category learning. *Perception & Psychophysics*, 68, 1013–1026. doi:10.3758/BF03193362

Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning.

Journal of Experimental Psychology: General, 127, 107-140.

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple

sequences. *Statistical Science*, 7, 457–472. doi:10.1214/ss/1177011136

Gureckis, T. M., James, T. W., & Nosofsky, R. M. (2011). Re-evaluating dissociations between implicit and explicit category learning: An event-related fMRI study.

Journal of Cognitive Neuroscience, 23, 1697-1709. doi: 10.1162/jocn.2010.21538

Homa, D., Sterling, S., & Trepel, L. (1981). Limitations of exemplar-based generalization and the abstraction of categorical information. *Journal of Experimental*

Psychology: Human Learning and Memory, 7, 418-439. doi: 10.1037/0278-

7393.7.6.418

Kruschke, J. K. (2012, July 9). Bayesian estimation supersedes the *t* test. *Journal of*

Experimental Psychology: General. 142(2), 573-603. doi: 10.1037/a0029146

Maddox, W. T., & Ashby, F. G. (2004). Dissociating explicit and procedural-learning

based systems of perceptual category learning. *Behavioural Processes*, 66, 309–

32. doi:10.1016/j.beproc.2004.03.011

Maddox, W. T., Ashby, F. G., & Bohil, C. J. (2003). Delayed feedback effects on rule-

based and information-integration category learning. *Journal of Experimental*

Psychology: Learning, Memory, and Cognition, 29, 650-662.

Maddox, W. T., Ashby, F. G., Ing, A. D., & Pickering, A. D. (2004). Disrupting feedback

processing interferes with rule-based but not information-integration category

learning. *Memory & Cognition*, 32, 582-591.

- Maddox, W. T., Bohil, C. J., & Ing, A. D. (2004). Evidence for a procedural-learning-based system in perceptual category learning. *Psychonomic Bulletin & Review*, *11*, 945-952.
- Maddox, W. T., Glass, B. D., O'Brien, J. B., Filoteo, J. V., & Ashby, F. G. (2010). Category label and response location shifts in category learning. *Psychological Research/Psychologische Forschung*, *74*, 219-236.
- Maddox, W. T., Ing, A. D., & Lauritzen, J. S. (2006). Stimulus modality interacts with category structure in perceptual category learning. *Perception & Psychophysics*, *68*(7), 1176–90.
- Maddox, W. T., Lauritzen, J. S., & Ing, A. D. (2007). Cognitive complexity effects in perceptual classification are dissociable. *Memory & Cognition*, *35*, 885-894.
- Newell, B. R., & Dunn, J. C. (2008). Dimensions in data: Testing psychological models using state-trace analysis. *Trends in Cognitive Sciences*, *12*, 285-290.
- Newell, B. R., Dunn, J. C., & Kalish, M. (2010). The dimensionality of perceptual category learning: a state-trace analysis. *Memory & Cognition*, *38*, 563-581. doi: 10.3758/MC.38.5.563
- Newell, B. R., Dunn, J. C., & Kalish, M. (2011). Systems of Category Learning: Fact or Fantasy? *Psychology of Learning and Motivation: Advances in Research and Theory*, Vol 54, *54*, 167-215. doi: Doi 10.1016/B978-0-12-385527-5.00006-1
- Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review*, *101*, 53-79. doi: 10.1037/0033-295x.101.1.53

- Nosofsky, R.M, & Palmeri, T.J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review*, 104, 266-300.
- Nosofsky, R. M., Stanton, R. D., & Zaki, S. R. (2005). Procedural interference in perceptual classification: Implicit learning or cognitive complexity? *Memory & Cognition*, 33, 1256-1271.
- Nosofsky, R. M., & Zaki, S. R. (1999). Math modeling, neuropsychology, and category learning: Response to B. Knowlton (1999). *Trends in Cognitive Sciences*, 3, 125-126.
- Plummer, M. (2003). JAGS: A program for analysis of bayesian graphical models using Gibbs sampling. *proceedings of the 3rd international workshop on distributed statistical computing (DSC 2003)*, March 20–22, Vienna, Austria. <http://mcmc-jags.sourceforge.net>
- Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences*, 17, 367-447.
- Shepard, R. N., Hovland, C. I., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs: General and Applied*, 75, 1-42, doi:10.1037/h0093825
- Spiering, B. J., & Ashby, F. G. (2008). Response processes in information-integration category learning. *Neurobiology of Learning and Memory*, 90, 330–8. doi:10.1016/j.nlm.2008.04.015
- Stanton, R. D., & Nosofsky, R. M. (2007). Feedback interference and dissociations of classification: Evidence against the multiple-learning-systems hypothesis. *Memory & Cognition*, 35, 1747-1758.

