Category Learning Deficits in Parkinson's Disease

F. Gregory Ashby and Sharon Noble University of California, Santa Barbara

Elliott M. Waldron Sightward Inc. J. Vincent Filoteo University of California, San Diego, and Veterans Administration San Diego Healthcare System

> Shawn W. Ell University of California, Santa Barbara

Sixteen patients with Parkinson's disease (PD), 15 older controls (OCs), and 109 younger controls (YCs) were compared in 2 category-learning tasks. Participants attempted to assign colored geometric figures to 1 of 2 categories. In rule-based tasks, category membership was defined by an explicit rule that was easy to verbalize, whereas in information-integration tasks, there was no salient verbal rule and accuracy was maximized only if information from 3 stimulus components was integrated at some predecisional stage. The YCs performed the best on both tasks. The PD patients were highly impaired compared with the OCs, in the rule-based categorization task but were not different from the OCs in the information-integration task. These results support the hypothesis that learning in these 2 tasks is mediated by functionally separate systems.

Categorization is the act of responding differently to objects and events in the environment that belong to separate classes or categories. It is a critical skill that every organism must possess in at least a rudimentary form because it allows them to respond differently, for example, to nutrients and poisons and to predators and prey. There is recent evidence that human category learning relies on multiple systems (e.g., Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Ashby & Ell, 2001, 2002a, 2002b; Erickson & Kruschke, 1998; Pickering, 1997; Smith, Jonides, & Koeppe, 1996; Smith, Patalano, & Jonides, 1998; Waldron & Ashby, 2001). In all cases in which multiple categorylearning systems have been proposed, it has been hypothesized that one system uses explicit (i.e., rule-based) reasoning and at least one involves some form of implicit learning.

Ashby et al. (1998; see also Ashby, Isen, & Turken, 1999) proposed that the explicit system is mediated by frontal-cortical structures (i.e., prefrontal cortex and anterior cingulate) and by the head of the caudate nucleus, whereas learning in the implicit system is largely mediated within the tail of the caudate (in the case of visual stimuli). In this model, dopamine facilitates the operations of both

This research was supported in part by National Science Foundation Grant BCS99-75037.

systems. Parkinson's disease (PD) causes the death of dopamine-producing cells, most notably in the substantia nigra but also in the ventral tegmental area (VTA). As a result, PD patients have decreased dopamine levels in the frontal cortex and in the caudate nucleus (and other substantia nigra and VTA projection sites). For this reason, Ashby et al. (1998) predicted category-learning deficits in PD patients. On the other hand, these predictions are complicated somewhat by large individual differences in the amount of damage sustained in these patients within the tail of the caudate, the head of the caudate, and the frontal cortical areas (e.g., van Domburg & ten Donkelaar, 1991, pp. 71-72). In general however, of these three areas, damage caused by PD usually occurs first, and is most severe, in the head of the caudate (e.g., van Domburg & ten Donkelaar, 1991, pp. 71-72). Thus, Ashby et al. (1998) predicted overall category-learning deficits in PD patients, with the most severe deficits coming in tasks in which some salient explicit rule determines category membership (i.e., explicit tasks).

Many studies have shown that PD patients are impaired in explicit category learning (e.g., Brown & Marsden, 1988; Cools, van den Bercken, Horstink, van Spaendonck, & Berger, 1984; Downes et al., 1989), and some studies have reported that PD patients are also impaired in tasks in which there is no salient explicit rule that determines category membership (Filoteo, Maddox, & Davis, 2001a; Knowlton, Mangels, & Squire, 1996; Knowlton, Squire, et al., 1996; Maddox & Filoteo, 2001). To our knowledge, however, no studies have compared the performance of the same PD patients in these two types of tasks. This article reports such a study. Relative to a group of age-matched controls, we show that PD patients are more severely impaired when learning explicit category structures than when learning implicit category structures. Our results also provide support for the theory that there are functionally separate explicit and implicit category-learning systems.

F. Gregory Ashby and Shawn W. Ell, Department of Psychology, University of California, Santa Barbara; Sharon Noble, Counseling/Clinical/School Psychology Program, Gevirtz Graduate School of Education, University of California, Santa Barbara; J. Vincent Filoteo, Department of Psychiatry, University of California, San Diego, and Psychology Service, Veterans Administration San Diego Healthcare System, San Diego, California; Elliott M. Waldron, Sightward Inc., Seattle, Washington.

Correspondence concerning this article should be addressed to F. Gregory Ashby, Department of Psychology, University of California, Santa Barbara, California 93106. E-mail: ashby@psych .ucsb.edu

Two Different Category-Learning Tasks

Much of the evidence for multiple category-learning systems (summarized in the Discussion section) comes from two different types of category-learning tasks. Rule-based *tasks* are those in which participants can learn the category structures via some explicit reasoning process. Frequently, the rule that maximizes accuracy (i.e., the optimal rule) is easy to describe verbally (Ashby et al., 1998). As a result, participants can learn the category structures via an explicit process of hypothesis testing (Bruner, Goodnow, & Austin, 1956) or theory construction and testing (Murphy & Medin, 1985). In the most common applications, only one stimulus dimension is relevant, and the participant's task is to discover this relevant dimension and then to map the different dimensional values to the relevant categories. Figure 1 shows the stimuli and category structure of a rule-based task used in our experiment. The categorization stimuli were colored geometric figures presented on a colored background. The stimuli varied on four binary-valued dimensions: background color (blue or yellow, depicted as light or dark gray, respectively), embedded symbol color (red or green, depicted as black or white, respectively), number of symbols (1 or 2), and symbol shape (square or circle). This yields a total of 16 possible stimuli. To create rule-based category structures, we arbitrarily selected one dimension to be relevant. The two values on that dimension were then assigned to the two contrasting categories. In Figure 1, this process results in a task in which perfect accuracy is achieved with the following rule: Respond A if the background color is blue (depicted as light gray), and respond B if the background color is yellow (depicted as dark gray). Not surprisingly, participants are usually able to describe the rule they used in rule-based tasks quite accurately by the end of training. Virtually all categorization tasks used in neuropsychological assessment are rule-based, including the well-known Wisconsin Card Sorting Test (WCST; Heaton, 1981).

