Available online at www.sciencedirect.com

SCIENCE DIRECT®

Behavioural Processes 66 (2004) 309–332

BEHAVIOURAL PROCESSES

www.elsevier.com/locate/behavproc

Dissociating explicit and procedural-learning based systems of perceptual category learning

W. Todd Maddox[a,*], F. Gregory Ashby[b]

[a] Department of Psychology, 1 University Station A8000, University of Texas, Austin, TX 78712, USA
[b] University of California, Santa Barbara, CA, USA

Abstract

A fundamental question is whether people have available one category learning system, or many. Most multiple systems advocates postulate one explicit and one implicit system. Although there is much agreement about the nature of the explicit system, there is less agreement about the nature of the implicit system. In this article, we review a dual systems theory of category learning called competition between verbal and implicit systems (COVIS) developed by Ashby et al. (1998). The explicit system dominates the learning of verbalizable, rule-based category structures and is mediated by frontal brain areas such as the anterior cingulate, prefrontal cortex (PFC), and head of the caudate nucleus. The implicit system, which uses procedural learning, dominates the learning of non-verbalizable, information-integration category structures, and is mediated by the tail of the caudate nucleus and a dopamine-mediated reward signal. We review nine studies that test six a priori predictions from COVIS, each of which is supported by the data.

Keywords: Categorization; Classification; Memory; Caudate nucleus; Working memory; Prefrontal cortex

1. Introduction

Categorization is a vitally important skill. The feeding deer must categorize a sound as "friend", "foe", "a gust of wind", etc. with a foe judgment inducing the deer to cease feeding and to run. The expedition doctor camped out near the summit of Lhotse must categorize a climber's breathing difficulty, level of dizziness, etc. as a sign of "pulmonary edema", "exhaustion", etc. with a pulmonary edema diagnosis leading to immediate retreat. These are categorization problems because in both cases there are many (generally an infinite number of) information states, but only a few courses of action. Categorization performance is governed by the organism's evolutionary history, experience with the environment and the reinforcing consequences of the decisions that they make.

Category learning involves laying down a memory trace of some sort that can be used to improve the efficiency (i.e. accuracy and speed) of responding. It is now widely accepted that mammals have multiple memory systems (Schacter, 1987; Squire, 1992), and this fact alone makes it reasonable to postulate that multiple category learning systems might also exist. In fact, a growing body of research suggests that the learning of different types of category structures is mediated by different systems (Ashby and Ell, 2001, 2002; Ashby et al., 2002; Erickson and Kruschke, 1998; Pickering, 1997; Maddox et al., 2003a,b,c, 2004; Reber and Squire, 1994; Smith et al., 1998; however see Nosofsky and Johansen, 2000). Most multiple systems theorists have argued for one ex-

* Corresponding author. Tel.: +1-512-475-8494; fax: +1-512-471-5935.
E-mail address: maddox@psy.utexas.edu (W.T. Maddox).

0376-6357/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.beproc.2004.03.011

plicit, hypothesis-testing system that is tied to conscious awareness, and one implicit system that does not have full access to conscious awareness. Whereas most multiple systems theorists agree that one system is explicit and another is implicit, there is disagreement about the nature of the implicit system. Some argue for an exemplar-based system (e.g. Erickson and Kruschke, 1998), some for a perceptual representation system (e.g. Reber et al., 1998), and others for a procedural learning-based system (e.g. Ashby et al., 1998, 2003a,b; Ashby and Waldron, 1999). A likely possibility is that each proposal has some validity, and multiple implicit category learning systems may exist (Ashby and Casale, 2002). Although it is reasonable to suppose that exemplar-memory and perceptual-representation systems may exist, the implicit system focused on in this article is procedural learning-based.

One of the most successful multiple systems models of category learning, and the only one that specifies the underlying neurobiology, is the competition between verbal and implicit systems (COVIS) model proposed by Ashby et al. (1998) and Ashby and Waldron (1999). COVIS postulates two systems that compete throughout learning—an explicit, rule-based system that uses logical reasoning and depends on working memory and executive attention, and a procedural learning-based system. One intriguing aspect of the procedural learning-based system is its association with motor performance (e.g. Hazeltine and Ivry, 2001; Willingham et al., 1989; Willingham, 1998). Whereas there is no a priori reason to expect that exemplar-based or perceptual representation-based category learning should be closely linked to associated motor responses, categories learned via a procedural learning-based system should have a close link to the motor response.

Much of the evidence for multiple category learning systems comes from two different types of categorization tasks. Rule-based category learning tasks are those in which the category structures can be learned via some explicit reasoning process that treats each stimulus dimension separately. Frequently, the rule that maximizes accuracy (i.e. the optimal rule) is easy to describe verbally, and thus will often be referred to as verbalizable rule tasks (Ashby et al., 1998). In the most common applications, only one stimulus dimension is relevant, and the observer's task is to discover this relevant dimension and then to map the different dimensional values to the relevant categories. Examples of rule-based category learning tasks are outlined in Figs. 1a and 2a. In Fig. 1a, the stimuli vary along four binary-valued dimensions: background color (light gray versus dark gray), symbol shape (circle versus square), symbol number (one versus two), and symbol color (white versus black), and a small number of items are in each category (eight). In the Fig. 1a task, background color is relevant, with the light gray background being assigned to category A, the dark gray background being assigned to category B, and the other three dimensions being irrelevant. In Fig. 2a, the stimulus dimensions are continuous valued, and a large number of items are in each category. Each symbol in Fig. 2 denotes the spatial frequency and orientation of a single Gabor patch (see Fig. 9 for a sample stimulus). Also shown in Fig. 2 are the decision bounds that maximize categorization accuracy. In the rule-based task, the optimal bound requires observers to attend to spatial frequency and ignore orientation. The stimulus dimensions of Gabor patches are

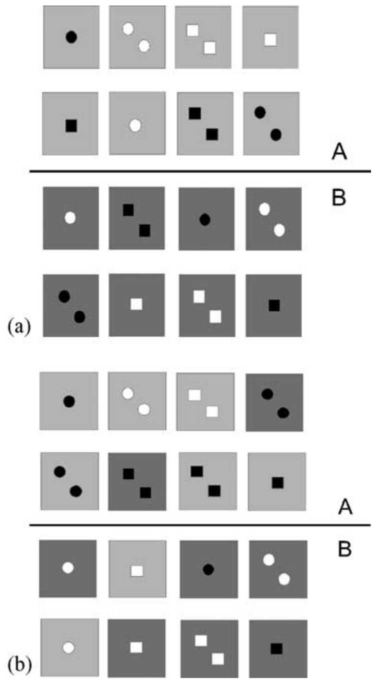


Fig. 1. (a) Verbalizable, rule-based and (b) non-verbalizable, information-integration category structures derived from a small number of stimuli that vary along four binary valued dimensions.

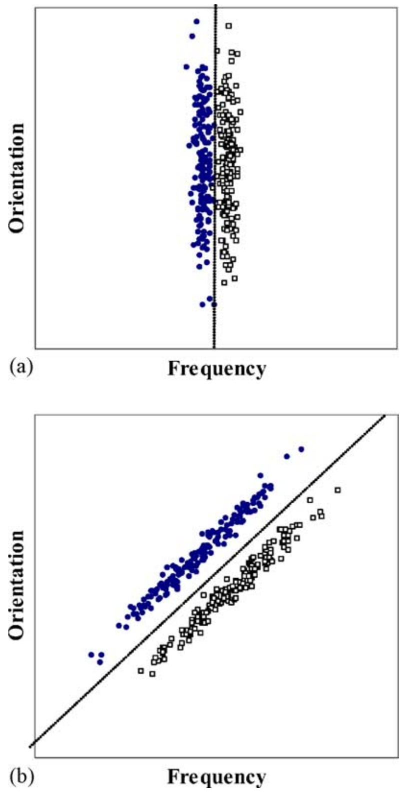


Fig. 2. (a) Rule-based and (b) information-integration category structures derived from a large number of stimuli that vary along two continuous-valued dimensions. Each circle denotes the spatial frequency and spatial orientation of a Gabor pattern from category A. Each square denotes the spatial frequency and spatial orientation of a Gabor pattern from category B. The broken line in each panel denotes the location of the optimal decision bound.

perceptually separable, have simple verbal labels (bar width and orientation), and have no emergent (or configural) features (e.g. all patches are exactly the same size and shape, regardless of frequency or orientation). For these reasons, there is a simple explicit rule that separates the contrasting categories. In particular, the vertical bound in Fig. 2a corresponds to the rule: "Respond A if the bars are thick and B if they are thin".

Information-integration category learning tasks are those in which accuracy is maximized only if information from two or more stimulus components (or dimensions) is integrated at some pre-decisional stage (Ashby and Gott, 1988). Perceptual integration could take many forms—from treating the stimulus as a Gestalt to computing a weighted linear combination of the dimensional values.[1] In many cases, the optimal rule in information-integration tasks is difficult or impossible to describe verbally, and thus will often be referred to as non-verbalizable rule tasks (Ashby et al., 1998). To create an information-integration structure with the Fig. 1 stimuli, one dimension is selected arbitrarily to be irrelevant. For example, in Fig. 1b, the irrelevant dimension is symbol shape. Next, one level on each relevant dimension is arbitrarily assigned a numerical value of +1 and the other level is assigned a value of 0. In Fig. 1, a background color of light gray, a symbol color of black, and a symbol number of 2 are all assigned a value of +1. Finally, the category assignments are determined by the following rule: The stimulus belongs to category A if the sum of values on the relevant dimensions > 1.5; Otherwise it belongs to category B. The information-integration task in Fig. 2b was generated by rotating the rule-based categories by 45°. These four category structures were used in the studies reviewed below.

This article summarizes the results from nine studies conducted in our laboratories that provide rigorous tests of several a priori predictions derived from COVIS. By a priori, we mean predictions that do not depend on specific parameter estimates, and therefore can be made before any data are collected. Predictions that depend on the numerical values of specific parameter estimates cannot be made until after the data have been collected, which makes them post hoc. The aim of these studies was two-fold. First, we wished to test the hypothesis that some category structures are learned via an implicit, procedural-learning based categorization system. Second, we wished to empirically dissociate the explicit system from the procedural learning-based systems. To achieve these goals we introduced experimental manipulations that were predicted a priori (often based on an examination of the proposed underlying neurobiology) to affect information-integration category learning, but not rule-based category learning, or were predicted

[1] A conjunction rule (e.g. respond A if the stimulus is small on dimension x and small on dimension y) is a dimensional task rather than a information-integration task because separate decisions are first made about each dimension (e.g. small or large) and then the outcome of these decisions is combined (integration is post-decisional, not pre-decisional).

a priori to affect rule-based, but not information-integration category learning.

The next (second) section provides an overview of the proposed neurobiological underpinnings of the explicit, hypothesis-testing, and implicit procedural learning-based systems of COVIS. The third section outlines three a priori predictions derived from COVIS, and five studies that test these predictions by empirically dissociating information-integration category learning from rule-based category learning. The fourth section outlines three additional a priori predictions derived from COVIS, and four studies that test these predictions by empirically dissociating rule-based category learning from information-integration category learning. The final section offers a general summary.

