The Neurobiology of Category Learning

F. Gregory Ashby
Brian J. Spiering
University of California, Santa Barbara

Many recent studies have examined the neural basis of category learning. Behavioral neuroscience results suggest that both the prefrontal cortex and the basal ganglia play important category-learning roles; neurons that develop category-specific firing properties are found in both regions, and lesions to both areas cause category-learning deficits. Similar studies indicate that the inferotemporal cortex does not mediate the learning of new categories. The cognitive neuroscience literature on category learning appears contradictory until the results are partitioned according to the type of category-learning task that was used. Three major tasks can be identified: rule based, information-integration, and prototype-distortion. Recent results are consistent with the hypotheses that (a) learning in rule-based tasks requires working memory and executive attention and is mediated by frontal-striatal circuits, (b) learning in information-integration tasks requires procedural memory and is mediated primarily within the basal ganglia, and (c) learning in prototype-distortion tasks depends on multiple memory systems, including the perceptual representation system.

Key Words: working memory, procedural learning, episodic memory, perceptual representation memory system, multiple category learning systems

Categorization is the act of responding differently to objects and events that belong to separate classes or categories. It is a critically important skill that all organisms must possess because it allows them to respond differently, for example, to nutrients and poisons and to predators and prey. Given its importance, it is not surprising that there is the huge and old literature on the perceptual and cognitive processes that mediate categorization. However, for many years, virtually no attention was paid to the neural basis of category learning. As in many other disciplines, however, the categorization community has embraced the cognitive neuroscience revolution. As a result, the past decade has seen an explosion of new results that collectively are beginning to paint a detailed picture of the neural mechanisms and pathways that mediate category learning. These results come from a wide variety of sources, including traditional animal lesion and single-cell recording studies, as well as the cognitive neuroscience fields of neuroimaging and neuropsychology. This article reviews that literature. We begin with behavioral neuroscience studies on nonhuman animals, and then we review cognitive neuroscience studies on human category learning. Finally, we discuss theoretical implications of these many results.

Before beginning, it is important to be explicit about what we will not be reviewing. First, our focus on learning prevents us from considering the categorization behavior of highly experienced experts. This distinction is important because there is good evidence that the neural mechanisms and pathways that mediate the learning of new categories are different from the neural structures that mediate the representation of highly learned categories. For example, many neuropsychological groups that are impaired in category learning (e.g., frontal patients and Parkinson's disease patients) do not lose old, familiar categories (e.g., fruits and tools). Similarly, there is no evidence that people who lose a familiar category (i.e., who develop a category-specific agnosia) develop any general category-learning deficits. Readers interested in the representation of highly learned categories are referred to any of several excellent reviews of the category representation literature (e.g., Cree & McRae, 2003; Humphreys & Forde, 2001; Joseph, 2001).

Second, because our focus is on the neural basis of category learning, we mostly neglect the voluminous cognitive literature on categorization. The major excep-

---

Authors' Note: This research was supported in part by Public Health Service Grant MH3760. Correspondence concerning this article should be addressed to F. Gregory Ashby, Department of Psychology, University of California, Santa Barbara, CA 93106; e-mail: ashby@psych.ucsb.edu

Behavioral and Cognitive Neuroscience Reviews
Volume 3 Number 2 , June 2004 101-113
DOI: 10.1177/1534582304270782

tion will be those (relatively few) cognitive studies that inform us about the underlying neurobiology. Interested readers are referred to any of a number of recent reviews of the cognitive literature (Ashby & Maddox, in press; Barsalou, 2003; Markman & Ross, 2003; Murphy, 2002).

BEHAVIORAL NEUROSCIENCE STUDIES

Most of the research on category learning in nonhumans has involved either single-cell recordings in monkeys or lesion studies in rats. The most natural way to organize this work is by the brain area that was the main target of the research. The three most widely studied areas have been the inferotemporal cortex (IT), the prefrontal cortex (PFC), and the basal ganglia.

IT

The IT is the terminal sensory association region on the so-called ventral (or "what") pathway out of the primary visual cortex (Ungerleider & Mishkin, 1982). A huge literature implicates IT in the high-level representation and processing of visual objects (for reviews, see Bullier, 2002; Tanaka, 1996). Interest in IT as a possible neural locus of category learning was sparked by reports of a variety of category-specific agnosias that result from lesions in IT and other high-level visual areas. Category-specific agnosia refers to the ability to perceive or categorize most visual stimuli normally but a reduced ability to recognize exemplars from some specific category, such as inanimate objects (e.g., tools or fruits). The most widely known of such deficits, which occur with human faces (i.e., prosopagnosia), are associated with lesions to the fusiform gyrus in IT.

Of course, a category-specific agnosia that results from an IT lesion does not logically imply that category learning occurs in IT. For example, although such agnosias are consistent with the hypothesis that category learning occurs in IT, they are also generally consistent with the hypothesis that visually similar objects are represented in nearby areas of visual cortex. In particular, it is well known that neighboring cells in IT tend to fire to similar stimuli. Thus, damage to some contiguous region of IT (or any other visual cortical area) is likely to lead to perception deficits within a class of similar stimuli.

In fact, there is now strong evidence that IT is not a site of category learning. For example, Rolls, Judge, and Sanghera (1977) recorded from cells in IT of monkeys. In these experiments, one visual stimulus was associated with reward and one with a mildly aversive taste. After training, the rewards were switched. Thus, in effect, the animals were taught two simple categories (i.e., "good" and "bad"), and then the category assignments were switched. If the categories were represented in the visual cortex, then the firing properties of visual cortical cells should have changed when the category memberships were switched. However, Rolls et al. found no change in the response of any of these cortical cells, although other similar studies have found changes in the responses of cells in other brain areas (e.g., orbitofrontal cortex).

More recent studies have found similar null results with more traditional categorization tasks (Freedman, Riesenhuber, Poggio, & Miller, 2003; Op de Beeck, Wagemans, & Vogels, 2001; Sigala, 2004; Thomas, Van Hulle, & Vogels, 2001; Vogels, 1999). In each of these studies, monkeys were taught to classify visual objects into one of two categories (e.g., tree versus nontree, two categories of arbitrary complex shapes). Single-cell recordings showed that the firing properties of IT cells did not change with learning. In particular, IT cells showed sensitivity to specific visual images, but category training did not make them more likely to respond to other stimuli in the same category or less likely to respond to stimuli belonging to the contrasting category.

On the other hand, under certain conditions, categorization training can change the firing properties of IT cells. Sigala and Logothetis (2002) trained two monkeys to classify faces into one of two categories and then in a separate condition to classify fish. In both conditions, some stimulus features were diagnostic and some were irrelevant to the categorization response. After categorization training, many neurons in IT showed enhanced sensitivity to the diagnostic features compared to the irrelevant features. Similar results were reported by Sigala (2004). Such changes are consistent with the widely held view that category learning is often associated with changes in the allocation of perceptual attention (Nosofsky, 1986).

In summary, the best evidence suggests that IT does not mediate the learning of new categories. It is crucial to the categorization process, however, because it appears to encode the highest level representation of the visual stimulus. Thus, IT could be seen as the terminal stage of the perceptual system. According to this view, our search for the neural locus of category learning should begin with those structures receiving direct projections from IT. There are three obvious candidates: the PFC, the medial temporal lobes (i.e., the hippocampal system), and the basal ganglia. As it happens, few behavioral neuroscience studies of category learning have targeted the medial temporal lobes (however, see Hampson, Pons, Stanford, & Deadwyler, 2004), but many relevant studies have examined the PFC and the basal ganglia.

PFC

The first evidence that the PFC is important for category learning came many years ago when Milner (1963) showed that patients with frontal lesions are impaired in the Wisconsin Card Sorting Test (WCST), a classic neuropsychological assessment that requires participants to learn a series of simple categories (Heaton, 1981). In fact, deficits on the WCST are perhaps the classic symptom of frontal lobe pathology (e.g., Kimberg, D'Esposito, & Farah, 1997). The critical role played by the PFC in the WCST has been more recently verified in controlled lesion studies in rats (Joel, Weiner, & Feldon, 1997).