Information-integration tasks are those in which accuracy is maximized only if information from two or more stimulus



Figure 1. Category structure of a rule-based category-learning task. The optimal explicit rule is as follows: Respond A if the background color is blue (depicted as light gray), and respond B if the background color is yellow (depicted as dark gray).

components (or dimensions) is integrated at some predecisional stage (Ashby & 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 (Ashby et al., 1998).

Figure 2 shows the stimuli and category structure of an information-integration task used in our experiment. The categorization stimuli are the same as in the rule-based task shown in Figure 1. To create information-integration category structures, we arbitrarily selected one dimension to be irrelevant. For example, in Figure 2 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 Figure 2, a background color of blue (depicted as light gray), a symbol color of red (depicted as black), and two symbols 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 is greater than 1.5; otherwise it belongs to Category B. This rule is readily learned by healthy young adults, but even after achieving perfect performance, they can virtually never accurately describe the rule they used (Waldron & Ashby, 2001).

A conjunction rule (e.g., respond A if the background is blue and the embedded symbol is round; otherwise respond B) is a rule-based task rather than an information-integration task because separate decisions are first made about each dimension (e.g., symbol is round or square) and then the outcome of these decisions is combined (integration is not predecisional). The critical criteria are that this rule is easy to describe verbally and easy to learn through an explicit reasoning process. Note that according to this criteria, there is no limit on the complexity of the optimal rule in rule-based tasks. However, as the complexity of the optimal rule increases, its salience decreases and it becomes less likely that observers will learn the associated categories through an explicit reasoning process. In fact, Alfonso-Reese (1997) found that even simple conjunction rules have far lower salience than unidimensional rules. This does not mean that people cannot learn conjunction rules-only that they are unlikely to experiment with such rules unless feedback compels them in this direction. This discussion should make it clear that the boundary between rule-based and information-integration tasks is fuzzy. Tasks in which the optimal rule is unidimensional are unambiguously rule based (at least with separable stimulus dimensions), and tasks in which the optimal rule is significantly more complex than a conjunction rule are almost never rule based. In between, the classification is not so clear-cut. For this reason, the rule-based tasks used in this article all have a

¹ Whereas a weighted linear combination of the dimensional values might be classified as a mathematical rule, in this article we use the term *rule* to mean a strategy that is easy to describe verbally.



Figure 2. Category structure of an information-integration categorization task. The irrelevant dimension is symbol shape (square or circle). A background color of blue (depicted as light gray), a symbol color of red (depicted as black), and two symbols are all assigned a value of 1. The category assignments are determined by the following rule: The stimulus belongs to Category A if the sum of values on the relevant dimensions is greater than 1.5; otherwise it belongs to Category B.

unidimensional optimal rule, whereas the information-integration tasks all have three relevant stimulus dimensions.

In the experiment described below, a group of PD patients learned two rule-based category structures of the type shown in Figure 1 and two information-integration structures like the one shown in Figure 2. On each trial, one of the 16 stimuli was randomly selected and shown to the participant, whose task was then to assign it to Category A or B by pressing the appropriate response key. Participants were then told whether their response was correct or incorrect, and learning proceeded until 10 correct responses in a row were given or until 200 trials from the same category structure had occurred without the criterion 10 correct responses in a row. On the first trial after one of these two criteria was met, the category structure was changed without warning. Unlike some versions of the WCST, however, all participants were told at the beginning of the experimental session that the category structures would change at some time during the session.

Compared with the WCST, we believe the task used here has several advantages. First, it requires participants to learn category structures of qualitatively different types (i.e., rule based and information integration), whereas all category structures in the WCST are of the same type (i.e., rule based). Second, the stimulus information is simpler in our task because the participant sees only a single "card" at a time. In contrast, the visual display in the WCST always includes at least five cards, so learning failures in our task are less likely to be caused by a participant being overwhelmed by visual information compared with the WCST.

The performance of the PD patients was compared with the performance of a group of age-matched controls. The mean age of the PD patients was 66.8 years, so the agematched controls could be considered as a group of healthy elderly people. Relative to a group of young adults however, we predict some deficits in the elderly control group because during the course of normal aging dopamine levels in the human brain decrease about 7% or 8% during each decade of life (e.g., van Domburg & ten Donkelaar, 1991, pp. 85–86). It is thought that there are no obvious motor effects of this decrease until overall dopamine levels are reduced by about 70%, at which point elderly people begin to experience symptoms of early PD. In fact, it has been hypothesized that PD occurs when the normal dopamine depletion process is accelerated for one reason or another (Calne & Langston, 1983). For this reason, the data of the PD and healthy elderly participants presented below will be compared with data from healthy young adults (run on the same category structures, but in a different experiment).

