Rule-Based Category Learning is Impaired in Patients with Parkinson's Disease but not in Patients with Cerebellar Disorders

W. Todd Maddox[1], Paul Aparicio[2], Natalie L. Marchant[2], and Richard B. Ivry[2]

Abstract

■ The basal ganglia and cerebellum have both been implicated in motor skill acquisition. Recent hypotheses concerning cognitive functions of the basal ganglia and cerebellum have emphasized that these subcortical structures may also contribute to nonmotor learning. To explore this issue, patients with Parkinson's disease (PD) and patients with cerebellar lesions (CB) were tested on two category-learning tasks. Identical stimulus displays were used for the two tasks, consisting of a reference line and target line. In the length task, the two categories were defined based on the length of the target line. In the distance task, the two categories were defined by the distance between the target and reference lines. Thus, both categories could be defined by a simple rule in which attention must be restricted to a single relevant dimension. Consistent with previous results, the patients with PD were impaired on both tasks compared with neurologically healthy controls. In contrast, the CB patients performed similar to the control participants. Model-based analyses indicate that the patients with PD were able to select the appropriate categorization rule, but that they adopted suboptimal category boundaries in both conditions and were more variable in the application of the selected rule. These results provide an important neuropsychological dissociation on a non-motor-learning task between the effects of basal ganglia and cerebellar lesions. Moreover, the modeling work suggests that at least part of the Parkinson patients' impairment on these tasks reflect a tendency to exhibit strong response biases. ■

INTRODUCTION

People make thousands of categorization judgments on a daily basis with remarkable accuracy. Extensive study over the past 50 years has led to sophisticated psychological models of category learning and representation (e.g., Estes, 1994; Ashby, 1992; Smith & Medin, 1981). More recently, researchers have begun to apply the methods of cognitive neuroscience, seeking new tools to assess the validity of the psychological models, as well as identify the neural substrates associated with these important cognitive processes (Maddox & Filoteo, 2001; Poldrack, Prabhakaran, Seger, & Gabrieli, 1999; Ashby, Alfonso-Reese, Turken, & Waldron, 1998). We extend this line of work in the current study, assessing the performance of patients with Parkinson's disease (PD) or cerebellar lesions (CB) in rule-based category learning.

The basal ganglia and cerebellum have long been hypothesized to be critical components of the neural networks involved in skill acquisition (see Doyon et al., 2003, for a concise review). For example, patients with basal ganglia degeneration due to PD or Huntington's disease show reduced learning on a variety of motor learning tasks such as pursuit tracking (Haaland, Harrington, O'Brien, & Hermanowicz, 1997; Soliveri, Brown, Jahanshahi, Caraceni, & Marsden, 1997), force field adaptation (Krebs, Hogan, Hening, Adamovich, & Poizner, 2001; Smith, Brandt, & Shadmerh, 2000), and the serial reaction time task (Helmuth, Mayr, & Daum, 2000; Jackson, Jackson, Harrison, Herderson, & Kennard, 1995; Pascual-Leone et al., 1993). Similarly, several studies have shown that patients with cerebellar degeneration or focal cerebellar lesions are severely impaired on the serial reaction time task (Shin & Ivry, 2003; Molinari et al., 1997; Pascual-Leone et al., 1993) as well as a host of other tasks requiring sensorimotor adaptation (e.g., Martin, Keating, Goodkin, Bastian, & Thach, 1996; Sanes, Dimitrov, & Hallett, 1990).

Over the past decade, considerable effort has been devoted to exploring whether these subcortical structures also contribute to nonmotor learning (see Packard & Knowlton, 2002; Ullman, 2001; Schmahmann, 1997). This effort reflects several converging lines of evidence suggesting that the functional domain of the basal ganglia and cerebellum be extended beyond sensorimotor control and learning. Detailed neuroanatomical studies, made possible through the use of multisynaptic

[1]University of Texas, [2]University of California

tracings, have clearly demonstrated that both the basal ganglia and cerebellum have extensive reciprocal connections with association cortices, including various regions of the prefrontal cortex (Middleton & Strick, 2001, 2002). In addition, neuroimaging studies consistently show that activation patterns within the basal ganglia and cerebellum cannot be accounted for by simply considering the motor requirements of different tasks (e.g., Shohamy et al., 2004; Desmond, Gabriele, Wagner, Ginier, & Glover, 1997). Whereas a wide range of functional hypotheses have been considered, a possible role in nonmotor learning is appealing given that both subcortical structures have been hypothesized to play a prominent role in reinforcement (reward or error-based) processes.

Patients with degenerative disorders of the basal ganglia have been tested on a range of category-learning tasks (for a review, see Keri, 2003; Ashby & Ell, 2001). This article focuses on rule-based category learning. Rule-based category-learning 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 the most common applications, only one stimulus dimension is relevant, and the subject's task is to discover this relevant dimension and then to map the different dimensional values to the relevant categories.[1]

Ashby et al. (2003) tested patients with PD on two rule-based categorization tasks. The stimulus set was created from the factorial combination of four binary–valued dimensions. In the first condition, one of the dimensions was selected to serve as the relevant dimension, with one value assigned to one response category and the other value assigned to the other response category; variation on the other three dimensions was irrelevant across trials. On each trial participants were shown a stimulus and categorized it by pressing one of two response keys. The response was followed by feedback to allow for trial-and-error learning. Once the participant generated 10 correct responses in a row or completed 200 trials a new dimension was chosen randomly to be relevant and the participant again continued until they generated 10 correct responses in a row or completed 200 trials. A learner was defined as a participant who learned both rules within 200 trials. Whereas nearly 90% of the healthy controls learned both rules within 200 trials, only 50% of the patients with PD met this criterion suggesting a severe impairment in rule-based category learning in PD.

The impairment on this task is consistent with a detailed neurobiological model of category learning proposed by Ashby et al. (1998). The competition between verbal and implicit systems or COVIS model, postulates an important role for the basal ganglia in rule-based category learning. In COVIS, rule-based category learning involves the conscious generation and testing of hypotheses, operations attributed to interactions between prefrontal cortex, anterior cingulate, and the head of the caudate nucleus.