- Van Orden, G. C., Pennington, B. F., & Stone, G. O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review*, 97, 488-522. doi: 10.1037/0033-295x.97.4.488
- 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, 168–76.
- Willingham, D. B. (1998). A neuropsychological theory of motor skill learning. *Psychological Review*, 105, 558-584. doi: 10.1037/0033-295x.105.3.558
- Willingham, D. B. (1999). The neural basis of motor-skill learning. *Current Directions in Psychological Science*, 8, 178-182. doi: 10.1111/1467-8721.00042
- Zaki, S. R. (2004). Is categorization performance really intact in amnesia? A meta-analysis. *Psychonomic Bulletin & Review*, 11, 1048-1054.
- Zaki, S. R., & Nosofsky, R. M. (2001). A single-system interpretation of dissociations between recognition and categorization in a task involving object-like stimuli. *Cognitive, Affective & Behavioral Neuroscience*, 1, 344-359.
- Zeithamova, D., & Maddox, W. T. (2006). Dual-task interference in perceptual category learning. *Memory & Cognition*, 34, 387–98.

Footnotes

¹ Because we used Gibbs sampling, a Markov Chain Monte Carlo (MCMC) technique to draw samples from the posterior distribution of parameters given the data, we refer to this p value as $pMCMC$. Specifically, it is the one minus the proportion of samples from the posterior distribution where the specified effect is greater than zero (or less, depending on the modal effect direction).

² The noise variance parameter of the decision bound models we fit can be interpreted in a variety of ways, including the learner's uncertainty about the percept or uncertainty/noise in the criterion value which defines the boundary. Either way, it controls the randomness of the model responses, and in the limit as it increases towards infinity the behavior of any of the decision bound models approaches that of a random guessing model. For a subject who truly is guessing randomly, with a reasonable sample size the likelihood of the decision bound models and the random guessing model will be essentially identical, and thus model selection with AIC (or any other similar method, like BIC) will come down to the number of free parameters. If the random guessing model is not considered, this means that the unidimensional model, with two free parameters, will be chosen as the best model.

³ To provide some context for these AIC values, if a particular model can predict a participant's responses exactly (i.e., the responses are perfectly systematic in a way

that can be captured by the model), then the log-likelihood of the data given that model will be 0 and the AIC will be equal to twice the number of parameters. The lowest AIC value actually observed for any participant in a given block was 15. The upper bound on the AIC of the winning model in this case is the AIC of the completely random guessing model, which is 111. Thus the median AIC value in the Yes-No condition was only slightly better than randomly guessing, suggesting that many participants were responding very noisily in this condition. These results confirm the greater proportion of delayed learners found in this condition in the Bayesian analysis.

⁴ As was discussed in the introduction, Maddox, Bohil, et al. (2004) hypothesized that the Yes-No response condition might place a greater burden on working memory, since subjects not only have to keep track of their response and the feedback they received, but the prompt as well. They attempt to control for this by including the Yes-No prompt along with the feedback, which did not qualitatively change their results. However, this manipulation does not address the additional steps the subject must go through to connect feedback, response, and prompt in order to deduce the correct category.

Appendix A: Bayesian Analysis

As described in the main text, traditional accuracy analyses do not account for the role of non-learners. Strategy-based analysis such as fitting decision-bound models is one way to address this, but these analyses depend on the range of models considered. An alternative is to use a more flexible descriptive model, and analyze accuracy using a model which allows for the possibility of non-learners. We did this using a logistic regression model with a learning delay which could vary across subjects.

Specifically, we model the log-odds accuracy in condition j by subject i as chance (0.5) before onset of learning, t_j , and as a linear function of delay-shifted block $b-t_j$,

$$\theta_{i,j}(b) = \beta_{j,0} + (b - t_j)\beta_{j,1}$$

In fitting the model, we used mean-centered block values, in order to remove correlation between the estimated slope and intercepts. Log-odds are converted to accuracy via the logistic function. The parameters of the model are the condition-specific regression coefficients (intercept $\beta_{j,0}$ and slope $\beta_{j,1}$) and the subject-specific learning delays t_j . The regression coefficients are given very weak Normal priors with mean 0 and variance 1000 (much larger than the scale of the coefficients, since a change in log-odds from 0 to 1 corresponds to a change of percentage from 50% to 73%). The delays t_j are given a uniform prior from 0 (no delay in learning) to 5 (the number of blocks, plus one). Any delay over 4 means the subject shows no learning, and so this prior on delay corresponds to a prior probability of being a non-learner of 20%, the same as the prior probability of learning beginning in any given block. We

used a binomial likelihood distribution on the number of correct and total responses for each subject and block. This takes into account the inherent uncertainty in the underlying accuracy given observed counts of correct responses.