Method

Participants

Sixteen patients with idiopathic PD and 15 older controls (OCs) participated in this study. Patients were diagnosed by a boardcertified neurologist on the basis of the presence of two of the following symptoms: (a) resting tremor, (b) rigidity, or (c) bradykinesia as well as a beneficial response to dopaminergic therapy. The patients had been diagnosed an average of 9.29 years (SD = 10.80) prior to their participation in this study. At the time of their participation, all 16 patients were taking some form of dopaminergic medication, and 12 patients were also taking an anticholinergic. Using Hoehn and Yahr's (1967) rating scale, 3 patients were classified as being in Stage I of the disease, 4 patients were in Stage II, 8 patients were in Stage III, 1 patient was in Stage IV, and 1 patient was in Stage V. OC participants were recruited from the community. All participants were screened for a history of neurological disease (other than PD for the patient group), psychiatric illness, and substance abuse. Table 1 shows the mean age, years of education, scores on the Dementia Rating Scale (DRS; Mattis, 1988), and scores on the Geriatric Depression Scale (GDS; Brink et al., 1982) for the PD patients and OC participants. All PD patients and OC participants had scores of 132 or greater on the DRS. This score on the DRS represents a one standard deviation cutoff from the average of controls provided in the DRS manual. The PD patients did not differ from the OC participants in terms of age, years of education, scores on the DRS, or scores on the GDS (all ps > .05). In addition to the PD patients and OC participants, 109 younger controls (YC) participated in the study. Unlike the PD and OC participants, however, each YC participant only learned one type of categorization rule. Specifically, 44 YC participants learned the explicit rules and 65 YC participants

Table 1

Demographic Characteristics, Dementia Rating Scale (DRS) Scores, and Geriatric Depression Scale (GDS) Scores of the Patients With Parkinson's Disease (PD) and Older Controls (OCs)

Variable	PD		OC	
	М	SD	М	SD
Age (years)	66.81	9.06	67.87	5.38
Education (years)	17.00	2.80	16.13	1.60
DRS	138.94	3.51	139.57	2.87
GDS	8.13	4.66	5.64	3.34

Note. Groups did not differ significantly on any variable.

learned the information-integration rules. All YCs participated as part of partial fulfillment of course requirements.

Stimuli and Apparatus

The stimuli and category structures are described above. There were a total of four possible rule-based category structures, one corresponding to each of the four binary-valued dimensions (back-ground color, symbol color, symbol shape, and number of symbols) that constituted the stimuli. There were also a total of four possible information-integration category structures (corresponding to each of the four possible irrelevant dimensions). In all cases, there were eight exemplars in Category A and eight in Category B. The presentation order of both the rule-based and information-integration category structures was random across participants.

The stimuli were computer generated and displayed on an Apple Multiple Scan cathode-ray tube (CRT) Model No. M2943 (15 in., using a resolution of 832×624 pixels) in a dimly lit room. Each categorization stimulus subtended a visual angle of 9.5 degrees (300×300 pixels). The stimuli were presented on a black background using Macintosh PowerPC computers running the Psychophysics Toolbox software (Brainard, 1997; Pelli, 1997) in the Matlab (MathWorks, Sherborn, MA) environment.

Procedure

Prior to performing the categorization tasks, participants were shown a series of sample stimuli. During this time the dimensions that constituted the stimuli (background color, background shape, number of symbols, and symbol shape) were shown to the observers. In the experiment, participants classified each stimulus into Category A or Category B by pressing the appropriate key on the keyboard after the categorization stimulus appeared on the CRT display. The stimulus presentation was response terminated. A brief (1 s) high-pitched tone (500 Hz) was presented if the response was correct, and a low-pitched tone (200 Hz) was presented if the response was incorrect. Participants were given examples of the high-pitched tone and low-pitched tone prior to the administration of the experimental trials. The experimenter did not proceed to the practice trials until it was evident that the participant understood which tone corresponded to a correct or incorrect response. The response-stimulus interval was 1 s. The criterion for learning the category structures was 10 consecutive correct responses. After the criterion was met, or if the criterion was not met within 200 trials, the categorization rule was changed without warning, although participants were told at the beginning of the experiment that the categorization rules would occasionally change during the course of the experiment.

The experiment was conducted over two sessions, separated on average by less than a week. Over the course of these two sessions participants learned three explicit rules and two information-integration rules, never learning the same rule twice. Explicit rules are generally easier for healthy young adults to learn than informationintegration rules (e.g., Waldron & Ashby, 2001). Consequently, in an effort to avoid overwhelming the PD patients and OCs during the first experimental session, participants learned three explicit rules during the first session and two information-integration rules during the second session.

All stimuli, category structures, and the general procedures for the YC participants were identical to those described above. The only differences were that the design of the experiment in which the data of the YC participants were collected was betweenparticipants (i.e., different participants learned the explicit and information-integration rules) as opposed to the within-participants design used for the PD and OC groups and that the YC group learned only two explicit rules (in contrast to the three learned by the PD and OC groups). To facilitate comparison between all groups, we analyzed only the first two explicit rules learned by the PD and OC groups.

Results

Participants were classified as *learners* or *nonlearners* on the basis of their ability to learn the two rules in the two conditions. If a participant was able to learn both rules in a condition (i.e., they were able to achieve 10 correct responses in a row within 200 trials for both rules), they were classified as a learner for that condition. However, if a participant was unable to learn at least one of the two rules in a condition, they were classified as a nonlearner for that condition. Note that the classification of learner and nonlearner was made separately for the rule-based and information-integration conditions. Thus, a participant could be a learner in one condition and a nonlearner in the other condition. The proportion of nonlearners for each group and rule type is shown in Figure 3.

