Journal of Experimental Psychology:
Learning, Memory, and Cognition
2012, Vol. 38, No. 4, 840–859

0278-7393/12/$12.00 DOI: 10.1037/a0027867

The Effect of Feedback Delay and Feedback Type on Perceptual Category Learning: The Limits of Multiple Systems

John C. Dunn
University of Adelaide

Ben R. Newell
University of New South Wales

Michael L. Kalish
University of Louisiana at Lafayette

Evidence that learning rule-based (RB) and information-integration (II) category structures can be dissociated across different experimental variables has been used to support the view that such learning is supported by multiple learning systems. Across 4 experiments, we examined the effects of 2 variables, the delay between response and feedback and the informativeness of feedback, which had previously been shown to dissociate learning of the 2 types of category structure. Our aim was twofold: first, to determine whether these dissociations meet the more stringent inferential criteria of state-trace analysis and, second, to determine the conditions under which they can be observed. Experiment 1 confirmed that a mask-filled feedback delay dissociated the learning of RB and II category structures with minimally informative (yes/no) feedback and also met the state-trace criteria for the involvement of multiple latent variables. Experiment 2 showed that this effect is eliminated when a less similar, fixed pattern mask is presented in the interval between response and feedback. Experiment 3 showed that the selective effect of feedback delay on II learning is reduced with fully informative feedback (in which the correct category is specified after an incorrect response) and that feedback type did not dissociate RB and II learning. Experiment 4 extended the results of Experiment 2, showing that the differential effect of feedback delay is eliminated when a fixed pattern mask is used. These results pose important challenges to models of category learning, and we discuss their implications for multiple learning system models and their alternatives.

Keywords: categorization, category learning, state-trace analysis, functional dissociation, multiple learning systems

Perceptual categorization refers to the capacity to organize different perceptual objects into groups. Many different models have been proposed to account for this capacity (Pothos & Wills, 2011). Some of these, such as the generalized context model (Nosofsky, 1986, 2011) and the prototype model (Minda & Smith, 2002, 2011), explore the extent to which a single representational system or strategy can account for human categorization. Other hybrid models examine the ways in which different processes or strategies are brought to bear on different categorization problems (cf. Kruschke, 2011). One of these is the COVIS model proposed by Ashby, Maddox, and colleagues (Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Ashby & O'Brien, 2005; Ashby, Paul, & Maddox, 2011). This model draws a strong distinction between two different ways in which categorization tasks can be solved and associates each of these with a distinct neurobiological system in the brain. According to COVIS, there exists a verbal system that attempts to solve categorization tasks by generating and testing simple verbalizable hypotheses, or rules, and depends on structures in the anterior cingulate, the prefrontal cortices, the medial temporal lobe, and the head of the caudate nucleus (Ashby & Ell, 2001; Ashby & Spiering, 2004; Nomura et al., 2007; Nomura & Reber, 2008). In addition, there also exists a procedural system that solves categorization tasks by learning to associate a response with regions of perceptual space based on reinforcement (Ashby et al., 2011) and depends on neural structures in the tail of the caudate nucleus (Ashby et al., 1998; Nomura & Reber, 2008).

A feature of the COVIS model is that the verbal and procedural systems compete to determine the response to any one categorization judgment. Which system dominates is partly determined by their relative past success in generating the correct response. According to the model, despite an initial bias in favor of the verbal system, the procedural system will, in time, come to determine the response if the categorization task cannot be solved using a simple verbalizable rule. This has led to an extensive series of studies that have compared the learning of *rule-based* (RB) and *information-integration* (II) category structures. RB structures define category membership according to values on salient stimulus

John C. Dunn, School of Psychology, University of Adelaide, Adelaide, South Australia, Australia; Ben R. Newell, School of Psychology, University of New South Wales, Sydney, New South Wales, Australia; Michael L. Kalish, Institute of Cognitive Science, University of Louisiana at Lafayette.

The research reported in this article was supported by Australian Research Council Discovery Grant DP0877510. We thank Emily Adcock, Carissa Bonner, and Rachel Stephens for assistance with data collection.

Correspondence concerning this article should be addressed to John C. Dunn, School of Psychology, University of Adelaide, South Australia, Australia 5005. E-mail: john.c.dunn@adelaide.edu.au

dimensions. The verbal system is able to learn to categorize such structures by identifying the relevant dimensions and discovering the appropriate rule (e.g., if the size of a circle is greater than *x* and the orientation of a line is greater than *y,* then the stimulus is a member of category *A*). In contrast, II structures define category membership according to the conjoint values on two or more dimensions using rules that are not easily verbalizable. Consequently, such structures cannot be learnt by the verbal system, which must eventually yield control of the response to the procedural system.

Given that categorization of RB and II structures depends primarily on the verbal and procedural systems, respectively, it is possible to test two kinds of prediction made by COVIS. The first is the fundamental claim that there are two distinct systems underlying RB and II learning. Although the definition of what constitutes a *system* is not well settled, it is generally agreed that a minimal requirement is that systems be functionally distinct (Sherry & Schacter, 1987). Because the verbal and procedural systems operate on different principles, it should be possible to find variables that differentially affect their functions. This has most often taken the form of finding variables that dissociate RB and II learning. Variables that selectively affect the verbal system should affect RB learning but have little or no effect on II learning. Similarly, variables that selectively affect the procedural system should affect II learning but have little or no effect on RB learning. The second kind of prediction concerns the properties of the verbal and procedural systems proposed by COVIS. As well as predicting that RB and II learning can be dissociated, COVIS also predicts the kinds of variables that produce this result and the conditions under which it occurs. Tests of this prediction have tended to be investigated hand-in-hand with the attempts to dissociate RB and II learning. For example, because the procedural system learns an association between a region of perceptual space and an overt response, if the nature of this response is changed, performance should suffer. Consistent with this prediction, Ashby, Ell, and Waldron (2003) found that a reassignment of the response buttons associated with each of two categories impaired the categorization of II structures following learning but had little effect on the categorization of RB structures. Conversely, because the verbal system depends upon working memory, adding a cognitive load or limiting the time available for processing feedback should selectively impair this system. Consistent with these predictions, Maddox, Ashby, Ing, and Pickering (2004) found that reducing the time available to process feedback on each trial affected RB learning but had little effect on II learning. Similarly, Zeithamova and Maddox (2006) found that the addition of a working memory load impaired RB learning but had little effect on II learning.

Our aim in the present article is to examine the two kinds of prediction made by COVIS in relation to two reported dissociations involving the selective effects of feedback delay and feedback type on learning II structures. COVIS predicts that a delay of even a few seconds between the categorization response and presentation of feedback should impair learning by the procedural system. This prediction is derived from the neurobiological mechanism that is the basis of this system. This mechanism depends upon local reward-mediated learning within the tail of the caudate nucleus and requires that the pattern of activation associated with the response be maintained until the occurrence of a dopamine-mediated reward signal (Maddox, Ashby, & Bohil, 2003). However, the morphology of cells in the tail of the caudate nucleus allows activation to be maintained for only a few seconds. If reward is delayed beyond this, learning will be impaired. Maddox et al. (2003) tested this prediction by comparing RB and II learning with immediate feedback and with feedback following a delay of either 2.5 s, 5 s, or 10 s. They found that when feedback was delayed, essentially no II learning occurred.

COVIS also predicts that the type of feedback should differentially affect RB and II learning. Feedback may either be minimal, when the participant is informed only if he or she is correct or incorrect, or it may be full, when the participant is informed of the correct response if his or her response is incorrect. According to COVIS, RB learning benefits from full feedback because the verbal system uses this additional information to generate and test more appropriate rules. In contrast, because the procedural system relies on local reward-based learning, it is indifferent to the type of feedback. While this may suggest that II learning should also be unaffected, Maddox, Love, Glass, and Filoteo (2008) argued instead that full feedback impairs II learning because this type of feedback encourages the verbal system to maintain its control of the response, despite the fact that it is unable to learn the II structure.

COVIS is supported in two ways by the reported effects of feedback delay and feedback type. First, the fact that they differentially affect RB and II learning in itself supports the concept of multiple category learning systems. Second, the nature of these differential effects flows directly from the properties that COVIS attributes to each of the two systems. Despite this, we believe that there are good reasons to question both conclusions (Newell, Dunn, & Kalish, 2011). In the section to follow, we present two arguments. First, we show that functional dissociations cannot, in themselves, support the inference of multiple latent variables and illustrate this point in relation to the selective effect of changing response assignments on II learning (Ashby et al., 2003; Nosofsky, Stanton, & Zaki, 2005). We propose instead that state-trace analysis provides the appropriate analytic tool to examine these effects (Newell & Dunn, 2008). Second, we show that even when dissociations support the inference of multiple latent variables, consistent with COVIS, this pattern may depend on features of the task that, according to COVIS, should be irrelevant. We illustrate this point in relation to the selective effect of feedback disruption on RB learning (Maddox et al., 2004; Stanton & Nosofsky, 2007). On the basis of these arguments, we conducted four experiments examining whether variations in feedback delay and feedback type dissociate RB and II learning according to the more stringent criteria of state-trace analysis and, if they do, whether this depends on features that, according to COVIS, should be irrelevant.

The Limitation of Functional Dissociations

Functional dissociations provide only weak evidence in favor of multiple processing systems since it is possible to observe a dissociation (including a double dissociation) when performance on both tasks (in the present case, RB and II learning) depend upon a single intervening process or latent variable (Dunn & Kirsner, 1988; Newell & Dunn, 2008). This is illustrated in Figure 1 in relation to the effect of response assignment on categorization performance for RB and II structures (Ashby et al., 2003; Nosofsky et al., 2005). Figure 1a shows the results found by Nosofsky et

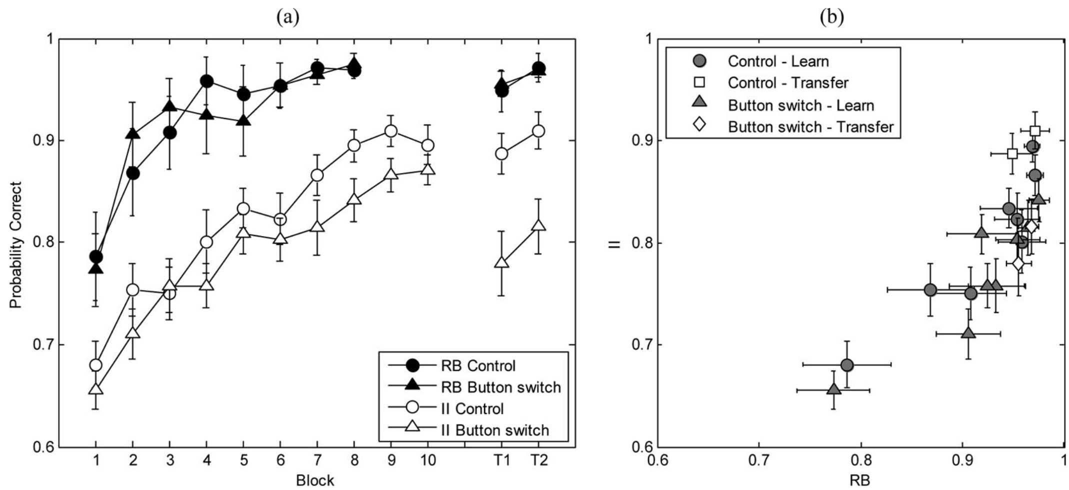

Figure 1. Data from Nosofsky, Stanton, and Zaki (2005, Experiment 1). a: Mean proportion correct on RB and II category structures for each block of trials in the learning or preswitch phase (Blocks 1–8 for the RB structure and Blocks 1–10 for the II structure) and in the postswitch or transfer phase (T1 and T2). In the control condition, the same response assignment was maintained across both phases. In the button switch condition, the response assignment was switched between learning and transfer phases. Error bars indicate standard errors. b: The same data in the form of a state-trace plot. Filled symbols correspond to performance in the preswitch phase. Unfilled symbols correspond to performance in the postswitch phase. Error bars indicate standard errors. II = information-integration; RB = rule-based. Panel a adapted from "Procedural Interference in Perceptual Classification: Implicit Learning or Cognitive Complexity?", by R. M. Nosofsky, R. D. Stanton, and S. R. Zaki, 2005, *Memory & Cognition, 33,* p. 1261. Copyright 2005 by the Psychonomic Society.

al. (2005) in their Experiment 1, which replicated the results found earlier by Ashby et al. (2003).[1] In the study by Nosofsky et al., participants learned to classify either RB or II category structures by pressing one of two designated keys on a computer keyboard. For the RB structure, training consisted of eight blocks of 50 trials (400 trials total), while for the II structure (called the Diagonal structure in the study), training consisted of 10 blocks of trials (500 trials total). Following training, participants were instructed that the response assignments they had learned were now reversed. They then completed two blocks of trials under this regimen.