Single-unit recording studies in the PFC have demonstrated two different phenomena that are particularly relevant to category learning. First, one set of studies reported neurons in the lateral PFC that show category-specific responding. For example, Freedman et al. (Freedman et al., 2001; Freedman, Riesenhuber, Poggio, & Miller, 2002, 2003) taught monkeys to categorize computer-generated visual images as dogs or cats. The images included prototypes of each category and many ambiguous morphs. The animals were trained to a high level of accuracy, and then single-unit recordings were collected in the lateral PFC region receiving direct input from the IT. Many neurons were found that seemed to respond to the category membership of the visual stimulus; that is, each of these cells responded almost equally to almost all members of one category and showed little or no response to members of the contrasting category. Importantly, the firing properties of these cells were better predicted by category membership than by visual similarity.

A second series of studies established that the PFC plays an important role in learning and applying categorization rules. In a typical study, monkeys are taught to classify objects by applying either one rule (e.g., spatial) or another (e.g., associative). Each trial begins with a cue that signals the animal which rule to use with the ensuing stimulus. Several studies using a paradigm of this type have reported that many PFC cells show rule-specific activity; that is, they fire during application of one of the rules (but not during the other), regardless of which stimulus is shown (Asaad, Rainer, & Miller, 2000; Hoshi, Shima, & Tanji, 1998; I. M. White & Wise, 1999). Many lesion studies have verified the important role played by the PFC in rule learning and use (for a review, see Roberts & Wallis, 2000).

Basal Ganglia

Early interest in the basal ganglia as a possible locus of category learning came from a long series of lesion studies in rats and monkeys showing that the tail of the caudate nucleus is both necessary and sufficient for visual discrimination learning. Many studies have shown that lesions of the tail of the caudate nucleus impair the ability of animals to learn visual discriminations that require one response to one stimulus and a different response to some other stimulus (e.g., McDonald & White, 1993, 1994; Packard, Hirsch, & White, 1989; Packard & McGaugh, 1992). For example, in one study, rats with lesions in the tail of the caudate nucleus 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 & McGaugh, 1992). The same animals learned normally, however, when the cues signaling which platform was safe were spatial. Because the visual cortex is intact in these animals, it is unlikely that their difficulty is in perceiving the stimuli. Rather, it appears that their difficulty is in learning to associate an appropriate response with each stimulus alternative, and in fact, many researchers have hypothesized that this is the primary role of the neostriatum (e.g., Rolls, 1994; Wickens, 1993). Technically, such studies are categorization tasks with one exemplar per category. It is difficult to imagine how adding more exemplars to each category could alleviate the deficits caused by caudate lesions, and it is for this reason that the caudate lesion studies support the hypothesis that the caudate contributes to normal category learning.

The sufficiency of the caudate nucleus for visual discrimination learning was shown in a series of studies by Gaffan and colleagues that lesioned all pathways out of visual cortex except into the tail of the caudate (e.g., projections into the PFC were lesioned by Eacott & Gaffan, 1991, and Gaffan & Eacott, 1995; projections to the hippocampus and amygdala were lesioned by Gaffan & Harrison, 1987). None of these lesions affected visual discrimination learning.

More recent single-unit recording studies confirm the neostriatal contribution to categorization. In a series of studies, Romo and his colleagues taught monkeys to classify a vibrotactile stimulus (i.e., a rod vibrating against the monkey's finger) as either "low speed" or "high speed" (Merchant, Zainos, Hernandez, Salinas, & Romo, 1997; Romo, Merchant, Ruiz, Crespo, & Zainos, 1995). A large number of cells in the putamen showed learning-related changes in their firing properties. For example, after training, many cells fired to any vibrational frequency in the low-speed category but not to frequencies in the high-speed category (or vice versa). These same neurons were not active during passive experience with the stimuli or during a control motor task. Furthermore, the activity of these neurons predicted the behavior of the monkeys.

Thus, many studies implicate the neostriatum in discrimination and category learning. One feature that makes the neostriatum an attractive candidate for such learning is the dopamine input it receives from the substantia nigra (pars compacta), which is widely thought to provide a reward signal that is critical for reward-mediated learning (e.g., Beninger, 1983; Miller, Sanghera, & German, 1981; Montague, Dayan, & Sejnowski, 1996; N. M. White, 1989; Wickens, 1993).

A separate line of research implicates the dorsal striatum (e.g., head of the caudate nucleus) in another skill that could be important in many forms of category learning, namely, task and rule switching. For example, injections of a glutamate agonist directly into the striatum increase the frequency with which cats switch from one motor activity to another in a task where food rewards are delivered for such switching behaviors (Jaspers, de Vries, & Cools, 1990a, 1990b). More relevant to category learning is a report that lesioning the dopamine fibers that project from the ventral tegmental area into the PFC improves the performance of monkeys in an analogue of the WCST, even though it impairs their spatial working memory (Roberts et al., 1994). If switching occurs in the PFC, then such lesions should impair performance (as seen, e.g., in Parkinson's disease patients). If the switching occurs in the basal ganglia, then one's first thought might be that lesioning dopamine fibers into the PFC should have no direct effect on switching. However, it turns out that such lesions increase dopamine levels in the basal ganglia (Roberts et al., 1994). Therefore, if the basal ganglia are responsible for switching, and if switching is enhanced by dopamine, then lesioning dopamine fibers into the PFC should improve switching, which is exactly what Roberts et al. (1994) found.

COGNITIVE NEUROSCIENCE STUDIES

Three Different Category Learning Tasks

On first examination, the cognitive neuroscience literature on category learning appears confusing and even contradictory. For example, various published articles have reported that the category-learning ability of Parkinson's disease patients is either normal, mildly impaired, or profoundly impaired. It turns out, however, that much of this confusion disappears when these reports are partitioned according to the type of category-learning task that was used. Although many possible classifications are possible, an especially useful scheme identifies three major types of category-learning task: rule-based tasks, information-integration tasks, and prototype-distortion tasks.

Rule-based tasks are those in which the categories can be learned via some explicit reasoning process. Frequently, the rule that maximizes accuracy (i.e., the optimal strategy) is easy to describe verbally. In one common application, only one stimulus dimension is relevant. Even so, there is no requirement that the rule that maximizes accuracy (i.e., the optimal rule) in rule-based tasks is one dimensional. For example, a conjunction rule (e.g., respond A if the stimulus is small on dimension x and small on dimension y) is a rule-based task because it is easy to describe verbally. The WCST and all other category-learning tests commonly used in neuropsychological assessment use rule-based tasks.

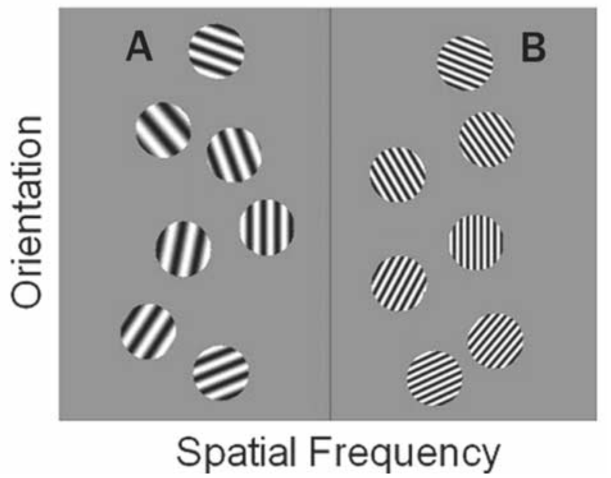

Figure 1: Stimuli That Might Be Used in a Rule-Based Category-Learning Task.