Figure 3 shows a number of results of interest. First, although the YC group failed slightly more often on the information-integration categories than on the rule-based categories, this difference was not significant, $\chi^2(1, N_{\rm II} =$ 65, $N_{\rm RB} = 44$) = 3.15, p = .12 (where II = information integration and RB = rule based). The PD patients also failed equally often on the two types of category structures, $\chi^2(1, N = 16) = 0.00, p = 1.00$. In contrast, the OCs failed more frequently on the information-integration categories than on the rule-based categories, $\chi^2(1, N = 15) = 3.97$, p < .05. Second, in the rule-based categories there was no difference in the proportion of nonlearners for the YC and OC groups, $\chi^2(1, N_{\rm YC} = 44, N_{\rm OC} = 15) = 1.37, p > .05,$ but there were significantly more PD nonlearners than either YC nonlearners, $\chi^2(1, N_{\rm PD} = 16, N_{\rm YC} = 44) = 17.46, p < .01$, or OC nonlearners, $\chi^2(1, N_{\rm PD} = 16, N_{\rm OC} = 15) = 4.76$, p < .05. Third, in the information-integration categories the OC group failed as often as the PD patients, $\chi^2(1,$ $N_{\rm OC} = 15$, $N_{\rm PD} = 16$) = 0.03, p > .50, but both of these groups failed significantly more often than the YC group: compared with the OC group, $\chi^2(1, N_{\rm YC} = 65, N_{\rm OC} =$ 15) = 8.90, p < .01: compared with the PD group, $\chi^2(1, N_{\rm YC} = 65, N_{\rm PD} = 16) = 10.07, p < .01$.

To investigate possible differences among the learners in the PD and control groups further, we averaged the number of trials to criterion for the two rules in both the rule-based and information-integration conditions separately. These data can be seen in Figure 4. A one-way analysis of variance (ANOVA) of the data in the rule-based condition indicated that the three groups differed in their averaged trials to criterion, F(2, 60) = 4.70, p < .05. Follow-up contrasts using Tukey's honestly significant difference (HSD) test indicated that the OC group required a greater number of trials to criterion than the YC group (p < .05), whereas the PD group did not differ from the OC or the YC groups (see Figure 4). In contrast, a one-way ANOVA of the data in the information-integration condition revealed that the learners in the three groups did not differ in their averaged trials to



Figure 3. Proportion of nonlearners in each group for both types of category structures. Error bars represent standard errors of proportions. YC = younger control group; OC = older control group; PD = Parkinson's disease group.

criterion, F(2, 67) = 0.38, p > .50. We also estimated learning curves for each PD and OC learner (i.e., proportion correct plotted against trial number). These were identical for the two groups in both the rule-based and informationintegration conditions. Thus, for those participants who learned both category structures in a condition, there was no difference between the PD patients and the OC group either in the number of trials required to reach criterion accuracy or in the rate at which this categorical knowledge was acquired.

As a follow-up to the proportional analyses presented at the beginning of this section, we also contrasted the learners and nonlearners in the PD and OC groups on several variables to determine if these subgroups differed. For the PD group, learners and nonlearners in the rule-based condition did not differ in terms of age, years of education, Hoehn and Yahr (1967) scores, or scores on the GDS (all ts < 1.20); however, there was a trend for nonlearners in the PD group to have lower DRS scores than learners, t(14) = 2.14, p = .05. The learners and nonlearners of the explicit rules in the OC group did not differ reliably on any of the above variables (all ts < 0.41). The PD learners and nonlearners in the number of years of education, with learners having a greater number



Figure 4. Trials required of learners to reach criterion (i.e., 10 correct responses in a row) for each group with both types of category structures. Error bars represent standard errors of mean trials to criterion. YC = younger control group; OC = older control group; PD = Parkinson's disease group.

of years than nonlearners, t(14) = 3.15, p < .01, but the two groups did not differ on any of the other variables (all ts < 1.22). The OC learners and nonlearners did not differ on any of the above variables (all ts < 0.75). Finally, to determine if performance in the two categorization tasks was associated with disease severity, we correlated trials to criterion in the two conditions with the Hoehn and Yahr (1967) scores of the PD learners. The results of these correlations were not significant (p > .50 for both correlations).

Discussion

The results clearly indicate that, as a group, the PD patients were severely impaired, relative to the OC group, when learning rule-based category structures but not when learning information-integration category learning. In fact, fully half of the PD group failed to learn at least one of the two rule-based category structures, despite the apparent simplicity of the rule-based categories and the extensive experience participants were allowed (i.e., 200 trials). In contrast, only 14% of the OC group failed on at least one rule-based category structure. On the other hand, there was no difference between these two groups in the more difficult information-integration condition, either in the proportions who failed to learn both structures or in the performance of those who did learn. Thus, compared with the OC group, our results show that the PD patients were selectively impaired on the simpler rule-based category structures.

A comparison of the YC and OC groups yields a different conclusion. The OC group failed more frequently than the YC group on the difficult information-integration categories, but they were just as successful as the YC group on the rule-based categories. Thus, compared with the YC group, the OC group was selectively impaired on the informationintegration categories. This conclusion must be interpreted with caution, however, because of the small, but possibly important, design differences under which the two groups were run (i.e., see the Method section).