The conclusion that patients with PD are impaired in rule-based category learning must be qualified. Maddox and Filoteo (2001) created a stimulus set in which each stimulus was composed of a vertical and horizontal line. In the rule-based condition, one categorization response was associated with stimuli for which the horizontal segment was longer than the vertical segment, and the other categorization response was associated with stimuli for which the horizontal segment was shorter than the vertical segment. The patients with PD performed as well as the controls across an entire session of 600 trials. Taken together, the results from these two studies indicate that patients with PD may have problems with rule-based categorization, but the conditions under which such deficits become apparent remains to be determined.

To date, neuropsychological studies of category learning have focused on the basal ganglia. No studies have examined the performance of patients with CB on category-learning tasks. In part, this reflects the fact that models of categorization such as COVIS only include the basal ganglia as a subcortical component of the learning network. Pragmatically, patients with PD are generally more prevalent and easier to recruit for neuropsychological studies than patients with cerebellar disorders.

Nonetheless, there are at least two important reasons for testing patients with CB on tasks similar to those used in the PD studies reported above. First, if mechanisms for cognitive learning are similar to those for motor learning, then given the role of the cerebellum in motor learning, one might expect this structure to provide a similar contribution during non-motor-learning tasks, especially when considering the extensive projections from and to the prefrontal cortex (Middleton & Strick, 2002). Indeed, the effects of cerebellar damage on nonmotor associative learning have been examined in a number of studies. Canavan, Sprengelmeyer, Diener, and Homberg (1994) and Bracke-Tolkmitt et al. (1989) used a paired-associate task in which participants were required to learn an arbitrary S:R mapping between words and colors. Patients with either cerebellar degeneration or focal cerebellar lesions performed quite poorly on this task. Similarly, Fiez, Petersen, Cheney, and Raichle (1992) report a case study of an individual who suffered a focal lesion of the right cerebellar hemisphere who was impaired on a similar task (but see Helmuth, Ivry, & Shimizu, 1997). Paired-associate learning tasks place considerable demands on working memory processes for maintaining the arbitrary S:R associations. Such processes are likely to involve the generation of explicit hypotheses, similar to what is assumed to occur during rule-based categorization.

Second, the focus on patients with PD to assess subcortical contributions to category learning involves a strategy of seeking confirmatory evidence. As with all

neuropsychological research, dissociations provide much more compelling evidence. It is possible that the categorization deficits in the patients with PD are specifically related to the pathology associated with this disease (and assumed to be related to the dopamine loss in the basal ganglia). However, it is possible that the deficits reflect a more generic reduction in cognitive function related to neurological disease. Patients with CB provide an opportunity to evaluate the performance of a different neurological group on the same set of tasks. Obtaining similar patterns of deficits in the CB group would constrain the interpretation of patient data regarding the role of the basal ganglia in category learning.

Alternatively, obtaining dissociations between the performance of patients with PD and CB lesions would help specify the manner in which these two subcortical structures contribute to category learning. Indeed, a direct comparison of patient groups has proven useful in previous work from our laboratory (Diedrichsen, Ivry, & Pressing, 2003; Ravizza & Ivry, 2001). For example, separate lines of research suggested that a nonmotor function of the basal ganglia and cerebellum involved coordinating rapid shifts of attention (i.e., task switching). By evaluating patients with either PD or cerebellar pathology on the same tasks, Ravizza and Ivry (2001) showed that PD patients' deficit on such tasks was specifically related to shifting attention, whereas the cerebellar patients' deficit was related to a general resource issue when response requirements were high. We apply a similar strategy in the current study to investigate the role of the basal ganglia and cerebellum in category learning.

Two rule-based category learning conditions were examined in the current study. In both conditions, the stimulus display consisted of two horizontal lines, a ''target'' line and a ''reference'' line (Figure 1). Two parameters were varied, the length of the target line and the vertical distance between the target and reference lines. In the length condition, the optimal decision rule required participants to learn to respond ''A'' when the target line length was shorter than the criterion and respond ''B'' when the target line length was longer than the criterion. The vertical distance of the target line from the reference line was irrelevant. In the distance condition, the optimal decision rule required participants to learn to respond ''A'' when the distance between the target and reference lines was less than the criterion and respond ''B'' when this distance was greater than the criterion, ignoring the variation in the length of the target line (see Figure 2). Through trial and error, participants had to identify the relevant dimension and learn the optimal decision criterion dividing the two categories.

The target line was always centered on the vertical meridian of the computer monitor. In contrast, the vertical position of the reference line was varied across trials. Thus, in the distance condition, the target line might appear at the same absolute vertical position on different trials but require different classifications depending on the position of the reference line. In this case, the participant must compare the perceived distance with some internally represented critical distance. The horizontal center of the target line, on the other hand, was fixed across trials. Thus, the participant could use an external reference: the absolute position of the endpoints of the target line, or they could compare the perceived target length with an internally represented critical length.

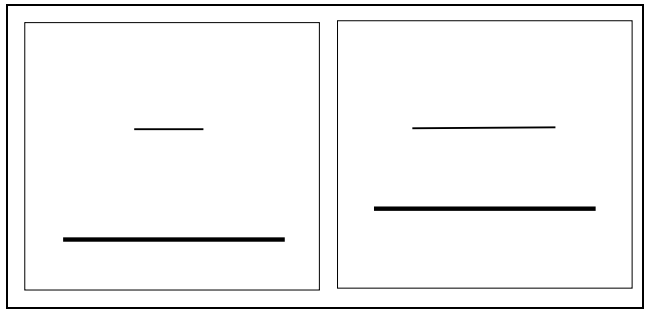

Figure 1. Two representative stimulus displays. The reference line was always positioned below the target line and double the thickness of the target line. Note that in these examples the absolute position of the target line is the same. However, in the distance condition, a correct response would be ''far'' for the display on the left and ''near'' for the display on the right. Thus, the categorization requires reference to an internal, distance criterion. In the length condition, the judgment could be based on an internal reference (of length) or an external reference based on the absolute position of the endpoints because the target line was always centered on the screen.