Some stimuli that might be used in a rule-based task are shown in Figure 1. In this case, two contrasting categories are each composed of circular sine-wave gratings (i.e., disks in which luminance varies sinusoidally). The disks are all of equal diameter, but they differ in spatial frequency (i.e., the frequency of the sine wave) and sine-wave orientation. The task is rule based because a simple verbal rule partitions the disks belonging to Categories A and B (i.e., disks in Category A have wide bars, whereas disks in Category B have narrow bars).

Information-integration tasks are those in which accuracy is maximized only if information from two or more stimulus components (or dimensions) is integrated at some predecisional stage (Ashby & Gott, 1988). Perceptual integration could take many forms: from computing a weighted linear combination of the dimensional values to treating the stimulus as a Gestalt. In many cases, the optimal strategy in information-integration tasks is difficult or impossible to describe verbally. Healthy young adults eventually learn to respond accurately in rule-based tasks, but afterward, they are poor at describing their decision strategy (Ashby et al., 1998; Ashby & Maddox, 1992). Real-world examples of infor-

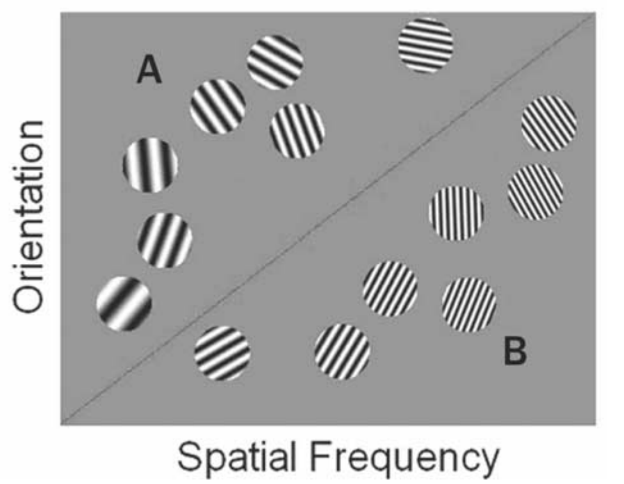

Figure 2: **Stimuli That Might Be Used in an Information-Integration Category-Learning Task.**

mation-integration tasks are common. For example, deciding whether an X ray shows a tumor requires years of training, yet expert radiologists are only partially successful at describing their categorization strategies.

An example of stimuli that might be used in an information-integration task is shown in Figure 2. In this case, the same sine-wave gratings are used as in the Figure 1 rule-based task, but the disks are assigned to the two categories in such a way that there is no simple verbal description of the optimal decision bound (depicted by the diagonal line in Figure 2).

Prototype-distortion tasks are a third type of category-learning task in which each category is created by first constructing a category prototype (Posner & Keele, 1968, 1970). The other exemplars of the category are then created by randomly distorting the prototype. In the most popular prototype distortion task, each stimulus is a random pattern of dots. One pattern is selected as the prototype, and then the other category exemplars are created by randomly perturbing the location of each dot in the prototype pattern. Two versions of this task are popular. In (A, not A) tasks, the participant must decide whether each stimulus is or is not a member of Category A. The not-A stimuli are random patterns with no coherent structure (i.e., they are not created by distorting a prototype). An example of some stimuli that might be used in an (A, not A) prototype distortion task is shown in Figure 3. In (A, B) tasks, two patterns are selected as the Category A and B prototypes, and then the other members of these categories are created by randomly distorting the respective prototypes. The participant's task is to determine whether each stimulus is a member of Category A or B (see Figure 4 for an example).

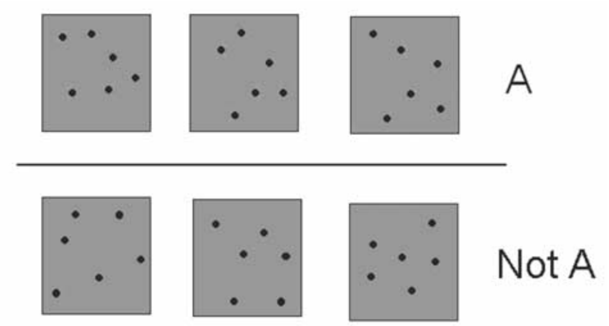

Figure 3: **Stimuli That Might Be Used in an (A, not A) Prototype Distortion Task.**

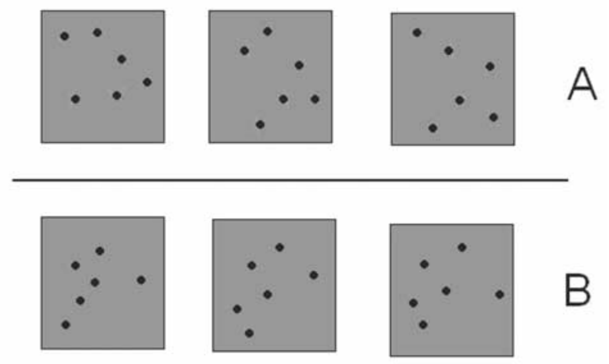

Figure 4: **Stimuli That Might Be Used in an (A, B) Prototype-Distortion Task.**

It is important to emphasize that the terms *rule based*, *information integration*, and *prototype distortion* make no assumptions about how people learn these different category structures in any particular application. For example, there is evidence that pigeons can learn both rule-based and information-integration category structures (Herbranson, Fremouw, & Shimp, 1999), but no one would claim that they learn rule-based categories via an explicit reasoning process. The question of how people learn these different types of category structures is strictly empirical. As such, this particular classification of categorization tasks is useful only because there are many interesting empirical dissociations among the tasks (e.g., Ashby, Maddox, & Bohil, 2002; Ashby, Noble, Filoteo, Waldron, & Ell, 2003; Ashby, Queller, & Berretty, 1999; Ashby & Waldron, 1999; Maddox, Ashby, & Bohil, 2003).

It should also be noted that there are other popular category-learning tasks. One prominent example is the "weather-prediction task" (Eldridge, Masterman, & Knowlton, 2002; Knowlton, Mangels, & Squire, 1996; Knowlton, Squire, & Gluck, 1994; Knowlton, Squire et al., 1996; Reber, Knowlton, & Squire, 1996; Reber & Squire, 1999). Stimuli in this task are tarot cards that each display a unique geometric pattern. The participant's task is to decide if the particular constellation of cards that are shown signals "rain" or "sun." The optimal strategy requires integrating information across cards, so technically, this is an information-integration task. Even so, in the original version, a single-cue strategy in which the participant gives one response if one card is present and the other response if that same card is absent yields an accuracy only 1% less than the optimal information-integration strategy. In addition, because the task uses only a few cards, explicit memorization is difficult to rule out. Because a variety of different strategies are all about equally effective, it is especially important to determine what strategy each participant is using before interpreting his or her data. Gluck, Shohamy, and Myers (2002) provided a strategy analysis of data collected in the weather-prediction task to address this issue. Their analysis suggested that participants adopt a variety of different strategies and that the most popular choice appears to be explicit memorization. The large individual differences seen in the strategies that participants adopt in the weather-prediction task make it difficult to draw strong inferences from data collected with this task. For this reason, the remainder of this article will focus on the rule-based, information-integration, and prototype-distortion tasks.

Rule-Based Tasks

NEUROPSYCHOLOGICAL PATIENT DATA

As mentioned previously, perseverative responding on the WCST is among the most classic of all signs of frontal lobe damage. It is not surprising then that many studies have shown that frontal patients are impaired at rule-based category learning (see, e.g., Kimberg et al., 1997; Robinson, Heaton, Lehman, & Stilson, 1980). Another group with well-known deficits in rule-based tasks is Parkinson's disease patients (e.g., Ashby, Noble, et al., 2003; Brown & Marsden, 1988; Cools, van den Bercken, Horstink, van Spaendonck, & Berger, 1984; Downes et al., 1989). Although later in the disease Parkinson's patients have frontal damage (primarily the result of cell death in the ventral tegmental area), the disease mainly targets the basal ganglia. Within the caudate nucleus, the head tends to be more adversely affected than the body or tail (van Domburg & ten Donkelaar, 1991), and because the head is reciprocally connected to the PFC, the rule-based category-learning deficits of frontal and Parkinson's disease patients are consistent with the hypothesis that rule-based category learning is mediated, in part, by frontal-striatal circuits (Ashby et al., 1998).