Our results also indicate that there may be distinct subpopulations of PD patients, a possibility that has been suggested by other investigators (El-Awar, Becker, Hammond, Nebes, & Boller, 1987; Filoteo et al., 1997; Maddox, Filoteo, Delis, & Salmon, 1996; Mortimer, Jun, Kuskowski, & Webster, 1987). In the rule-based condition, half the PD patients were nonlearners but the other half (i.e., the PD learners) performed no worse than the OC group. The PD nonlearners in the rule-based condition scored marginally worse on the DRS than the PD learners (p = .05), so one possibility is that the nonlearners had more pathology than the learners, and given that damage to frontal-subcortical systems might underlie cognitive alterations in PD (e.g., Dubois & Pillon, 1997; Owen et al., 1992), it is possible that the PD nonlearners had more pathology in these brain regions. As we will show shortly, there is strong evidence that the frontal cortex is critical in rule-based category learning (especially the prefrontal cortex and the anterior cingulate).

One surprising result of our study was a failure to find an information-integration category-learning deficit in the PD

group. There are several reasons why such a deficit is expected. First, there have been previous reports of deficits in PD patients in information-integration category learning (Knowlton, Mangels, & Squire, 1996; Maddox & Filoteo, 2001). Second, as we argue below, there is reason to believe that information-integration category learning uses the procedural memory system, which is thought to be damaged in PD (Soliveri, Brown, Jahanshahi, Caraceni, & Marsden, 1997; Thomas-Ollivier et al., 1999). One possibility is that such damage exists in the PD group but that our information-integration categorization task was not difficult enough to show any effects of such damage. The other reports of PD impairment in information-integration category learning used tasks that were considerably more difficult than the task used here, either because the rule determining category membership was more complex (Maddox & Filoteo, 2001) or because the category assignments were probabilistic (Knowlton, Mangels, & Squire, 1996). This hypothesis is supported by the results of Maddox and Filoteo (2001). In two separate information-integration categorization experiments, they found that PD patients were impaired in the more difficult task but not in the easier task.²

Single Versus Multiple Category-Learning Systems

An important question we have not yet considered is whether learning of the rule-based and information-integration category structures was mediated by separate systems. Several lines of evidence point in this direction. First, a wide variety of data collected from younger observers in rule-based and information-integration tasks support the multiple-systems hypothesis (for a review, see Ashby & Ell, 2002b). For example, in the absence of trial-by-trial feedback people can learn some rule-based categories, but there is no evidence that it is possible to learn informationintegration categories without feedback (Ashby, Queller, & Berretty, 1999). 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 s after the response. In contrast, such delays have no effect on rulebased category learning (Maddox, Ashby, & Bohil, 2002). Similar results are obtained when observational learning is compared with traditional feedback learning. Ashby, Maddox, and Bohil (2002) trained participants on rule-based and information-integration categories using an observational training paradigm in which participants were informed of the category membership of the ensuing stimulus. Following stimulus presentation, participants then pressed the appropriate response key. Traditional feedback training was as effective as observational training with rule-based categories, but with information-integration categories, feedback training was significantly more effective than observational

² The interpretation of this finding is complicated somewhat by the fact that the optimal rule in the simpler information-integration task used by Maddox and Filoteo (2001) was easy to describe verbally. Thus, although accurate responding required integrating stimulus information at a predecisional stage, it is not clear whether participants used explicit rules in this condition.

training. Finally, switching the location of the response keys after training is complete has no effect on rule-based categorization, but it impairs information-integration categorization (Ashby, Ell, & Waldron, 2002), thereby suggesting that information-integration category learning is more closely tied to motor outputs than rule-based category learning.

One criticism of all these results is that informationintegration tasks are usually more difficult than rule-based tasks. For example, in our study the performance of the YC and the OC groups clearly shows that the rule-based task is simpler than the information-integration task-for both groups, the proportion of nonlearners and the trials to criterion of learners were lower in the rule-based condition. Because of this difficulty difference, one concern is that collectively these studies might show only that there are many ways to disrupt learning of difficult tasks compared with simpler tasks. However, several results argue strongly against this hypothesis. First, as can be seen in Figure 3, compared with the OC group, the PD patients were more impaired on the simple rule-based categories than on the difficult information-integration categories. When a single learning system is damaged, we expect to see the effects first in difficult tasks. The opposite pattern of results observed here thereby suggests that learning in the two tasks was mediated by separate category-learning systems and that the pattern of results was not due to difficulty per se.

Other evidence against the difficulty hypothesis comes from an experiment by Waldron and Ashby (2001). In this study, which used exactly the same stimuli and category structures as the present study (e.g., as in Figures 1 and 2), a group of younger adults learned rule-based or information-integration categories either under typical single-task conditions or when simultaneously performing a numerical Stroop task. On each trial of the dual-task conditions, the categorization stimulus was flanked by two numbers, which differed across trials in numerical value and physical size. Participants were required to maintain this information in memory until after their categorization response, at which time they were given a cue to report either the numerically larger digit or the digit that was physically larger. Results showed that the concurrent Stroop task dramatically impaired learning of the simple rule-based categories but did not significantly delay learning of the more difficult information-integration categories. Such results are highly problematic for single-system accounts of category learning.

Other evidence supporting separate systems comes from quantitative modeling of category learning data. For example, arguably the most popular single-system model of category learning is Kruschke's (1992) attentional learning covering map (ALCOVE) model. Ashby and Ell (2002a) 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 rulebased structures, participants would 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 the one dissociation reported by Waldron and Ashby (2001).