The motivation for these two tasks comes from previous theorizing in the motor control literature regarding the differential role of the basal ganglia and cerebellum in the control of externally and internally guiding movements. Goldberg (1985) introduced this idea based on his evaluation of the gross neuroanatomy of the motor systems. In his model, actions could be controlled through two primary circuits. Visually guided (external) actions primarily rely upon the parietal cortex, cerebellum, lateral premotor cortex, and motor cortex, whereas internally guided actions involve a circuit composed of the prefrontal cortex, supplementary motor area, basal ganglia, and motor cortex. Although anatomical work demonstrates that the basal ganglia are broadly innervated by the cerebral cortex, including the parietal and lateral premotor regions (reviewed in Middleton & Strick, 2000; Selemon & Goldman-Rakic, 1985), neuroimaging studies suggest that the basal ganglia are disproportionately engaged during internally guided actions (e.g., Debaere, Wenderoth, Sunaert, Van Hecke, & Swinnen, 2003; Grafton et al., 1992). Moreover, behavioral studies suggest that movement disorders associated with PD are ameliorated when salient visual cues are provided to guide the actions (Flowers, 1976).

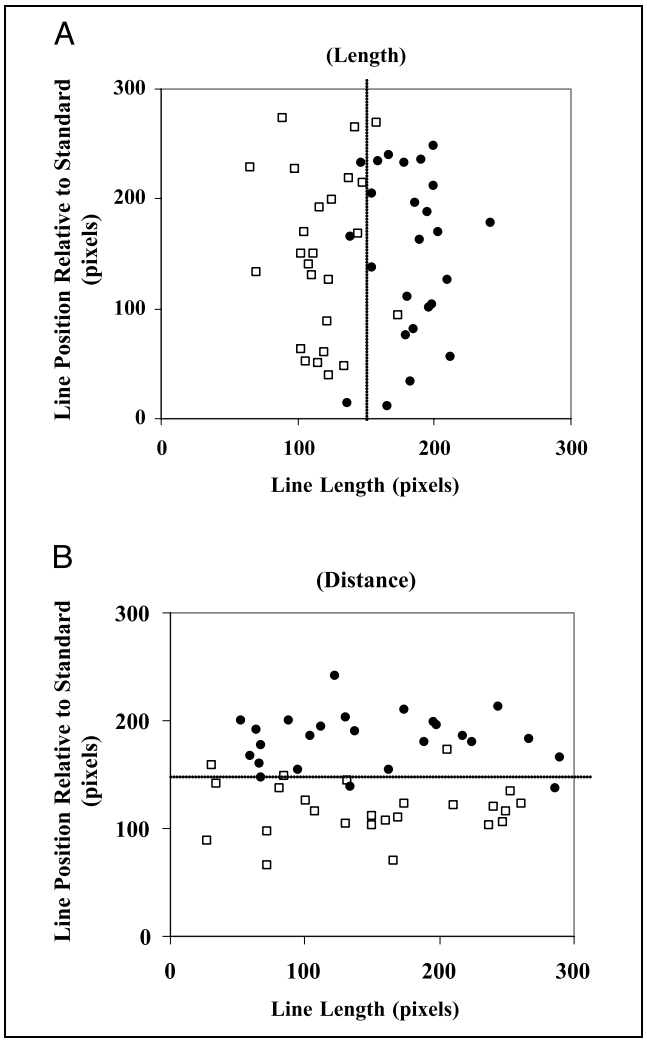

Figure 2. (A) Length-relevant and (B) distance-relevant condition stimuli. Each open square denotes the length and vertical position relative to a standard from Category A. Each filled circle denotes the length and vertical position relative to a standard from Category B. The broken line in each panel denotes the location of the optimal decision bound.

The internal/external distinction has helped account for some discrepant findings in studies of visual attention, with PD patients showing deficits when attention is internally driven but not when attention is externally driven (Brown & Marsden, 1988). If cognitive functions of the basal ganglia and cerebellum parallel motor functions, then we might expect to observe a similar distinction in category learning. Specifically, patients with PD may be disproportionately impaired on the distance condition given that the categorization judgments require the use of an internal referent. Patients with CB may be disproportionately impaired on the length condition given that these judgments can be based solely on information provided in the stimulus display.

An optimal experimental design would have involved two additional conditions: length judgments requiring an internal reference (e.g., varying the horizontal position of the target line) and distance judgments that could be based on an external reference (e.g., fixing the vertical position of the reference line). However, given that we did not want to subject the patient participants to excessive testing on these demanding (and frustrating) tasks, we opted to confound the two categorization rules with the internal/external variable. Nonetheless, the tasks were designed to make the two conditions as similar as possible. In addition to using the same stimulus dimensions in both conditions, we equated several structural aspects of the categories in an effort to make the conditions of comparable complexity (e.g., optimal accuracy, within-cluster scatter, and cluster coherence). This form of stimulus control has not been used in previous neuropsychological studies of category learning. By using stimulus sets in which the value on each dimension varies continuously (Figure 2), we can apply a rich set of model-based techniques to analyze the data.

There are several distinct reasons why performance might be suboptimal on categorization tasks. First, a participant might fail to adopt the appropriate categorization rule. With the tasks used in the current study, these rules require that attention be focused on a single dimension. Influence of the irrelevant dimension on performance would suggest a failure of selective attention, and therefore would be a suboptimal strategy. Second, the appropriate rule might be selected, but the decision bound might be suboptimal. Such a result would suggest that the impairment reflects a form of bias such as a preference for one response over the other. Third, even if the person uses the appropriate rule and decision bound, success requires that these representations be applied in a consistent manner. To date, neuropsychological studies of category learning have focused on the question of whether or not a particular patient group (and associated neural structure) is impaired at these tasks. Our model-based analyses offer the opportunity to determine the source of any observed deficits, thus developing functional accounts for the contribution of either the basal ganglia or cerebellum to category learning.