The technical details of Bayesian models are beyond the scope of this paper, and there are many good introductions (e.g., Kruschke 2012), so we give only a brief overview of our methods. We fit this model via Gibbs sampling, a Markov Chain Monte Carlo (MCMC) technique which approximates the full, joint distribution over parameters given the data by drawing a series of samples from the conditional distribution of each individual parameter, assuming all other parameters and data are fixed. We used the freely-available JAGS software package (Plummer, 2003) to perform Gibbs sampling, drawing 5000 samples from each of three chains after 5000 burn-in samples (to remove any influence of random starting point), and thinned the chains to 1000 samples in order to ensure reasonably independent samples. We checked for convergence using the Rhat statistic (Gelman & Rubin, 1992), which measures the degree to which the chains end up in the same region of parameter space, despite their random starting points. For all parameters, the value of Rhat was less than 1.1 indicating good convergence.

Figure A1 shows the posterior distribution over learning curve slopes and intercepts, and subject learning delays, and Figure A2 shows the corresponding posterior predictive distribution (the distribution of accuracy expected of a new, non-delayed learner given the available data, calculated on the basis of the posterior parameter samples). A few aspects of these posterior distributions are notable. First, the distributions for AB and Yes-No-Extra conditions are clearly highly similar, while the Yes-No condition stands out. Second, the parameter estimates clearly reflect the

ambiguity between non-learning and low/slow learning: the variance of the estimates of the slope in the Yes-No condition is slightly larger, and possibly bimodal (notice the slight "shoulder" on the lower flank of both distributions, Figure A1, left-middle). Finally, the posterior distribution over delays shows some "spikiness", which results from the fact that accuracy jumps after the delay from chance to the start of the learning curve, which in all conditions is substantially greater than chance (generally 65% or higher, see Figure A2).

Specifying and fitting the model to the data are only the first steps of a Bayesian analysis. Once the distribution of model parameters which are credible given the data is obtained, they must be interpreted in the context of the experimentally relevant hypotheses. For instance, we might wonder whether two conditions differ in terms of learning curve slope or intercept. In order to test this, we can use the samples drawn from the posterior distribution over parameters to estimate the posterior distribution over the *difference* between two of these parameters (Kruschke, 2012). In the text we report two summaries of these distributions. The first, the Highest Posterior Interval (HPI), is the range of the parameter (or difference, etc.) which covers 95% of the posterior probability. This is superficially similar to a classical confidence interval, but unlike a classical confidence interval, we can truly say that, given our data and the assumptions of the model, there is a 95% probability that the parameter/effect lies in this interval (Kruschke, 2012). Because we are using MCMC method to fit the model, the HPI was constructed by finding the smallest interval containing 95% of the sampled parameter values. Difference HPIs were constructed in the same way, but based on the difference between pairs of parameters calculated from each sample.

The second summary of the posterior distribution is the probability that the parameter (or difference, etc.) lies to one side of zero, which we denote p_{MCMC} . This is a measure of the confidence in the *sign* of an effect, and is superficially similar to a classical p value, but, again, it reflects an actual probability that a statement about the model parameters given the data is true. For the sake of clarity, we report the probability that the sign of the mean value is *wrong*, which approaches zero as the confidence in the sign increases, much like a classical p value goes to zero for strong effects. Like the HPI, this too was calculated using the MCMC sample of parameter values (or differences between parameter values), by calculating the proportion of samples/differences less than zero.