2. Neurobiology of the explicit and procedural-learning based category learning systems

2.1. The explicit system

Patients with frontal or basal ganglia dysfunction are impaired in verbalizable, rule-based tasks (e.g. Brown and Marsden, 1988; Cools et al., 1984; Kolb and Whishaw, 1990; Robinson et al., 1980), but patients with medial temporal lobe damage are normal in this type of category learning (e.g. Janowsky et al., 1989; Leng and Parkin, 1988). Thus, an obvious first hypothesis is that the prefrontal cortex (PFC) and the basal ganglia participate in this type of learning, but the medial temporal lobes do not. Converging evidence for the hypothesis that these are important structures in rule-based category learning comes from several sources. First, an fMRI study of a rule-based task similar to the Wisconsin Card Sorting Test showed activation in the right dorsal-lateral prefrontal cortex, the anterior cingulate, and the (head of the) right caudate nucleus (a major input structure within the basal ganglia) (among other regions) (Rao et al., 1997). Second, many studies have implicated these structures as key components of executive attention (Posner and Petersen, 1990) and working memory (e.g. Fuster, 1989; Goldman-Rakic, 1987), both of which are likely to be critically important to the explicit processes of rule formation and testing that are assumed to mediate rule-based category learning. Third, neuroimaging studies have identified the (dorsal) anterior cingulate as the site of hypothesis generation in a rule-based category learning task (Elliott and Dolan, 1998). Fourth, lesion studies in rats implicate the dorsal caudate nucleus in rule switching (Winocur and Eskes, 1998).

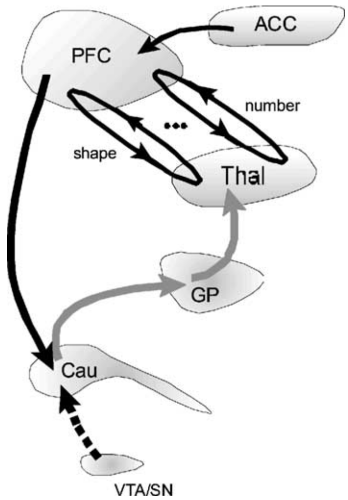


Fig. 3. A model of the COVIS explicit category learning system. Note: black projections are excitatory, gray projections are inhibitory, and dashed projections are dopaminergic. PFC: prefrontal cortex; ACC: anterior cingulate cortex; Thal: thalamus; GP: globus pallidus; Cau: caudate nucleus; VTA: ventral tegmental area; SN: substantia nigra.

Fig. 3 describes the COVIS explicit system (Ashby et al., 1998, 1999a,b). The key structures are the anterior cingulate, the prefrontal cortex, and the head of the caudate nucleus. Fig. 3 shows the model during a trial of the rule-based category learning task illustrated in Fig. 1. Various salient explicit rules reverberate in working memory loops between prefrontal cortex and thalamus (Alexander et al., 1986). In Fig. 3, one such loop maintains the representation of a rule focusing on the shape of the symbols and one loop maintains a rule focusing on symbol number. An excitatory projection from the PFC to the head of the caudate nucleus prevents the globus pallidus from interrupting these loops. The anterior cingulate selects new explicit rules to load into working memory, and the head of the caudate nucleus mediates the switch from one active loop to another (facilitated by dopamine projections from the ventral tegmental area and the substantia nigra).

The Fig. 3 model is consistent with the neuroimaging data described above, and it accounts for the existing neuropsychological data on rule-based category learning. First, of course, it is obvious that the model predicts that patients with lesions of the prefrontal cortex will be impaired on rule-based category learning tasks. It also predicts that the deficits seen in Parkinson's disease are due to dysfunction in the head of the caudate nucleus. Postmortem autopsy reveals that damage to the head of the caudate is especially severe in Parkinson's disease (van Domburg and ten Donkelaar, 1991), so the model predicts that this group should show widespread and profound deficits on rule-based categorization tasks. The neuropsychological evidence strongly supports this prediction (e.g. on the WCST; Brown and Marsden, 1988; Cools et al., 1984). In fact, the model described in Fig. 3 predicts that, because of its reciprocal connection to the prefrontal cortex, many of the well documented "frontal-like" symptoms of Parkinson's disease might actually be due to damage in the head of the caudate nucleus.

2.2. *The procedural-learning based system*

In non-verbalizable, information-integration tasks, patients with basal ganglia dysfunction are impaired (Filoteo et al., 2001a; Maddox and Filoteo, 2001), but medial temporal lobe patients are normal (Filoteo et al., 2001b). So, first hypothesis should be that the basal ganglia are critical in this task, but the medial temporal lobes are not. This hypothesis was supported by Poldrack et al. (1999), who used fMRI to measure neural activation at four different time points of learning in an information-integration task. They reported learning-related changes within prefrontal cortex and in the right caudate nucleus. Interestingly, they also reported a simultaneous suppression of activity within the medial temporal lobes. Thus, the available neuroimaging data predict that the deficits of basal ganglia disease patients in information-integration tasks may arise from dysfunction in the caudate nucleus.

Lesions of the tail of the caudate, in both rats and monkeys, impair the ability of the animal to associate one motor response with one visual stimulus and a different response with some other stimulus (e.g. vertical versus horizontal lines; McDonald and White, 1993, 1994; Packard et al., 1989; Packard and McGaugh, 1992). For example, in one study, rats with lesions in the tail of the caudate could not learn to discriminate between safe and unsafe platforms in the Morris water maze when the safe platform was marked with horizontal lines and the unsafe platform was marked with vertical lines (Packard and McGaugh, 1992). The same animals learned normally, however, when the cues signaling which platform was safe were spatial. Because the tail of the caudate nucleus is not a classic visual area, it is unlikely that these animals have an impaired ability to perceive the stimuli. Rather, it seems more likely that their deficit is in learning the appropriate stimulus-response associations. Similarly, the basal ganglia have also been proposed as the main locus of procedural learning in humans (e.g. Mishkin et al., 1984).

Fig. 4 illustrates the procedural learning-based categorization system of COVIS (Ashby et al., 1998; Ashby and Waldron, 1999). The most important neural region in this model is the caudate nucleus. In primates, all of extrastriate visual cortex projects directly to the tail of the caudate nucleus, with about 10,000 visual cortical cells converging on each caudate cell (Wilson, 1995). Cells in the tail of the caudate (i.e. medium spiny cells) then project to prefrontal and premotor regions (especially the supplementary motor area; via the globus pallidus and thalamus; e.g. Alexander et al., 1986). The model assumes that, through a procedural learning process, each unit in caudate serves to link a large group of visual cortical cells (i.e. all that project to it) with an abstract motor program, which is most likely represented in supplementary motor area.

It has also been hypothesized that this learning is facilitated by a dopamine-mediated reward signal from the substantia nigra (pars compacta) (e.g. Wickens, 1993). There is a large literature linking dopamine and reward, and many researchers have argued that a primary function of dopamine is to serve as the reward signal in reward-mediated learning (e.g. Beninger, 1983; Miller et al., 1981; Montague et al., 1996; White, 1989; Wickens, 1993). For example, it has been shown that rewards, and events that signal reward, elicit release of dopamine from several brainstem sites (for reviews, see, e.g. Bozarth, 1994; Pfaus and Phillips, 1991; Phillips et al., 1992), and it is well known that dopamine antagonists (i.e. neuroleptics) disrupt the reward signal and render reinforcement ineffective (e.g. Ataly and Wise, 1983).

Fairly specific neurobiological models of this learning process have been developed (e.g. Wickens, 1993).

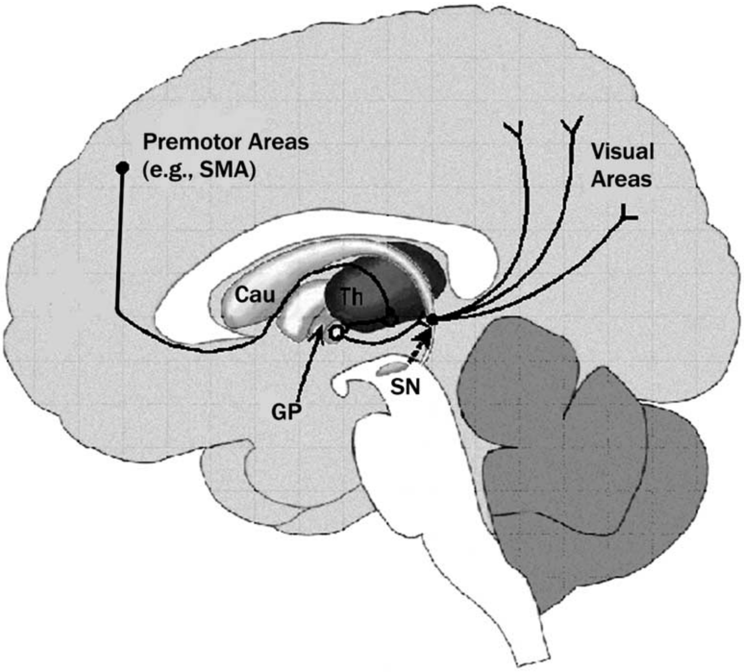


Fig. 4. The COVIS procedural-memory-based category learning system. Excitatory projections end in solid circles, inhibitory projections end in open circles, and dopaminergic projections are dashed. PFC: prefrontal cortex; Cau: caudate nucleus; GP: globus pallidus, and Th: thalamus.

Fig. 5 shows a close-up view of a synapse between the axon of a pyramidal cell originating in visual cortex and the dendrite of a medium spiny cell in the caudate nucleus. Note that glutamate projections from visual cortex and dopamine projections from the substantia nigra both synapse on the dendritic spines of caudate medium spiny cells (DiFiglia et al., 1978; Freund et al., 1984; Smiley et al., 1994). A cortical signal causes an influx of free Ca^{2+} into the spines (through NMDA receptors). The main effects of Ca^{2+} entering the cell are to activate Ca-dependent protein kinases, which then perform a number of cellular functions, including depolarizing the cell and strengthening (long term potentiation (LTP)) the synapse (e.g. Cooper et al., 1991; Lynch et al., 1983; Wickens, 1993). Because the spines are somewhat separated from the bulk of the intracellular medium, the spine remains depolarized for several seconds after the cell fires (Gamble and Koch, 1987; MacDermott et al., 1986). Under ideal conditions, the dopamine-mediated reward signal will arrive during this time, and there is substantial evidence that it will interact with the glutamate signal. The most popular model of this interaction assumes that after dopamine binds to the D_1 receptor and activates its associated G protein, a sequence of chemical reactions result that ultimately inhibit the deactivation of the Ca-dependent protein kinases that are activated after glutamate binds to the NMDA receptor (Nairn et al., 1988; Pessin et al., 1994; Wickens, 1990, 1993). The effect of this inhibition is that dopamine locks the glutamate second messenger in the "on" position, thereby potentiating the learning effect. Thus, the presence of dopamine strengthens the synapses that were active on a trial when reward was delivered (e.g. Huang and Kandel, 1995).

The Fig. 4 model accounts for the category learning deficits of Parkinson's and Huntington's disease patients in information-integration tasks because both of these populations suffer from caudate dysfunction. It also explains why frontal patients and medial temporal lobe amnesiacs are relatively normal in these tasks—that is, because neither prefrontal cortex nor medial temporal lobe structures play a prominent role in the Fig. 4 model.