In contrast to frontal and basal ganglia disease patients, several studies have reported that amnesiacs with medial temporal lobe damage are normal in rule-based category learning (Janowsky, Shimamura, Kritchevsky, & Squire, 1989; Leng & Parkin, 1988). An obvious possibility is that many rule-based tasks are simple enough (e.g., the WCST) that working memory is sufficient for participants to keep track of which alternative rules they have tested and rejected. If so, then a natural prediction is that medial temporal lobe amnesiacs should be impaired in complex rule-based tasks (e.g., when the optimal rule is disjunctive).

NEUROIMAGING DATA

A number of neuroimaging studies have used the WCST or a rule-based task similar to the WCST. All of these have reported task-related activation in the PFC, most have reported activation in the head of the caudate nucleus, and at least one has also reported task-related activation in the anterior cingulate (Konishi et al., 1999; Lombardi et al., 1999; Rao et al., 1997; Rogers, Andrews, Grasby, Brooks, & Robbins, 2000; Volz et al., 1997). Converging evidence for the hypothesis that these are important structures in rule-based category learning comes from several sources. First are the many studies that have implicated these structures as key components of executive attention (Posner & Petersen, 1990) and working memory (Goldman-Rakic, 1987, 1995), 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. Second, a recent neuroimaging study identified the (dorsal) anterior cingulate as the site of hypothesis generation in a rule-based category-learning task (Elliott, Rees, & Dolan, 1999). Third, of course, are the neuropsychological data reviewed above, which show that patient groups with damage to any of these structures are impaired in rule-based tasks.

Information-Integration Tasks

NEUROPSYCHOLOGICAL PATIENT DATA

Filoteo, Maddox, and Davis (2001b) tested the ability of amnesiacs to learn a difficult (i.e., nonlinear) information-integration rule when the categories were normally distributed and a large number of unique stimuli were sampled from each category. The amnesiacs and controls performed equally over the full 600 trials of the experiment. One patient and 1 control returned for a second session on the following day. During the first block of trials on the 2nd day, the amnesiac and control again showed equivalent performance, and in fact, performance during the first block of the second session was slightly better than during the final block of trials from the first session. Note that this excellent 2nd-day performance falsifies the hypothesis that amnesiacs perform well in category-learning tasks because they substitute working memory for their damaged episodic and semantic (i.e., exemplar-based) memory systems (as proposed, e.g., by Nosofsky & Zaki, 1998; Palmeri & Flanery, 1999). A participant relying on working memory to learn on the 1st day would be at chance on the first few trials of Day 2.

Filoteo, Maddox, and Davis (2001a) and Maddox and Filoteo (2001) tested the ability of Huntington's disease and Parkinson's disease patients to learn the same category structures used by Filoteo et al. (2001b). Relative to controls, both groups showed a consistent performance decrement, which implies a role for the neostriatum in nonlinear information-integration category learning. On the other hand, Ashby, Noble, et al. (2003) found that Parkinson's disease patients learned as well as an age-matched control group in an information-integration task with linearly separable categories. More recently, Filoteo, Maddox, Salmon, and Song (in press) compared the ability of Parkinson's disease patients to learn a linear and a nonlinear information-integration rule. The linear results replicated the Ashby, Noble, et al. (2003) results; that is, the Parkinson's disease patients were not impaired in learning linearly separable categories. On the other hand, the same patients were impaired in the nonlinear condition, but only later in training. Thus, these studies suggest that Parkinson's disease patients are impaired in information-integration tasks but only if the category structures are complex (as, e.g., when the categories are nonlinearly separable).

NEUROIMAGING DATA

To date, only one neuroimaging study of information-integration category learning has been reported. Seger and Cincotta (2002) gave participants extensive training with the categories before scanning, and they reported significant striatal and lateral occipital activation.

Empirical Dissociations Between Rule-Based and Information-Integration Tasks

A number of recent empirical dissociations between performance in rule-based and information-integration tasks collectively provide strong evidence that learning in these two types of tasks is mediated by separate systems, and they provide important clues as to the underlying neural systems that mediate rule-based and information-integration learning.

One set of results show that the nature and timing of trial-by-trial feedback about response accuracy is critical with information-integration categories but not with rule-based categories. First, in the absence of any trial-by-trial feedback, people can learn some rule-based categories, but there is no evidence that they can learn information-integration categories (Ashby et al., 1999). Second, even when feedback is provided on every trial, information-integration category learning is impaired if the feedback signal is delayed by as little as 5 seconds after the response. In contrast, such delays have no effect on rule-based category learning (Maddox et al., 2003). Third, training in which participants observe exemplars from each category is much less effective than traditional feedback training (i.e., in which participants see a stimulus, respond, and then receive feedback) with information-integration categories, but observational and feedback training are equally effective with rule-based categories (Ashby et al., 2002).

A second set of studies established that information-integration categorization uses procedural learning, whereas rule-based category learning does not. First, Ashby, Ell, and Waldron (2003) had participants learn either rule-based or information-integration categories using traditional feedback training. Next, some participants continued as before, some switched their hands on the response keys, and for some, the location of the response keys was switched (so that the Category A key was assigned to Category B and vice versa). For those participants learning rule-based categories, there was no difference among any of these transfer instructions, thereby suggesting that abstract category labels are learned in rule-based categorization. In contrast, for those participants learning information-integration categories, switching hands on the response keys caused no interference, but switching the locations of the response keys caused a significant decrease in accuracy. Thus, it appears that response locations are learned in information-integration categorization, but specific motor programs are not. Further evidence supporting this hypothesis was reported recently by Maddox, Bohil, and Ing (in press). These information-integration results essentially replicate results found with traditional procedural-learning tasks (Willingham, Wells, Farrell, & Stemwedel, 2000).

A third set of studies established the importance of working memory and executive attention in rule-based category learning and simultaneously showed that executive function is not critical in the learning of information-integration categories. First, Waldron and Ashby (2001) had participants learn rule-based and information-integration categories under typical single-task conditions and when simultaneously performing a secondary task that required working memory and executive attention. The dual task had a massive detrimental effect on the ability of participants to learn the simple one-dimensional rule-based categories (trials-to-criterion increased by 350%) but had no significant effect on the ability of participants to learn the complex information-integration categories. Second, Maddox et al. (in press) tested the prediction that feedback processing requires attention and effort in rule-based category learning but not in information-integration category learning. In this study, participants alternated a trial of categorization with a trial of Sternberg's (1966) memory scanning. Two conditions were identical except for the durations of the intertrial intervals (ITIs). In one, a short ITI followed categorization and a long ITI followed memory scanning, whereas these two durations were flipped in the other condition. Information-integration category learning was the same in both conditions, whereas rule-based category learning was significantly impaired when participants had only a short time to process the categorization feedback (i.e., when the short ITI followed categorization).

It is important to realize that these dissociations are not driven simply by differences in the difficulty of rule-based versus information-integration tasks. First, in several cases, the experimental manipulation interfered more with the learning of the simple rule-based categories than with the more difficult information-integration strategies (Maddox et al., in press; Waldron & Ashby, 2001). Second, most of the studies explicitly controlled for difficulty differences either by decreasing the separation between the one-dimensional rule-based categories or by using a more complex two-dimensional conjunction rule in the rule-based conditions. Both manipulations increase the difficulty of rule-based categorization, yet in no case did such increases in rule-based difficulty affect the qualitative dissociations described above.