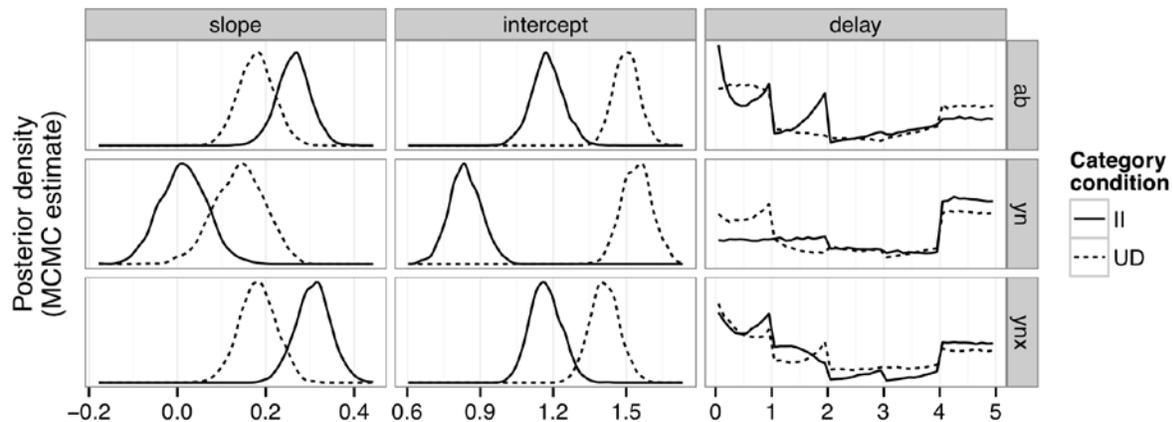


Figure A1: Posterior distributions over model parameters (from 3000 MCMC samples).

Left: Learning curve slopes (increase in log-odds of correct response after one block of learning). Middle: Learning curve intercepts (log-odds of correct response at block $b = 2.5$, or, equivalently average log-odds of correct response for non-delayed learner).

Log-odds of 1 corresponds to 73%. Right: Subject-specific learning delays, collapsed across subjects in each condition.

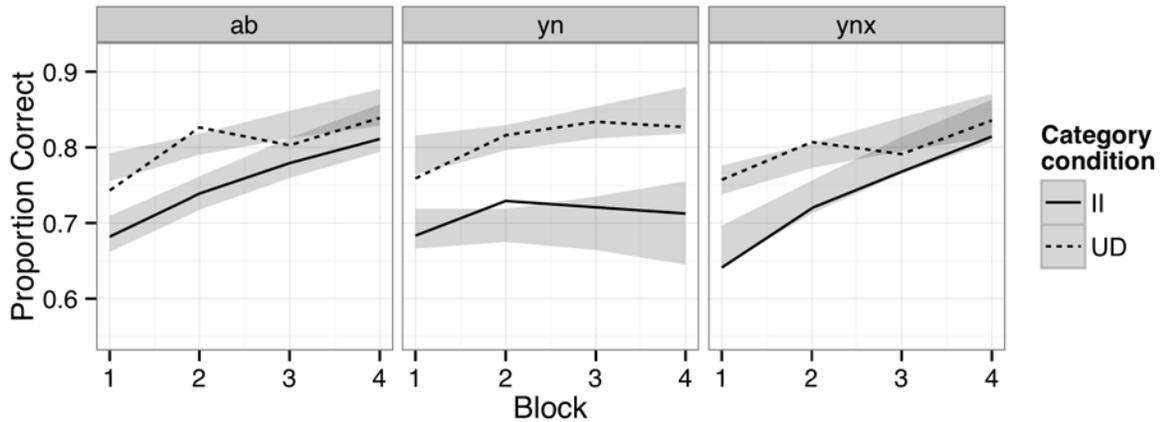


Figure A2: HPIs of the posterior predictive distribution over learning curves in each condition of Experiment 1 (shaded areas), which is the probability distribution over learning curves expected for a new (non-delayed) subject in each condition. Each condition is shown with average accuracy of non-delayed learners (delay < 1) in that condition.

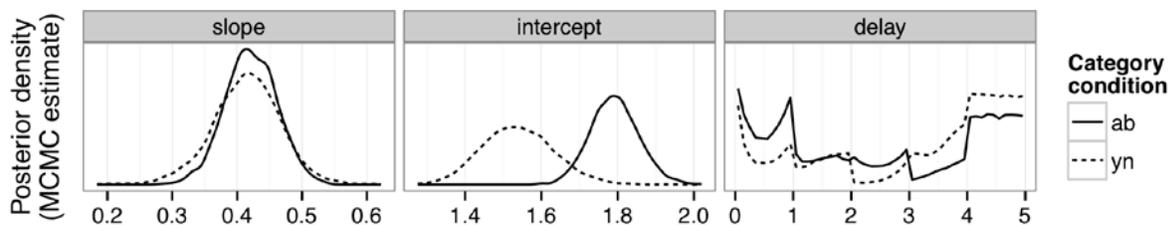


Figure A3: Posterior distributions over model parameters for Experiment 2, the biconditional task. Note the substantially decreased intercept (overall accuracy) and substantially increased delays in the Yes-No condition.