As can be seen from Figure 1a, the change in response assignment yielded a single dissociation—performance on the II structure was impaired to a greater extent than performance on the RB structure. This outcome is consistent with COVIS, which proposes that the procedural system forms an association between regions of perceptual space and a specific response. Changing the nature of the response assigned to each region of space, as in the transfer phase, therefore impairs the function of this system. In contrast, the verbal system is relatively unaffected by the response switch because this system has learned an abstract rule that does not depend upon specific response assignments.

Although COVIS successfully predicts a greater effect of response reassignment on II categorization, the data do not support the inference that more than one latent variable, process, or processing system is involved. This is demonstrated in Figure 1b, which plots the data from Figure 1a in the form of a *state-trace plot* (Bamber, 1979). This is a parametric plot of performance on one outcome measure (level of II learning) as a function of performance on the other outcome measure (level of RB learning) across the set of experimental conditions. The state-trace plot provides a means of determining if changes in the outcomes, despite evidence of functional dissociations, are nevertheless consistent with a single latent variable (Newell & Dunn, 2008). This would be the case if learning of both RB and II structures depended upon the same kind of mechanism and differed only in how quickly they reached asymptote. This idea is captured in state-trace analysis by supposing that categorization performance for RB and II structures can be considered as two different but unknown monotonic functions of a single latent variable (such as amount of training). If so, the data points corresponding to the different experimental conditions will necessarily fall on a monotonically increasing curve in the state-trace plot (Bamber, 1979). This is called a *one-dimensional* state-trace. On the other hand, if the learning of RB and II category structures depends upon different latent variables or processes or processing systems and these are differentially affected by the experimental variables (training blocks and change in response assignment in the present case), then the data will not, in general, fall on a single monotonic curve in state-trace space. This is called a *two-dimensional* state-trace. It is this pattern, rather than the occurrence of dissociations, that provides decisive evidence for multiple latent variables, processes, or systems (Newell & Dunn, 2008).

Figure 1b shows that a functional dissociation can be consistent with a one-dimensional state-trace. The filled data points show average RB and II performance in each block of the preswitch

[1] We are grateful to Roger Stanton for providing the raw data from this experiment as well as those for Experiments 1 and 2 from Stanton and Nosofsky (2007).

training phase. Here and in following figures, points in the state-trace plot are colored gray to suggest that they reflect both RB and II category structures that are identified by black and white symbols, respectively, in Figure 1a. It is apparent that the data from the preswitch phase are well described by a single monotonically increasing curve, which is only to be expected since the two groups (control and button switch) are tested under identical conditions during this phase. The unfilled data points show average RB and II performance in each block of the postswitch transfer phase. Crucially, these points show little evidence of departing from the monotonic curve that describes the training data and therefore provide little evidence for a two-dimensional state-trace. This conclusion is supported by the statistical analysis that we describe in the following section. The data are therefore consistent with response reassignment impairing categorization performance for both RB and II structures in essentially the same way. The fact that the effect of switching response buttons is more apparent for the II structure is due to the nature of the monotonic relationship between RB and II performance across training. Although the data dissociate RB and II learning, there is no evidence of the involvement of distinct category learning systems. This is also consistent with the conclusion reached by Nosofsky et al. (2005) that the observed dissociation was due to how much was learned about each category structure rather than to the ways in which the structures were learned.

State-Trace Analysis

State-trace analysis provides insights that more traditional analyses, such as analysis of variance (ANOVA), do not. In the experiments to follow, we used state-trace analysis to determine whether the effects of pairs of variables on learning RB and II category structures lead to a one-dimensional or two-dimensional state-trace. This, in turn, required the application of a tailored statistical procedure first described by Newell, Dunn, and Kalish (2010). The aim of this section is to briefly describe this procedure and to contrast it with ANOVA.

As noted above, a functional dissociation does not necessarily imply a two-dimensional state-trace plot. Dissociations are often identified by examining relevant interaction terms using ANOVA, and this is precisely what Nosofsky et al. (2005) did in their experiment. The claim that there was a selective effect of switching responses on II learning was supported by a significant interaction between blocks (preswitch vs. postswitch), condition (control vs. button switch), and category structure (RB vs. II). However, such interactions are neither necessary nor sufficient for a two-dimensional state-trace (Loftus, Oberg, & Dillon, 2004). This is because the aims of the two procedures are different. ANOVA is concerned with detecting departures from a particular model of effects, called the general linear model, while state-trace analysis is concerned with detecting departures from a monotonically ordered configuration of points in outcome space. In the application of ANOVA to the present data, category structure is treated as an independent variable with two levels (e.g., RB vs. II), which, like the remaining independent variables, is assumed to be linearly related to a single dependent variable (categorization performance). In the present application of state-trace analysis, the two category structures define two different dependent variables that need not be linearly related to the other independent variables.

Although ANOVA and state-trace analysis make different assumptions and have different goals, they are not totally unrelated. In the Appendix, we show that if ANOVA reveals one of several patterns of effects, then the corresponding state-trace must be one-dimensional. However, the converse is not true—even if ANOVA reveals significant interactions, the corresponding state-trace may still be one-dimensional. This is the case for the data in Figure 1b. Even though Nosofsky et al. (2005) found a significant three-way interaction, the state-trace is still one-dimensional, as determined by the procedure introduced by Newell et al. (2010) and outlined below.

The statistical procedure introduced by Newell et al. (2010) consists of two parts: a model fitting part and a model testing part. In the model fitting part, two kinds of model are fit to the observed data. The first is an *order-restricted two-dimensional model,* which allows the state-trace to be two-dimensional (i.e., nonmonotonic) but specifies a prior ordering of some of the data points. This model is fit to exclude departures from monotonicity that reflect unexpected experimental or measurement error (Prince, Brown, & Heathcote, 2012). For example, in Figure 1, the two-dimensional model may impose the restriction that performance on both RB and II tasks should not decrease across training blocks. This means that random departures from this expectation will not counted as evidence against monotonicity. The fit of this model is then compared to the fit of an *order-restricted one-dimensional model* that is nested within the two-dimensional model with the additional constraint that the data points must be ordered in the same way on both dependent variables.[2] The greater the difference in the goodness of fit between the two models (measured by the difference in the G^2 statistic, ΔG^2), the greater is the evidence for a two-dimensional state-trace. The assessment of this evidence is formalized in the model testing part where the empirical distribution of ΔG^2 is estimated using a Monte Carlo simulation based on the parametric bootstrap procedure developed by Wagenmakers, Ratcliff, Gomez, and Iverson (2004). At each iteration of the procedure, a bootstrap sample of the data is fit by the order-restricted one-dimensional model, which is used to generate a new sample that is fit by both the (order-restricted) one-dimensional and two-dimensional models. The ΔG^2 obtained from this comparison can be considered as a sample from the unknown empirical distribution of ΔG^2 under the hypothesis that the one-dimensional model is true. Using this distribution, we tested the hypothesis that the value of ΔG^2 obtained from the original data is also a sample from this distribution (i.e., that the data form a one-dimensional state-trace) by calculating a p value in the normal way. We typically based our estimate on 10,000 samples.

We applied the above statistical analysis to a subset of the data shown in Figure 1 consisting of the last two preswitch training blocks (Blocks 7–8 for the RB structure and Blocks 9–10 for the II structure) and the following two postswitch transfer blocks.

[2] We also considered a third model called the *order-restricted one-dimensional nonoverlap model*. This is used to exclude the possibility that the data trivially satisfy monotonicity due to the fact that the means of the conditions fail to overlap in state-trace space. This was not a factor in the present series of experiments, and so we have not discussed it further.

Although ANOVA revealed a significant interaction between condition (control vs. button switch) and category structure (RB vs. II), we were unable to reject the one-dimensional model ($\Delta G^2 = 2.16$, $p = .23$). This illustrates the point that a functional dissociation supported by relevant significant interaction term does not guarantee a two-dimensional state-trace.

Disappearing Dissociations

COVIS predicts the conditions under which categorization performance depends on the involvement of multiple category learning systems. State-trace analysis can be used to either confirm or disconfirm these predictions. Figure 2 shows the results of two experiments conducted by Stanton and Nosofsky (2007) that investigated the effect of disrupting the processing of feedback on RB and II learning. According to COVIS, the verbal system uses working memory to process feedback. If working memory is directed to another task, then feedback processing is impaired in this system, leading to a decrement in RB learning. In contrast, the procedural system does not depend on working memory and so is less affected by additional processing following feedback, resulting in little or no effect on II learning. The effect of disrupting feedback on category learning was first investigated by Maddox et al. (2004) and subsequently replicated by Stanton and Nosofsky. In both experiments, after being given feedback on their categorization response, participants were required to perform a memory scanning task that occupied working memory. The manipulation of interest was whether this task occurred immediately following feedback or after a 2.5-s delay.

Figure 2a shows the results from Stanton and Nosofsky's (2007) Experiment 1 presented in the form of a state-trace plot. These data show that across the four training blocks, RB performance was selectively impaired by immediate presentation of the memory scanning task. Importantly, the state-trace appears two-dimensional since the data points do not clearly fall on a monotonically increasing curve. However, when we conducted a statistical analysis of these data, although the result approached significance, it was not possible to reject the one-dimensional model ($\Delta G^2 = 4.92$, $p = .083$). Given the apparent pattern in Figure 2a, we suspected that this may have been due to the precision of the data, which we increased by combining the first two blocks and the last two blocks of trials. In this case, it was possible to reject the one-dimensional model ($\Delta G^2 = 4.98$, $p = .014$).[3] In other words, these data demonstrate the involvement of two or more latent variables affecting RB and II performance and differentially affected by feedback disruption. This is consistent with COVIS, which interprets these latent variables as the verbal and procedural category learning systems.

COVIS assumes that when performance depends on the verbal system, feedback disruption will impair performance. It follows that if the RB structure in Figure 2a is replaced by a different RB structure, the same pattern of results should occur since COVIS assumes that all RB structures defined by simple, verbalizable rules are learned by the verbal system. Figure 2b shows the state-trace plot that results from combining performance on the II structure from Stanton and Nosofsky's (2007) Experiment 1 with performance on the RB structure from their Experiment 2. The RB structure in Experiment 1 consisted of stimuli varying on two dimensions, one of which determined category membership, and is labeled RB(2D) in Figure 2a. The RB structure in Experiment 2 consisted of stimuli varying on four dimensions of which one determined category membership. This is labeled RB(4D) in Figure 2b. What is striking about these data is that the two-dimensional pattern and corresponding dissociation apparent in Figure 2a are no longer present. Instead, the data points appear to fall on a single monotonically increasing curve, consistent with a one-dimensional state-trace. Formal analysis supports this intuition. The one-dimensional model could not be rejected both when all four blocks were included ($\Delta G^2 = 0.48$, $p = .361$) and when adjacent blocks were combined in the same way as the previous analysis ($\Delta G^2 = 0$, $p = 1$). This change in the dimensionality of the state-trace poses a challenge to COVIS since there is no particular reason why the verbal system should not be involved in the learning of both RB structures and therefore be equally impaired by the disruption of feedback processing. Although COVIS can explain the two-dimensional state-trace in Figure 2a by postulating the involvement of different category learning systems, it can only explain the one-dimensional state-trace in Figure 2b by postulating the involvement of a single category learning system, presumably the procedural system.[4] However, the RB(4D) structure is no less verbalizable than the RB(2D) structure, and so it is difficult to see why it cannot also be learned by the verbal system.

Outline of Experiments

In the four experiments following, we examined the effects of delaying feedback following a categorization response and the type of feedback provided—minimal or yes/no feedback versus full or corrective feedback. Table 1 presents an outline of the structure of these experiments. In each experiment, participants learned to categorize either an RB or II category structure under either immediate or delayed feedback. These conditions differed in the duration of the interval between a response and the provision of feedback; 0.5 s in the No Delay condition and 5 s in the Delay condition. In both cases, a stimulus mask was presented during the interval between response and feedback. In Experiments 1 and 3, this was a Gabor patch that was both similar to the stimuli to be categorized and varied from trial to trial. In Experiments 2 and 4, the mask was a cross-hatch pattern that was dissimilar to the stimuli and did not vary across trials. In Experiments 1 and 2, feedback was minimal, while in Experiments 3 and 4, it was full.

The four experiments allowed us to investigate the following questions: (a) Do feedback delay and feedback type dissociate RB and II learning under the more stringent criteria of state-trace analysis—that is, does variation in these factors lead to a two-

[3] The results of such averaging need to be treated with some caution as it may introduce distortions from combining data from different parts of the outcome space. In the present case, the averaged data points are relatively close together, which will tend to militate against these effects.

[4] Alternatively, both systems may be in play but affected by the independent variables in the same way (as discussed in the Appendix). In this case, it just happens that learning the RB(4D) structure in Experiment 2 is affected by variation in training and feedback interference in exactly the same way as learning the II structure in Experiment 1. While possible, this seems very unlikely.