Alternative Hypotheses

Our results are consistent with the hypothesis that rulebased and information-integration category learning are mediated by separate systems and that PD causes greater damage to the system mediating rule-based learning. Even so, there are two competing interpretations that are important to consider. The first is that PD causes a general impairment in cognition, rather than a deficit in a specific aspect of category learning such as we have proposed. Note that this hypothesis predicts that the greatest category-learning deficits in PD should occur in the most difficult categorization tasks. Therefore, the finding that our PD patients were more impaired on the simple rule-based tasks than on the difficult information-integration tasks is strong evidence against this general impairment hypothesis.

Another important possibility to consider is that because each of our categories contained only eight exemplars, participants might have learned the information-integration categories by memorizing responses to each individual exemplar. If so, then the rule-based and information-integration tasks would both depend on explicit memory. The dual-task results of Waldron and Ashby's (2001) study, however, argue strongly against this possibility. Recall that this study used exactly the same stimuli and category structures as in the present study. If observers learned the information-integration categories by memorizing individual exemplars, they should have experienced massive interference from the simultaneous numerical Stroop task used by Waldron and Ashby (2001) because the working memory load imposed by this dual task should have greatly interfered with active memorization. Therefore, the absence of any significant interference on information-integration category learning in the Waldron and Ashby study suggests that memorization strategies could have played, at most, only a minor role in the information-integration conditions.

Neural Basis of Rule-Based and Information-Integration Category Learning

If rule-based and information-integration category learning are largely mediated by separate systems, then it is natural to ask about the neural basis of these systems. Although much more work is needed on this problem, some strong clues come from neuropsychological and neuroimaging studies of category learning. For example, many studies in addition to ours have shown that patients with frontal or basal ganglia dysfunction are impaired in rulebased tasks (e.g., Brown & Marsden, 1988; Cools et al., 1984; Downes et al., 1989; Janowsky, Shimamura, Kritchevsky, & Squire, 1989; Leng & Parkin, 1988; Robinson, Heaton, Lehman, & Stilson, 1980). In addition, patients with medial temporal lobe damage are normal in this type of category learning (e.g., Janowsky et al., 1989; Leng & Parkin, 1988). Thus, an obvious first hypothesis is that the prefrontal cortex and the basal ganglia participate in rulebased category learning but the medial temporal lobes do not.

To a remarkable degree, these conclusions agree with existing neuroimaging data. For example, a functional magnetic resonance imaging (fMRI) study of a rule-based task similar to the WCST showed activation in the right dorsallateral prefrontal cortex, the anterior cingulate, and the head of the right caudate nucleus (among other regions; Rao et al., 1997). Similar results were obtained in a recent fMRI study of the WCST (Monchi, Petrides, Petre, Worsley, & Dagher, 2001). 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 & Dolan, 1998). Third, lesion studies in rats implicate the dorsal caudate nucleus in rule switching (Winocur & Eskes, 1998), which is believed to be an important aspect of rule-based category learning (Ashby et al., 1998).

Note that these conclusions suggest that the rule-based deficits seen in PD are due primarily to dysfunction in the head of the caudate nucleus. This conclusion is consistent with postmortem autopsy data, which reveal that damage to the head of the caudate is especially severe in PD (van Domburg & ten Donkelaar, 1991, pp. 71–72). In fact, because of its reciprocal connections to the prefrontal cortex, many of the well-documented "frontal-like" symptoms of PD might actually be due to damage in the head of the caudate nucleus (e.g., Owen et al., 1992).

Similarly, a wide variety of evidence implicates the basal ganglia in information-integration category learning. First, despite the absence of a PD deficit in the present study, there are a variety of reports that patients with basal ganglia dysfunction, including both PD and Huntington's disease patients, are impaired in difficult information-integration tasks (Filoteo et al., 2001a; Knowlton, Mangels, & Squire, 1996; Knowlton, Squire, et al., 1996; Maddox & Filoteo, 2001). In contrast, medial temporal lobe amnesic patients are normal (Filoteo, Maddox, & Davis, 2001b), except when memorization is a feasible strategy, in which case they show late-training deficits (Knowlton, Squire, & Gluck, 1994). In addition, there are reports that frontal patients are impaired in rule-based tasks but not in information-integration tasks (Knowlton, Mangels, & Squire, 1996). An obvious first hypothesis, therefore, is that the basal ganglia are critical for information-integration category learning but frontal and medial temporal lobe structures are not.

The evidence reviewed here implicates the basal ganglia in both rule-based and information-integration category learning. Even so, given the striking difference between the performance of our PD group on these tasks, it seems plausible that different regions within the basal ganglia may be implicated in these two types of category learning. We argued above that the head of the caudate nucleus is critical in rule-based learning. Neuroanatomical and behavioral neuroscience data indicate that the tail of the caudate nucleus may be critical in information-integration tasks (at least with visual stimuli).