RESULTS AND THEORETICAL ANALYSES

Accuracy

Control Participants

We begin by examining category learning for both the length and distance conditions in the 14 control participants. Not only is it important to verify that neurologically healthy participants can learn these tasks, but we also wished to assess whether the two category-learning tasks were of comparable difficulty. These data are displayed in Figure 3. Notice that control participants performance never approached the optimal level of 90%; asymptotic performance was around 75% for both conditions. Nonetheless, there is clear evidence of learn-

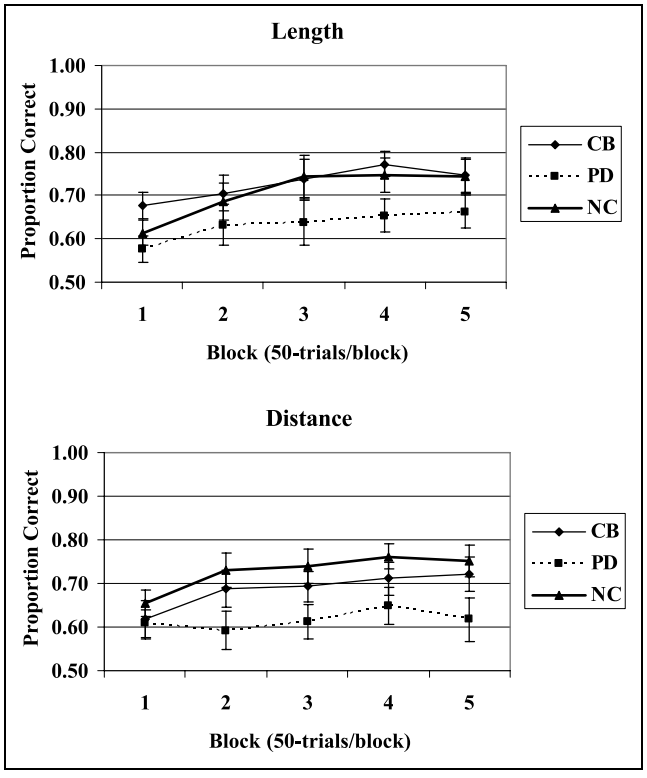

Figure 3. Proportion correct for the length and distance category structures for the CB, PD, and normal control (NC) participant groups for each block (standard error bars included).

ing as evidenced by a significant block effect, $F(4,52) = 9.00$, $p < .001$. The effects of condition and the Condition × Block interaction did not approach significance (Fs <1). Thus, although learning occurred for both tasks, the results confirm the participants' phenomenal reports that the task was very challenging. The group data obscure the fact that there were rather large individual differences. We adopted a criterion of 55% correct over the final 100 trials as a way to make a binary discrimination between learners and nonlearners. By this definition, three control participants failed to show learning in the length condition, one of whom also failed to learn in the distance condition (Table 1). In contrast, three participants approached optimal performance (85% correct on final 100 trials) on both tasks and an additional two participants approached optimal performance on just the length task and one participant on just the distance task.

We also correlated performance over the final 100 trials across length and distance conditions. A significant positive correlation emerged, $r(12) = .57$, $p < .05$.

Patient Participants

Given that we did not have strong a priori hypotheses concerning differential performance of the PD and CB groups, we chose to compare each patient group with

Table 1. Participant Accuracy and Response Bias

Participant No.	*Acc-L*	*Acc-D*	*Freq Diff-L*	*Freq Diff-D*
Patients with PD				
PD03	0.59	0.52	46	44
PD05	0.56	0.63	32	14
PD06	0.84	0.47	22	70
PD09	0.78	0.64	6	0
PD11	0.58	0.76	20	38
PD12	0.84	0.86	26	50
PD14	0.71	0.77	20	38
PD17	0.50	0.57	70	68
PD18	0.79	0.85	46	12
PD19	0.50	0.35	62	114
PD20	0.48	0.57	28	4
PD23	0.78	0.74	18	40
PD28	0.61	0.49	54	20
Patients with CB				
LC01	0.82	0.87	62	12
LC02	0.84	0.71	36	8
LC03	0.78	0.50	14	10
LC04	0.89	0.87	20	8
LC06	0.65	0.68	56	24
LC07	0.45	0.51	18	22
AC01	0.85	0.84	22	20
AC02	0.68	0.50	18	34
AC04	0.85	0.88	32	14
AC05	0.80	0.70	12	12
AC08	0.61	0.87	32	16
AC09	0.83	0.70	4	26
AC10	0.76	0.68	34	72
AC11	0.83	0.72	26	34
Control participants				
MP01	0.86	0.87	14	24
MP04	0.50	0.77	22	14
MP06	0.80	0.81	32	38
MP07	0.84	0.83	22	28
MP24	0.85	0.56	2	18
OP02	0.58	0.67	18	8
OP04	0.73	0.79	0	12
OP08	0.52	0.68	10	22

Table 1. (*continued*)

Participant No.	*Acc-L*	*Acc-D*	*Freq Diff-L*	*Freq Diff-D*
OP09	0.84	0.87	26	4
OP11	0.86	0.87	22	32
OP14	0.84	0.80	42	8
OP15	0.51	0.48	2	6
OP26	0.85	0.85	8	6
OP27	0.85	0.74	0	48

Acc-L and Acc-D denote accuracy during the final 100 trials in the length and distance conditions, respectively. Freq Diff-L and Freq Diff-D denote the absolute difference between the number of A and B responses across the full 250-trial session in the length and distance conditions, respectively.

the control participants. In this manner, we ask whether patients with either PD or CB lesions are impaired on category learning. If we were to find such differences, then post hoc comparisons between the patient groups could be pursued.