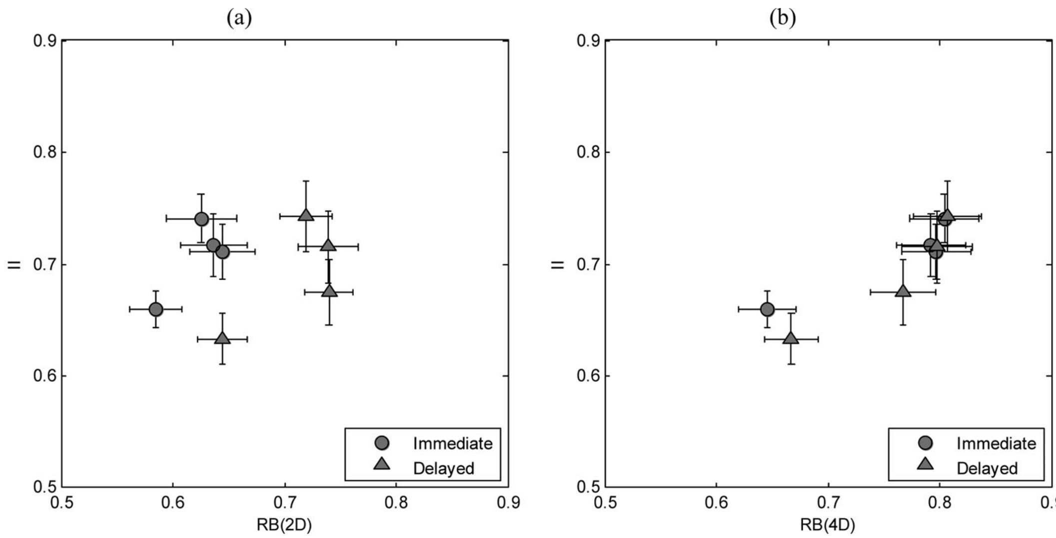

Figure 2. Data from Stanton and Nosofsky (2007, Experiments 1 and 2). a: State-trace plot of mean proportion correct from Experiment 1 (two-dimensional RB category structure). b: State-trace plot of mean II performance from Experiment 1 against mean RB performance from Experiment 2 (four-dimensional category structure). Error bars indicate standard errors. II = information-integration; RB = rule-based.

dimensional state-trace? (b) If so, does this depend upon the type of mask? In other words, does the two-dimensional state-trace become one-dimensional under conditions that should be irrelevant according to COVIS? We discuss the reasons why we examined this possibility after considering the results of Experiment 1.

Experiment 1

Experiment 1 examined the differential effects of feedback delay and training block on RB and II learning under minimal feedback and used a Gabor mask. It was based on the stimuli and procedures used by Maddox and Ing (2005) in their replication of the results found by Maddox et al. (2003). To address a potential confound in the earlier study, Maddox and Ing compared an RB category structure and an II category structure in which both dimensions were relevant to the categorization response. We used the same category structures in Experiment 1. We also replicated their use of a Gabor mask between response and feedback, and although it is not stated in the report, we presumed that this feedback was minimal.

According to COVIS, feedback delay should primarily affect II learning and have little or no effect on RB learning. Maddox and Ing (2005) found that the presence of a 5-s delay selectively impaired, but did not eliminate, II learning while having little effect on RB learning. This conclusion was based on a marginally significant interaction between feedback delay and category structure (p = .05). The aim of Experiment 1 was therefore to determine if this result meets the more stringent criteria of state-trace analysis. If so, then it is consistent with the involvement of multiple learning systems as postulated by COVIS.

Table 1
Structure of Experiments 1–4

Experiment	Conditions	Feedback type	Mask
1	Delay/no delay	Minimal	Gabor
2	Delay/no delay	Minimal	Pattern
3	Delay/no delay	Full	Gabor
4	Delay/no delay	Full	Pattern

Method

Participants. One hundred and thirty undergraduate students from the University of Adelaide (Adelaide, South Australia, Australia) participated in return for course credit or a payment of AUD12. Each participant completed one experimental condition. There were 34 participants assigned to the RB No Delay condition, 34 to the RB Delay condition, and 30 and 32 participants assigned to the II No Delay and II Delay conditions, respectively. In all of the experiments reported here, we followed Maddox and Ing (2005) and our own previous work (Newell et al., 2010) by adopting a learning criterion. Participants who did not exceed 27% correct responses in the final block of training were excluded. We used this criterion because the performance of those excluded was not reliably different from chance (25%). Applying this criterion excluded four and 10 participants from the RB No Delay and RB Delay conditions, respectively, and four and 16 participants from the II No Delay and II Delay conditions, respectively.

Stimuli and apparatus. The categorization stimuli were generated using the same procedures used by Maddox and Ing (2005). The stimuli were sine wave gratings that varied in spatial frequency and orientation. Twenty stimuli in each of the four categories were generated by sampling randomly from the same four parameter distributions used by Maddox and Ing. These parameters are shown in Table 2. Actual values of spatial frequency (f) and orientation (o) were generated from a random sample (x, y) from these distributions using the following transformations: $f = 0.25 + x/50$, $o = y.\pi/500$. The stimuli were generated

Table 2
Category Distribution Parameters for the Stimuli Used in Experiments 1–4 and in Maddox and Ing (2005)

Category structure	μ_x	μ_y	σ_x^2	σ_y^2	cov_{xy}
Rule-based					
Category A	268	93	75	75	0
Category B	268	157	75	75	0
Category C	332	93	75	75	0
Category D	332	157	75	75	0
Information-integration					
Category A	268	125	75	75	0
Category B	300	157	75	75	0
Category C	300	93	75	75	0
Category D	332	125	75	75	0

using MATLAB (Mathworks, Natick, MA) routines from the Psychophysics Toolbox (Brainard, 1997). When presented, each stimulus was 200 × 200 pixels and centered on the computer screen.

Procedure. The experiment consisted of four 80-trial blocks. Within each block, all 80 stimuli were presented in a random order (with different orders for each subject). Participants were told to learn which of four categories (labeled, 1, 2, 3, and 4) each stimulus belonged to. On each trial, a stimulus was presented, and participants terminated the display by pressing one of the keys labeled *1–4* on the computer keyboard corresponding to Categories 1–4, respectively. Following the response, a mask appeared for either 0.5-s (No Delay condition) or 5-s (Delay condition). The mask was a Gabor patch that was twice the dimensions of the stimulus (i.e., 400 × 400 pixels) and had frequency and orientation values drawn at random from within the range of stimulus values. Following presentation of the mask, feedback appeared on the computer screen for 0.75 s. If the response was correct, the word "Correct" was presented; otherwise, the word "Incorrect" was presented. Following presentation of feedback, the screen was blank for either 5 s (No Delay condition) or 0.5 s (Delay condition) before the next trial commenced. The sequence and timing of these events were same as those used by Maddox and Ing (2005).

At the end of each block of trials, participants were given feedback on the number of correct responses in the previous block and reminded of the level of chance performance (i.e., 25%).

Results and Discussion

The mean accuracy rates averaged across participants are presented in Figure 3. Figure 3a plots the data as a function of block (1–4), category structure (RB vs. II), and feedback delay (No Delay vs. Delay). Figure 3b plots the same data in the form of a state-trace plot with axes defined by category structure. The inset graph in Figure 3b shows the state-trace plot corresponding to all participants (i.e., averaged over both learners and nonlearners). We have shown previously that the inclusion of different proportions of nonlearners across conditions can increase the dimensionality of the state-trace (Newell et al., 2010) and include this additional information to check that the apparent dimensionality of the state-trace does not depend upon the exclusion of nonlearners. In neither this nor subsequent experiments does the overall shape of the state-trace change substantially if nonlearners are included.

We analyzed the data in two ways. First, a propos of the manner of plotting the data in Figure 3a, we conducted a 2 (category structure) × 2 (feedback delay) × 4 (block) mixed-design ANOVA. This revealed a main effect of block, $F(3, 282) = 74.78$, $p < .001$, indicating learning; a main effect of category structure, $F(1, 94) = 6.55$, $p = .012$, indicating superior accuracy overall for RB learning compared to II learning; no main effect of feedback delay, $F(1, 94) = 2.91$, $p = .091$; but significant interactions between block and feedback delay, $F(3, 282) = 3.30$, $p = .021$; between block and category structure, $F(3, 282) = 2.83$, $p = .039$; between category structure and feedback delay, $F(1, 94) = 4.13$, $p = .045$; and between block, category structure, and feedback delay, $F(3, 282) = 2.82$, $p =$

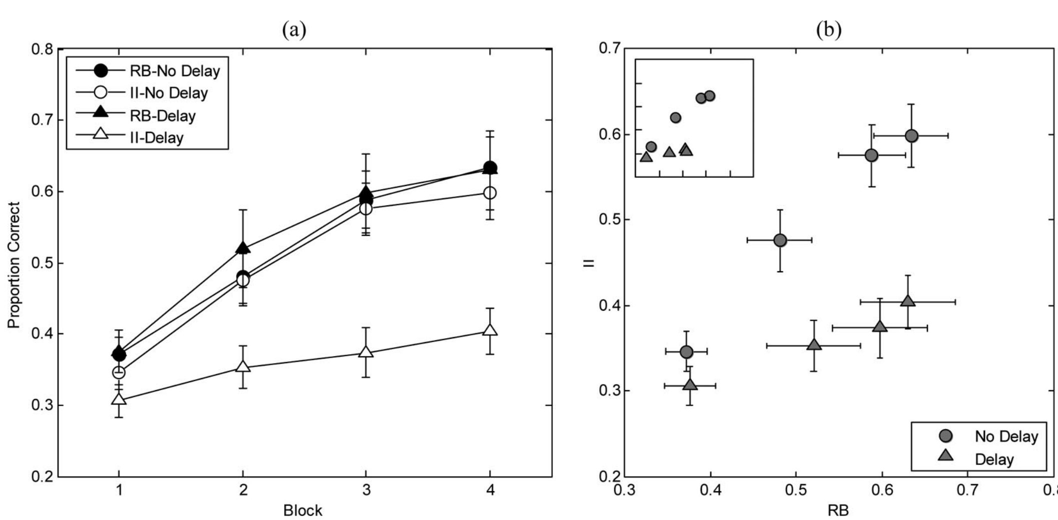

Figure 3. Results of Experiment 1. a: Mean proportion correct as a function of category structure, feedback delay, and block. Data are from learners who achieved greater than 27% correct in the final block of trials. b: The same data in the form of a state-trace plot. Inset graph shows state-trace plot of all participants (learners and nonlearners). Error bars in both panels indicate standard errors. II = information-integration; RB = rule-based.

.039. The interactions with category structure indicate that feedback delay had a greater effect on II learning than RB learning and that this difference increased across blocks.

Second, a propos of the manner of plotting the data in Figure 3b, we conducted a formal state-trace analysis as described earlier. In this and subsequent analyses, we imposed the following order restrictions: (a) that performance does not decrease across training blocks and (b) that performance does not increase with increased feedback delay. The result was that we were able to reject the one-dimensional model ($\Delta G^2 = 6.88$, $p = .039$). Although a similar pattern of data was found when all participants were included (see inset in Figure 3b), the result was not statistically significant due to the inclusion of what amounts to error data from nonlearners ($\Delta G^2 = 3.03$, $p = .183$).

The results of Experiment 1 replicate those found by Maddox et al. (2003) and Maddox and Ing (2005). In the No Delay condition, participants were able to learn both the RB and II category structures. However, in the Delay condition, while RB learning was relatively unaffected, II learning was considerably impaired. This is both consistent with the predictions of COVIS and inconsistent with the hypothesis that only one category learning system or latent variable is involved.

The role of the mask. As discussed earlier, previous research has shown that relevant dissociations (and the dimensionality of the state-trace) can be modulated by variables that ought to be irrelevant according to COVIS (Newell et al., 2010; Nosofsky & Kruschke, 2002; Nosofsky et al., 2005; Stanton & Nosofsky, 2007). A theme that emerges from this literature is that the outcomes predicted by COVIS depend critically on levels of perceptual and criterial noise. Perceptual noise refers to uncertainty in the location of a stimulus in perceptual space. Criterial noise refers to uncertainty in the number and placement of category boundaries in this space. Both forms of noise combine to impair learning by increasing uncertainty about the location of the stimulus in relation to the regions of perceptual space that map onto the response categories.

Both perceptual and criterial noise motivated the choice of manipulations used by Stanton and Nosofsky (2007) in their investigation of feedback interference. As shown in Figure 2b, above, by increasing the number of dimensions in the RB category structure, the differential effect of feedback interference could be reduced if not eliminated. In the same study, a similar outcome was obtained by manipulating perceptual noise, this time in relation to the II category structure (Stanton & Nosofsky, 2007, Experiment 3). The moderating effect of criterial noise has also been investigated in a recent study by Ell, Ing, and Maddox (2009) that also focused on the effect of feedback delay. They compared several two-dimensional RB category structures where category membership depended on the value of one relevant dimension. The structures differed only in the number of response categories. When the number of categories increased to three, Ell et al. observed a significant effect of feedback delay on learning even though the category structure was RB. They interpreted this result within the COVIS framework by postulating an effect on working memory. As the number of boundaries increases, participants find it increasingly difficult to remember their locations (i.e., there is an increase in criterial noise), and this effect is exacerbated by delaying feedback.