Prototype Distortion Tasks

NEUROPSYCHOLOGICAL PATIENT DATA

Prototype distortion tasks are particularly important because the neuropsychological patient data are profoundly different than in rule-based or information-integration tasks. In particular, a variety of patient groups that are known to have deficits in rule-based and information-integration tasks show apparently normal prototype-distortion learning, at least in (A, not A) designs. This includes patients with Parkinson's disease (Reber & Squire, 1999), schizophrenia (Kéri, Kelemen, Benedek, & Janka, 2001), or Alzheimer's disease (Sinha, 1999; although see Kéri et al., 1999). Normal (A, not A) performance has also been shown in patients with amnesia (Knowlton & Squire, 1993; Squire & Knowlton, 1995). On the other hand, more research is needed here because several studies have shown that if Category A is created from low-level distortions of the Category A prototype, then healthy adults can learn in (A, not A) tasks without any feedback (Homa & Cultice, 1984; Palmeri & Flanery, 1999). Thus, it is not yet clear that all these patient groups would learn normally in a difficult (A, not A) task (i.e., one that requires feedback for optimal performance).

At least two studies have compared (A, not A) and (A, B) prototype distortion learning on the same patients, and both studies report the same striking dissociation. Specifically, Sinha (1999) reported normal (A, not A) performance in Alzheimer's disease patients but impaired (A, B) performance, and Zaki, Nosofsky, Jessup, and Unverzagt (2003) reported this same pattern of results with amnesiacs. Sinha (1999) also reported deficits in (A, B) prototype-distortion learning in patients with amnesia. On the other hand, Kolodny (1994) reported intact performance of amnesic patients in a difficult (A, B, C) prototype distortion task (i.e., controls were only 10% above chance on transfer items).

NEUROIMAGING DATA

All neuroimaging studies that used (A, not A) prototype-distortion tasks have reported learning-related changes in the occipital cortex (Aizenstein et al., 2000; Reber, Stark, & Squire, 1998a, 1998b); in general, reduced occipital activation was found in response to Category A exemplars, although Aizenstein et al. (2000) found this reduction only under implicit learning conditions. When participants were given explicit instructions to learn the A category, increased occipital activation was observed.

Studies that used (A, B) tasks have reported quite different results. Seger et al. (2000) did report categorization-related activation in the occipital cortex, but they also found significant learning-related changes in prefrontal and parietal cortices. Vogels, Sary, Dupont, and Orban (2002) reported results from a hybrid task in which participants were to respond “A,” “B,” or “neither.” Thus, stimuli were created either from distortions of an A prototype or a B prototype or were just random patterns. Like Seger et al. (2000), Vogels et al. (2002) found prefrontal and parietal activation (although in different foci). However, they also reported task-related activation in the orbitofrontal cortex and the neostriatum, and they failed to find any task-related activation in occipital cortex.

SINGLE VERSUS MULTIPLE SYSTEMS OF HUMAN CATEGORY LEARNING

The data reviewed in this article are generally consistent with the hypothesis that human category learning is mediated by multiple systems that are essentially equivalent to the known memory systems that have been proposed. This makes sense from a logical perspective because learning is a process of laying down some kind of memory trace, and there seems no good reason why any memory system should be prevented from learning about categories. Even so, it is important to acknowledge that the question of whether human category learning is mediated by one or by multiple systems remains contentious. For example, Nosofsky and his colleagues have argued that single-system exemplar models can account for many of the phenomena that have been used to support the notion of multiple systems (Nosofsky & Johansen, 2000; Nosofsky & Kruschke, 2002; Nosofsky & Zaki, 1998).

Every year, more studies appear that present evidence purporting to support the multiple-systems hypothesis. As the number of these studies increases, it becomes increasingly important to evaluate all available evidence simultaneously. No single study can resolve the single versus multiple systems debate. The important question should be, Does the model that best accounts for all existing data postulate one or multiple systems of learning and memory?

Although single-system accounts of some of the results described in this article are possible, no existing single-system model can account for them all simultaneously. And at least some of the individual results reviewed above are highly problematic for unified accounts of rule-based and information-integration categorization. For example, consider the Waldron and Ashby (2001) report that a dual-task interfered with rule-based category learning but not with information-integration category learning. Arguably the most successful existing single-process model of category learning is Kruschke's (1992) exemplar-based 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 one-dimensional rule-based structures, participants would have no idea that only one dimension was relevant in the dual-task conditions. Ashby and Ell (2002) reported empirical evidence that strongly disconfirmed this prediction. Thus, the best available single-system model fails to account even for the one dissociation reported by Waldron and Ashby (2001). Of course, this does not mean that single-system accounts of the Waldron and Ashby results, or of any of the results described in this article, are impossible. Clearly, more work is needed before consensus will be reached in the single- versus multiple-systems debate.

If one accepts the hypothesis that the various memory systems that have been proposed all contribute to category learning, then it is plausible to assume that learning in rule-based tasks is mediated by working memory and possibly also by other declarative memory systems (i.e., episodic and / or semantic memory). The idea is that participants generate an explicit hypothesis about category membership, which they hold in working memory until the hypothesis is accepted or rejected. If the task is complex and many hypotheses must be tested before the correct strategy is discovered, then declarative memory strategies are recruited.

This hypothesis is consistent with virtually all of the behavioral and cognitive neuroscience data reviewed above. First, a huge literature implicates the PFC in working memory (e.g., Fiez, Raichle, Balota, Tallal, & Petersen, 1996; Funahashi, Bruce, & Goldman-Rakic, 1989; Fuster, 1989; Fuster & Alexander, 1971; Goldman-Rakic, 1987, 1995; Goldman-Rakic, Funahashi, & Bruce, 1990; Jonides et al., 1993; McCarthy et al., 1994; Petrides, 1991). The PFC involvement in working memory accounts for the rule-based category-learning deficits of frontal patients (e.g., Kimberg et al., 1997; Robinson et al., 1980) and for the frontal activation reported in neuroimaging studies of rule-based tasks (Konishi et al., 1999; Lombardi et al., 1999; Rao et al., 1997; Rogers et al., 2000; Volz et al., 1997).

Second, there is evidence that the basal ganglia also help mediate working memory (e.g,. Ellis & Nathan, 2001; Fournet, Moreaud, Roulin, Naegele, & Pellat, 2000; Gabrieli, Singh, Stebbins, & Goetz, 1996; Jahanshahi et al., 2002; Lewis et al., 2002; Postle, Jonides, Smith, Corkin, & Growdon, 1997; Postle, Locascio, Corkin, & Growdon, 1997; Robbins et al., 1995) and the switching of executive attention (e.g., Brown & Marsden, 1988; Jaspers, de Vries, & Cools, 1990a, 1990b; Roberts et al., 1994; van Golf Racht-Delatour & El Massioui, 1999). For these reasons, the working memory hypothesis is consistent with the reported rule-based deficits of basal ganglia disease patients (e.g., Ashby, Noble, et al., 2003; Brown & Marsden, 1988; Cools et al., 1984; Downes et al., 1989).

Third, the working memory account of rule-based category learning is also consistent with the dissociation data reviewed in the preceding section. This includes the report that a dual task that used working memory and executive attention interfered with rule-based category learning much more than with information-integration learning (Waldron & Ashby, 2001) and the report that feedback processing requires attention and effort in rule-based learning but not in information-integration learning (Maddox et al., in press).

The collection of results on information-integration category learning is generally consistent with the hypothesis that performance in these tasks is mediated primarily by a procedural-memory-based system (Ashby & Waldron, 1999). The best direct evidence supporting this idea are the reports that response positions are learned in information-integration tasks but abstract category labels are learned in rule-based tasks (Ashby, Ell, et al., 2003; Maddox et al., in press). Also, as we would expect in procedural learning, information-integration (but not rule-based) category learning is sensitive to the nature and timing of feedback (Ashby et al., 1999, 2002; Maddox et al., 2003).