In primates, all of the extrastriate visual cortex projects directly to the tail of the caudate nucleus, and the cells in this area then project, via the globus pallidus (the output portion of the basal ganglia) and thalamus, to the prefrontal and premotor cortices. These projections place the caudate in an ideal position to link percepts and actions (e.g., Rolls, 1994; Wickens, 1993), which is one reason many theorists believe that the basal ganglia mediate procedural learning (Jahanshahi, Brown, & Marsden, 1992; Mishkin, Malamut, & Bachevalier, 1984; Saint-Cyr, Taylor, & Lang, 1988; Willingham, Nissen, & Bullemer, 1989). In support of this hypothesis, a long series of lesion studies in rats and monkeys shows that the tail of the caudate nucleus is both necessary and sufficient for normal visual-discrimination learning (e.g., Eacott & Gaffan, 1992; Gaffan & Eacott, 1995; Gaffan & Harrison, 1987; McDonald & White, 1993, 1994; Packard, Hirsh, & White, 1989; Packard & McGaugh, 1992). For example, many of these studies showed 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. 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. Technically, such studies are categorization tasks with one exemplar per category, but it is difficult to imagine how adding more exemplars to each category could alleviate the deficits caused by caudate lesions. For these reasons, Ashby et al. (1998) proposed that informationintegration category learning is dominated by a proceduralmemory based system that is largely mediated within the tail of the caudate nucleus (see also Ashby & Waldron, 1999; Ashby, Waldron, Lee, & Berkman, 2001).

Taken together, the present results and those of past studies strongly suggest that different brain systems mediate different types of category learning. In the case of rulebased category learning, it is proposed that the prefrontal cortex and head of the caudate are primarily involved, whereas in information-integration category learning, the tail of the caudate may play a crucial role.

Conclusion

To our knowledge, this is the first attempt to compare the ability of the same PD patients to learn rule-based (i.e., explicit) and information-integration (i.e., implicit) category structures. When compared with a group of OCs, PD patients were highly impaired when category membership was defined by an explicit rule but were not different from the OC group when membership was defined by a rule that was difficult to verbalize and required integrating information about the stimulus components at a predecisional stage. This dissociation, and especially the fact that the PD patients were impaired on the easier rule-based task but not on the more difficult information-integration task, is consistent with the hypothesis that (a) there are multiple categorylearning systems and (b) the brain regions affected in PD (i.e., frontal cortex and caudate nucleus) are likely part of these multiple systems.

We believe these results are intriguing and highly suggestive. Even so, this was a single study with relatively limited samples of highly educated PD patients and OCs. Therefore, before any stronger conclusions can be drawn, replications are needed with other diverse groups of participants and with categories containing qualitatively different types of stimuli (e.g., auditory stimuli or visual stimuli that vary on different dimensions than those used here). In addition, calibration studies are needed that directly compare and contrast the task used here with standard neuropsychological assessments (e.g., the WCST).

References

- Alfonso-Reese, L. (1997). Dynamics of category learning. Dissertation Abstracts International: B. The Physical Sciences and Engineering, 57 (9-B), 5929. (UMI No. AAM9704301)
- Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, 105, 442–481.
- Ashby, F. G., & Ell, S. W. (2001). The neurobiology of human category learning. *Trends in Cognitive Sciences*, 5, 204–210.
- Ashby, F. G., & Ell, S. W. (2002a). Single versus multiple systems of category learning: Reply to Nosofsky and Kruschke (2001). *Psychonomic Bulletin and Review*, 9, 175–180.
- Ashby, F. G., & Ell, S. W. (2002b). Single versus multiple systems of learning and memory. In J. Wixted & H. Pashler (Eds.), *Stevens' handbook of experimental psychology: Vol. 4. Methodology in experimental psychology* (3rd ed., pp. 655–691). New York: Wiley.
- Ashby, F. G., Ell, S. W., & Waldron, E. M. (2002). Procedural learning in perceptual categorization. Manuscript submitted for publication.
- 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., Isen, A. M., & Turken, A. U. (1999). A neuropsychological theory of positive affect and its influence on cognition. *Psychological Review*, 106, 529–550.
- Ashby, F. G., Maddox, W. T., & Bohil, C. J. (2002). Observational versus feedback training in rule-based and information-integration category learning. *Memory and Cognition*, 30, 666–677.
- Ashby, F. G., Queller, S., & Berretty, P. T. (1999). On the dominance of unidimensional rules in unsupervised categorization. *Perception and Psychophysics*, 61, 1178–1199.
- Ashby, F. G., & Waldron, E. M. (1999). On the nature of implicit categorization. *Psychonomic Bulletin and Review*, 6, 363–378.
- Ashby, F. G., Waldron, E. M., Lee, W. W., & Berkman, A. (2001). Suboptimality in human categorization and identification. *Journal of Experimental Psychology: General*, 130, 77–96.
- Brainard, D. H. (1997). Psychophysics software for use with MATLAB. Spatial Vision, 10, 433–436.
- Brink, T. L., Yesavage, J. A., Lum, O., Heersema, P. H., Adey, M., & Rose, T. S. (1982). Screening tests for geriatric depression. *Clinical Gerontologist*, 1, 27–43.