On the basis of the finding by Ell et al. (2009), we hypothesized that if participants attempt to learn the II structure using multiple decision bounds, then their performance may be susceptible to the effects of feedback delay (as was found in Experiment 1). We further hypothesized that this detrimental effect could be reduced, or even eliminated, by reducing the level of perceptual noise. One way of achieving this would be to increase the discriminability of the categories in the II condition. However, we were concerned that this might change the ways in which participants may approach the task (i.e., affecting the nature of the rules they may generate and test) and that changing the conditions for one category structure and not the other would make the resulting pattern of results difficult to interpret. For these reasons, we chose to reduce perceptual noise by changing the nature of mask presented between response and feedback.

In Experiment 1, this mask was a Gabor patch randomly drawn from the range of frequency and orientation values of the stimuli themselves (see Figure 4a for an example). Previous studies have shown that memory for the properties of Gabor patches is disrupted if a similar stimulus is presented during the retention

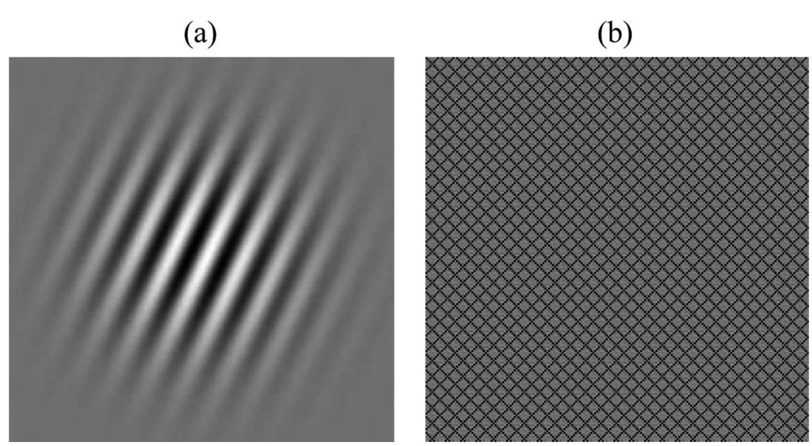

Figure 4. Example of (Panel a) a Gabor mask used in Experiments 1 and 3 and (Panel b) the cross-hatch pattern mask used in Experiments 2 and 4.

interval (Magnussen, 2000), thereby increasing the level of perceptual noise. To reduce this level, we replaced the variable Gabor mask with a fixed, cross-hatched, pattern mask (as shown in Figure 4b). If the effect of feedback delay on II learning is due to the passage of time, as proposed by COVIS, then although the overall level of learning may increase (or decrease) using this mask, II learning should still be selectively impaired. If, on the other hand, the relative impairment of II learning is due to the levels of perceptual and/or criterial noise, then a reduction in perceptual noise may be sufficient to overcome the selective effect of feedback delay.

Experiment 2

This experiment was identical to Experiment 1 except that the mask, presented between response and feedback, was replaced by a fixed cross-hatch pattern.

Method

Participants. Eighty undergraduate students from the University of New South Wales (Sydney, New South Wales, Australia) participated in return for course credit. There were initially 20 participants assigned to each experimental group, but after applying the learning criterion (>27% accuracy in the final block), two participants were excluded from the II No Delay condition and four and three participants from the II Delay and RB Delay conditions, respectively.

Stimuli, design, and procedure. All aspects of the stimuli, design, and procedure were identical to Experiment 1 with the exception that a cross-hatched pattern was used to fill the 5-s interval between making a response and receiving feedback in the Delay conditions and the 0.5-s interval between response and feedback in the No Delay conditions. The mask consisted of a set of closely spaced black diagonal lines on a gray background—matching the field on which the stimulus was presented (see Figure 4b). The same mask was used on every trial and was the same size as the Gabor mask used in Experiment 1.

Results and Discussion

The mean accuracy rates averaged across participants are presented in Figure 5. Figure 5a plots the data as a function of block (1–4), category structure (RB vs. II), and feedback delay (No Delay vs. Delay). Figure 5b plots the same data in the form of a state-trace plot with axes defined by category structure. The inset graph in Figure 5b shows the state-trace plot averaged over all participants. Both Figures 5a and 5b demonstrate an improvement in performance across blocks in all conditions, but the effect of delay on II learning is much reduced compared to that observed in Experiment 1.

A 2 (category structure) $\times$ 2 (feedback delay) $\times$ 4 (block) mixed-design ANOVA revealed a main effect of block, $F(3, 201) = 72.61$, $p < .001$, indicating learning, and a main effect of category structure, $F(1, 67) = 7.80$, $p = .007$, indicating superior RB accuracy overall, but there was no main effect of feedback delay ($F < 1$), and no interaction between feedback delay and category structure ($F < 1$) or between feedback delay, category structure, and block ($F < 1$).

The ANOVA results do not indicate any selective effect of feedback delay on II learning. This was supported by formal state-trace analysis, which failed to reject the one-dimensional model ($\Delta G^2 = 1.05$, $p = .243$). Analysis of all participants revealed a similar result ($\Delta G^2 = 0.50$, $p = .364$). Despite this, the form of the state-trace in Figure 5b appears to suggest a two-dimensional structure—the means for the Delay condition are displaced relative to the corresponding means for the No Delay condition, suggesting a small but possibly real effect. This impression, however, does not take into account the high level of correlation in performance across blocks. The apparent consistency is a necessary consequence of this correlation. The important point is

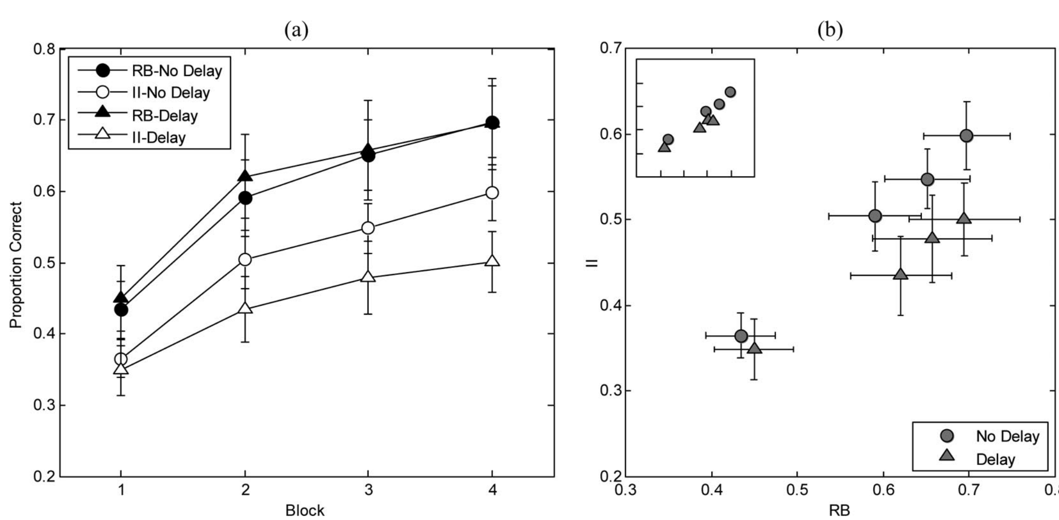

Figure 5. Results of Experiment 2. a: Proportion correct as a function of category structure, feedback delay, and block. Data are from learners who achieved greater than 27% correct in the final block of trials. b: The same data in the form of a state-trace plot. Inset graph shows state-trace plot of all participants (learners and nonlearners). Error bars in both panels indicate standard errors. II = information-integration; RB = rule-based.

that the effect, if it exists, is too small to be detected either by state-trace analysis or by ANOVA. Nevertheless, bearing in mind the admonition that the lack of a statistical effect in Experiment 2 does not mean that there is a statistically significant difference between Experiments 1 and 2 (cf. Gelman & Stern, 2006; Nieuwenhuis, Forstmann, & Wagenmakers, 2011), we directly compared the relative fits of the one-dimensional model to the data from each experiment. Essentially, we asked the question whether the ΔG^2 observed in Experiment 1, which allowed us to reject the one-dimensional model, was greater than the ΔG^2 observed in Experiment 2, which did not. To answer this question, we drew random samples from each of the empirical distributions derived from the two experiments and estimated the distribution of differences in fits. The observed difference in the fits for Experiments 1 and 2 (i.e., a difference of differences) was 5.83. The probability of observing a difference at least this large given the null hypothesis that the state-trace plots from both experiments are one-dimensional is .039. We therefore conclude that the departure from monotonicity observed in Experiment 1 is significantly greater than that observed in Experiment 2.

A major prediction of COVIS, based on fundamental neurobiological properties of the procedural system, is that learning by this system should be impaired if feedback is delayed for as little as 2.5 s (Maddox et al., 2003). Although we confirmed that learning of the II category structure was impaired by a 5-s delay (Experiment 1), we also found that this effect was largely eliminated if a variable Gabor patch mask, intervening between response and feedback, was replaced by a fixed cross-hatch mask (Experiment 2). This poses a challenge for COVIS since the nature of the mask should have no bearing on the effect of delay.

Our use of the Gabor mask followed the procedure used by Maddox and Ing (2005) and Maddox et al. (2003) although the function of the mask in those studies was not explained. One possibility is that it serves to prevent visual processing of the stimulus extending into the delay interval. If a visual representation of the stimulus were to persist during this interval, then it is possible that activation may persist in relevant cells in the tail of the caudate nucleus. If this activation were to last until the occurrence of the delayed dopamine-mediated reinforcement, some learning by the procedural system could still take place. On this view, given the selective effect of feedback delay, the Gabor mask appears to be effective. The lack of a similar effect using the cross-hatch mask might therefore be due to the ineffectiveness of this mask in curtailing perceptual processing. Studies of forms of visual persistence have identified at least three forms of visual persistence: visible persistence, informational persistence, and retinal afterimages (Coltheart, 1980). The time course of visible persistence is of the order of 100 ms (Di Lollo, Clark, & Hogben, 1988), while that of informational persistence, or iconic memory, is typically less than a second (Sperling, 1960). Consistent with this, Bennett and Cortese (1996) suggested that the storage of visual information from a Gabor patch was essentially complete within 0.5 s. Given that Maddox et al. found impaired learning of the II structure with delays of 2.5 s it is unlikely, even without a mask, that these forms of visual persistence would have been sufficient to allow the procedural system access to a stimulus representation across the 5-s delay used in the present series of experiments. This leaves the possibility that information may be extracted from a prolonged retinal afterimage. Although afterimages may last many seconds, their occurrence depends upon properties of the stimulus and the viewing conditions. Since participants viewed the relatively low-luminance stimuli in a well-lit environment, the viewing conditions were not conducive to the formation of an afterimage, and subjectively, no such images are apparent. In addition, it is well known that any form of subsequent retinal stimulation, consistent with both the Gabor and pattern masks, is sufficient to eradicate any coherent afterimage.

Given that the two kinds of mask affect visual processing in similar ways, the most likely difference between them lies in the extent to which they differentially affect memory for the categorized stimulus (Magnussen, 2000), which we have referred to as perceptual noise. If memory for the stimulus is impaired, then learning will be affected simply because participants will be unsure of the location of the stimulus in perceptual space. We return to this point in the general discussion after first examining the effect of feedback type.

Effect of feedback type. A second aim of the present study was to examine the effect of feedback type on learning RB and II category structures. Maddox et al. (2008) presented participants with the same category structures used by Maddox and Ing (2005) and varied the informativeness of feedback. In the minimal feedback condition, participants were told only that their response was either correct or incorrect. In the full feedback conditions, participants were additionally told the correct response if their response was incorrect. COVIS predicts a differential effect of feedback type on RB and II learning—in particular, full feedback should benefit RB learning compared to minimal feedback, but it should have little or no effect on II learning. As a result, this should lead to a two-dimensional state-trace.

Figure 6 shows the results found by Maddox et al. (2008) in the form of a state-trace plot.[5] The filled data points in each condition correspond to the mean levels of RB and II performance in each of the six training blocks. This reveals a complex picture. Consistent with the prediction of COVIS, full feedback improved RB learning compared to minimal feedback. However, inconsistent with the straightforward prediction of COVIS, full feedback impaired II learning compared to minimal feedback. To account for these results, Maddox et al. proposed that under full feedback, the verbal system fails to relinquish control to the procedural system during II learning despite the fact that it performs less well on this structure.

Although Maddox et al. (2008) interpreted their results in terms of the COVIS model, it is not clear that the data shown in Figure 6 describe a two-dimensional state-trace consistent with the involvement of different category learning systems. In fact, despite the appearance of a two-dimensional state-trace in Figure 6, we were unable to reject the one-dimensional model[6] ($\Delta G^2 = 18.33$, $p = .102$). This mirrors the conclusions drawn from our earlier analysis of the results found by Nosofsky et al. (2005). In both cases, a significant interaction with category structure did not imply a two-dimensional state-trace.

[5] We are grateful to Todd Maddox for providing these data.

[6] For this analysis, since feedback type is not expected to have the same effect on RB and II learning, we specified an order restriction only on blocks.