The first analysis involved the comparison of the patients with PD and the controls. Whereas the effect of block was significant, $F(4,100) = 10.24$, $p < .01$, indicative of learning, the patients with PD were significantly less accurate than the controls, $F(1,25) = 4.71$, $p < .05$; mean percent correct: PD, 62.4%; controls, 71.7%. The main effect of category learning condition, the Group × Category Learning Condition interaction, the Category Learning Condition × Block interaction, and the three-way interaction were not reliable (all Fs <1). The Participant Group × Block interaction was also nonsignificant, $F(4,100) = 1.75$, $p = .14$, a result that is somewhat surprising given the nearly flat performance profile for the patients with PD in the distance condition.

Using the 55% criterion, three patients with PD failed to learn in the length condition, a number equal to that observed for the control group. However, four patients with PD failed to reach this criterion on the distance task compared with a single control participant. If we raise this criterion to 65% the difference is magnified. With this criterion, seven patients with PD failed to learn the length discrimination compared with four controls (nonsignificant based on a Fisher's exact test). In the distance condition, eight patients with PD failed to reach 65% correct compared with two controls ($p < .05$ based on a Fisher's exact test). Thus, both tasks proved extremely difficult for many of the patients with PD. We did not observe a difference between the PD patients' performance on the two tasks, although there was a tendency for the deficit to be more pronounced in the distance condition, especially on the binary classification of nonlearners. Unlike the controls, performance did not correlate significantly across the two conditions, although the effect was nearly significant, $r(11) = .52$, $p = .068$.

A very different pattern of results was observed in the performance of the CB group when compared with the control participants. Only the effect of block was significant, $F(4,104) = 14.35$, $p < .01$. The main effect of participant group did not approach significance, $F(1,26) < 1$; indeed, in terms of overall accuracy, the CB participants (70.7%) performed at a level equivalent to that of controls (71.7%). The category learning condition, the Group × Block, category Learning Condition × Block, and three-way interactions all yielded F values less than 1. Only the Participant Group × Learning Condition interaction yielded an F value greater than 1, $F(1,26) = 2.04$, $p = .17$. In terms of the binary classification of nonlearners based on the 55% criterion, one patient with CB failed to learn the length task and three failed on the distance task. When this criterion was raised to 65%, one additional patient with CB was classified as a nonlearner on the length task. Thus, on measures of overall accuracy and our arbitrary classification of nonlearners, the CB group performed comparable to the control participants. Like the patients with PD, performance did not correlate significantly across the two conditions, although the effect was nearly significant, $r(12) = .50$, $p = .066$.

Neuropsychological and Neurological Assessment Correlates with PD Category Learning

Although the patients with PD as a group were impaired on both category-learning tasks, there were considerable individual differences within the PD group. To look at these differences in more detail, we examined whether performance on the category-learning tasks was correlated with measures of cognitive or motor dysfunction as measured by the standardized assessment tools we used to evaluate neuropsychological and neurological impairments. These data are presented in Table 2 and are described in the Methods section. Correlations between these tests and accuracy on each of the two categorization tasks is displayed in Table 3. The only correlation that reached significance was between digit span and the distance condition accuracy, $r(11) = .56$, $p < .05$. This correlation might reflect the involvement of working memory in rule-based category learning given that digit span is thought to assess working memory span (Waldron & Ashby, 2001; Maddox, Ashby, Ing, & Pickering, 2004). However, the correlation between digit span and the length condition was essentially zero, $r(11) = .03$, $p > .05$. A priori, we would expect that the two category-learning tasks place similar demands on working memory.

The patients with PD had also been administered the Hoehn and Yahr Scaling and the Unified Parkinson's Disease Rating Scale (UPDRS). Neither of these measures correlated with accuracy on the category-learning task. The largest correlation was between the Hoehn and Yahr measure and distance condition ac-

Table 2. Participant Demographic Information, Neuropsychological Assessment, and Severity Measures

Participant No.	Age (years)	Handedness	Education (years)	MMSE	NART	WAIS-III		UPDRS/H & Y
						Digit Span[a]	Matrix Reasoning[a]	
Patients with PD								
PD03	74	Right	13	26	106	10	7	–
PD05	88	Right	18	28	122	12	15	48/2.5
PD06	75	Right	16	29	123	9	12	35/1.5
PD09	57	Right	19	30	124	11	12	9/1.0
PD11	74	Right	18	29	125	12	16	42/2.0
PD12	77	Right	20	30	127	13	13	39/2.0
PD14	56	Right	19	30	129	15	13	8/1.5
PD17	66	Right	23	29	128	11	13	30/2.5
PD18	71	Right	16	29	123	10	15	61/3.5
PD19	72	Left	16	29	–	10	17	45/1.5
PD20	71	Right	20	29	130	14	8	25/3.0
PD23	61	Left	21	30	130	14	8	51/3.0
PD28	59	Right	19	29	114	8	7	–
Patients with CB								
LC01	54	Right	13	30	117	13	10	10.00
LC02	66	Right	14	30	111	11	8	4.25
LC03	58	Right	12	30	89	4	10	19.25
LC04	45	Left	18	30	124	9	15	11.00
LC06	77	Right	16	29	112	12	9	34.00
LC07	58	Left	10	25	87	6	7	23.25
AC01	56	Right	18	30	128	14	16	29.00
AC02	79	Right	16	30	121	12	17	22.00
AC04	48	Right	18	28	116	7	12	49.75
AC05	45	Right	13	30	–	–	–	31.00
AC08	50	Right	15	29	122	9	11	20.75
AC09	63	Right	22	29	116	7	11	17.75
AC10	72	Right	12	29	100	6	12	45.00
AC11	43	Right	16	30	109	7	10	35.00
Control participants								
MP01	58	Right	18	30	122	–	–	–
MP04	60	Right	16	30	126	12	12	–
MP06	57	Right	17	30	–	–	–	–
MP07	63	Right	17	30	125	9	13	–
MP24	60	Right	12	28	–	–	–	–
OP02	65	Right	14	30	119	8	9	–

Table 2. (*continued*)

Participant No.	Age (years)	Handedness	Education (years)	MMSE	NART	WAIS-III		UPDRS/H & Y
						Digit Span[a]	Matrix Reasoning[a]	
OP04	68	Right	15	29	124	14	16	–
OP08	60	Right	20	30	116	9	7	–
OP09	64	Right	22	30	121	13	8	–
OP11	68	Right	16	30	127	12	16	–
OP14	74	Right	16	30	125	16	16	–
OP15	68	Right	14	28	102	11	14	–
OP26	75	Right	20	29	126	11	15	–
OP27	71	Right	17	27	127	10	15	–

Dashes indicate data that are not available or not applicable. MMSE = Mini Mental State Exam; NART = National Adult Reading Test; WAIS-III = Wechsler Adult Intelligence Scale – Third Edition; UPDRS = Unified Parkinsonism Disease Rating Scale; H & Y = Hoehn and Yahr Staging Score; International Cooperative Ataxia Rating Scale (ICARS).