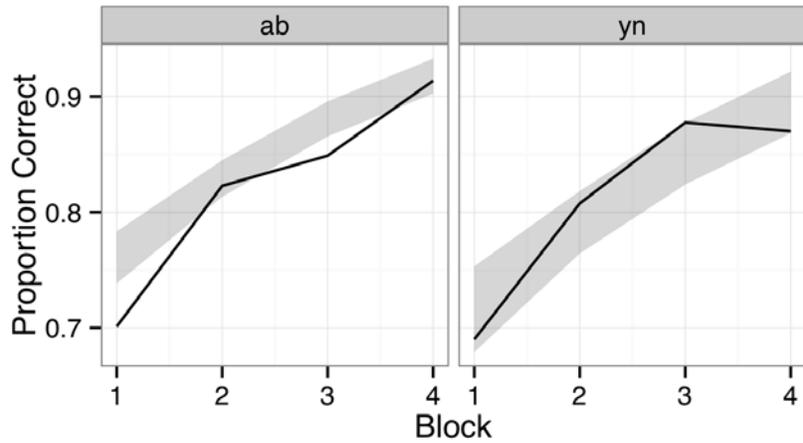


Figure A4: Posterior predictive distributions for Experiment 2 (the biconditional task), with data from subjects who show no delay in learning. Note that the predicted learning curve in the AB condition overshoots the actual accuracy of subjects who show no delay. This illustrates the role that subjects whose learning is delayed but nevertheless reach a high degree of accuracy play in the average accuracy curves in the AB condition.

Appendix B: Decision-Bound Model Analysis

As in previous work comparing UD and II tasks, we fit a variety of decision bound models in order to investigate the strategies that subjects use to perform these categorization tasks. We follow the approach of Maddox et al. (2004), with two exceptions: we included a random guessing model, and we did not fit the non-parametric SPC model (because it was applied only to a conjunctive task which we did not investigate here).

The first two models are consistent with rule-based strategies, where decisions are made on each dimension and if any integration is required it is done post-decisionally. The *unidimensional model* has a single decision bound on one dimension. This model has two parameters (decision bound location and perception noise variance). We fit both horizontal and vertical decision bounds. The *conjunctive model* has two decision bounds, one on each dimension, and assigns one corner of perceptual space to one category and the rest to the other. This model is generally considered to have three parameters: two boundary locations and one noise variance. We fit four versions of this model, one for each quadrant.

The information integration model we fit is the *General Linear Classifier* (GLC), which partitions the stimulus space using a single, diagonal decision bound. This model has three parameters: slope and intercept of the boundary, and noise variance. The GLC contains the unidimensional model as a special case.

The final model we fit is a *random guessing* model, which only has a single parameter, the probability of guessing 'A', which allows for bias towards one category.

For the second experiment, we fit three additional models, all of which divide the stimulus space into more than two regions. First, we fit a biconditional or disjunctive rule-based model, which divides the stimulus space into four corners with a decision boundary on each dimension, and labels opposite corners as the same category. This model has three parameters: the locations of the two boundaries, and the response noise parameter, and can achieve perfect accuracy on this task.

Second, we fit an information-integration model which has two parallel decision boundaries, and assigns the region between them to one category and the two outer regions to the other. This model has four parameters: the shared slope of the boundaries, the two boundary intercepts, and the noise variance. This model can achieve very high accuracy (above 95%), but not perfect accuracy. Third, we also fit a unidimensional rule-based version of this model, where the two parallel boundaries were constrained to be orthogonal to one of the stimulus dimensions, which has three parameters (since the angle is constrained). Since very few participants' blocks were classified as using such a strategy (fewer than 7% overall, and fewer than 5% in the final two blocks) these were grouped with participants using one of single-boundary unidimensional models in the results below.

Models were fit to each subject's data, block-by-block, using Maximum Likelihood estimation of parameters. We used Akaike's Information Criterion (AIC; Akaike, 1974) to compare models, which penalizes models for free parameters, preferring the simplest of a set of models fitting the data equally well. AIC is calculated based on the number of free parameters r and likelihood L as $AIC=2r-2\log L$; smaller AIC values indicate a better and/or more parsimonious fit.