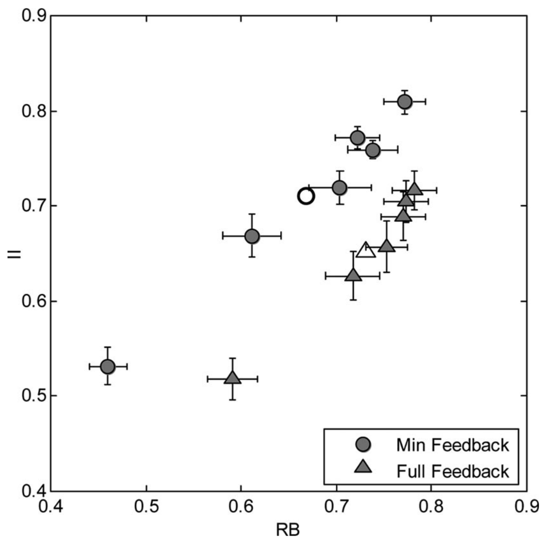

Figure 6. State-trace plot of the results found by Maddox, Love, Glass, and Filoteo (2008). Filled symbols indicate mean proportion correct for each training block. Error bars indicate standard errors. Unfilled symbols indicate mean proportion correct averaged over training blocks (the circle corresponds to the minimum [min] feedback condition; the triangle corresponds to the full feedback condition). II = information-integration; RB = rule-based.

The apparently different results found by state-trace analysis and ANOVA reflect differences in their assumptions and aims rather than any lack of statistical power in the former. ANOVA assumes a strong linear relationship between RB and II learning, while state-trace analysis allows this relationship to be nonlinear but monotonic. This difference can be illustrated in relation to the significant interaction between feedback type and category structure found by Maddox et al. (2008). This interaction is represented by the two unfilled symbols in Figure 6 showing RB and II performance averaged over all six blocks for the minimal feedback condition (the unfilled circle) and for the full feedback condition (the unfilled triangle). ANOVA detects a significant interaction if the change in RB performance between minimal and full feedback conditions differs from the change in II performance. This is equivalent to the requirement that the slope of the line connecting the two summary data points in Figure 6 differs from 1 (reflecting a change in RB performance equal to the change in II performance). Since the observed slope of this line is clearly not 1 (it is closer to -1), ANOVA detects a significant interaction. However, as pointed out by Loftus (1978), this ignores the possibility of a nonlinear relationship between RB and II performance—as appears ubiquitous in category learning data. In contrast, this assumption is fundamental to state-trace analysis, which detects a departure from monotonicity if the slope of the line connecting the two data points is negative—a more conservative and appropriate test.[7] As it happens, when this test is applied to these data, the one-dimensional model can be rejected ($\Delta G^2 = 9.70$, $p = .013$). However, given the foregoing arguments, it is also possible that averaging data from different parts of a nonlinear outcome space may create a negative relationship where none exists. For this reason, we place greater reliance on the state-trace analysis of all the data in Figure 6 rather than that averaged over all training blocks.

In summary, the results found by Maddox et al (2008) do not lead to a clear conclusion. On the one hand, the claim that different category learning systems are involved implies a two-dimensional state-trace that, despite some suggestions in the data, cannot be claimed to have been observed. On the other hand, when averaged across training blocks, full feedback was found to improve RB performance and impair II performance. To help resolve this question, we repeated Experiment 1 under full feedback conditions.

The following experiment, Experiment 3, is identical to Experiment 1 with the exception that participants were provided with full rather than minimal feedback. This experiment tested two implications of COVIS as interpreted by Maddox et al. (2008). First, to the extent that the procedural system is involved in II learning under full feedback, it should be differentially affected by feedback delay, and we should observe a two-dimensional state-trace similar to that found in Experiment 1. However, if full feedback leads to the greater involvement of the verbal system in II learning, as proposed by Maddox et al., the two-dimensional structure may be reduced.[8] In the limit, if the verbal system completely governs II learning, then only one system (or latent variable) is in play, and the resulting state-trace will be one-dimensional. Second, consistent with the results found by Maddox et al., comparison of the No Delay conditions of Experiments 1 and 3 should replicate the effects of feedback type on RB and II learning, resulting in a two-dimensional state-trace similar to that shown in Figure 6.

Experiment 3

Experiment 3 was identical to Experiment 1 except that participants were provided with full rather than minimal feedback on each learning trial. The same variable Gabor mask was used.

Method

Participants. One hundred and fifty-nine undergraduate students from the University of Adelaide participated in return for course credit or a payment of AUD12. Each participant completed one experimental condition, with 39 participants assigned to the RB No Delay condition, 35 assigned to the II No Delay condition, and 39 and 46 participants assigned to the RB Delay and II Delay conditions, respectively. Applying the learning criterion (>27%

[7] In our view, rejecting the hypothesis of a single category learning system based on a failure of the linear model assumed by ANOVA is analogous to rejecting the hypothesis that the earth goes around the sun by finding that the trajectory is not a perfect circle. In both cases, there are theoretically viable alternatives consistent with the data.

[8] This means that the state-trace, while still two-dimensional, will appear closer to being one-dimensional. The configuration of a pair of scissors provides a concrete analogy to this idea. Points on the blades of the scissors correspond to differences on the two experimental variables. When the scissors are maximally open and the blades are at right angles, intrinsic two-dimensionality is apparent. As the blades close, while technically remaining two-dimensional, the pattern is less obvious. In the limit, the blades finally close, and the structure becomes one-dimensional.

accuracy in the final block) excluded two participants from the RB No Delay condition and six and three participants from the II Delay and RB Delay conditions, respectively.

Stimuli, design, and procedure. All aspects of the stimuli, design, and procedure were identical to Experiment 1 with the exception that the nature of the corrective feedback was changed. After a correct response, the following feedback appeared on the computer screen: "Correct, that was Category *x*," where *x* was the corresponding key number (1–4) of the correct category. Similarly, after an incorrect response the following feedback appeared: "No, that was Category *x*."

Results and Discussion

The mean accuracy rates averaged across participants who met the learning criterion are shown in Figure 7. Figure 7a plots the data as a function of block (1–4), category structure (RB vs. II), and feedback delay (No Delay vs. Delay). Figure 7b plots the same data in the form of a state-trace plot with axes defined by category structure. The inset in this figure shows the state-trace plot for all participants. These data show that there is an improvement across blocks in all conditions but that the effect of delaying feedback is still mixed. Although it is generally deleterious to both RB and II learning, the effect of delay affects RB learning to a greater extent early in training but affects II learning to a greater extent later in training.

A 2 (category structure) × 2 (feedback delay) × 4 (block) mixed-design ANOVA revealed a main effect of block, $F(3, 432) = 130.09$, $p < .001$, indicating learning; a main effect of category structure, $F(1, 144) = 22.38$, $p < .001$, indicating superior accuracy overall for the RB structure; and a main effect of feedback delay, $F(1, 144) = 6.51$, $p = .012$, indicating superior accuracy in the No Delay conditions. Although the Category Structure × Feedback Delay interaction was not significant, the three-way interaction between category structure, feedback delay, and block was, $F(3, 432) = 4.24$, $p = .006$.

Dimensionality of the feedback delay state-trace. COVIS predicts that the state-trace shown in Figure 7b should be two-dimensional although, because of the potentially greater involvement of the verbal system in II learning, this may be less apparent than that found in Experiment 1. While the overall shape of the state-trace is consistent with this, formal analysis revealed that although the result approached significance, the one-dimensional model could not be rejected ($\Delta G^2 = 4.67$, $p = .075$). Analysis of all participants revealed a similar result ($\Delta G^2 = 2.15$, $p = .202$). However, direct comparison of the relative fits of the one-dimensional model between Experiments 1 and 3 failed to reveal a significant difference ($\Delta G^2 = 2.21$, $p = .142$). This result is best considered indeterminate—we cannot reject the one-dimensionality of the feedback delay state-trace, but neither can we reject the hypothesis that the fit of the one-dimensional model is significantly worse than that found for Experiment 1.

Dimensionality of the feedback type state-trace. Figure 8 shows mean accuracy rates averaged across participants who met the learning criterion for the No Delay conditions of Experiments 1 and 3. Figure 8a plots the data as a function of block (1–4), category structure (RB vs. II), and feedback type (minimal vs. full). Figure 8b shows the same data in the form of a state-trace plot with axes defined by category structure. The inset in this figure shows the state-trace plot for all participants. We conducted a formal state-trace analysis of the data shown in Figure 8b and were unable to reject the one-dimensional model ($\Delta G^2 = 3.97$, $p = .199$). Analysis of all participants revealed a similar result ($\Delta G^2 = 0.37$, $p = .587$).

Effect of feedback type on RB and II learning. The data from the No Delay conditions of Experiments 1 and 3 were submitted to a 2 (category structure) × 2 (feedback type) × 4 (block) mixed-design ANOVA. This revealed a main effect of block, $F(3, 378) = 149.18$, $p < .001$, indicating learning; a significant effect of category structure, $F(1, 126) = 4.46$, $p = .037$, indicating higher performance for RB than II structures; and an

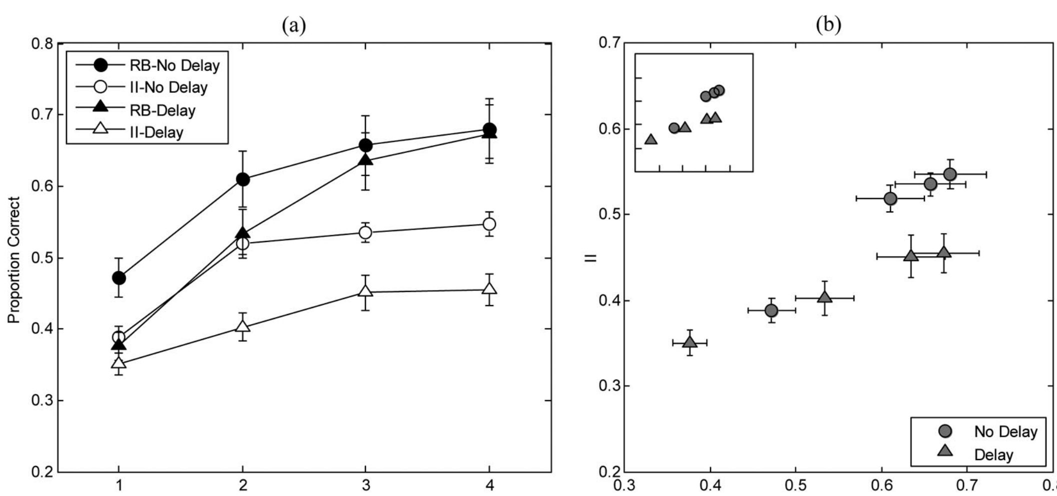

Figure 7. Results of Experiment 3. a: Mean proportion correct as a function of category structure, feedback delay, and block. Data are from learners who achieved greater than 27% correct in the final block of trials. b: The same data in the form of a state-trace plot. Inset graph shows state-trace plot of all participants (learners and nonlearners). Error bars in both panels indicate standard errors. II = information-integration; RB = rule-based.

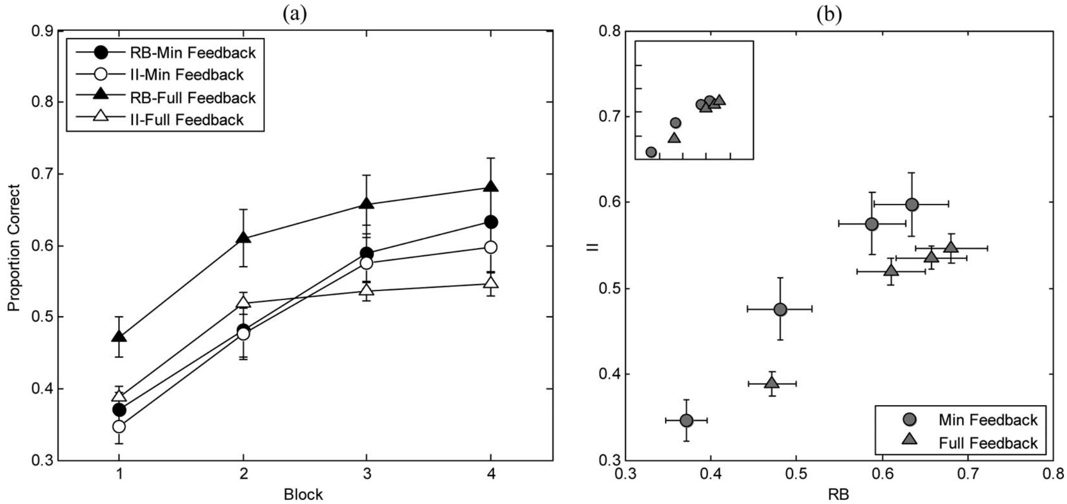

Figure 8. The effect of feedback type in the No Delay conditions of Experiments 1 and 3. a: Mean proportion correct as a function of category structure, feedback type, and block. Data are from learners who achieved greater than 27% correct in the final block of trials. Note the change of scale on the *y*-axis compared to previous graphs. b: The same data in the form of a state-trace plot. Inset graph shows state-trace plot of all participants (learners and nonlearners). Error bars in both panels indicate standard errors. II = information-integration; RB = rule-based.

interaction between block and feedback type, $F(3, 378) = 7.10$, $p < .001$, indicating that the detrimental effect of minimal feedback reduced as training progressed. In contrast to the results found by Maddox et al. (2008), the interaction between feedback type and category structure was not significant, $F(1, 126) = 2.12$, $p = .148$. Nevertheless, the results do suggest that whereas full feedback improves RB learning, it tends to impair II learning, at least during the later training blocks. Separate analysis of RB learning revealed a significant effect of block, $F(3, 201) = 70.98$, $p < .001$, and a significant improvement in overall performance under full feedback (mean difference = 0.09), $t(67) = 1.73$, one-tailed $p = .044$. Analysis of II learning also revealed a significant effect of block, $F(3, 201) = 89.47$, $p < .001$, and although there was no main effect of feedback type ($F < 1$), the Block $\times$ Feedback Type interaction was significant, $F(3, 177) = 6.83$, $p < .001$, consistent with improved performance with full feedback in the first two blocks coupled with impaired performance in the last two blocks of trials.