[a]Scaled score equivalents of raw scores.

curacy (r = .35), but even here the p value was high (p = .29).

In summary, the correlational analyses suggest that there are no clear predictors from our battery of standardized assessments on category learning performance. There was the one significant correlation between digit span and the performance on the distance condition, and this may reflect a common dependency on working memory processes. However, the fact that a similar pattern was not observed for both categorization tasks is problematic for this hypothesis. It is important to keep in mind that this form of analysis must be treated cautiously here, given the relatively small number of PD participants.

Model-Based Analyses of Rule-Based Category Learning Deficit in PD

Quantitative models have been developed to clarify the operation of various cognitive processes required for successful category learning (see Maddox & Ashby, 1993; Ashby, 1992). Inaccurate performance could result from deficits in a number of different processes. *Categorization rule learning* centers on the participants' ability to learn to partition the perceptual space in accordance with the optimal categorization rule. In the current study, this involves determining whether the participant was able to ignore irrelevant information and apply a rule-based strategy along the relevant dimension. A failure to filter irrelevant information would result in a rule that integrates information from both stimulus dimensions.

Table 3. Neuropsychological Assessment and Rule-Based Category Learning Accuracy Score Correlations in PD

	Acc-L	*Acc-D*
MMSE	0.44	0.41
Digit Span	0.03	0.56*
Matrix Reasoning	−0.21	0.18
NART	0.14	0.48

*p < .05.

To assess rule learning we fit three decision bound models to the data. These decision bound models are derived from general recognition theory (GRT; Ashby & Townsend, 1986), which is a multivariate generalization of signal detection theory (e.g., Green & Swets, 1967). The fundamental assumption of GRT is that there is trial-by-trial variability in the perceptual representation of the stimulus, even with protracted viewing conditions as in the current study. On each trial, it is assumed that the percept can be represented as a point in a multidimensional psychological space. Decision bound theory assumes the participant partitions perceptual space into response regions. On each trial, the participant determines which region the percept is in and then produces the associated response. Despite this deterministic decision rule, decision bound models predict probabilistic responding because of trial-by-trial perceptual and criterial noise. Details of the theory are provided elsewhere (e.g., Maddox & Filoteo, 2001; Ashby, 1992); here we provide a brief overview.

The *optimal rule-based model* assumes that the participant adopts the experimenter-defined (optimal) rule. This is the rule displayed in Figure 2A (for the length condition) and B (for the distance condition). In

the optimal model, the decision process ignores the irrelevant dimension and sets a criterion of 150 pixels (in length for the length condition or distance for the distance condition) along the relevant dimension. The model has one free parameter that denotes trial-by-trial variability in perceptual and criterial noise. The *suboptimal rule-based model* assumes that the participant also selectively attends to the relevant dimension, but the data are used to estimate the participant's decision criterion. Thus, this model has two parameters: one to reflect trial-by-trial variability in perceptual and criterial noise and a second to estimate the decision boundary. The *information-integration model* assumes that the participant's decision on each trial is based on information from both dimensions, although the weighting given to these may be unequal. The model assumes that this integration is linear, and estimates the participant's decision bound slope and intercept (two parameters) from the data. The model also has one free parameter that denotes trial-by-trial variability in perceptual and criterial noise. Because only one stimulus dimension was relevant in each task, information integration is suboptimal because the decision bound is not orthogonal to the relevant dimension.

We first compared the three models to determine which provided the best fit of the data for each participant. The best fit was defined as the simplest model for which a more general model did not provide a statistically significant improvement. As a quantitative index of categorization rule learning we examined how well the optimal rule-based model captured performance using a goodness-of-fit measure for this model (−lnL; negative log likelihood). The smaller the fit, the better the rule describes the data.

We also evaluated the participants' ability to identify the optimal decision criterion. The modeling results from the suboptimal rule-based model were used to estimate each participant's decision criterion and these values were compared with the optimal decision criterion. Individual criteria that deviate from the optimal criterion likely reflect a response bias given that, a priori, we have no reason to expect a preference for one value (small or large in length; near or far in distance) over the other value.

In addition to estimating the rules derived by the participants, we also examined how consistently the participants used their selected rule. Trial-by trial variability in the application of each participant's categorization rule was measured by estimating a parameter in the suboptimal rule-based model called the *rule application variability* parameter (Ashby, 1992). The smaller the magnitude of this parameter, the less variable the participant is in his or her application of the adopted rule. Restricting this analysis to the suboptimal rule is sufficient because the optimal rule is a specific variant (i.e., no weight is assigned to the irrelevant dimension) of the suboptimal rule.

For simplicity, we report only the modeling results for the PD and control participants. Similar analyses were performed on the data from the CB patients, and in all cases, their results were similar to those obtained for the controls. In addition, we only report the modeling results for the PD and control participants who met the 55% performance criterion described above. We would be modeling noise if we used the data from people who never performed above chance, although we did examine their data to see if we could derive any qualitative insights into their performance. A stricter criterion of 65% might be more appropriate. However, this would have led us to exclude almost 50% of the patients with PD. Given this, the 55% criterion is a compromise between modeling ''learners'' and modeling as much of the data as possible. We did replicate all the analyses presented below with the 65% criterion nonetheless, and found that the general pattern continued to hold although the mean differences were somewhat smaller.