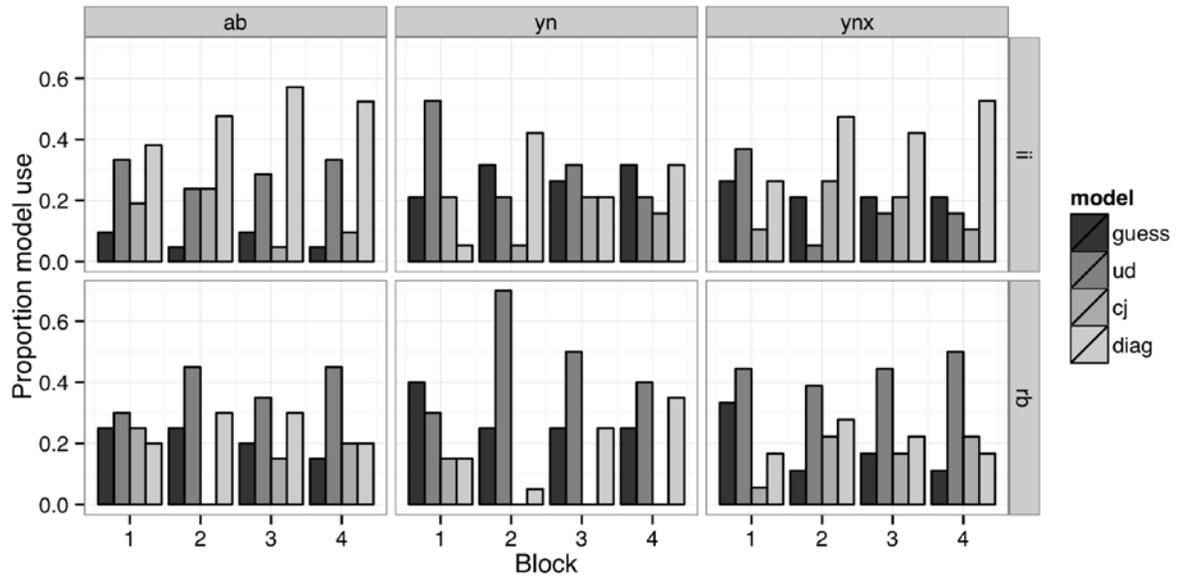


Figure B1: Proportion of participants in each block and condition best fit by each type of model in Experiment 1. For each block, bars correspond, from left to right/darkest to lightest, to the guessing model, the unidimensional model, the conjunctive model, and the GLC (diagonal) model.

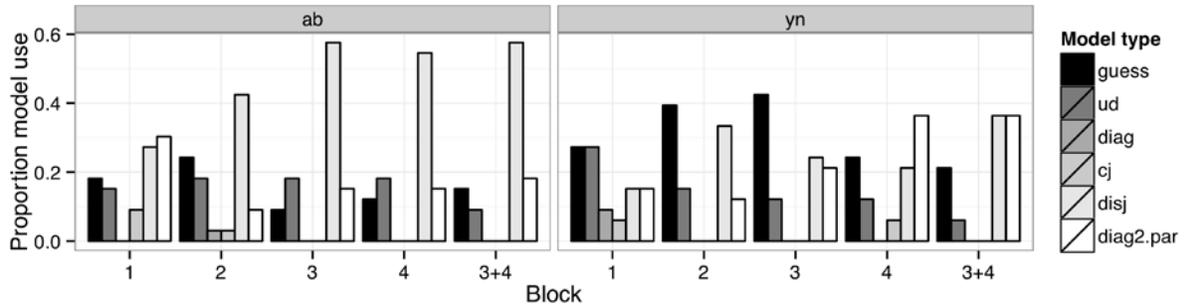


Figure B2: Proportion of participants best fit by each model in each block (including when both the third and fourth blocks are fit together) in the AB and Yes-No conditions in Experiment 2. In each block, from left to right/darkest to lightest, are the proportion of subjects classified as using the guessing, unidimensional rule-based, single-diagonal information-integration, conjunctive rule-based, disjunctive rule-based, and two-parallel-diagonal information-integration strategies.

Table B1: Percent model use in final block by condition in Experiment 1.

Category Structure	Condition	Guess	UD	CJ	GLC
Information-integration	AB	5	33	10	52
	Yes-No	32	21	16	32
	Yes-No-Extra	21	16	11	53
Unidimensional	AB	15	45	20	20
	Yes-No	25	40	0	35
	Yes-No-Extra	11	50	22	17

Note: Guess= random guessing model, UD=unidimensional rule-based model, CJ= conjunctive rule-based, and GLC=diagonal model.

Table B2: Percentage of subjects in each condition best fit by each model in Experiment 2, blocks three and four fit separately and together.

Condition	Block	Guess	UD	GLC	CJ	Disj	2-Diag
ab	3	9	18	0	0	58	15
ab	4	12	18	0	0	55	15
ab	3+4	15	9	0	0	58	18
yn	3	42	12	0	0	24	21
yn	4	24	12	0	6	21	36
yn	3+4	21	6	0	0	36	36

Note: Guess=random-guessing model, UD=unidimensional rule-based model, CJ=conjunctive rule-based model, Disj=optimal disjunctive rule-based model, and 2-Diag= 2-diagonal model.

Table 1. Parameters of bivariate normal distributions of the Gabor patch stimuli in Experiment 1.