In summary, the results of Experiments 1 and 3 offer mixed support for the variant of the COVIS model proposed by Maddox et al. (2008). This model predicts that the state-trace for feedback delay should be two-dimensional under full feedback, although the ability to detect this should be reduced. The observed pattern of results is consistent with this prediction, although we were unable to formally reject the one-dimensional model. The model also predicts that the state-trace under feedback type should also be two-dimensional. However, in this case, we were unable to formally reject the one-dimensional model. Finally, the model predicts that RB learning should be enhanced and II learning impaired under full feedback. Consistent with this, we found a small positive effect of full feedback on RB learning, and although there was no overall effect on II learning, full feedback improved II learning early in training but impaired it later in training.

Although the results of Experiment 3 provide some support for the COVIS model, the results of Experiment 2 (using the pattern mask) suggest that much of the observed pattern of data may depend on the type of mask used. For this reason, we repeated Experiment 2 under full feedback conditions. Since, according to COVIS, the type of mask is irrelevant to the effects of feedback delay and, presumably, the effects of feedback type, the outcome should be the same as found for Experiment 3. Alternatively, if the results of Experiment 2 are replicated, there should be no differential effects of feedback delay and feedback type on RB and II learning. In other words, both kinds of state-trace should be one-dimensional.

Experiment 4

This experiment was identical to Experiment 2 except that the pattern mask shown in Figure 4 was used.

Method

Eighty undergraduate students from the University of New South Wales participated in return for course credit. There were initially 20 participants assigned to each experimental group, but after applying the learning criterion (>27% accuracy in the final block), two participants were excluded from each of the II conditions, and one participant was excluded from each of the RB conditions.

All aspects of the stimuli, design, and procedure were identical to Experiment 3 (i.e., full corrective feedback was given) with the exception that the fixed pattern mask was used instead of the variable Gabor mask.

Results and Discussion

The mean accuracy rates averaged across participants are presented in Figure 9. Figure 9a plots the data as a function of block (1–4), category structure (RB vs. II), and feedback delay (No Delay vs. Delay). Figure 9b shows the same data in the form of a state-trace plot with axes defined by category structure. The inset graph shows the state-trace plot averaged over all participants. The data show that there is an improvement across blocks in all conditions but that the effect of delay on II learning is much reduced.

A 2 (category structure) × 2 (feedback delay) × 4 (block) mixed-design ANOVA revealed a main effect of block, $F(3, 210) = 77.79$, $p < .001$, indicating learning; a main effect of category structure, $F(1, 70) = 12.35$, $p < .001$, indicating superior RB accuracy overall; but no main effect of feedback delay ($F < 1$) and no interaction between feedback delay and category structure ($F < 1$) or between feedback delay, category structure, and block ($F = 1.14$). There was a significant Block × Category Structure interaction, $F(3, 210) = 3.63$, $p = .014$.

Dimensionality of the feedback delay state-trace. Figure 9b shows little evidence of two-dimensionality. Formal state-trace analysis revealed that the fit of the one-dimensional model was identical to the fit of the order-restricted two-dimensional model ($\Delta G^2 = 0$) and thus could not be rejected ($p = 1$). Analysis of all participants revealed a similar result ($\Delta G^2 = 0.31$, $p = .345$). The apparent violations of monotonicity in Figure 9b do not signify bidimensionality. Instead, they correspond to violations of the order restrictions of both the one-dimensional and two-dimensional models due to feedback delay increasing rather than decreasing RB performance in Blocks 2 and 3 (see Figure 9a).

Dimensionality of the feedback type state-trace. Figure 10 shows mean accuracy rates averaged across participants who met the learning criterion for the No Delay conditions of Experiments 2 and 4. Figure 10a plots the data as a function of block (1–4), category structure (RB vs. II), and feedback type (minimal vs. full). Figure 10b shows the same data in the form of a state-trace plot with axes defined by category structure. The inset in this figure shows the state-trace plot for all participants. There is little or no indication of a two-dimensional state-trace ($\Delta G^2 = 0.01$, $p = 0.874$). Analysis of all participants revealed a similar result ($\Delta G^2 = 0.16$, $p = .744$).

Effect of feedback type on RB and II learning. The data from the No Delay conditions of Experiments 2 and 4 were submitted to a 2 (category structure) × 2 (feedback type) × 4 (block) mixed-design ANOVA. This revealed a main effect of block, $F(3, 213) = 89.72$, $p < .001$, indicating learning, and an effect of category structure, $F(1, 71) = 6.78$, $p = .011$, indicating greater accuracy for RB structures. No other effect was significant. Separate analysis of RB and II learning revealed no effect of feedback type or of a Block × Feedback Type interaction (all Fs < 1).

The results of Experiment 4 replicate those of Experiment 2 and generalize them to full feedback conditions. When the variable Gabor mask is replaced by an invariant pattern mask, there is no evidence of a differential effect of either feedback delay or feedback type on RB and II learning.

General Discussion

We conducted four experiments that examined the effects of feedback delay using different types of feedback and masks. The results of these experiments are clear. When a variable Gabor patch mask was used, there was a selective effect of feedback delay on II learning that was greatest under minimal feedback and reduced, but still detectable, under full feedback. These effects were predicted by and are thus consistent with COVIS. According to this model, any delay of more than a few seconds between response and feedback should impair the procedural system that usually determines II learning, and based on the argument pro-

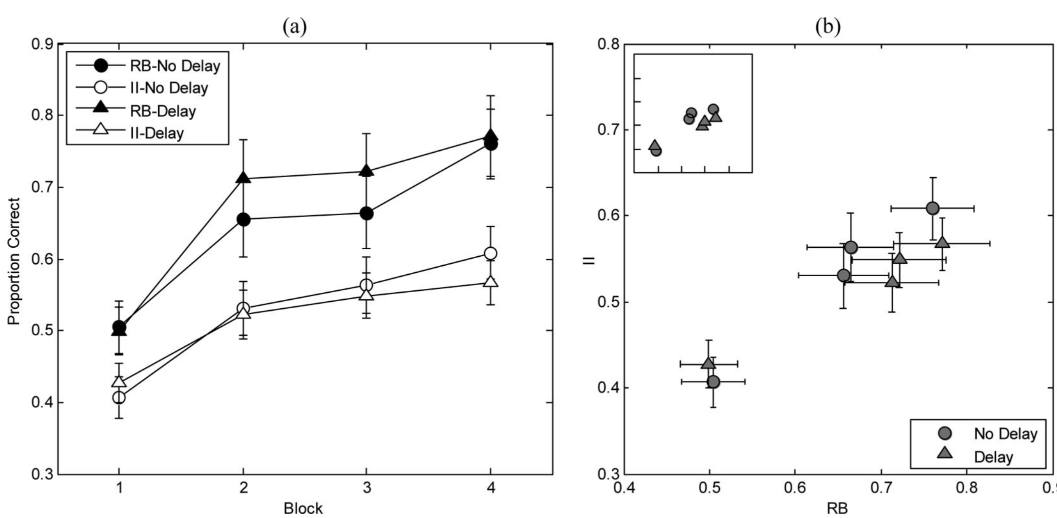

Figure 9. Results of Experiment 4. a: Mean proportion correct as a function of category structure, feedback delay, and block. Data are from learners who achieved greater than 27% correct in the final block of trials. Note the change of scale on the *x*-axis compared to previous graphs. b: The same data in the form of a state-trace plot. Inset graph shows state-trace plot of all participants (learners and nonlearners). Error bars in both panels indicate standard errors. II = information-integration; RB = rule-based.

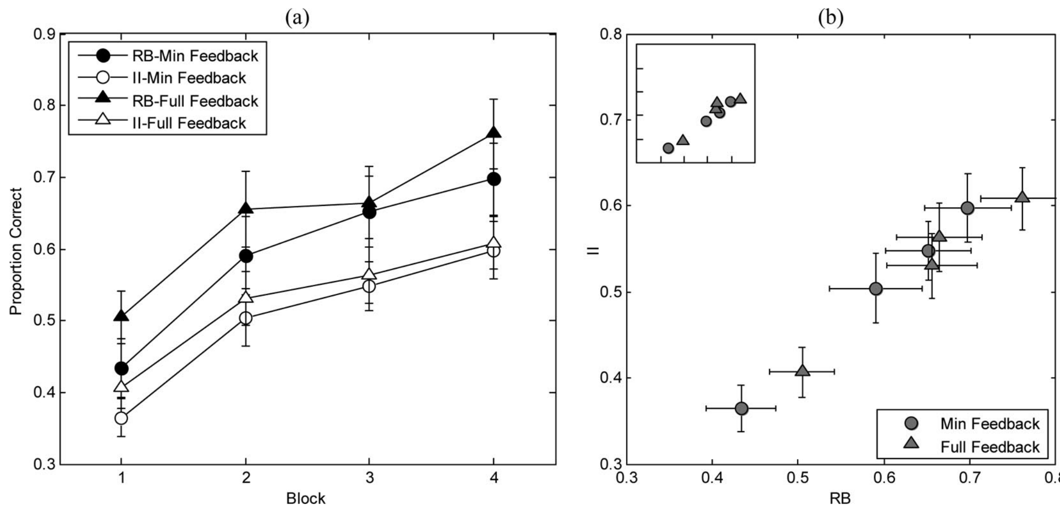

Figure 10. The effect of feedback type in the No Delay conditions of Experiments 2 and 4. a: Mean proportion correct as a function of category structure, feedback type, and block. Data are from learners who achieved greater than 27% correct in the final block of trials. b: The same data in the form of a state-trace plot. Inset graph shows state-trace plot of all participants (learners and nonlearners). Error bars in both panels indicate standard errors. II = information-integration; RB = rule-based.

posed by Maddox et al. (2008), this effect would be stronger under minimal feedback than under full feedback. Both these predictions were confirmed.

State-trace analysis also revealed that it was possible to reject the hypothesis that RB learning and II learning in Experiment 1 could be attributed to a single latent variable. This is also consistent with COVIS, which predicts differential involvement of the verbal and procedural learning systems as a function of feedback delay. Although a similar pattern of results was also found in Experiment 3, it was not possible to reject the one-dimensional model. Nevertheless, this is also consistent with the prediction from COVIS that the effects of the verbal and procedural systems would be less differentiated under these conditions.

Despite the success of COVIS in predicting the results of Experiments 1 and 3, it was unable to account for the results of Experiments 2 and 4. When the delay interval was filled by a fixed pattern mask, there was no effect of feedback delay on either RB or II learning. Since COVIS attributes the effect of delay to the passage of time and the consequent disruption of neurobiological processes in the tail of the caudate nucleus, it cannot readily account for the elimination of the selective effect on II learning under conditions that ought to be irrelevant. Within the COVIS framework, these results can only be interpreted as reflecting the singular involvement of the verbal system in both RB and II learning, with the procedural system being somehow rendered inoperative. There are three problems with this explanation. First, it is obviously ad hoc. Second, there is no mechanism within COVIS that would lead to control by the verbal system under these circumstances. Third, it supposes that the verbal system can learn the II category structure to a high level of accuracy. If this were the case, it is unclear why the verbal system was unable to intervene to improve II learning when this was impaired in Experiments 1 and 3 using the Gabor patch mask.