Using a 55% criterion resulted in an equal number of PD and control participants being excluded in the length condition. However, this balance is lost in the distance condition: here, four patients with PD are excluded and only one control participant. There are two ways to consider the consequences of excluding nonlearners. First, if one assumes that nonlearners reflect a distinct subpopulation, then the modeling approach examines whether patients with PD who learn perform differently from control participants who learn. Alternatively, the nonlearners may reflect the tail end of a normal distribution. If this is correct, then excluding more patients with PD than control participants results in an unfair statistical comparison. The comparison is of two distributions, but the poorest performing samples from one distribution are excluded. A more appropriate comparison here would be to exclude an equal number of poor-performing control participants. To this end, we perform a secondary statistical comparison of the modeling data by excluding the four controls with the lowest percent correct on the last 100 trials (one of whom failed to reach the 55% criterion). For these comparisons, we refer to the controls as the trimmed control group.

Categorization Rule Learning

Rule-Based or Information-Integration Strategy

The first modeling question centers on whether the participants derived the appropriate unidimensional rule (optimal or suboptimal) or whether they (inappropriately) integrated information from the two stimulus dimensions. To this end, we identified the best fitting model for each participant using likelihood ratio tests for nested models. The results were clear. In the length condition, a rule-based model provided the best fit of the data sets for 84% of the patients with PD and 80% of the control participants. Values for the distance condi-

tion were 84% and 82% for the PD and control participants, respectively. These results suggest that patients with PD were as proficient as the control participants in deriving a categorization rule that was based solely on the relevant dimension, although almost 20% of the participants in each group failed to adopt a unidimensional rule.[2]

Categorization Rule Learning Index

Figure 4 displays the categorization rule learning index for the length (panel a) and distance (panel b) conditions for the PD and control participants (averaged across observers) for each of the five blocks of trials. Analyses were performed separately on the length and distance condition data. To determine how well participants learned the optimal rule, we focused on the goodness-of-fit values (or categorization rule learning index) for the optimal rule-based model. A 2 (group, PD vs. normal controls [NC]) × 5 (50-trial) block mixed-design ANOVA was performed on these values. For the length condition, the main effect of group was nonsignificant, $F(1,19) = 3.71$, $p = .07$, the main effect of block was significant, $F(4,76) = 19.83$, $p < .001$, and the participant group by block interaction was significant, $F(4,76) = 2.99$, $p < .05$. *t* tests performed on a block-by-block basis suggested that the control participants were better at using the optimal categorization rule during blocks 3 ($p = .07$), 4 ($p < .05$), and 5 ($p < .05$).

For the distance condition, the main effect of block was significant, $F(4,80) = 10.17$, $p < .001$, but the participant group effect, $F(1,20) = 2.17$, $p = .16$, and the interaction were nonsignificant, $F(4,80) = 1.08$, $p > .10$. When we compare the patients with PD with the trimmed control group (see Figure 4B), the effect of participant group, $F(1,17) = 8.66$, $p < .01$, and block, $F(4,68) = 12.07$, $p < .001$, were both significant, but the interaction was nonsignificant, $F(4,68) = 2.07$, $p = .09$.

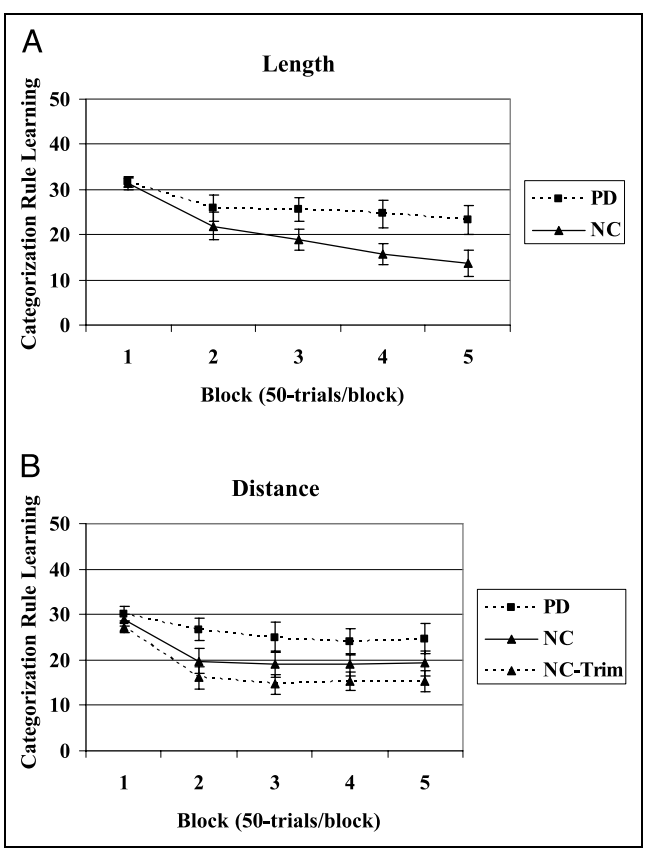

Figure 4. Categorization rule learning index for the (A) length and (B) distance category structures for the PD, NC, and NC-Trim (see text for details) participant groups for each block (standard error bars included).

Optimality of Decision Criterion Learning

The preceding analyses suggest that patients with PD were as likely to select a task-relevant rule as control participants, but their use of this rule was less optimal, especially in the distance condition. One way in which the rule might be used nonoptimally is in the placement of the decision criterion. To assess this question, we examined the decision criterion estimates from the suboptimal rule-based model. Specifically, we investigated the absolute deviation between the best fitting decision criterion and the optimal decision criterion. By using the absolute deviation, the analyses assess the extent of a response bias independent of the direction of that bias. These values are plotted separately for the length and distance conditions in Figure 5.