	Category	μ_1	μ_2	σ_1^2	σ_2^2	covariance
UD	A	278.7	125.5	75.26	9307	-123.8
	B	320.6	121.9	87.6	9633	-5.546
II	A	217.0	154.3	4567	4816	4587
	B	326.5	94.11	4738	4799	4301

Note: Each pair of parameters sampled from the bivariate normal distributions above was converted to a Gabor patch by deriving the frequency $f = 0.25 + x_1/50$ and orientation $\pi/600$ in radians. The covariance between spatial frequency and orientation in the unidimensional task is non-zero because we estimated them from the actual stimulus values used in Maddox, Bohil, et al. (2004) (which were graciously provided by the authors). These variances/covariances were somewhat different from those reported by Maddox, Bohil, et al. (2004), and so we report the sample covariance here.

Table 2. Parameters of bivariate normal distributions used to generate the biconditional rule based structure in Experiment 2.

Category	μ_L	μ_A	σ_L^2	σ_A^2	covariance
A1	310	190	180	160	0
A2	390	110	180	160	0
B1	310	110	180	160	0
B2	390	190	180	160	0

Note: μ_L = mean length, μ_A =mean angle, σ_L^2 =variance length and σ_A^2 =variance angle. Line length was measured in pixels and angle values were multiplied by a scale factor of $\pi/600$.

Figure Captions

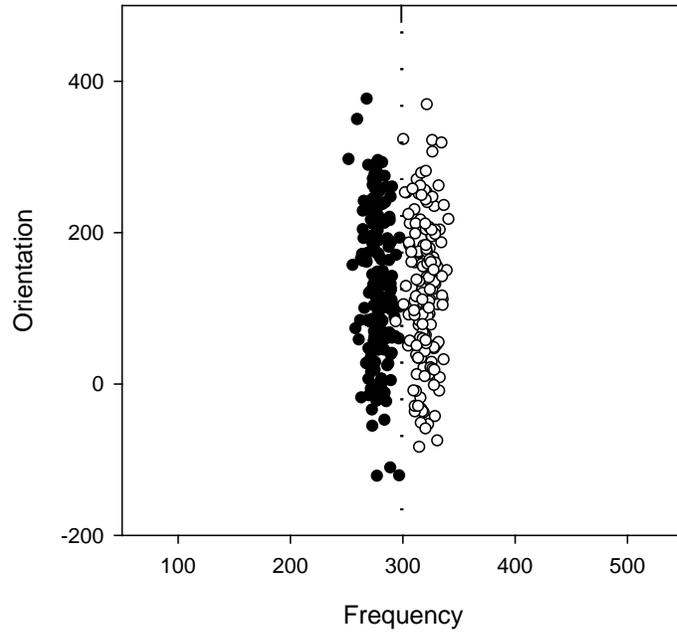
Figure 1. Category structures used in Experiment 1. Top panel: unidimensional rule-based category structure. Bottom panel: information-integration category structure. Each circle represents a stimulus in the experiment with a particular orientation and frequency. The dotted lines represent the optimal decision bound.

Figure 2. Results of Experiment 1. Left panel: Observed probability correct for participants in the four Maddox, Bohil, et al. (2004) replication conditions as a function of the number of training blocks. Right panel: Observed probability correct for the participants in the critical Yes-No condition in which extra feedback is added. Error bars represent ± 1 SEM.

Figure 3. Biconditional rule-based category structure used in Experiment 2. Each circle represents a line stimulus in the experiment with a particular angle and length. The dotted lines represent the optimal decision bound.

Figure 4. Observed probability correct for participants in the Experiment 2 as a function of the number of training block. Error bars represent ± 1 SEM.