One caveat to these conclusions arises from the fact that Experiments 1 and 3 tested more participants than Experiments 2 and 4 and so had greater power to detect a two-dimensional state-trace if one was present. It is therefore possible that the failure to detect a two-dimensional state-trace in Experiments 2 and 4 could be attributed to this lack of power. To examine this question, we increased the total power to detect an effect by combining Experiments 1 and 3 (total n = 246) and Experiments 2 and 4 (total n = 145). Consistent with previous analyses, the one-dimensional state-trace model could be rejected for Experiments 1 and 3 ($\Delta G^2 = 11.67$, $p = .015$), but not for Experiments 2 and 4 ($\Delta G^2 = 0.34$, $p = .358$). We then drew 10,000 random samples (without replacement) of a subset of 145 participants from the combined Experiments 1 and 3 and calculated the corresponding ΔG^2 values. If the difference between the Gabor and pattern masks is attributable to differences in statistical power, the observed ΔG^2 found for Experiments 2 and 4 ought to be typical of that observed for a random sample of the same size drawn from Experiments 1 and 3. The mean ΔG^2 of the 10,000 samples was 6.72. The observed ΔG^2 of 11.67 (from all participants in Experiments 1 and 3) was greater than 88% of these samples and is to be expected given the larger n. Importantly, the observed ΔG^2 of 0.34 (from Experiments 2 and 4) was greater than only 1.2% of the samples and therefore hardly typical. We therefore conclude that it is unlikely that the size of the effect observed in Experiments 2 and 4 is the same as that observed in Experiments 1 and 3 but masked by a lower n.

Rather than being attributable to the mere passage of time, we have suggested that the differential effect of feedback delay is due to added perceptual noise that makes it difficult to locate the categorized stimulus in perceptual space. Since presentation of a Gabor patch during a retention interval interferes with memory for the attributes of a similar, previously presented, Gabor patch, it is likely that feedback processing would be disrupted in this way (Magnussen, Greenlee, Asplund, & Dyrnes, 1991). For any categorization strategy, uncertainty concerning the identity of the stimulus should impair learning because it would make it difficult to accurately update knowledge of the relationship between the stimulus and the relevant category structures. When a less similar and fixed pattern mask is used, memory for the attributes of the stimulus is likely to be less impaired, and learning should therefore be relatively unaffected. In fact, the present results suggest that under these conditions, learning is not affected at all.

While the proposed effect of perceptual noise on feedback processing accounts for the difference between the two types of mask, it does not directly account for the selective effect of the Gabor patch mask on II learning (as observed in Experiments 1 and 3). Similarly, while COVIS accounts for this selective effect, it does not directly account for the moderating effect of mask type. The challenge for category learning models, including COVIS, is to account for both effects simultaneously. We discuss this challenge in a later section.

A second aim of the present study was to examine the differential effect of feedback type (minimal vs. full) on RB and II learning. In their study, Maddox et al. (2008) found that full feedback facilitated RB learning but impaired II learning. They accounted for this outcome by proposing, within the COVIS model, that full feedback encouraged the verbal system to maintain control of the categorization response in the II condition to a greater degree under full feedback than under minimal feedback. Comparison of the No Delay conditions of Experiments 1 and 3, which used the variable Gabor patch mask and were most directly comparable to the stimuli and procedure used by Maddox et al., also revealed a differential effect of feedback type on RB and II learning. Although the relevant interactions between feedback type and category structure were not significant using ANOVA and the one-dimensional state-trace could not be rejected, when RB and II learning were analyzed separately, we found a small effect of feedback type on RB learning. Consistent with Maddox et al., this revealed improved RB learning under full feedback. We also found an interaction between feedback type and block for II learning, which, broadly consistent with the results found by Maddox et al., showed improved learning for the first two blocks of trials and impaired learning in the last two blocks of trials.

However, in contrast to the pattern of results found in Experiments 1 and 3, we found no evidence of a differential effect of feedback type in Experiments 2 and 4. Comparison of the No Delay conditions using the fixed pattern mask (although only present for 0.5 s) revealed no effect of feedback type on either RB or II learning. Once again, the predictions of COVIS were restricted to the use of a variable, and confusable, mask.

The results of Experiments 1–4 may be summarized as follows. When a variable Gabor patch mask was used to fill the interval between response and feedback, consistent with the various predictions derived from COVIS, we found evidence of a differential effect of feedback delay and, to a lesser extent, feedback type on RB and II learning. When a fixed pattern mask was used to fill the interval between response and feedback, contrary to the predictions by COVIS, we found no evidence of a differential effect of either feedback delay or feedback type on RB and II learning. We have hypothesized that the moderating effect of mask type may be understood in terms of the effect of perceptual noise on category learning. The implications of these results for models of categorization are discussed next.

State-Trace Analysis and Models of Category Learning

We used state-trace analysis to determine a lower limit to the number of latent variables required to account for the effects of feedback delay and feedback type on RB and II learning. Depending on the substantive theory used to account for our results, the latent variables revealed by state-trace analysis may be interpreted in several different ways. In the case of formal models of categorization, such as ALCOVE (Kruschke, 1992), GCM (Nosofsky, 1986), SUSTAIN (Love, Medin, & Gureckis, 2004), or even COVIS (Ashby et al., 2011), the latent variables may be interpreted as corresponding to different parameters of the models. However, some models demarcate sets of parameters that reflect, according to the model, functionally distinct components. For example, ATRIUM (Erickson & Kruschke, 1998) implements two different learning strategies—one based on rules, the other on exemplars—which may be described as corresponding either to different *modules* (Kruschke, 2011) or to a *mixture of experts* (Kalish, Lewandowsky, & Kruschke, 2004). In such instances, depending upon the results of formal modeling, a two-dimensional state-trace may indicate the differential involvement of parameters reflecting these structures rather than parameters common to both. Other models propose an even sharper distinction between component structures. COVIS for example refers to its functionally distinct components as *systems,* which are explicitly linked to anatomically distinct systems in the brain (Ashby & Ell, 2001; Ashby & Spiering, 2004). In this case, the latent variables identified by state-trace analysis may be interpreted as corresponding to parameters or sets of parameters characterizing these different systems.

The terminology used by models to describe their functional components, whether as parameters or modules or systems, is not directly relevant to state-trace analysis. A two-dimensional state-trace no more implies the existence of separate systems than it implies the existence of separate parameters or modules or any other componential term. On the other hand, models such as COVIS or ATRIUM that commit to a functional distinction between major constituent elements differentially involved in learning different kinds of category structure predict that it should be possible to differentially affect the functions of these elements consistent with a two-dimensional state-trace. This is a natural interpretation of the results of Experiment 1. On this view, RB learning depends more strongly on an RB strategy, while II learning may depend more strongly on an exemplar-based (or non-RB) strategy. A two-dimensional state-trace results because feedback delay has a greater effect on the exemplar-based strategy than on the RB strategy, at least given high levels of perceptual noise.

The present results pose a challenge for all current models of categorization. They reveal a complex interaction between feedback delay, feedback type, mask type, and category structure. COVIS fails to predict this interaction because the mechanism it proposes for some of these effects cannot account for the moderating effect of mask type. According to COVIS, any delay between response and feedback will impair the procedural system and, hence, II learning. While a reduction in perceptual noise may improve performance for both category structures, it is difficult to see how this could overcome the selective deficit posited by COVIS. On the other hand, it is also likely that other models will encounter difficulties in explaining other features of the data. The results of Experiment 1, in particular, reveal a qualitative difference between the RB and II category structures. While we have suggested that any reasonably complex model can account for such differences through ad hoc changes in parameters, this is unlikely to generate a satisfactory theoretical account. Models that treat RB and II structures in essentially the same way are likely to have difficulties accounting for these results in a natural and compelling way. For example, we have suggested that the effect of feedback delay using the variable Gabor patch mask may be attributable to increased perceptual noise. This can be incorporated into a formal model of category learning by adding different degrees of random error to the representation of the stimulus prior to feedback-driven learning. SUSTAIN is an example of a model in which the present RB and II category structures are learned in essentially the same way through the generation of clusters that link similar stimuli to responses (Love et al., 2004). While SUSTAIN does possess a mechanism to allocate attention to a single dimension when this is sufficient to learn an RB category structure, this cannot be used to differentiate RB and II learning in the present series of experiments since the RB structure is defined in terms of the conjunction of values on two dimensions (so attention must be paid to both). To confirm this intuition, we investigated the capacity of SUSTAIN to simulate the results of Experiment 1, adding perceptual noise corresponding to the effect of the variable Gabor patch mask. We found that perceptual noise had equivalent effects on both RB and II learning and led to a one-dimensional state-trace. The same was also true when we examined ALCOVE's predictions in a similar way. The challenge of these data for such models is thus comparable to that posed by other examples of the coexistence of RB and exemplar-based learning (Erickson & Kruschke, 1998, 2002).

Evidence for Multiple Category Learning Systems

The present set of results is consistent with recent evidence that questions whether human category learning is best characterized in terms of distinct, neurobiological systems. As outlined earlier and discussed in detail by Newell et al. (2011), much of the evidence based on dissociations that has appeared to support the multiple learning systems view is challenged when viewed through the lens of the more stringent state-trace analysis. Several dissociations that have been thought to be central to the argument for multiple systems do not survive this reanalysis. One such dissociation concerns the selective effect of response reassignment on II categorization (Ashby et al., 2003; Maddox, Glass, O'Brien, Filoteo, & Ashby, 2010; Nosofsky et al., 2005), which, as Figure 1 shows, does not meet state-trace criteria for more than one latent variable. Instead, these data suggest that a common learning mechanism drives both RB and II learning but at different rates at the time of transfer. Consistent with this, when these rates are equated, either by making the RB structure more complex (Nosofsky et al., 2005, Experiment 3) or by allowing participants extended practice with the task (Hélie, Waldschmidt, & Ashby, 2010), the level of impairment following response reassignment is found to be the same for both structures.

The reassessment of dissociations also applies to evidence that has been used to support the neurobiological aspects of COVIS. According to this model, the verbal system depends upon a network of brain regions that includes the head of the caudate nucleus, while the procedural system depends upon neural structures in the tail of the caudate nucleus. Evidence supporting this view is derived from studies of patients with Parkinson's disease (PD) and Huntington's disease (HD). PD patients often have damage to the head of the caudate nucleus and so should be primarily impaired on RB learning. HD patients have more extensive damage to both the head and tail of the caudate nucleus and so should be impaired on both RB and II learning. Thus, while both groups should show impaired learning of RB structures (which they do), PD patients should be relatively unimpaired learning at least some II structures. Consistent with this, Filoteo, Maddox, Salmon, and Song (2005) found that PD patients were impaired on the categorization of complex (nonlinear) II category structures while they were relatively unimpaired on the categorization of less complex (linear) II structures. In contrast, Filoteo, Maddox, and Davis (2001) found that HD patients were impaired on both linear and nonlinear II tasks. This dissociation between the complexity of the II category structure and level of impairment to the caudate nucleus (as indexed by patient group) has been used to support the neurobiological distinctions made by COVIS (Ashby & Ennis, 2006). However, when these same data are presented in the form of a state-trace plot, it is apparent that the levels of performance of the patient groups and their respective normal controls fall on a single monotonically increasing curve (Newell et al., 2011, Figure 7). As with the response-switching paradigm, the reported dissociations are a consequence of the shape of this curve (relatively flat in some sections, less so in others) and provide no evidence either for the existence of multiple learning systems or for the differential involvement of the caudate nucleus in category learning.

Conclusion

A major strength of COVIS is that it attempts to integrate biological and psychological data. In our view, the relationship between these different sets of data is likely to be complex and to evolve over time. The patterns of behavioral data that we have presented and briefly reviewed pose a challenge for COVIS, at least in its current form. At the same time, they also pose a challenge for other models of categorization. Few of these models are in a form that allows researchers to derive specific predictions concerning the effects of different variables on learning different kinds of category structures. While we have suggested how such models may accommodate the interacting effects of feedback delay and mask type, a formal model has not yet been developed. We have argued that behavioral dissociations, in the form of state-trace plots, do not selectively support a multiple systems view but are instead consistent with a variety of architectures. In our view, this insight provides a means of integrating computational, behavioral, and neurophysiological models that, while differing in their terminology, may not in fact differ greatly in their underlying functionality. The challenge for all models is to account for patterns of observed human performance rather than debating the relative merits of terms such as systems, modules, and parameters. The development of a comprehensive model of perceptual categorization that accounts for the empirical phenomena revealed by this and other studies remains a challenge to the field.

References

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. doi:10.1037/0033-295X.105.3.442

Ashby, F. G., & Ell, S. W. (2001). The neurobiology of human category learning. *Trends in Cognitive Sciences, 5,* 204–210. doi:10.1016/S1364-6613(00)01624-7

Ashby, F. G., Ell, S. W., & Waldron, E. M. (2003). Procedural learning in perceptual categorization. *Memory & Cognition, 31,* 1114–1125. doi:10.3758/BF03196132

Ashby, F. G., & Ennis, J. M. (2006). The role of the basal ganglia in category learning. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 46, pp. 1–36). San Diego, CA: Elsevier.

Ashby, F. G., & O'Brien, J. B. (2005). Category learning and multiple memory systems. *Trends in Cognitive Sciences, 9,* 83–89. doi:10.1016/j.tics.2004.12.003

Ashby, F. G., Paul, E. J., & Maddox, W. T. (2011). COVIS. In E. M. Pothos & A. J. Wills (Eds.), *Formal approaches in categorization* (pp. 65–87). New York, NY: Cambridge University Press.