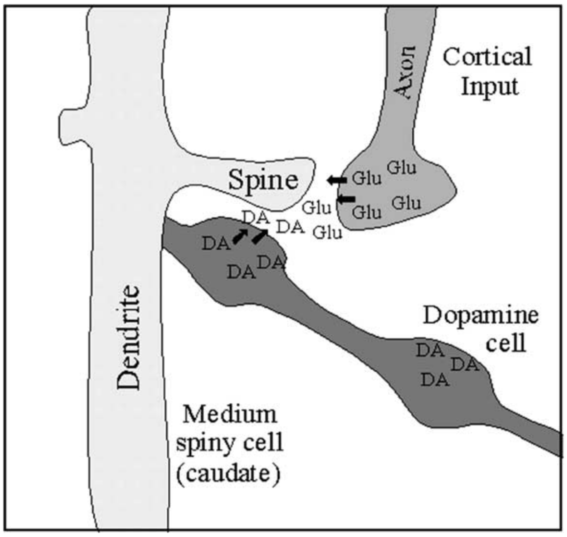


Fig. 5. A closer view of a cortical-striatal synapse. Here, a cortical cell terminal releases glutamate (Glu) onto the dendritic spine of a medium spiny cell of the caudate nucleus. Dopamine cells of the substantia nigra also project onto medium spiny cells and upon presentation of reward, release dopamine (DA) into the same synapse.

Although we describe the explicit and implicit systems in separate sections we do not wish to imply that they are neurobiologically and cognitively independent. Both systems simultaneously attempt to learn the categorization rule on each trial, and there is some neurobiological overlap. Even so, we propose that the each system has its own strengths and weaknesses and that each is better suited to specific environmental conditions.

3. Empirical tests of six a priori predictions derived from the proposed neurobiological underpinnings of COVIS

As outlined above, this article reviews a number of studies that test a priori predictions derived from an examination of the neurobiological underpinnings of COVIS. In this section, we outline each of these a priori predictions. The first set of predictions are based on experimental manipulations that should interfere with information-integration learning more than with rule-based learning, whereas the second set are based on experimental manipulations that should interfere with rule-based learning more than with information-integration learning.

3.1. Predicted effects on non-verbalizable, information-integration but not verbalizable, rule-based category learning

3.1.1. A priori Prediction 1

Unsupervised training should disrupt information-integration more than rule-based category learning: The COVIS explicit system uses working memory and executive attention and has access to conscious awareness. As a result, it is quite flexible in its use of feedback. Under some conditions, it might even be able to generate its own feedback signal. On the other hand, learning can occur in the procedural-learning system of COVIS only if correct responses are followed immediately by a (dopamine-mediated) reward signal. Therefore, in the absence of trial-by-trial feedback—what we refer to as unsupervised conditions—COVIS predicts that only the explicit system has the potential to learn. Further, since the two systems compete, COVIS predicts that people should use explicit, rule-based strategies during unsupervised categorization.

Ashby et al. (1999a,b) tested these predictions by having observers attempt to learn two categories that were each composed of single lines that differed in length and orientation (an example can be found in Fig. 7). Training was completely unsupervised, in the sense that trial-by-trial feedback was absent. At the beginning of the experiment, observers were told only that there were two categories, and that perfect performance was possible. They were also told that the category labels were arbitrary but that they should be consistent in their category assignment. Observers learned either the Fig. 2a rule-based category structures, or the Fig. 2b information-integration category structures.

The learning curves (averaged across observers) are displayed in Fig. 6a for the rule-based and information-integration category learning tasks. First note that learning did occur in the rule-based condition, in the sense that accuracy improved with practice. In fact, by the end of the training period, performance in the rule-based condition was nearly perfect. On the other hand, performance reached asymptote during the first block of trials in the information-integration

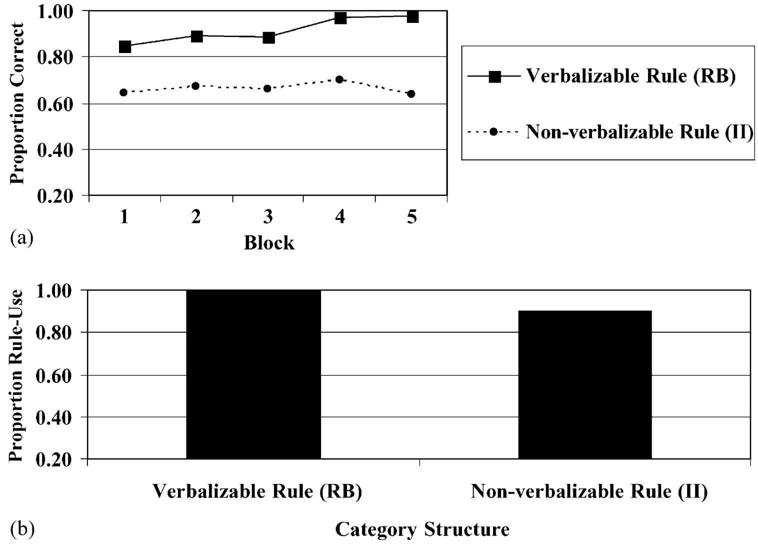


Fig. 6. (a) Proportion correct during the final response block for the verbalizable rule (RB) and non-verbalizable rule (II) category structures (standard error bars included). (b) Proportion rule-use during the final response block for these same conditions.

condition. Accuracy in the last block was no higher than accuracy in the first block. In a follow-up experiment, observers were given trial-by-trial feedback while attempting to learn the same information-integration category structures. With feedback, observers achieved an average of 93.3% accuracy during the final block of trials. Taken together, these results support the a priori prediction that the absence of feedback disrupts information-integration more than rule-based category learning.

The accuracy-based analyses provide important information regarding overall performance but they tell us little about the types of strategies that observers might use to solve these tasks. An understanding of strategy use and of how these strategies might be affected by the different training procedures is of critical importance to a complete understanding of category learning. Our focus is on two qualitatively different types of strategies. Rule-based strategies are those that involve the application of verbalizable rules to separate the category exemplars. Information-integration strategies are those that involve a pre-decisional integration of the stimulus information. The explicit, hypothesis-testing system in COVIS is assumed to use rule-based strategies, whereas the implicit, procedural learning system is assumed to use information-integration strategies. As suggested earlier, because the explicit system is assumed to dominate early, and because the implicit system should be unable to learn without trial-by-trial feedback, COVIS predicts that, in the absence of feedback, rule-based strategies should be applied to both rule-based and information-integration category structures. To test these predictions, a number of different decision bound models were fit to the final block of trials from each observer. Decision bound theory assumes each observer partitions the perceptual space into response regions by constructing a decision bound (Ashby, 1992a; Maddox and Ashby, 1993). On each trial, the observer determines which region the percept is in, and then emits the associated response.

Two different types of decision bound models were fit to each observer's responses (the details of the models are provided in the Appendix A). One type is compatible with the assumption that observers used a rule-based strategy and one type assumes an information-integration strategy. These models make no detailed process assumptions in the sense that a number of different process accounts are compatible with each of the models (e.g. Ashby, 1992a; Ashby and Waldron, 1999). For example, if an information-integration model fits significantly better than a rule-based model, then we can be reasonably confident that observers did not use a rule-based strategy, but we learn little about which information-integration strategy might have been used (e.g. decision bound, exemplar, or prototype interpretations would all be compatible with such results). In contrast, if a rule-based model fits significantly better than an information-integration model, then we gain confidence that observers used a rule-based strategy but we cannot rule out all information-integration strategies, because some of these can mimic rule-based responding. Thus, these models provide a useful tool for testing hypotheses about the decision strategies used by observers, but say little about psychological process.

For each observer's final block of trials, Ashby et al. (1999a,b) determined which model type (i.e. rule-based, or information-integration) provided the best account of the data. The proportion of observers who used a rule-based strategy is displayed in Fig. 6b. Note that a rule-based model provided the best account of the data for all observers learning rule-based categories and for 90% of the observers learning information-integration categories. Taken together with the accuracy data, these results strongly suggest that rule-based category learning can occur without feedback, whereas they fail to show any evidence of information-integration category learning.

3.1.2. A priori Prediction 2

Observational training and delayed feedback should interfere with information-integration category learning more than with rule-based category learning: Within the tail of the caudate nucleus, a reward-mediated feedback signal is thought to be provided by dopamine released from the substantia nigra pars compacta shortly after the animal receives an unexpected reward (Hollerman and Schultz, 1997; Schultz, 1992; Wickens, 1993). The presence of this dopamine is widely thought to strengthen recently active synapses (which presumably are responsible for the animal obtaining the reward) (e.g. Arbuthnott et al., 2000; Calabresi et al., 1996). In reward-mediated learning, it is essential to strengthen those (and only those) synapses that actively participated in the response that elicited the reward. Because there is necessarily some delay between response and reward delivery, this means that some trace must be maintained that signals which synapses were recently active. In the case of the medium spiny cells in the caudate nucleus, the morphology of the dendritic spines allows this trace to exist for about 2.5 s after the response is initiated (Gamble and Koch, 1987; MacDermott et al., 1986). If the reward is delayed by more than this amount, then the ensuing dopamine release will strengthen inappropriate synapses and learning will be adversely affected.

The fact that learning in the implicit system requires an unexpected reward suggests that observational training procedures—that is, situations in which the category label is presented prior to the stimulus—should adversely affect information-integration category learning. In addition, the fact that this system requires a fairly quick feedback presentation suggests that increasing the delay between the person's response and presentation of the feedback should adversely affect information-integration category learning. On the other hand, since the explicit system has full access to working memory, observational training and delayed feedback should minimally affect rule-based category learning.

Ashby et al. (2002) tested the hypothesis that observational training will disrupt information-integration category learning more than rule-based category learning by having a large number of observers attempt to learn the same rule-based and information-integration category structures (with the same stimuli) that were used by Ashby et al. (1999a,b). However, rather than unsupervised training, Ashby et al. (2002) contrasted observational training with standard feedback training. In an attempt to equate rule-based and information-integration category learning in the control (i.e. feedback) condition, category discriminability in the rule-based conditions was slightly lower than that shown in Fig. 2a. A sample trial for the observational and feedback training procedures is outlined in Fig. 7. Note that the timing of each observational and feedback learning trial is identical except that the location of the category label is reversed.

The proportion correct during the final response block (averaged across observers) for the rule-based and information-integration category structures under observational and feedback-learning conditions is displayed in Fig. 8a. Two comments are in order. First, rule-based and information-integration category learning were equivalent with feedback training. Second, observational learning was significantly worse than feedback learning with the information-integration categories, but was equivalent for the rule-based

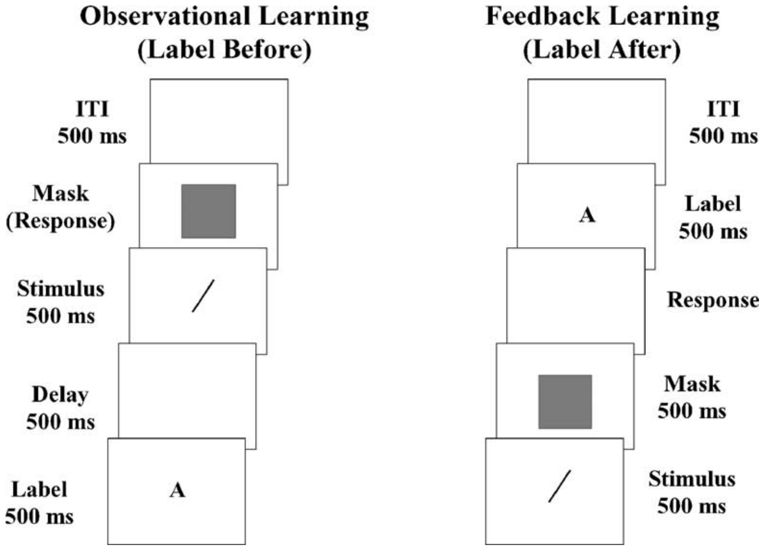


Fig. 7. Basic design for the observational and feedback training procedures.

categories. These results support the a priori prediction that observational training disrupts information-integration more than rule-based category learning.