Unidimensional Rule-Based Categories



Information-Integration Categories

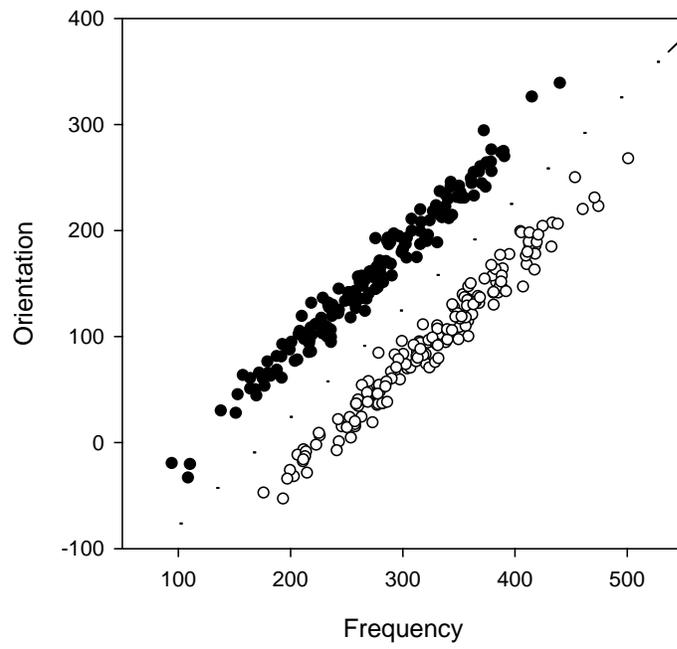


Figure 2

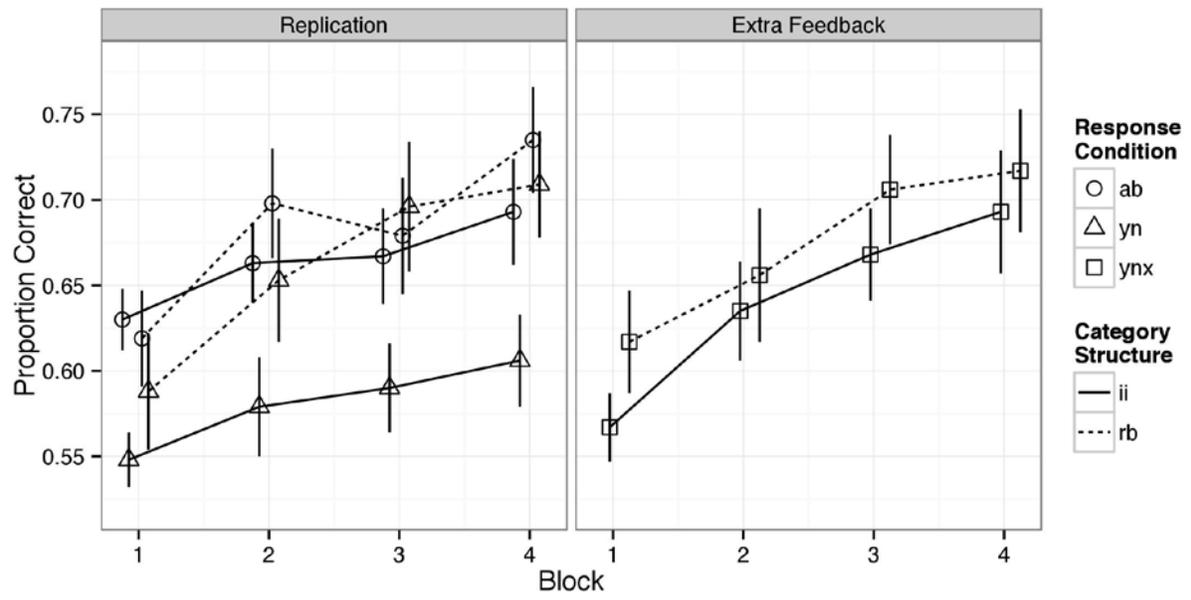


Figure 3

Biconditional rule-based category structure

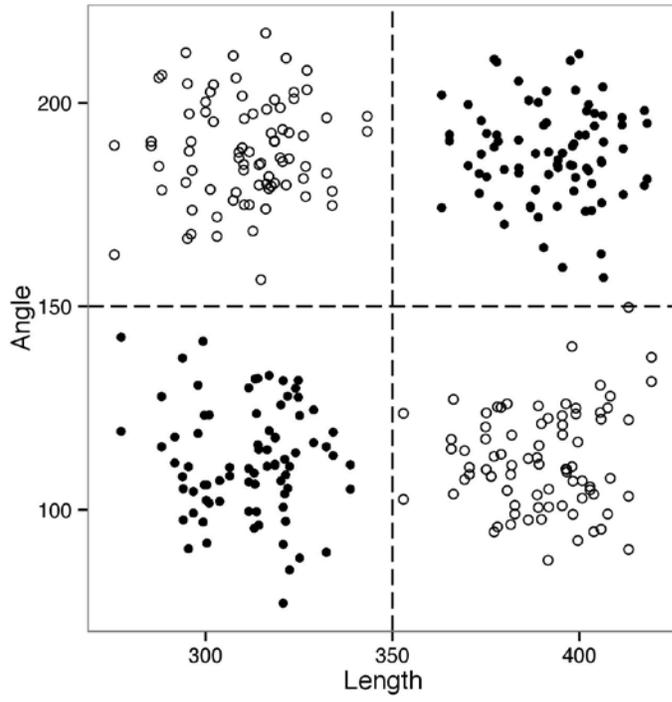


Figure 4