Many studies have argued that procedural memory is mediated largely within the basal ganglia (e.g., Jahanshahi, Brown, & Marsden, 1992; Mishkin, Malamut, & Bachevalier, 1984; Saint-Cyr, Taylor, & Lang, 1988; Willingham, 1998; Willingham, Nissen, & Bullemer, 1989). Thus, the procedural-learning account of information-integration categorization is consistent with reports that basal ganglia disease patients are impaired in difficult information-integration tasks (Filoteo, Maddox, & Davis, 2001a; Filoteo et al., in press; Maddox & Filoteo, 2001) and with reports of striatal activation in neuroimaging studies that used information-integration tasks (Seger & Cincotta, 2002).

Finally, the data suggest that a variety of different memory systems may contribute to learning in prototype-distortion tasks. For example, the frontal activation seen in (A, B) versions of prototype distortion suggests that people may formulate and test explicit hypotheses about category membership. However, we also reviewed evidence that in (A, not A) versions of the task, perceptual learning may be important. This hypothesis accounts for the learning-related changes in the occipital cortex seen in functional magnetic resonance imaging studies of (A, not A) prototype distortion, as well as for the spared performance of patient groups who are impaired in other types of category learning.

CONCLUSIONS

Category learning is a critically important skill that everyone performs countless times each day. Research into its neural basis is appearing at an ever increasing rate. Besides its intrinsic scientific value, this work has the potential to make important practical contributions. First, it can facilitate the design of more efficacious training procedures. For example, the research reviewed above stresses the importance of immediate feedback following a judgment by a radiologist in training as to whether an X ray shows a tumor. Second, research into the neural basis of category learning has important consequences for a variety of special neuropsychological populations—most important, to suggest possible interventions to alleviate some of the cognitive deficits associated with the condition.

REFERENCES

Aizenstein, H. J., MacDonald, A. W., Stenger, V. A., Nebes, R. D., Larson, J. K., Ursu, S., et al. (2000). Complementary category learning systems identified using event-related functional MRI. *Journal of Cognitive Neuroscience*, *12*, 977-987.

Asaad, W. F., Rainer, G., & Miller, E. K. (2000). Task-specific neural activity in the primate prefrontal cortex. *Journal of Neurophysiology*, *84*, 451-459.

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

Ashby, F. G., & Ell, S. W. (2002). Single versus multiple systems of category learning: Reply to Nosofsky and Kruschke. *Psychonomic Bulletin & Review*, *9*, 175-180.

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

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

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

Ashby, F. G., & Maddox, W. T. (in press). Human category learning. *Annual Review of Psychology*.

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

Ashby, F. G., Noble, S., Filoteo J., Waldron, E. M., & Ell, S. W. (2003). Category learning deficits in Parkinson's disease. *Neuropsychology*, *17*, 115-124.

Ashby, F. G., Queller, S., & Berretty, P. M. (1999). On the dominance of unidimensional rules in unsupervised categorization. *Perception & Psychophysics*, *61*, 1178-1199.

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

Barsalou, L. W. (2003). Situated simulation in the human conceptual system. *Language & Cognitive Processes*, *18*, 513-562.

Beninger, R. J. (1983). The role of dopamine in locomotor activity and learning. *Brain Research*, *287*, 173-196.

Brown, R. G., & Marsden, C. D. (1988). Internal versus external cues and the control of attention in Parkinson's disease. *Brain*, *111*, 323-345.

Bullier, J. (2002). Neural basis of vision. In H. Pashler & S. Yantis (Eds.), *Steven's handbook of experimental psychology, Vol. 1: Sensation and perception* (3rd ed., pp. 1-40). New York: John Wiley.

Cools, A. R., van den Bercken, J. H. L., Horstink, M. W. I., van Spaendonck, K. P. M., & Berger, H. J. C. (1984). Cognitive and motor shifting aptitude disorder in Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, *47*, 443-453.

Cree, G. S., & McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *Journal of Experimental Psychology: General*, *132*, 163-201.

Downes, J. J., Roberts, A. C., Sahakian, B. J., Evenden, J. L., Morris, R. G., & Robbins, T. W. (1989). Impaired extra-dimensional shift performance in medicated and unmedicated Parkinson's disease: Evidence for a specific attentional dysfunction. *Neuropsychologia*, *27*, 1329-1343.

Eacott, M. J., & Gaffan, D. (1991). The role of monkey inferior parietal cortex in visual discrimination of identity and orientation of shapes. *Behavioural Brain Research*, *46*, 95-98.

Eldridge, L. L., Masterman, D., & Knowlton, B. J. (2002). Intact implicit habit learning in Alzheimer's disease. *Behavioral Neuroscience*, *116*, 722-726.

Elliott, R., Rees, G., & Dolan, R. J. (1999). Ventromedial prefrontal cortex mediates guessing. *Neuropsychologia*, *37*, 403-411.

Ellis, K. A., & Nathan, P. J. (2001). The pharmacology of human working memory. *International Journal of Neuropsychopharmacology*, *4*, 299-313.

Fiez, J. A., Raichle, M. E., Balota, D. A., Tallal, P., & Petersen, S. E. (1996). PET activation of posterior temporal regions during auditory word presentation and verb generation. *Cerebral Cortex*, *6*, 1-10.

Filoteo, J. V., Maddox, W. T., & Davis, J. D. (2001a). 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.

Filoteo, J. V., Maddox, W. T. & Davis, J. D. (2001b). Quantitative modeling of category learning in amnesic patients. *Journal of the International Neuropsychological Society, 7,* 1-19.

Filoteo, J. V., Maddox, W. T., Salmon, D. P., & Song, D. D. (in press). Information-integration category learning in patients with striatal dysfunction. *Neuropsychology.*

Fournet, N., Moreaud, O., Roulin, J. L., Naegele, B., & Pellat, J. (2000). Working memory functioning in medicated Parkinson's disease patients and the effects of withdrawal of dopaminergic medication. *Neuropsychology, 14,* 247-253.

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2001). Categorical representation of visual stimuli in the primate prefrontal cortex. *Science, 291,* 312-316.

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2002). Visual categorization and the primate prefrontal cortex: Neurophysiology and behavior. *Journal of Neurophysiology, 88,* 914-928.

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2003). A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *Journal of Neuroscience, 23,* 5235-5246.

Funahashi, S., Bruce, C. J., & Goldman-Rakic, P. S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of Neurophysiology, 61,* 331-349.

Fuster, J. M. (1989). *The prefrontal cortex* (2nd ed.). New York: Raven.

Fuster, J. M., & Alexander, G. E. (1971). Neuron activity related to short-term memory. *Science, 173,* 652-654.

Gabrieli, J. D., Singh, J., Stebbins, G. T., & Goetz, C. G. (1996). Reduced working memory span in Parkinson's disease: Evidence for the role of frontostriatal system in working and strategic memory. *Neuropsychology, 10,* 321-332.

Gaffan, D., & Eacott, M. J. (1995) Visual learning for an auditory secondary reinforcer by macaques is intact after uncinate fascicle section: Indirect evidence for the involvement of the corpus striatum. *European Journal of Neuroscience, 7,* 1866-1871.

Gaffan, D., & Harrison, S. (1987). Amygdalectomy and disconnection in visual learning for auditory secondary reinforcement by monkeys. *Journal of Neuroscience, 7,* 2285-2292.

Gluck, M. A., Shohamy, D., & Myers, C. (2002). How do people solve the "weather prediction" task? Individual variability in strategies for probabilistic category learning. *Learning & Memory, 9,* 408-418.

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

Goldman-Rakic, P. S. (1995). Cellular basis of working memory. *Neuron, 14,* 477-485.

Goldman-Rakic, P. S., Funahashi, S., & Bruce, C. J. (1990). Neocortical memory circuits. *Cold Spring Harbor Symposia on Quantitative Biology, 55,* 1025-1038.

Hampson, R. E., Pons, T. P., Stanford, T. R., & Deadwyler, S. A. (2004). Categorization in the monkey hippocampus: A possible mechanism for encoding information into memory. *Proceedings of the National Academy of Sciences of the United States of America, 101,* 3184-3189.

Heaton, R. K. (1981). *A manual for the Wisconsin Card Sorting Test.* Odessa, FL: Psychological Assessment Resources.