Rule-based and information-integration decision bound models were fit to these data using the same procedures as in Ashby et al. (1999a,b). Fig. 8b shows the proportion of observers whose final response block data were best fit by a rule-based model. The results can be summarized as follows. First, and most importantly, the proportion of observers using a rule-

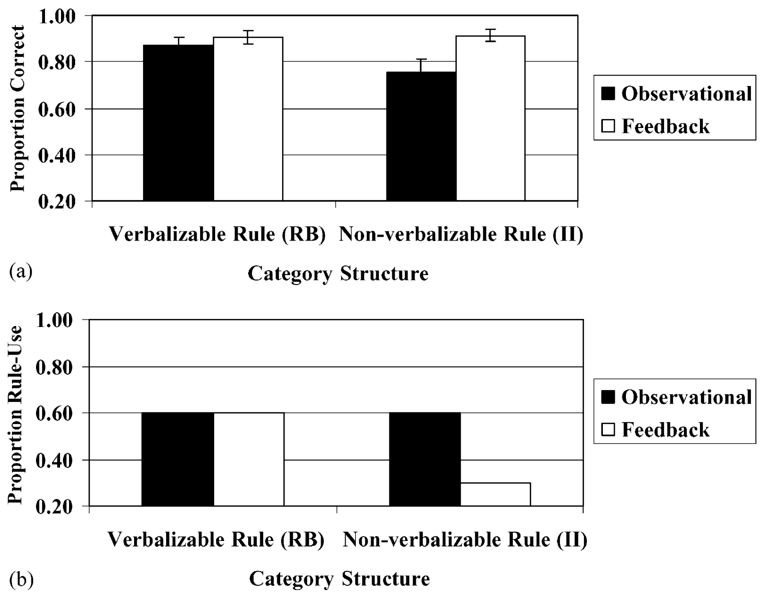


Fig. 8. (a) Proportion correct during the final block of trials for the verbalizable rule (RB) and non-verbalizable rule (II) category structures under observational and feedback training procedures (standard error bars included). (b) Proportion rule-use during the final block of trials for these same conditions.

based strategy with the non-verbalizable, information-integration category structures was high with observational training, but low with feedback training. Second, the proportion of observers using a rule-based strategy with the verbalizable, rule-based categories was high under both observational and feedback learning conditions. In other words, when feedback was provided, observers used the appropriate strategy for the task (i.e. a rule-based strategy with verbalizable, rule-based categories and an information-integration strategy with non-verbalizable, information-integration categories), but under observational training, people tended to use rule-based strategies with both types of category structures.

Maddox et al. (2003a,b,c, 2004) tested the hypothesis that delaying the feedback signal should disrupt information-integration category learning more than rule-based category learning. In this study, observers attempted to learn the rule-based and information-integration category structures outlined in Fig. 2 (except with the Gabor patch stimuli) with either immediate or delayed feedback. As in Ashby et al. (2002), category discriminability in the rule-based conditions was slightly lower than that shown in Fig. 2a. A sample trial for the immediate and delayed feedback training procedures is outlined in Fig. 9 for the case of a 5 s delay. Delays of 2.5 and 10 s were also examined and yielded similar results.

The proportion of correct responses during the final block of trials (averaged across observers) for the rule-based and information-integration category structures under delayed- and immediate-feedback conditions is shown in Fig. 10a. The models outlined above were also applied to these data, and the proportion of observers who used a rule-based strategy during the final block of trials is displayed in Fig. 10b. The pattern of results is similar to those obtained with observational versus feedback training. Specifically, delayed feedback led to lower accuracy rates and higher rule-use for the information-integration category structures, but had no effect on performance, relative to the immediate feedback condition, for the rule-based category structures. These results support the a priori prediction that delayed feedback adversely affects information-integration, but not rule-based category learning. Taken together with the results from Ashby et al. (1999a,b, 2002), these findings suggest that the nature and timing of the feedback has a strong effect on the efficiency of the procedural-learning system, and thus on the ability of people to learning information-integration category structures.

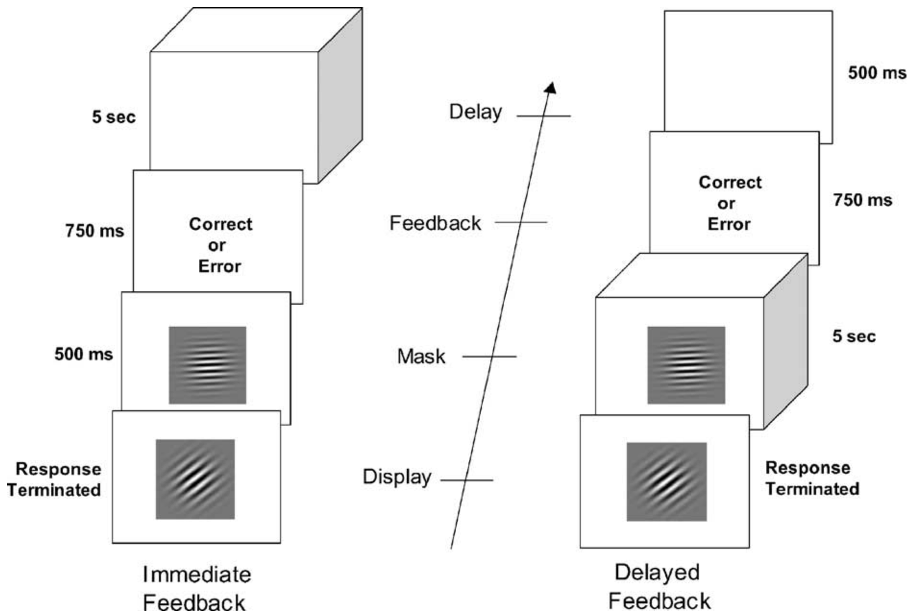
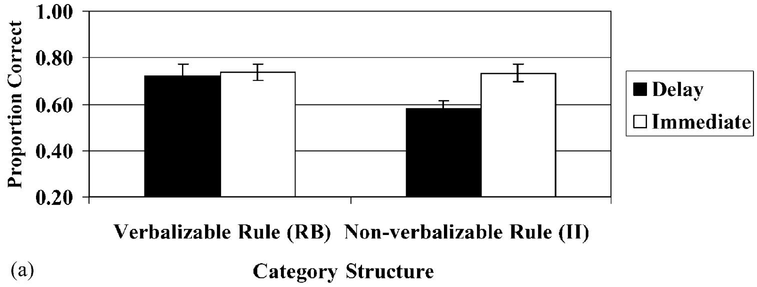


Fig. 9. Basic design for the immediate feedback and delayed feedback training procedures.



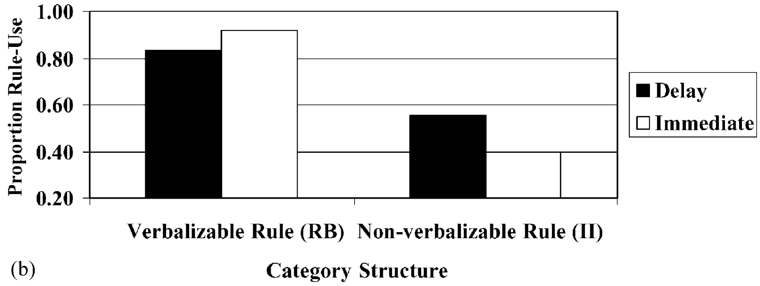


Fig. 10. (a) Proportion correct during the final block of trials for the verbalizable rule (RB) and non-verbalizable rule (II) category structures under immediate feedback and delayed feedback training procedures (standard error bars included). (b) Proportion rule-use during the final block of trials for these same conditions.

3.1.3. A priori Prediction 3

Changes in the motor requirements should affect information-integration category learning more than rule-based category learning: Procedural learning is closely related to motor learning, so category learning in the COVIS procedural-learning system should include some associated motor component, and this system should flourish under training procedures that preserve a close mapping between the category label and response location. Explicit reasoning, on the other hand, is abstract and not typically linked to a specific motor response. Therefore, the COVIS explicit system should not be especially sensitive to procedures that change the mapping between category label and response location. Two studies tested this hypothesis.

Ashby et al. (2003a) trained observers on the Fig. 2a and b rule-based and information-integration category structures (with the single line stimuli) using a training-transfer procedure. There were three conditions: control, hand-switch, and button-switch (see Fig. 11 for details). In the control condition, the response key assigned to category A was pressed with the left index finger and the response key assigned to category B and was pressed with the right index finger during both training and transfer. In the hand-switch condition, the hands were crossed during training so that the response key assigned to category A was pressed with the right index finger and the response key assigned to category B was pressed with the left index finger. During transfer, the hands were uncrossed on the response keys. In the button-switch condition, training was identical to that in the control condition, but during transfer the location of the buttons was switched. In addition, in all three conditions a 1.5 s response deadline was introduced at transfer.

The proportion correct during the final training block and the two transfer blocks (averaged across observers) is shown in Fig. 12 for the control, hand-switch, and button-switch conditions with rule-based (panel a) and information-integration (panel b) category structures. The results can be summarized as follows. For the rule-based category structures, neither switching the hands nor buttons had any effect on performance. For the information-integration category structures, there are two separate effects. In the control condition, the mild speed stress caused by introducing the 1.5 s response deadline caused a temporary drop in accuracy (of about 8%) that disappeared by the second transfer block. The hand-switch data did not differ significantly from the control

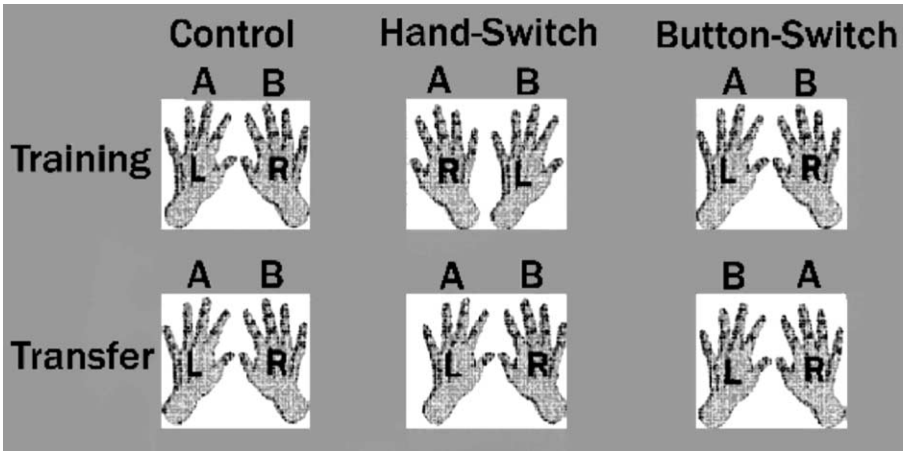


Fig. 11. Basic design for the hand-switch/button-switch training procedures.

data, which suggests that no additional interference was caused by switching the hands on the response keys. In the button-switch condition, however, the accuracy loss during the first transfer block was significantly larger than in either other condition, and it only partially recovered during the second transfer block. Thus, switching hands on the response keys introduced an additional interference, over and above the interference caused by the introduction of mild speed stress. These results suggest that the explicit system learns abstract category labels, whereas the procedural-learning system learns response positions (but not motor programs else the hand-switch would have caused an interference).