Ashby, F. G., & Spiering, B. J. (2004). The neurobiology of category learning. *Behavioral and Cognitive Neuroscience Reviews, 3,* 101–113. doi:10.1177/1534582304270782

Bamber, D. (1979). State-trace analysis: A method of testing simple theories of causation. *Journal of Mathematical Psychology, 19,* 137–181. doi:10.1016/0022-2496(79)90016-6

Bennett, P. J., & Cortese, F. (1996). Masking of spatial frequency in visual memory depends on distal, not retinal, frequency. *Vision Research, 36,* 233–238. doi:10.1016/0042-6989(95)00085-E

Brainard, D. H. (1997). Psychophysics software for use with MATLAB. *Spatial Vision, 10,* 433–436. doi:10.1163/156856897X00357

Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics, 27,* 183–228. doi:10.3758/BF03204258

Di Lollo, V., Clark, C. D., & Hogben, J. H. (1988). Separating visible persistence from retinal afterimages *Attention, Perception, & Psychophysics, 44,* 363–368. doi:10.3758/BF03210418

Dunn, J. C., & James, R. N. (2003). Signed difference analysis: Theory and application. *Journal of Mathematical Psychology, 47,* 389–416. doi:10.1016/S0022-2496(03)00049-X

Dunn, J. C., & Kirsner, K. (1988). Discovering functionally independent mental processes: The principle of reversed association. *Psychological Review, 95,* 91–101. doi:10.1037/0033-295X.95.1.91

Ell, S. W., Ing, A. D., & Maddox, W. T. (2009). Criterial noise effects on rule-based category learning: The impact of delayed feedback. *Attention, Perception, & Psychophysics, 71,* 1263–1275.

Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General, 127,* 107–140. doi:10.1037/0096-3445.127.2.107

Erickson, M. A., & Kruschke, J. K. (2002). Rule-based extrapolation in perceptual categorization. *Psychonomic Bulletin & Review, 9,* 160–168. doi:10.3758/BF03196273

Filoteo, J. V., Maddox, W. T., & Davis, J. D. (2001). A possible role of the striatum in linear and nonlinear categorization rule learning: Evidence from patients with Huntington's disease. *Behavioral Neuroscience, 115,* 786–798. doi:10.1037/0735-7044.115.4.786

Filoteo, J. V., Maddox, W. T., Salmon, D. P., & Song, D. D. (2005). Information-integration category learning in patients with striatal dysfunction. *Neuropsychology, 19,* 212–222. doi:10.1037/0894-4105.19.2.212

Gelman, A., & Stern, H. (2006). The difference between "significant" and "not significant" is not itself statistically significant. *American Statistician, 60,* 328–331. doi:10.1198/000313006X152649

Hélie, S., Waldschmidt, J., & Ashby, F. (2010). Automaticity in rule-based and information-integration categorization. *Attention, Perception, & Psychophysics, 72,* 1013–1031.

Kalish, M. L., Lewandowsky, S., & Kruschke, J. K. (2004). Population of linear experts: Knowledge partitioning and function learning. *Psychological Review, 111,* 1072–1099. doi:10.1037/0033-295X.111.4.1072

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review, 99,* 22–44. doi:10.1037/0033-295X.99.1.22

Kruschke, J. K. (2011). Models of attentional learning. In E. M. Pothos & A. J. Wills (Eds.), *Formal models in categorization* (pp. 120–152). Cambridge, England: Cambridge University Press.

Loftus, G. R. (1978). On interpretation of interactions. *Memory & Cognition, 6,* 312–319. doi:10.3758/BF03197461

Loftus, G. R., Oberg, M. A., & Dillon, A. M. (2004). Linear theory, dimensional theory, and the face-inversion effect. *Psychological Review, 111,* 835–865. doi:10.1037/0033-295X.111.4.835

Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A network model of category learning. *Psychological Review, 111,* 309–332. doi:10.1037/0033-295X.111.2.309

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. doi:10.1037/0278-7393.29.4.650

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. doi:10.3758/BF03195849

Maddox, W. T., Glass, B., O'Brien, J., Filoteo, J., & Ashby, F. (2010). Category label and response location shifts in category learning. *Psychological Research, 74,* 219–236. doi:10.1007/s00426-009-0245-z

Maddox, W. T., & Ing, A. D. (2005). Delayed feedback disrupts the procedural-learning system but not the hypothesis-testing system in perceptual category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 100–107. doi:10.1037/0278-7393.31.1.100

Maddox, W. T., Love, B. C., Glass, B. D., & Filoteo, J. V. (2008). When more is less: Feedback effects in perceptual category learning. *Cognition, 108,* 578–589. doi:10.1016/j.cognition.2008.03.010

Magnussen, S. (2000). Low-level memory processes in vision. *Trends in Neurosciences, 23,* 247–251. doi:10.1016/S0166-2236(00)01569-1

Magnussen, S., Greenlee, M. W., Asplund, R., & Dyrnes, S. (1991). Stimulus-specific mechanisms of visual short-term memory. *Vision Research, 31,* 1213–1219. doi:10.1016/0042-6989(91)90046-8

Minda, J. P., & Smith, J. D. (2002). Prototypes in category learning: The effects of category size, category structure, and stimulus complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 275–292. doi:10.1037/0278-7393.28.2.275

Minda, J. P., & Smith, J. D. (2011). Prototype models of categorization: Basic formulation, predictions, and limitations. In E. M. Pothos & A. J. Wills (Eds.), *Formal approaches in categorization* (pp. 40–64). Cambridge, England: Cambridge University Press.

Newell, B. R., & Dunn, J. C. (2008). Dimensions in data: Testing psychological models using state-trace analysis. *Trends in Cognitive Sciences, 12,* 285–290. doi:10.1016/j.tics.2008.04.009

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? In B. H. Ross (Ed.), *Psychology of learning and motivation* (Vol. 54, pp. 167–215). San Diego, CA: Academic Press.

Nieuwenhuis, S., Forstmann, B. U., & Wagenmakers, E. J. (2011). Erroneous analyses of interactions in neuroscience: A problem of significance. *Nature Neuroscience, 14,* 1105–1107. doi:10.1038/nn.2886

Nomura, E. M., Maddox, W. T., Filoteo, J. V., Ing, A. D., Gitelman, D. R., Parrish, T. B., . . . Reber, P. J. (2007). Neural correlates of rule-based and information-integration visual category learning. *Cerebral Cortex, 17,* 37–43. doi:10.1093/cercor/bhj122

Nomura, E. M., & Reber, P. J. (2008). A review of medial temporal lobe and caudate contributions to visual category learning. *Neuroscience and Biobehavioral Reviews, 32,* 279–291. doi:10.1016/j.neubiorev.2007.07.006

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General, 115,* 39–57. doi:10.1037/0096-3445.115.1.39

Nosofsky, R. M. (2011). The generalized context model: An exemplar model of classification. In E. M. Pothos & A. J. Wills (Eds.), *Formal approaches in categorization* (pp. 18–39). Cambridge, England: Cambridge University Press.

Nosofsky, R. M., & Kruschke, J. K. (2002). Single-system models and interference in category learning: Commentary on Waldron and Ashby (2001). *Psychonomic Bulletin & Review, 9,* 169–174. doi:10.3758/BF03196274

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. doi:10.3758/BF03193227

Pothos, E. M., & Wills, A. J. (Eds.). (2011). *Formal approaches in categorization.* Cambridge, England: Cambridge University Press.

Prince, M., Brown, S., & Heathcote, A. (2012). The design and analysis of state-trace experiments. *Psychological Methods, 17,* 78–99. doi:10.1037/a0025809

Sherry, D. F., & Schacter, D. L. (1987). The evolution of multiple memory systems. *Psychological Review, 94,* 439–454. doi:10.1037/0033-295X.94.4.439

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied, 74*(11), 1–29. doi:10.1037/h0093759

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. doi:10.3758/BF03193507

Wagenmakers, E.-J., Ratcliff, R., Gomez, P., & Iverson, G. J. (2004). Assessing model mimicry using the parametric bootstrap. *Journal of Mathematical Psychology, 48,* 28–50. doi:10.1016/j.jmp.2003.11.004

Zeithamova, D., & Maddox, W. T. (2006). Dual-task interference in perceptual category learning. *Memory & Cognition, 34,* 387–398. doi:10.3758/BF03193416

Appendix

State-Trace Analysis and the General Linear Model

In this section, we briefly examine the relationship between state-trace analysis (STA) and the general linear model (GLM) as assumed by the analysis of variance (ANOVA). To illustrate this, we consider a simplified version of the experiments conducted in the present article. In this simplified version, there are three independent variables, labeled *b, c,* and *d*. In relation to the current set of experiments, *b* corresponds to block, *c* to category type, and *d* to delay. We assume that each variable has two levels (although our conclusions do not depend on this) and, without loss of generality, that the two levels are coded as 0 and 1, respectively.

Let y be the average outcome across the conditions of the experiment. According to the GLM, this can be written as

$$y = a_0 + a_1b + a_2d + a_3bd + a_4c + a_5bc + a_6dc + a_7bdc. \tag{A1}$$

Here, the coefficients a_0 to a_7 correspond to the relevant effect sizes including that for the intercept (a_0). In the current application of STA, we analyze the two levels of c separately. Let y_{c_i} be average outcome corresponding to the ith level of c. Then, from Equation A1, we have the following:

$$y_{c_i} = a_0 + a_1b + a_2d + a_3bd + a_4c_i + a_5bc_i + a_6dc_i + a_7bdc_i, \tag{A2}$$

where the c_i are now constants rather than variables. Substituting into Equation A2 $c_1 = 0$ and $c_2 = 1$, we get

$$\begin{aligned} y_0 &= a_0 + a_1b + a_2d + a_3bd, \\ y_1 &= a_0 + a_1b + a_2d + a_3bd + a_4 + a_5b + a_6d + a_7bd, \\ &= (a_0 + a_4) + (a_1 + a_5)b + (a_2 + a_6)d + (a_3 + a_7)bd. \end{aligned} \tag{A3}$$

The plot of y_1 against y_0 for each combination of variables, *b* and *d,* is called the *state-trace on (variable) c*. Under what conditions is this state-trace one-dimensional? This question was examined by Dunn and James (2003), who showed that the answer depends on the nature of the function that maps the parameters of the relevant model, *b* and *d* in the present case, onto the outcomes, y_0 and y_1. Specifically, the dimension of the state-trace is determined by the rank of the Jacobian matrix of this function.[A1] The Jacobian matrix, *Dy,* is the matrix of partial derivatives of a multivariate function with respect to each of the parameters of the function. From Equation A3, this is given by

[A1] Strictly speaking, it depends upon the rank of the *integral Jacobian matrix,* formed by integrating the Jacobian matrix along a path between two arbitrary points in the domain of the function. However, Dunn and James (2003) also showed that if the function is linear, as in the present case, then the integral Jacobian matrix and the Jacobian matrix are identical.

(*Appendix continues*)

$$Dy = \begin{bmatrix} \dfrac{\partial y_0}{\partial b} & \dfrac{\partial y_0}{\partial d} \\ \dfrac{\partial y_1}{\partial b} & \dfrac{\partial y_1}{\partial d} \end{bmatrix}$$

$$= \begin{bmatrix} a_1 + a_3 d & a_2 + a_3 b \\ (a_1 + a_5) + (a_3 + a_7)d & (a_2 + a_6) + (a_3 + a_7)b \end{bmatrix}. \tag{A4}$$

Dunn and James showed that the state-trace on c is one-dimensional if Dy is less than full rank for all b and d. If a matrix is square, as in Equation A4, then it is less than full rank if and only if its determinant is equal to zero. The determinant of Dy, $|Dy|$, is given by

$$|Dy| = a_1a_6 - a_2a_5 + (a_1a_7 - a_3a_5)b + (a_3a_6 - a_2a_7)d,$$

which may go to zero (for all values of b and d) in any of several different ways. Some of these are the following:

1. No main effect of or interaction with b ($a_1 = a_3 = a_5 = a_7 = 0$).

2. No main effect of or interaction with d ($a_2 = a_3 = a_6 = a_7 = 0$).

3. No main effects of b or d and no two-way interaction of either variable with c ($a_1 = a_2 = a_5 = a_6 = 0$).

4. No main effects of b and d and no interaction ($a_1 = a_2 = a_3 = 0$).

5. No interaction of either b or d with c ($a_5 = a_6 = a_7 = 0$).

6. All effects equal in size ($a_1 = a_2 = a_3 = a_4 = a_5 = a_6 = a_7$).

In practical terms, this means that if we conduct an ANOVA on the data and it turns out that the pattern of effects is consistent with the Jacobian matrix being less than full rank (e.g., if any of the outcomes listed above occur), then the corresponding state-trace on c will be one-dimensional. However, the opposite is not true. If the pattern of effects is not of this sort, then it cannot be concluded that the state-trace must be two-dimensional. In other words, the fact that c interacts with one or more other variables is not sufficient to conclude that the state-trace on c is two-dimensional.

Received May 31, 2011
Revision received January 27, 2012
Accepted February 2, 2012 ■