- Brown, R. G., & Marsden, C. D. (1988). Internal versus external cures and the control of attention in Parkinson's disease. *Brain*, 111, 323–345.
- Bruner, J. S., Goodnow, J., & Austin, G. (1956). A study of thinking. New York: Wiley.
- Calne, D. B., & Langston, J. W. (1983). Aetiology of Parkinson's disease. *Lancet*, 2, 1457–1459.
- 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.
- 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.
- Dubois, B., & Pillon, B. (1997). Cognitive deficits in Parkinson's disease. *Journal of Neurology*, 244, 2–8.
- Eacott, M. J., & Gaffan, D. (1992). Inferotemporal-frontal disconnection: The uncinate fascicle and visual associative learning in monkeys. *European Journal of Neuroscience*, 4, 1320–1332.
- El-Awar, M., Becker, J. T., Hammond, K. M., Nebes, R. D., & Boller, F. (1987). Learning deficit in Parkinson's disease: Comparison with Alzheimer's disease and normal aging. *Archives of Neurology*, 44, 180–184.
- 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. *Journal of Experimental Psychology: Gen*eral, 127, 107–140.
- 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., Rilling, L. M., Cole, B., Williams, B. J., Davis, J. D., & Roberts, J. W. (1997). Variable memory profiles in Parkinson's disease. *Journal of Clinical and Experimental Neuropsychology*, 19, 878–888.
- 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 striataum. *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.
- Goldman-Rakic, P. S. (1987). Circuitry of the prefrontal cortex and the regulation of behavior by representational knowledge. In F. M. Plum, (ED.), *Handbook of physiology: Vol. 5. The nervous* system: Higher functions of the brain (pp. 373–417). Bethesda, MD: American Psychological Society.
- Goldman-Rakic, P. S. (1995). Cellular basis of working memory. *Neuron*, 14, 477–485.
- Heaton, R. K. (1981). Wisconsin Card Sorting Test. Odessa, FL: Psychological Assessment Resources.
- Hoehn, M. M., & Yahr, M. D. (1967). Parkinsonism: Onset, progression, and mortality. *Neurology*, 17, 427–442.
- 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.

- 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.
- Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996, September 6). A neostriatal habit learning system in humans. *Science*, 273, 1399–1402.
- Knowlton, B. J., Squire, L. R., & Gluck, M. A. (1994). Probabilistic classification learning in amnesia. *Learning and Mem*ory, 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. *Neuropsychol*ogy, 10, 538–548.
- 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.
- Maddox, W. T., Ashby, F. G., & Bohil, C. J. (2002). *Delayed feedback effects on rule-based and information-integration category learning*. Manuscript submitted for publication.
- 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 Soci*ety, 7, 710–727.
- Maddox, W. T., Filoteo, J. V., Delis, D. C., & Salmon, D. P. (1996). Visual selection attention deficits in patients with Parkinson's disease: A quantitative model-based approach. *Neuropsychology*, 10, 187–218.
- Mattis, S. (1988). Dementia Rating Scale. Odessa, FL: Psychological Assessment Resources.
- 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.
- 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.
- Monchi, O., Petrides, M., Petre, V., Worsley, K., & Dagher, A. (2001). Wisconsin card sorting revisited: Distinct neural circuits participating in different stages of the task identified by eventrelated functional magnetic resonance imaging. *Journal of Neuroscience*, 21, 7733–7741.
- Mortimer, J. A., Jun, S.-P., Kuskowski, M. A., & Webster, D. D. (1987). Subtypes of Parkinson's disease defined by intellectual impairment. *Journal of Neural Transmission*, 24, 101–104.
- Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*, 289–316.
- Owen, A. M., James, M., Leigh, P. N., Summers, B. A., Marsden, C. D., Quinn, N. P., et al. (1992). Fronto-striatal cognitive deficits at different stages of Parkinson's disease. *Brain*, 115, 1727–1751.

- 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. *Behavioral Neuroscience*, 106, 439–446.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442.
- Pickering, A. D. (1997). New approaches to the study of amnesic patients: What can a neurofunctional philosophy and neural network methods offer? *Memory*, 5, 255–300.
- Posner, M. I., & Petersen, S. E. (1990). Attention systems in the human brain. Annual Review of Neuroscience, 13, 25–42.
- Rao, S. M., Bobholz, J. A., Hammeke, T. A., Rosen, 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.
- 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. *Journal of Consulting and Clinical Psychology*, 48, 605–614.
- Rolls, E. T. (1994). Neurophysiology and cognitive functions of the striatum. *Revue Neurologique*, 150, 648-660.
- Saint-Cyr, J. A., Taylor, A. E., & Lang, A. E. (1988). Procedural learning and neostriatal dysfunction in man. *Brain*, 111, 941– 959.
- Smith, E. E., Jonides, J., & Koeppe, R. A. (1996). Dissociating verbal and spatial working memory using PET. *Cerebral Cortex*, 6, 11–20.
- Smith, E. E., Patalano, A. L., & Jonides, J. (1998). Alternative strategies of categorization. *Cognition*, 65, 167–196.
- Soliveri, P., Brown, R. G., Jahanshahi, M., Caraceni, T., & Marsden, C. D. (1997). Learning manual pursuit tracking skills in patients with Parkinson's disease. *Brain*, 120, 1325–1337.
- Thomas-Ollivier, V., Reymann, J. M., Le Moal, S., Schueck, S., Lieury, A., & Allain, H. (1999). Procedural memory in recentonset Parkinson's disease. *Dementia and Geriatric Cognitive Disorders*, 10, 172–180.
- van Domburg, P. H. M. F., & ten Donkelaar, H. J. (1991). *The human substantia nigra and ventral tegmental area*. Berlin, Germany: Springer-Verlag.
- Waldron, E. M., & Ashby, F. G. (2001). The effects of concurrent task interference on category learning. *Psychonomic Bulletin* and Review, 8, 168–176.
- Wickens, J. (1993). A theory of the striatum. New York: Pergamon Press.
- Willingham, D. B., Nissen, M. J., & Bullemer, P. (1989). On the development of procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 1047–1060.
- Winocur, G., & Eskes, G. (1998). Prefrontal cortex and caudate nucleus in conditional associative learning: Dissociated effects of selective brain lesions in rats. *Behavioral Neuroscience*, 112, 89–101.

Received October 16, 2001 Revision received July 3, 2002 Accepted July 25, 2002