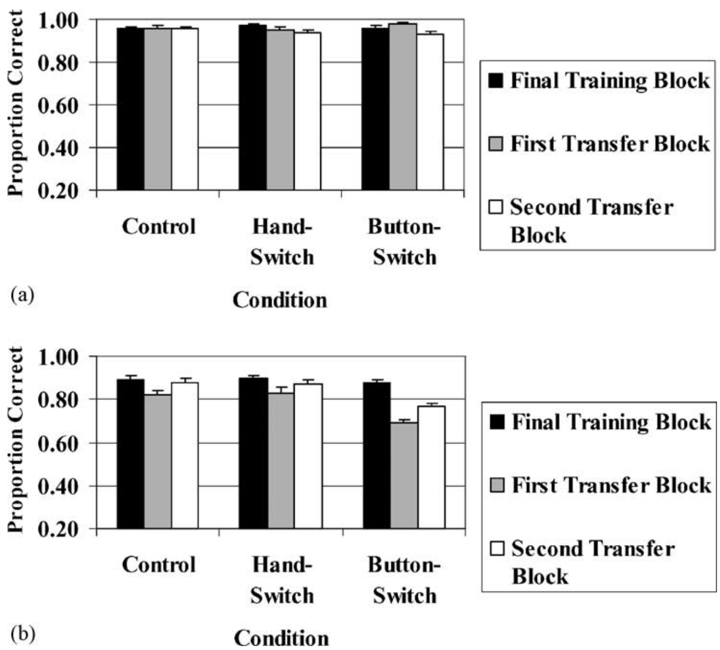


Fig. 12. Proportion correct during the final training, first transfer, and second transfer block for the control, hand-switch, and button-switch conditions for the (a) verbalizable rule (RB) and (b) non-verbalizable rule (II) conditions (standard error bars included).

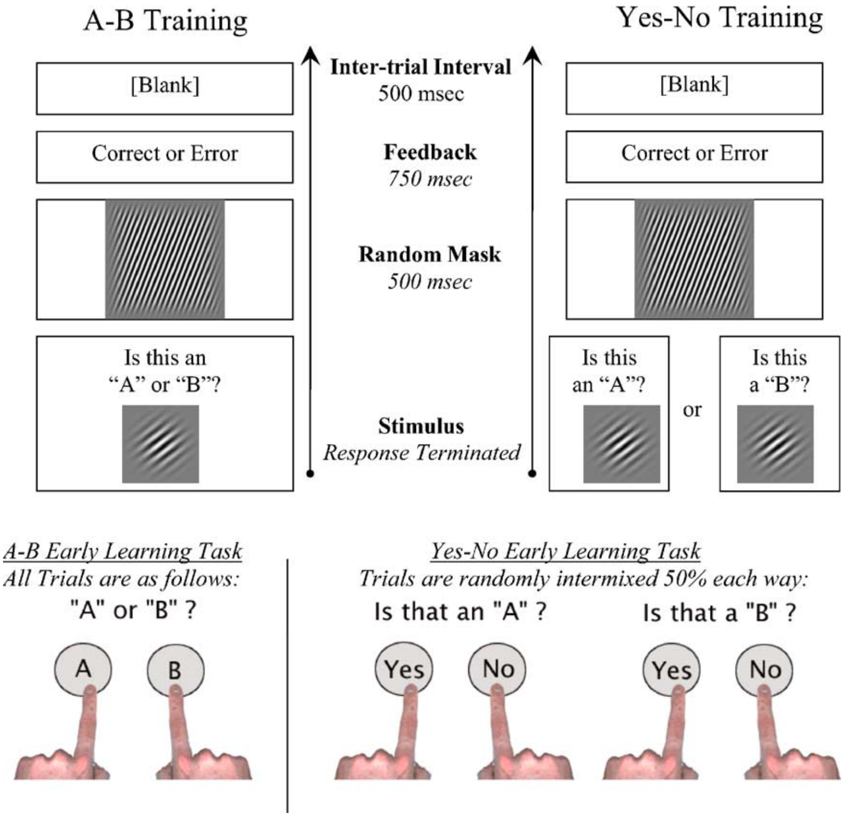


Fig. 13. Basic design for the A–B and yes–no training procedures.

Maddox et al. (2003a,b,c, 2004) examined rule-based and information-integration category learning under two different, but related training procedures (the Gabor patches and lower discriminability rule-based category structures were used in this study). The sequence of events on a single trial is illustrated in the top panel of Fig. 13, and the finger and button placement is illustrated in the bottom panel of Fig. 13. In the A–B training condition, the stimulus was displayed on each trial along with the query "Is this an A or B?" The observer pressed one key for category A and a separate key for category B, followed by corrective feedback. In the yes–no training condition, on half the trials (selected randomly) the stimulus was displayed along with the query "Is this an A?" and on the other half of the trials the stimulus was displayed along with the query "Is this a B?" The observer pressed one key to respond "No", and a separate key to respond "Yes". Notice that in the A–B condition, each category label is associated with a different response location, whereas in the yes–no condition, each category label is not associated with a different response location. Otherwise, the A–B and yes–no training procedures are identical. If rule-based categories are learned via an explicit reasoning process that learns abstract category labels, then performance should not differ across A–B and yes–no training procedures. If information-integration categories are learned via a procedural learning-based process that learns response positions, then learning should be poor in the yes–no conditions, relative to the A–B conditions.

The proportion correct during the final block of trials (averaged across observers) for the rule-based and information-integration category structures under A–B and yes–no training procedures is displayed in Fig. 14a. The models outlined above were also ap-

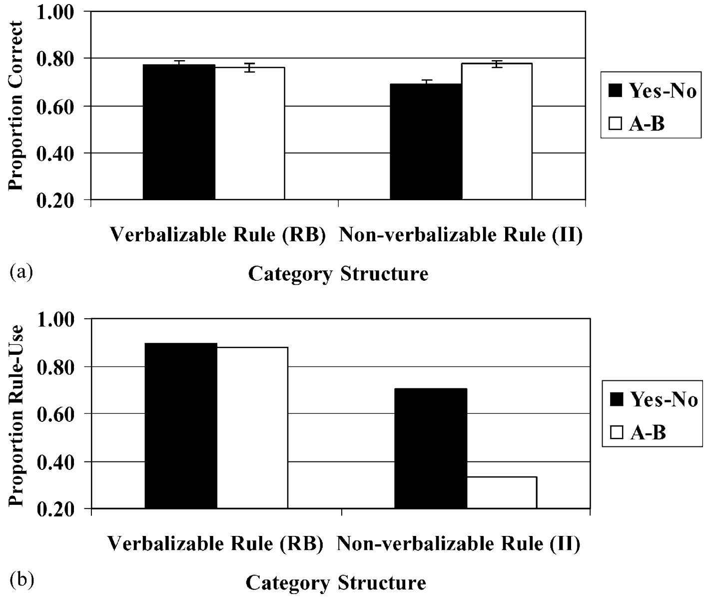


Fig. 14. (a) Proportion correct during the final block of trials for the verbalizable rule (RB) and non-verbalizable rule (II) category structures under A–B and yes–no training procedures (standard error bars included). (b) Proportion rule-use during the final block of trials for these same conditions.

plied to these data, and the proportion of observers who used a rule-based strategy during the final block of trials is shown in Fig. 14b. The pattern of results is similar to those obtained with observational versus feedback training, and delayed versus immediate feedback training. Specifically, yes–no training led to lower accuracy rates and higher rule-use for the information-integration category structures, but had no effect on performance, relative to the A–B training procedure, for the rule-based category structures.

3.2. Brief summary

This collection of five studies provides evidence in support of COVIS and in support of the three a priori predictions regarding deficits in non-verbalizable, information-integration category learning that were outlined above. Collectively, these studies show the importance of the nature and timing of feedback on information-integration learning, and they also establish an intimate link between motor processes and learning of these ill-formed categories. In particular, information-integration category learning is adversely affected by unsupervised, observational, and delayed feedback training, and by situations in which there is no direct correspondence between each category label and a unique response location.

At first glance, one might conclude from these results that the explicit system is a superior system because of its greater flexibility. However, this flexibility comes at a cost. First, the explicit system can only learn a limited class of category structures; namely, those for which the optimal strategy is rule-based and verbalizable. Many perceptual categories do not meet these requirements. Second, learning in the explicit system requires attention, effort, and working memory. On the other hand, the procedural learning system is not constrained to learn rules of any particular type (within reason; Ashby and Waldron, 1999; Ashby et al., 2001), and with the appropriate feedback, learning in this system is essentially automatic. The next section describes a series of studies that exploits some of the weaknesses of the explicit system. The result is a set of dissociations in which rule-based category learning is disrupted more than information-integration category learning.

3.3. Predicted effects on verbalizable, rule-based but not non-verbalizable, information-integration category learning

3.3.1. A priori Prediction 4

A concurrent task that loads on frontal cortex should interfere with rule-based category learning more than with information-integration category learning: The COVIS explicit system depends heavily on working memory and executive attention—processes that are mediated largely within frontal cortex. In contrast, the procedural-learning system is mediated primarily by neural structures outside of frontal cortex (e.g. basal ganglia). As a result, simultaneously performing a second task that recruits frontal cortex should interfere with rule-based category learning more than with information-integration category learning.

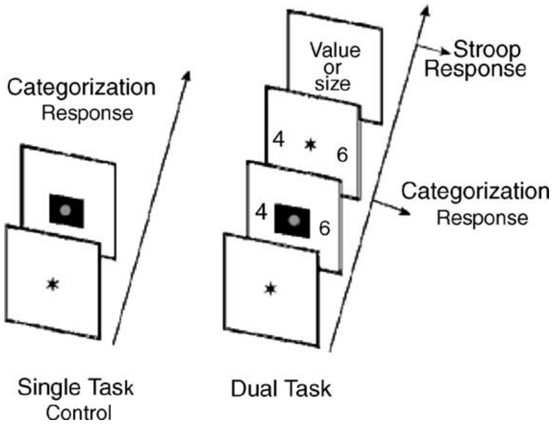


Fig. 15. Basic design for the single and dual task training procedures.

Waldron and Ashby (2001) tested this prediction by training observers on theFig. 1a and b rule-based and information-integration category structures under single-task control or dual-task conditions. The basic design is illustrated in Fig. 15. Each categorization stimulus was flanked by two numbers that varied in physical size and numerical value. For example, in Fig. 15 the physically larger number is 4 and the numerically larger number is 6. The observer first responded A or B to the categorization stimulus and then they were cued to report either the physically larger number or the numerically larger number. This dual task requires working memory and executive attention, and it is closely related to the classic Stroop (1935) task, which is known to activate frontal cortical regions (e.g. Bench et al., 1993). Each observer continued in the task until they achieved a criterion of eight correct responses in a row.

The trials-to-criterion (averaged across observers) for the rule-based and information-integration category structures under the single- and dual-task conditions are displayed in Fig. 16. As predicted, with the rule-based categories, the dual task caused a massive increase in the amount of training needed for category learning (i.e. of about 350%). In contrast, the effect of the dual task on information-integration category learning was modest.