Herbranson, W. T., Fremouw, T., & Shimp, C. P. (1999). The randomization procedure in the study of categorization of multidimensional stimuli by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 25,* 113-134.

Homa, D., & Cultice, J. C. (1984). Role of feedback, category size, and stimulus distortion on the acquisition and utilization of ill-defined categories. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 10,* 83-94.

Hoshi, E., Shima, K., & Tanji, J. (1998). Task-dependent selectivity of movement-related neuronal activity in the primate prefrontal cortex. *Journal of Neurophysiology, 80,* 3392-3397.

Humphreys, G. W., & Forde, E. M. E. (2001). Hierarchies, similarity and interactivity in object recognition: "Category-specific" neuropsychological deficits. *Behavioral and Brain Sciences, 24,* 453-509.

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

Jahanshahi, M., Rowe, J., Saleem, T., Brown, R., Limousin-Dowsey, P., Rothwell, J., et al. (2002). Striatal contribution to cognition: Working memory and executive function in Parkinson's disease before and after unilateral posteroventral pallidotomy. *Journal of Cognitive Neuroscience, 142,* 298-310.

Janowsky, J. S., Shimamura, A. P., Kritchevsky, M., & Squire, L. R. (1989). Cognitive impairment following frontal lobe damage and its relevance to human amnesia. *Behavioral Neuroscience, 103,* 548-560.

Jaspers, R. M. A., de Vries, T. J., & Cools, A. R. (1990a). Effect of intrastriatal apomorphine on changes in switching behaviour induced by the glutamate agonist AMPA injected into the cat caudate nucleus. *Behavioral Brain Research, 37,* 247-254.

Jaspers, R. M. A., de Vries, T. J., & Cools, A. R. (1990b). Enhancement in switching motor patterns following local application of the glutamate agonist AMPA into the cat caudate nucleus. *Behavioral Brain Research, 37,* 237-246.

Joel, D., Weiner, I., & Feldon, J. (1997). Electrolytic lesions of the medial prefrontal cortex in rats disrupt performance on an analog of the Wisconsin Card Sorting Test, but do not disrupt latent inhibition: Implications for animal models of schizophrenia. *Behavioural Brain Research, 85,* 187-201.

Jonides, J., Smith, E. E., Koeppe, R. A., Awh, E., Minoshima, S., & Mintun, M. A. (1993). Spatial working memory in humans as revealed by PET. *Nature, 363,* 623-625.

Joseph, J. E. (2001). Functional neuroimaging studies of category specificity in object recognition: A critical review and meta-analysis. *Cognitive, Affective & Behavioral Neuroscience, 1,* 119-136.

Kéri, S., Kalman, J., Rapcsak, S. Z., Antal, A., Benedek, G., & Janka, Z. (1999). Classification learning in Alzheimer's disease. *Brain, 122,* 1063-1068.

Kéri, S., Kelemen, O., Benedek, G., & Janka, Z. (2001). Intact prototype learning in schizophrenia. *Schizophrenia Research, 52,* 261-264.

Kimberg, D. Y., D'Esposito, M., & Farah, M. J. (1997). Frontal lobes: Neuropsychological aspects. In T. E. Feinberg & M. J. Farah (Eds.), *Behavioral neurology and neuropsychology* (pp. 409-418). New York: McGraw-Hill.

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science, 273,* 1399-1402.

Knowlton, B. J., & Squire, L. R. (1993). The learning of natural categories: Parallel memory systems for item memory and category-level knowledge. *Science, 262,* 1747-1749.

Knowlton, B. J., Squire, L. R., & Gluck, M. A. (1994). Probabilistic classification learning in amnesia. *Learning and Memory, 1,* 106-120.

Knowlton, B. J., Squire, L. R., Paulsen, J. S., Swerdlow, N. R., Swenson, M., & Butters, N. (1996). Dissociations within nondeclarative memory in Huntington's disease. *Neuropsychology, 10,* 538-548.

Kolodny, J. A. (1994). Memory processes in classification learning: An investigation of amnesic performance in categorization of dot patterns and artistic styles. *Psychological Science, 5,* 164-169.

Konishi, S., Karwazu, M., Uchida, I., Kikyo, H., Asakura, I., & Miyashita, Y. (1999). Contribution of working memory to transient activation in human inferior prefrontal cortex during performance of the Wisconsin Card Sorting Test. *Cerebral Cortex, 9,* 745-753.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review, 99,* 22-44.

Leng, N. R., & Parkin, A. J. (1988). Double dissociation of frontal dysfunction in organic amnesia. *British Journal of Clinical Psychology, 27,* 359-362.

Lewis, S. J. G., Cools, R., Robbins, T. W., Dove, A., Barker, R. A., & Owen, A. M. (2002). Using executive heterogeneity to explore the nature of working memory deficits in Parkinson's disease. *Neuropsychologia, 41,* 645-654.

Lombardi, W. J., Andreason, P. J., Sirocco, K. Y., Rio, D. E., Gross, R. E., Umhau, J. C., et al. (1999). Wisconsin Card Sorting Test performance following head injury: Dorsolateral fronto-striatal circuit activity predicts perseveration. *Journal of Clinical & Experimental Neuropsychology, 21,* 2-16.

Maddox, W. T., Ashby, F. G., & Bohil, C. J. (2003). Delayed feedback effects on rule-based and information-integration category learning. *Journal of Experimental Psychology: Learning, Memory and Cognition, 29*, 650-662.

Maddox, W. T., Bohil, C. J., & Ing, A. D. (in press). Evidence for a procedural learning-based system in perceptual category learning. *Psychonomic Bulletin & Review.*

Maddox, W. T., & Filoteo, J. V. (2001). Striatal contribution to category learning: Quantitative modeling of simple linear and complex non-linear rule learning in patients with Parkinson's disease. *Journal of the International Neuropsychological Society, 7,* 710-727.

Markman, A. B., & Ross, B. H. (2003). Category use and category learning. *Psychological Bulletin, 129*, 592-613.

McCarthy, G., Spicer, M., Adrignolo, A., Luby, M., Gore, J., & Allison, T. (1994). Brain activation associated with visual motion studied by functional magnetic resonance imaging in humans. *Human Brain Mapping, 2*, 234-243.

McDonald, R. J., & White, N. M. (1993). A triple dissociation of memory systems: Hippocampus, amygdala, and dorsal striatum. *Behavioral Neuroscience, 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. *Behavioral and Neural Biology, 61*, 260-270.

Merchant, H., Zainos, A., Hernandez, A., Salinas, E., & Romo, R. (1997). Functional properties of the primate putamen neurons during the categorization of tactile stimuli. *Journal of Neuroscience, 10*, 3032-3044.

Miller, J. D., Sanghera, M. K., & German, D. C. (1981). Mesencephalic dopaminergic unit activity in the behaviorally conditioned rat. *Life Sciences, 29*, 1255-1263.

Milner, B. (1963). Effects of different brain lesions on card sorting, the role of the frontal lobes. *Archives of Neurology, 9*, 90-100.

Mishkin, M., Malamut, B., & Bachevalier, J. (1984). Memories and habits: Two neural systems. In G. Lynch, J. L. McGaugh, & N. M. Weinberger (Eds.), *Neurobiology of human learning and memory* (pp. 65-77). New York: Guilford.

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of Neuroscience, 16*, 1936-1947.

Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.

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

Nosofsky, R. M., & Johansen, M. K. (2000). Exemplar-based accounts of "multiple-system" phenomena in perceptual cagteorization. *Psyconomic Bulletin & Review, 7*, 375-402.

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.

Nosofsky, R. M., & Zaki, S. R. (1998). Dissociations between categorization and recognition in amnesic and normal individuals: An exemplar-based interpretation. *Psychological Science, 9*, 247-255.

Op de Beeck, H., Wagemans, J., & Vogels, R. (2001). Inferotemporal neurons represent low-dimensional configurations of paramterized shapes. *Nature Neuroscience, 4*, 1244-1252.