Note that if rule-based and information-integration category structures were learned by the same system, then the dual task would be expected to have the opposite effects. This is because the rule-based categories used by Waldron and Ashby (2001) are simpler

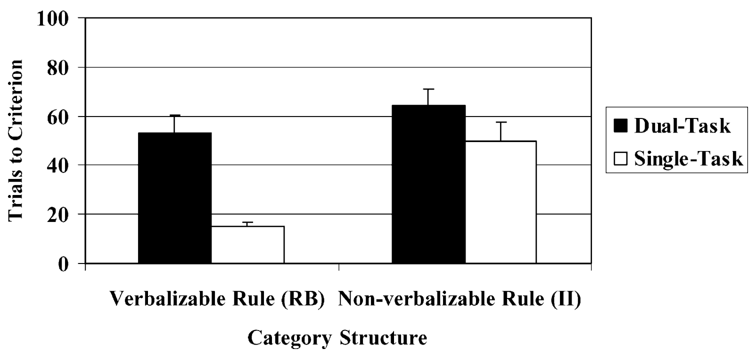


Fig. 16. Trials to criterion for the verbalizable rule (RB) and non-verbalizable rule (II) category structures under single- and dual-task conditions (standard error bars included).

to learn (i.e. under single-task conditions) than the information-integration categories, and in general, it is easier to perform two simple tasks simultaneously than two difficult tasks. In general, this intuition is correct. However, with extreme parameter settings, some single-system models can account for the control data shown in Fig. 16. Arguably, the most successful existing single-process model of category learning is Kruschke's (1992) ALCOVE model. Ashby and Ell (2002) showed that the only versions of ALCOVE that can fit the Waldron and Ashby (2001) data make the strong prediction that after reaching criterion accuracy on the simple rule-based structures, participants should have no idea that only one dimension was relevant in the dual-task conditions. Ashby and Ell reported empirical evidence that strongly disconfirmed this prediction of ALCOVE. Thus, the best available single-system model fails to account even for this one dissociation reported by Waldron and Ashby (2001).

Note also that the Fig. 16 data provide strong evidence against the hypothesis that observers learned the information-integration categories by actively memorizing the category memberships of the individual exemplars. Such a process should place high demands on working memory, and therefore should be extremely susceptible to interference from the dual task. The absence of much interference in the information-integration conditions suggests that observers did not use an active memorization strategy to learn these categories.

3.3.2. A priori Prediction 5

Decreasing the time to process the feedback signal should disrupt rule-based category learning more than information-integration category learning: The COVIS explicit system learns through a conscious process of hypothesis generation and testing. If feedback indicates that the most recent response was incorrect, then the observer must decide whether to apply the same rule (or hypothesis) again, or whether to switch to a new rule. In addition, if the latter decision is made, then a new rule must be selected and attention must be switched from the old rule to the new. These operations require time and attention. In contrast, in the procedural-learning system, learning depends on a reward signal that strengthens the appropriate (stimulus-category) associations in a relatively automatic fashion. Thus, according to COVIS, decreasing the time for feedback processing should disrupt rule-based category learning much more than information-integration category learning.

Maddox et al. (2003a,b,c, 2004) tested this hypothesis by having observers learn the rule-based and information-integration category structures outlined in Fig. 2 under reduced feedback processing or control conditions. The lower discriminability rule-based category structures were used again, along with the Gabor patches as stimuli. The basic design is illustrated in Fig. 17. In both conditions, on each trial the observer first performed a single categorization trial and then one trial of four-item memory scanning (Sternberg, 1966). In the control condition, the observer viewed the categorization stimulus, received 500 ms of feedback, was given a 2500 ms blank-screen delay, and then the memory scanning trial began. In the reduced feedback processing condition, the observer completed the categorization trial, was given 500 ms of feedback, and then was immediately presented with the memory-scanning trial, which was followed by a 2500 ms blank-screen delay. Because processing of the feedback requires attention and effort for the explicit system, but not for the implicit system, we predicted that reduced feedback processing should disrupt learning of the rule-based categories more than learning of the information-integration categories.

The proportion correct during the final block of trials (averaged across observers) for the rule-based and information-integration category structures under reduced feedback processing and control conditions is displayed in Fig. 18. Most importantly, as predicted, rule-based category learning was significantly worse in the reduced feedback processing condition than in the control condition, whereas information-integration category learning was unaffected.

3.3.3. A priori Prediction 6

Increasing the number of decision bounds should disrupt rule-based category learning whereas increasing the number of stimulus clusters should disrupt information-integration category learning: It is important to mention at the outset that this is the weakest of the six a priori predictions reviewed in this article, in the sense that the prediction regarding rule-based category learning is solidly grounded whereas the prediction regarding information-integration category learning is less clear. Because learning in the explicit

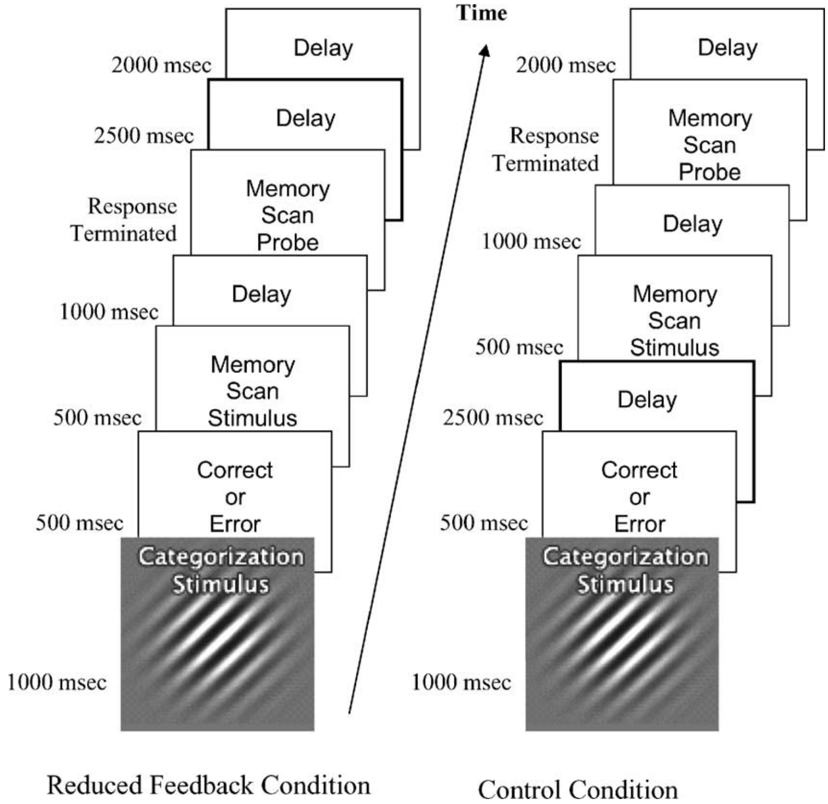


Fig. 17. Basic design for the feedback processing time experiment.

system requires working memory, attention, and effort, and since learning each additional decision bound increases the demand on working memory and attention, it is predicted that increasing the number of rule-based decision bounds should adversely affect rule-based category learning. Learning in the procedural-learning system involves linking each cluster of visual cortical cells with a specific category label. Assuming that each additional cluster-label linkage increases the complexity of the task, it is reasonable to predict that increasing the number of stimulus clusters will adversely affect information-integration category learning.

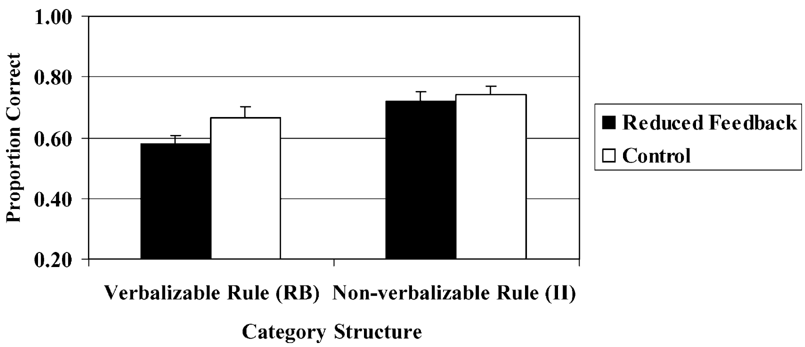


Fig. 18. Proportion correct during the final block of trials for the verbalizable rule (RB) and non-verbalizable rule (II) category structures under reduced feedback processing or and control conditions (standard error bars included).

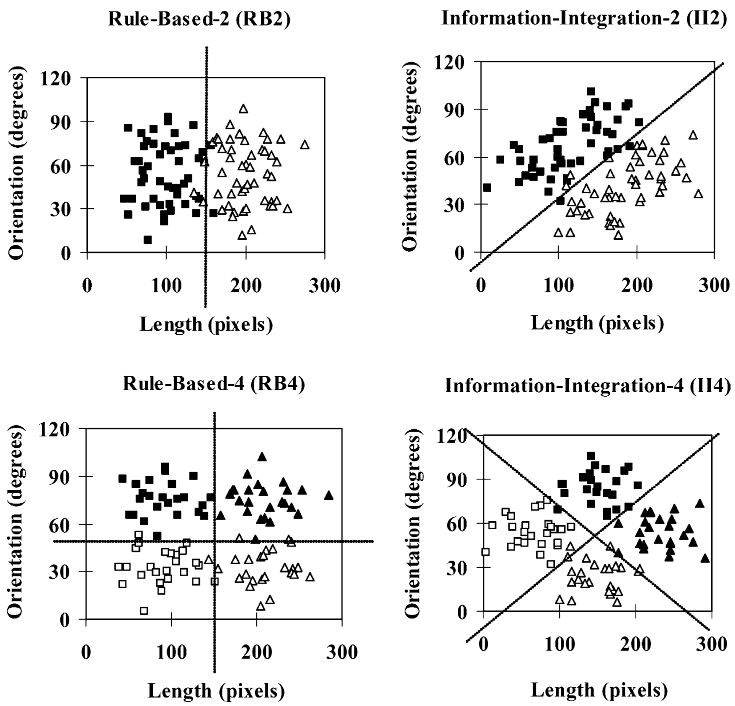


Fig. 19. Scatter-plots in the length and orientation space for the categorization stimuli from each of the four conditions. For the 2-category conditions, the filled squares denote stimuli from category A, and the open triangles denote stimuli from category B. For the four category conditions, the open squares denote stimuli from category A, the filled squares denote stimuli from category B, the open triangles denote stimuli from category C, and the filled triangles denote stimuli from category D. The broken lines in each plot denote the optimal decision bound(s).

Maddox et al. (2003a,b,c, 2004) tested this hypothesis in two experiments. In Experiment 1, observers completed four category learning tasks. Scatterplots of the stimuli used in the four conditions are depicted in Fig. 19. Notice that one decision bound is relevant in the two-category rule-based task, whereas two decision bounds are relevant in the four-category rule-based task. Thus, we predict that learning should be poorer in the four-category than in the two-category rule-based task. Notice also that there are four stimulus clusters in all experimental conditions despite the fact that the number of category labels changes. Thus, we predict that information-integration learning should be unaffected by the shift from two to four categories since the number of clusters remains constant.

The proportion correct during the final block of trials (averaged across observers) for the two- and four-category rule-based and information-integration category structures is displayed in Fig. 20 (the models outlined above were also applied to these data, but a detailed discussion of these findings is beyond the scope of this article). Two interesting results emerged. First, four-category rule-based category learning was significantly worse than two-category rule-based category learning, whereas two- and four-category information-integration learning was equivalent. Second, whereas two-category rule-based category learning was significantly better than two-category information-integration category learning, four-category rule-based category learning was significantly worse than four-category information-integration category learning. Taken together, these results support the prediction that the number of decision bounds affects rule-based category learning, whereas the number of stimulus clusters affects information-integration category learning.