Packard, M. G., Hirsh, R., & White, N. M. (1989). Differential effects of fornix and caudate nucleus lesions on two radial maze tasks: Evidence for multiple memory systems. *Journal of Neuroscience, 9,* 1465-1472.

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. *Behavior Neuroscience, 106*, 439-446.

Palmeri, T. J., & Flanery, M. A. (1999). Learning about categories in the absence of training: Profound amnesia and the relationship between perceptual categorization and recognition memory. *Psychological Science, 10*, 526-530.

Petrides, M. (1991). Learning impairments following excisions of the primate frontal cortex. In H. S. Levin, H. M. Eisenberg, & A. L. Benton (Eds.), *Frontal lobe function and dysfunction* (pp. 256-272). London: Oxford University Press.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology, 77,* 353-363.

Posner, M. I., & Keele, S. W. (1970). Retention of abstract ideas. *Journal of Experimental Psychology, 83*, 304-308.

Posner, M. I., & Petersen, S. E. (1990). Attention systems in the human brain. *Annual Review of Neuroscience, 13*, 25-42.

Postle, B. R., Jonides, J., Smith, E. E., Corkin, S., & Growdon, J. H. (1997). Spatial, but not object, delayed response is impaired in early Parkinson's disease. *Neuropsychology, 11*, 171-179.

Postle, B. R., Locascio, J. J., Corkin, S., & Growdon, J. H. (1997). The time course of spatial and object learning in Parkinson's disease. *Neuropsychologia, 35*, 1413-1422.

Rao, S. M., Bobholz, J. A., Hammeke, T. A., Tosen, A. C., Woodley, S. J., Cunningham, J. M., et al. (1997). Functional MRI evidence for subcortical participation in conceptual reasoning skills. *Neuroreport, 8*, 1987-1993.

Reber, P. J., Knowlton, B. J., & Squire, L. R. (1996). Dissociable properties of memory systems: Differences in the flexibility of declarative and nondeclarative knowledge. *Behavioral Neuroscience, 110*, 861-871.

Reber, P. J. & Squire, L. R. (1999). Intact learning of artificial grammars and intact category learning by patients with Parkinson's disease. *Behavioral Neuroscience, 113*, 235-242.

Reber, P. J., Stark, C. E. L., and Squire, L. R. (1998a). Contrasting cortical activity associated with category memory and recognition memory. *Learning & Memory, 5*, 420-428.

Reber, P. J., Stark, C. E. L., & Squire, L. R. (1998b). Cortical areas supporting category learning identified using functional MRI. *Proceedings of the National Academy of Sciences of the United States of America, 95*, 747-750.

Robbins, T. W., Shallice, T., Burgess, P. W., James, M., Rogers, R. D., Warburton, E., et al. (1995). Selective impairments in self-ordered working memory in a patient with a unilateral striatal lesion. *Neurocase, 1*, 217-30.

Roberts, A. C., De Salvia, M. A., Wilkinson, L. S., Collins, P., Muir, J. L., Everitt, B. J., et al. (1994). 6-hydroxydopamine lesions of the prefrontal cortex in monkeys enhance performance on an analog of the Wisconsin Card Sort Test: Possible interactions with subcortical dopamine. *Journal of Neuroscience, 14*, 2531-2544.

Roberts, A. C., & Wallis, J. D. (2000). Inhibitory control and affective processing in the prefrontal cortex: Neuropsychological studies in the common marmoset. *Cerebral Cortex, 10*, 252-262.

Robinson, A. L., Heaton, R. K., Lehman, R. A., & Stilson, D. W. (1980). The utility of the Wisconsin Card Sorting Test in detecting and localizing frontal lobe lesions. *Journal of Consulting and Clinical Psychology, 48*, 605-614.

Rogers, R. D., Andrews, T. C., Grasby, P. M., Brooks, D. J., & Robbins, T. W. (2000). Contrasting cortical and subcortical activations produced by attentional-set shifting and reversal learning in humans. *Journal of Cognitive Neuroscience, 12*, 142-162.

Rolls, E. T. (1994). Neurophysiology and cognitive functions of the striatum. *Revue Neurologique, 150*, 648-660.

Rolls, E. T., Judge, S. J., & Sanghera, M. (1977). Activity of neurons in the inferotemporal cortex of the alert monkey. *Brain Research, 130*, 229-238.

Romo, R., Merchant, H., Ruiz, S., Crespo, P., & Zainos, A. (1995). Neuronal activity of primate putamen during categorical perception of somaesthetic stimuli. *Neuroreport, 6*, 1013-1017.

Saint-Cyr, J. A., Taylor, A. E., & Lang, A. E. (1988). Procedural learning and neostriatal dysfunction in man. *Brain, 111*, 941-959.

Seger, C. A., & Cincotta, C. M. (2002). Striatal activity in concept learning. *Cognitive, Affective & Behavioral Neuroscience, 2*, 149-161.

Seger, C. A., Poldrack, R. A., Prabhakaran, V., Zhao, M., Glover, G. H., & Gabrieli, J. D. E. (2000). Hemispheric asymmetries and individual differences in visual concept learning as measured by functional MRI. *Neuropsychologia, 38*, 1316-1324.

Sigala, N. (2004). Visual categorization and the inferior temporal cortex. *Behavioural Brain Research, 149*, 1-7.

Sigala, N., & Logothetis, N. K. (2002). Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature, 415*, 318-320.

Sinha, R. R. (1999). Neuropsychological substrates of category learning (UMI No. AEH9932480). *Dissertation Abstracts International: Section B: The Sciences & Engineering, 60*(5-B), 2381.

Squire, L. R., & Knowlton, B. J. (1995). Learning about categories in the absence of memory. *Proceedings of the National Academy of Science of the United States of America, 92*, 12470-12474.

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

Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience, 19*, 109-139.

Thomas, E., Van Hulle, M. M., & Vogels, R. (2001). Encoding of categories by noncategory-specific neurons in the inferior temporal cortex. *Journal of Cognitive Neuroscience, 13*, 190-200.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549-586). Cambridge, MA: MIT Press.

van Domburg, P. H. M. F., & ten Donkelaar, H. J. (1991). *The human substantia nigra and ventral tegmental area*. Berlin: Springer-Verlag.

van Golf Racht-Delatour, B., & El Massioui, N. (1999). Rule-based learning impairment in rats with lesions to the dorsal striatum. *Neurobiology of Learning & Memory, 72*, 47-61.

Vogels, R. (1999). Complex visual images by rhesus monkeys. Part 2: single-cell study. *European Journal of Neuroscience, 11*, 1239-1255.

Vogels, R., Sary, G., Dupont, P., & Orban, G. A. (2002). Human brain regions involved in visual categorization. *Neuroimage, 16*, 401-414.

Volz, H.-P., Gaser, C., Haeger, F., Rzanny, R., Mentzel, H.-J., Kreitschmann-Andermahr, I., et al. (1997). Brain activation during cognitive stimulation with the Wisconsin Card Sorting Test: A functional MRI study on healthy volunteers and schizophrenics. *Psychiatry Research: Neuroimaging, 75*, 45-157.

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

White, N. M. (1989). A functional hypothesis concerning the striatal matrix and patches: Mediation of S-R memory and reward. *Life Science, 45*, 1943-1957.

White, I. M., & Wise, S. P. (1999). Rule-dependent neuronal activity in the prefrontal cortex. *Experimental Brain Research, 126*, 315-335.

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

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

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

Willingham, D. B., Wells, L. A., Farrell, J. M., & Stemwedel, M. E. (2000). Implicit motor sequence learning is represented in response locations. *Memory & Cognition, 28*, 366-375.

Zaki, S. R., Nosofsky, R. M., Jessup, N. M., & Unverzagt, F. W. (2003). Categorization and recognition performance of a memory-impaired group: Evidence for single-system models. *Journal of the International Neuropsychological Society, 9*, 394-406.