One weakness of Experiment 1 is that the two- and four-category rule-based tasks differ not only in the

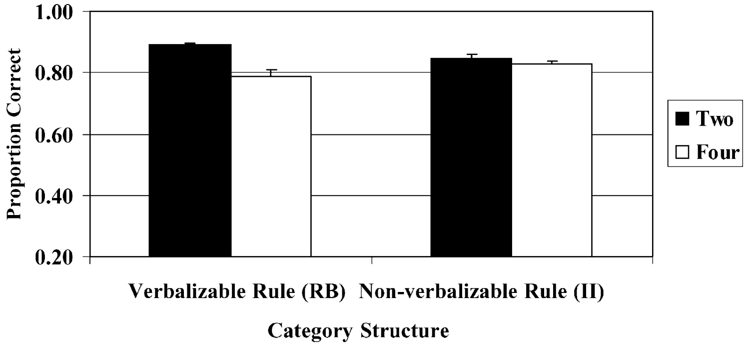


Fig. 20. Proportion correct during the final block of trials for the verbalizable rule (RB) and non-verbalizable rule (II) category structures with two or four categories (standard error bars included).

number of decision bounds, but also in the number of dimensions that are relevant to solving the task. In the two-category rule-based task only the length dimension is relevant, whereas in the four-category rule-based task both length and orientation are relevant. Experiment 2 teased apart these two factors by including a four-category rule task that required three decision bounds on the length dimension (and no decision bounds along orientation), along with the two- and four-category rule-based tasks from Experiment 1. Category learning in the four-category task with three decision bounds on length was intermediate between that in the original two- and four-category rule-based tasks, suggesting that both the number of decision bounds, and the number of relevant dimensions influence rule-based category learning.

3.4. Brief summary

This collection of four studies provides additional evidence in support of COVIS and the three a priori predictions regarding deficits in verbalizable, rule-based category learning. Collectively, these studies show the importance of attention, effort, and working memory on rule-based category learning. Rule-based category learning is adversely affected by the presence of a secondary task known to rely heavily upon frontal brain structures, changes in the number of decision bounds, changes in the number of relevant dimensions, and additional tasks that interfere with feedback processing.

4. General discussion

The article summarizes the results from nine experiments that test six a priori predictions derived from Ashby et al. (1998) COVIS theory of category learning. Importantly, each prediction was derived from a careful examination of the proposed neurobiological underpinnings of the COVIS explicit and implicit systems. The theory assumes an explicit category learning system that dominates the learning of verbalizable, rule-based category structures and is mediated by frontal structures that include the prefrontal cortex, anterior cingulate, and head of the caudate nucleus. The theory also assumes an implicit, procedural-learning system that dominates the learning of information-integration category structures and is mediated by the tail of the caudate nucleus, and depends heavily on a dopamine-mediated reward signal for learning.

The first three a priori predictions follow from the assumption that learning in the procedural-learning system is partly motor based and requires a dopamine-mediated reward signal. Learning in the explicit system, on the other hand, requires executive attention and working memory and is not directly linked to a motor response. In support of these predictions, removing or delaying the feedback, or substituting observational training for feedback training disrupted information-integration category learning but not rule-based category learning (Ashby et al., 1999a,b, 2002; Maddox et al., 2003a,b,c, 2004). In addition, training procedures for which there was not a one-to-one correspondence between each category label and a unique response location (i.e. button-switch: Ashby et al., 2003a; yes–no: Maddox et al., 2003a,b,c, 2004) adversely affected information-integration category learning while leaving rule-based category learning unaffected. In every case, the poor information-integration category learning was accompanied by an increase in the use of rule-based strategies. In short, when the procedural-learning system was not able to operate efficiently, due to the introduction of some sub-optimal training procedure, observers tended to rely on the more flexible rule-based system in an attempt to learn the categories.

The last three a priori predictions follow from the assumptions that learning in the explicit system requires attention and effort, and relies on frontal structures involved in working memory. Learning in the implicit system, on the other hand, is essentially automatic, and does not rely on frontal structures. In support of these predictions, rule-based but not information-integration category learning is disrupted by the presence of a concurrent Stroop-like task, by increases in the number of decision bounds and relevant dimensions, and by reducing the time available to process the feedback (Maddox et al., 2003a,b,c, 2004; Waldron and Ashby, 2001).

Taken together, this large body of research provides strong support for COVIS and for the more general hypothesis that human category learning is mediated by multiple distinct, but partially overlapping systems.

Acknowledgements

This research was supported in part by National Institute of Health Grant R01 MH59196 to WTM, Public Health Service Grant MH3760 to FGA, and a McDonnell-Pew Consortium Grant.

Appendix A. Details of the models

The following models were fit to the Ashby et al. (1999a,b) data, and to several other data sets described in this article (see Ashby, 1992a; Maddox and Ashby, 1993, for a more formal treatment of these models).

A.1. Rule-based models

Two models were compatible with the assumption that observers used an explicit rule-based strategy.

The unidimensional model assumes observers set a criterion on a single perceptual dimension and then make an explicit decision about the level of the stimulus on that dimension (Ashby and Gott, 1988; Shaw, 1982). For example, in the present experiments, observers might use the rule: respond A if the line length is short, and B if it is long. The unidimensional model has two free parameters: a decision criterion on the relevant perceptual dimension and the variance of internal (perceptual and criterial) noise (i.e. σ^2).

The two criterion unidimensional model assumes observers set two criteria on a single perceptual dimension and then assign one categorization response to stimuli falling between the criteria and the other response to stimuli falling outside the criteria. This model has three free parameters: two decision criteria on the relevant perceptual dimension and the variance of internal (perceptual and criterial) noise (i.e. σ^2).

The conjunction model assumes observers use a conjunction rule in which they make separate decisions about the levels on the two dimensions and then select a response based on the outcome of these two decisions. Two conjunction rules were examined:

(1) Respond A if length is short and orientation is large, otherwise respond B, and
(2) Respond B if length is long and orientation is small, otherwise respond A.

This strategy is rule-based because it is easy to describe verbally and it does not require perceptual integration of length and orientation. Conjunction models have three parameters (a criterion on each dimension, and σ^2).

A.2. Information-integration models

The general linear classifier (GLC) assumes that the decision bound between each pair of categories is linear. This produces an information-integration decision strategy because it requires linear integration of perceived length and orientation. The GLC has three parameters (slope and intercept of the linear bound and σ^2).

A.3. Model fits

Each of these models was fit separately to the data from the final block of trials for every observer. The model parameters were estimated using maximum likelihood (Ashby, 1992b; Wickens, 1982) and the goodness-of-fit statistic was

$$\text{AIC} = 2r - 2lnL,$$

where r is the number of free parameters and L is the likelihood of the model given the data (Akaike, 1974; Takane and Shibayama, 1992). The AIC statistic penalizes a model for extra free parameters in such a way that the smaller the AIC, the closer a model is to the "true model," regardless of the number of free parameters. Thus, to find the best model among a given set of competitors, one simply computes an AIC value for each model, and chooses the model associated with the smallest AIC value.

References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Auto. Control 19, 716–723.

Alexander, G.E., DeLong, M.R., Strick, P.L., 1986. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Ann. Rev. Neurosci. 9, 357–381.

Arbuthnott, G.W., Ingham, C.A., Wickens, J.R., 2000. Dopamine and synaptic plasticity in the neostriatum. J. Anat. 196, 587–596.

Ashby, F.G., 1992a. Multidimensional models of categorization. In: Ashby, F.G. (Ed.), Multidimensional Models of Perception and Cognition. Erlbaum, Hillsdale, NJ.

Ashby, F.G., 1992b. Multivariate probability distributions. In: Ashby, F.G. (Ed.), Multidimensional Models of Perception and Cognition. Erlbaum, Hillsdale, NJ, pp. 1–34.

Ashby, F.G., Gott, R.E., 1988. Decision rules in the perception and categorization of multidimensional stimuli. J. Exp. Psychol. Learn. Mem. Cogn. 14, 33–53.

Ashby, F.G., Waldron, E.M., 1999. The nature of implicit categorization. Psychol. Bull. Rev. 6, 363–378.

Ashby, F.G., Ell, S.W., 2001. The neurobiological basis of category learning. Trends Cogn. Sci. 5, 204–210.

Ashby, F.G., Casale, M.B., 2002. The cognitive neuroscience of implicit category learning. In: Jimenez, L. (Ed.), Attention and Implicit Learning. John Benjamins Publishing Company, Amsterdam.

Ashby, F.G., Ell, S.W., 2002. Single versus multiple systems of learning and memory. In: Wixted, J., Pashler, H. (Eds.), Stevens' Handbook of Experimental Psychology, Methodology in Experimental Psychology, vol. 4, third ed. Wiley, New York.

Ashby, F.G., Alfonso-Reese, L.A., Turken, A.U., Waldron, E.M., 1998. A neuropsychological theory of multiple systems in category learning. Psychol. Rev. 105, 442–481.

Ashby, F.G., Isen, A.M., Turken, A.U., 1999a. A neuropsychological theory of positive affect and its influence on cognition. Psychol. Rev. 106, 529–550.

Ashby, F.G., Queller, S., Berretty, P.T., 1999b. On the dominance of unidimensional rules in unsupervised categorization. Percept. Psychophys. 61, 1178–1199.

Ashby, F.G., Waldron, E.M., Lee, W.W., Berkman, A., 2001. Suboptimality in human categorization and identification. J. Exp. Psychol. Gen. 130, 77–96.

Ashby, F.G., Maddox, W.T., Bohil, C.J., 2002. Observational versus feedback training in rule-based and information-integration category learning. Mem. Cogn. 30, 666–677.

Ashby, F.G., Ell, S.W., Waldron, E.M., 2003a. Procedural learning in perceptual categorization, in press.

Ashby, F.G., Noble, S., Filoteo, J.V., Waldron, E.M., Ell, S.W., 2003b. Category learning deficits in Parkinson's Disease. Neuropsychology 17, 115–124.

Ataly, J., Wise, R.A., 1983. Time course of pimozide effects on brain stimulation reward. Pharm. Biochem. Behav. 18, 655–658.

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

Beninger, R.J., 1983. The role of dopamine in locomotor activity and learning. Brain Res. 287, 173–196.

Bozarth, M.A., 1994. Opiate reinforcement processes: reassembling multiple mechanisms. Addiction 89, 1425–1434.

Brown, R.G., Marsden, C.D., 1988. Internal versus external cures and the control of attention in Parkinson's disease. Brain 111, 323–345.

Calabresi, P., Pisani, A., Centonze, D., Bernardi, G., 1996. Role of Ca^{2+} in striatal LTD and LTP. Sem. Neurosci. 8, 321–328.

Cools, A.R., van den Bercken, J.H., Horstink, M.W., van Spaendonck, K.P., Berger, H.J., 1984. Cognitive and motor shifting aptitude disorder in Parkinson's disease. J. Neur. Neurosurg. Psychol. 47, 443–453.

Cooper, J.R., Bloom, F.E., Roth, R.H., 1991. The Biochemical Basis of Neuropharmacology, sixth ed. Oxford, New York.

DiFiglia, M., Pasik, T., Pasik, P., 1978. A Golgi study of afferent fibers in the neostriatum of monkeys. Brain Res. 152, 341–347.

Elliott, R., Dolan, R.J., 1998. Activation of different anterior cingulate foci in association with hypothesis testing and response selection. Neuroimage 8, 17–29.

Erickson, M.A., Kruschke, J.K., 1998. Rules and exemplars in category learning. J. Exp. Psychol. Gen. 127, 107–140.

Filoteo, J.V., Maddox, W.T., Davis, J.D., 2001a. A possible role of the striatum in linear and nonlinear category learning: evidence from patients with Huntington's disease. Behav. Neurosci. 115, 786–798.

Filoteo, J.V., Maddox, T.W., Davis, J.D., 2001b. Quantitative modeling of category learning in amnesic patients. J. Int. Neuropsychol. Soc. 7, 1–19.

Freund, T.F., Powell, J.F., Smith, A.D., 1984. Tyrosine hydroxylase-immunoreactive boutons in synaptic contact with identified striatonigral neurons, with particular reference to dendritic spines. Neuroscience 13, 1189–1215.

Fuster, J.M., 1989. The Prefrontal Cortex, second ed. Lippincott-Raven, Philadelphia.

Gamble, E., Koch, C., 1987. The dynamics of free calcium in dendritic spines in response to repetitive synaptic input. Science 236, 1311–1315.

Goldman-Rakic, P.S., 1987. Circuitry of the prefrontal cortex and the regulation of behavior by representational knowledge. In: Plum, F., Mountcastle, V. (Eds.), Handbook of Physiology. American Physiological Society, pp. 373–417.

Hazeltine, E., Ivry, R.B., 2001. Motor skill. In: Ramachandran, V.S. (Ed.), Encyclopedia of the Brain. Academic Press, San Diego.

Hollerman, J.R., Schultz, W., 1997. Dopamine neurons report an error in the temporal prediction of reward during learning. Nat. Neurosci. 1, 304–308.

Huang, Y.Y., Kandel, E.R., 1995. D1/D5 receptor agonists induce a protein synthesis-dependent late potentiation in the CA1 region of the hippocampus. Proc. Natl. Acad. Sci. U.S.A. 92, 2446–2450.

Janowsky, J.S., Kritchevsky, A.P., Squire, L.R., 1989. Cognitive impairment following frontal lobe damage and its relevance to human amnesia. Behav. Neurosci. 103, 548–560.

Kolb, B., Whishaw, I.Q., 1990. Fundamentals of Human Neuropsychology, third ed. Freeman, New York.

Leng, N.R., Parkin, A.J., 1988. Double dissociation of frontal dysfunction in organic amnesia. Brit. J. Clin. Psychol. 27, 359–362.

Lynch, G., Larson, J., Kelso, S., Barrionuevo, G., Schottler, F., 1983. Intracellular injections of EGTA block induction of hippocampal long-term potentiation. Nature 305, 719–721.

MacDermott, A.B., Mayer, M.L., Westbrook, G.L., Smith, S.J., Barker, J.L., 1986. NMDA-receptor activation increases cytoplasmic calcium concentration in cultured spinal cord neurons. Nature 321, 519–522 (Erratum in: Nature, 321: 888).

MacDermott, A.B., Mayer, M.L., Westbrook, G.L., Smith, S.J., Barker, J.L., 1986. NMDA-receptor activation increases cytoplasmic calcium concentration in cultured spinal cord neurones. Nature 321, 519–522.

Maddox, W.T., Ashby, F.G., 1993. Comparing decision bound and exemplar models of categorization. Percept. Psychophys. 53, 49–70.

Maddox, W.T., Filoteo, J.V., 2001. Striatal contributions to category learning: quantitative modeling of simple linear and complex nonlinear rule learning in patients with Parkinson's disease. J. Int. Neuropsychol. Soc. 7, 710–727.

Maddox, W.T., Ashby, F.G., Ing, A.D., Pickering, A.D., 2003a. Disrupting feedback processing interferes with rule-based, but not information-integration category learning, in press.

Maddox, W.T., Ashby, F.G., Bohil, C.J., 2003b. Delayed feedback effects on rule-based and information-integration category learning. J. Exp. Psychol. Learn. Mem. Cogn. 29, 650–662.

Maddox, W.T., Bohil, C.J., Ing, A.D., 2003c. Evidence for a procedural-learning based system in perceptual category learning. Psychol. Bull. Rev., in press.

Maddox, W.T., Filoteo, J.V., Hejl, K.D., Ing, A.D., 2004. Category number impacts rule-based but not information-integration category learning: further evidence for dissociable category learning systems. J. Exp. Psychol. Learn. Mem. Cogn. 30, 227–235.

McDonald, R.J., White, N.M., 1993. A triple dissociation of memory systems: hippocampus, amygdala, and dorsal striatum. Behav. Neurosci. 107, 3–22.

McDonald, R.J., White, N.M., 1994. Parallel information processing in the water maze: evidence for independent memory systems involving dorsal striatum and hippocampus. Behav. Neur. Biol. 61, 260–270.

Miller, J.D., Sanghera, M.K., German, D.C., 1981. Mesencephalic dopaminergic unit activity in the behaviorally conditioned rat. Life Sci. 29, 1255–1263.

Mishkin, M., Malamut, B., Bachevalier, J., 1984. Memories and habits: two neural systems. Neurobiology of Human Learning and Memory. Guilford, New York, pp. 65–77.

Montague, P.R., Dayan, P., Sejnowski, T.J., 1996. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. J. Neurosci. 16, 1936–1947.

Nairn, A.C., Hemmings Jr., H.C., Walaas, S.I., Greengard, P., 1988. DARPP-32 and phosphatase inhibitor-1, two structurally related inhibitors of protein phosphatase-1, are both present in striatonigral neurons. J. Neurochem. 50, 257–262.

Nosofsky, R.M., Johansen, M.K., 2000. Exemplar-based accounts of "multiple-system" phenomena in perceptual categorization. Psychol. Bull. Rev.

Packard, M.G., McGaugh, J.L., 1992. Double dissociation of fornix and caudate nucleus lesions on acquisition of two water maze tasks: further evidence for multiple memory systems. Behav. Neurosci. 106, 439–446.

Packard, A.B., Srivastava, S.C., Richards, P., Meinken, G.E., Ford, L., Benson, W.R., 1989. Synthesis and biological properties of the lipophilic technetium-99m complex 99mTcacac3. Int. J. Rad. Appl. Instrum. B 16, 291–294.

Pessin, M.S., Snyder, G.L., Halpain, S., Giraut, J.-A., Aperia, A., Greengard, P., 1994. DARPP-32/protein phosphatase-1/Na+/K+ ATPase system: a mechanism for bi-directional control of cell function. Trophic Regulation of the Basal Ganglia. Elsevier Science, New York.

Pfaus, J.G., Phillips, A.G., 1991. Role of dopamine in anticipatory and consummatory aspects of sexual behavior in the male rat. Behav. Neurosci. 105, 727–743.

Phillips, A.G., Blaha, C.D., Pfaus, J.G., Blackburn, J.R., 1992. Neurobiological correlates of positive emotional states: dopamine, anticipation and reward. International Review of Studies on Emotion. Wiley, New York.

Pickering, A.D., 1997. New approaches to study of amnesic patients: what can a neurofunctional philosophy and neural network methods offer? Memory 5, 255–300.

Poldrack, R.A., Prabhakaran, V., Seger, C.A., Gabrieli, J.D.E., 1999. Striatal activation during acquisition of a cognitive skill. Neuropsychology 13, 564–574.

Posner, M.I., Petersen, S.E., 1990. Attention systems in the human brain. Ann. Rev. Neurosci. 13, 25–42.

Rao, S.M., Bobholz, J.A., Hammeke, T.A., Rosen, A.C., Woodley, S.J., Cunningham, J.M., Cox, R.W., Stein, E.A., Binder, J.R., 1997. Functional MRI evidence for subcortical participation in conceptual reasoning skills. Neuroreport 27, 1987–1993.

Reber, P.J., Squire, L.R., 1994. Parallel brain systems for learning with and without awareness. Learn Mem. 1, 217–229.

Reber, P.J., Stark, C.E.L., Squire, L.R., 1998. Cortical areas supporting category learning identified using functional magnetic resonance imaging. Proc. Natl. Acad. Sci. U.S.A. 95, 747–750.

Robinson, A.L., Heaton, R.K., Lehman, R.A.W., Stilson, D.W., 1980. The utility of the Wisconsin Card Sorting Test in detecting and localizing frontal lobe lesions. J. Consult. Clin. Psychol. 48, 605–614.

Schacter, D.L., 1987. Implicit memory: history and current status. J. Exp. Psychol. Learn. Mem. Cogn. 13, 501–518.

Schultz, W., 1992. Activity of dopamine neurons in the behaving primate. Sem. Neurosci. 4, 129–138.

Shaw, M.L., 1982. Attending to multiple sources of information. I. The integration of information in decision making. Cogn. Psychol. 14, 353–409.

Smiley, J.F., Levey, A.I., Ciliax, B.J., Goldman-Rakic, P.S., 1994. D1 dopamine receptor immunoreactivity in human and monkey cerebral cortex: predominant and extrasynaptic localization in dendritic spines. Proc. Natl. Acad. Sci. U.S.A. 91, 5720–5724.

Smith, E.E., Patalano, A., Jonides, J., 1998. Alternative strategies of categorization. Cog. 65, 167–196.

Squire, L.R., 1992. Memory and the hippocampus: a synthesis from findings with rats, monkeys and humans. Psychol. Rev. 99, 195–231.

Sternberg, S., 1966. High-speed scanning in human memory. Science 153, 652–654.

Stroop, R.J., 1935. Studies of interference in serial verbal reactions. J. Exp. Psychol. 18, 643–662.

Takane, Y., Shibayama, T., 1992. Structures in stimulus identification data. In: Ashby, F.G. (Eds.), Multidimensional Models of Perception and Cognition. Erlbaum, Hillsdale, NJ, pp. 335–362.

van Domburg, P.H.M.F., ten Donkelaar, H.J., 1991. The Human Substantia Nigra and Ventral Tegmental Area. Springer-Verlag, Berlin.

Waldron, E.M., Ashby, F.G., 2001. The effects of concurrent task interference on category learning. Psychol. Bull. Rev. 8, 168–176.

White, N.M., 1989. A functional hypothesis concerning the striatal matrix and patches: mediation of S-R memory and reward. Life Sci. 45, 1943–1957.

Wickens, T.D., 1982. Models for Behavior: Stochastic Processes in Psychology. Freeman, San Francisco.

Wickens, J., 1990. Striatal dopamine in motor activation and reward-mediated learning: steps towards a unifying model. J. Neur. Trans. Gen. Sect. 80, 9–31.

Wickens, J., 1993. A Theory of the Striatum. Pergamon Press, New York.

Willingham, D.B., 1998. A neuropsychological theory of motor skill learning. Psychol. Rev. 105, 558–584.

Willingham, D.B., Nissen, M.J., Bullemer, P., 1989. On the development of procedural knowledge. J. Exp. Psychol. Learn. Mem. Cogn. 15, 1047–1060.

Wilson, C. J., 1995. The Contribution of Cortical Neurons to the Firing Pattern of Striatal Spiny Neurons. MIT Press, Cambridge, MA, pp. 29–50.

Winocur, G., Eskes, G., 1998. Prefrontal cortex and caudate nucleus in conditional associative learning: dissociated effects of selective brain lesions in rats. Behav. Neurosci. 112, 89–101.