For the length condition, the main effect of block was significant, $F(4,76) = 4.62$, $p < .01$, suggesting that the magnitude of the participants' response bias declined with experience, but the participant group effect, $F(1,19) = 1.91$, $p = .18$, and the interaction ($F < 1.0$) were both nonsignificant. For the distance condition, neither of the main effects, group: $F(1,20) = 1.29$, $p > .10$; block: $F(4,80) = 1.28$, $p > .01$; nor their interaction ($F < 1.0$) were significant. When the patients with PD were compared with the trimmed control group, the effect of participant group became significant, $F(1,17) = 6.86$, $p < .05$, yielding a response bias of 32 pixels for the patients with PD and 11 pixels for the controls. The block effect, $F(4,68) = 1.81$, $p > .01$, and interaction ($F < 1.0$) effects remained nonsignificant.

Including all participants who reached the 55% criterion, the response bias in the length condition was 36 pixels for the PD participants and 20 pixels for the control participants, and in the distance condition the bias was 32 pixels for the PD participants and 20 pixels for the control participants. Thus, patients with PD tended to exhibit larger response biases than the control participants. This response bias effect is also observed if the analysis is restricted to only include participants for whom the modeling results indicated use of a rule-based strategy.

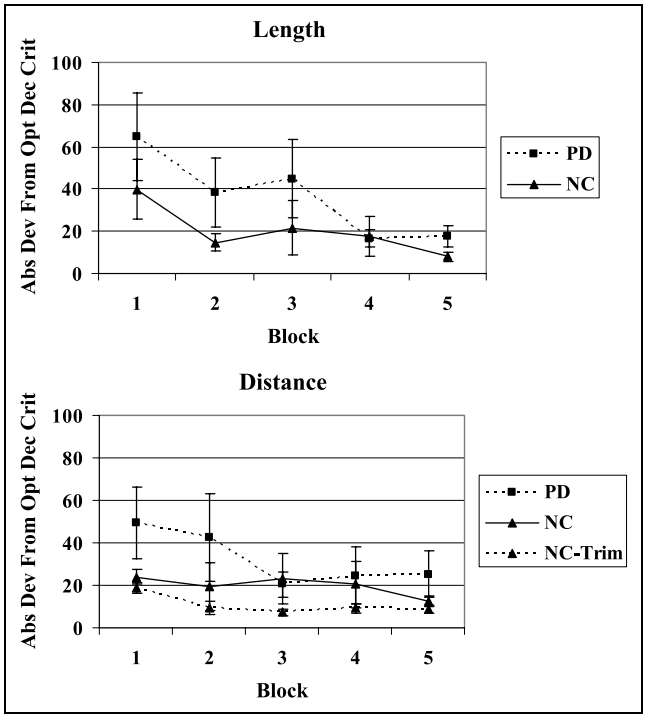

Figure 5. Absolute deviation from the optimal decision criterion for the length and distance category structures for the PD, NC, and NC-Trim (see text for details) participant groups for each block (standard error bars included).

We addressed two additional questions regarding the response bias. First, did the bias remain constant for each individual across blocks? We determined the direction of the decision criterion deviation for each participant on each block, then counted the number of times that the direction of the bias changed. This ''change'' score has a lower bound of 0 for a participant who consistently used a decision criterion smaller (or larger) than the optimal criterion, and an upper bound of 4 for a participant whose decision bound changed across every block. We computed the average change score across participants for each condition. For the controls, the average change score was 1.00 and 1.67 in the length and distance conditions, respectively. For the patients with PD, the average change score was 1.20 and 1.00 in the length and distance conditions, respectively. These scores suggest that biases tended to persist across blocks and, if anything, may have been more stable for the patients with PD than the controls in the distance condition. It should be noted, however, that the controls were close to the optimal decision bound. Given that the change score does not take into account the magnitude of the deviation, the score might be inflated by small fluctuations around the optimal value.

Second, was the bias consistently toward one response over the other? Six of the 10 PD learners showed a bias to respond ''longer'' on the length task, and 5 of the 9 PD learners showed a bias to respond ''farther'' on the distance task. Neither of these was significant based on a binomial test ($p > .05$), consistent with our assumption that the nonoptimal decision bounds reflect response biases rather than some sort of perceptual distortion.

In summary, patients with PD showed larger response biases than control participants. This finding explains, at least in part, their deficit on these categorization tasks. It is also possible that a response bias may account for why some participants fail to learn these tasks. In the extreme, a bias to use only one response key would result in overall performance of 50%. This was not the case for any of the nonlearners. Nonetheless, the patients with PD (and controls) who failed to learn the task did exhibit large response biases even though equal numbers of A and B stimuli were used in the task. To examine this possibility quantitatively, we computed the absolute difference between the A and B response frequencies over the full 250-trial session. This measure of response bias has a lower bound of 0 (when 125 A and 125 B responses are generated), and an upper bound of 250 (when either all A or all B responses are generated).

The response bias scores for all of the participants are presented in Table 1. For the length condition, the response bias scores averaged 53.3 for the three patients with PD who failed to reach the 55% criterion. This value is much larger than the 11.3 average for the three controls who failed to learn. The four PD nonlearners in the distance condition also exhibited a large response bias, with a mean score of 62.0 (compared with a response bias score of 6 for the one control participant who did not reach criterion in this condition). This simple measure of response bias also confirms the modeling results for the participants who reached the 55% learning criterion, although the effect is not as dramatic. For the learners, the mean response bias scores for the PD participants were 29.0 and 28.0 for the distance and length conditions, respectively. For the control participants, these values were 16.9 and 20.1. Thus, the results suggest that a susceptibility to response biases may be a core component of the category learning deficit in PD. Those who fail to learn exhibit pronounced biases, and those who did learn are suboptimal in their performance because they use an inappropriate (i.e., biased) decision bound.

Rule Application Variability

The application of an appropriate rule may also be suboptimal if it is applied inconsistently. Figure 6 displays the rule application variability index. For the length condition, the effect of block was significant, $F(4,76) = 4.64$, $p < .01$, but the effect of group, $F(1,19) = 1.23$, $p > .10$, and the group by block interaction ($F < 1.0$) were nonsignificant. For the distance condition, the effect of block, $F(4,80) = 4.97$, $p < .01$, and the interaction were significant, $F(4,80) = 3.50$, $p <$

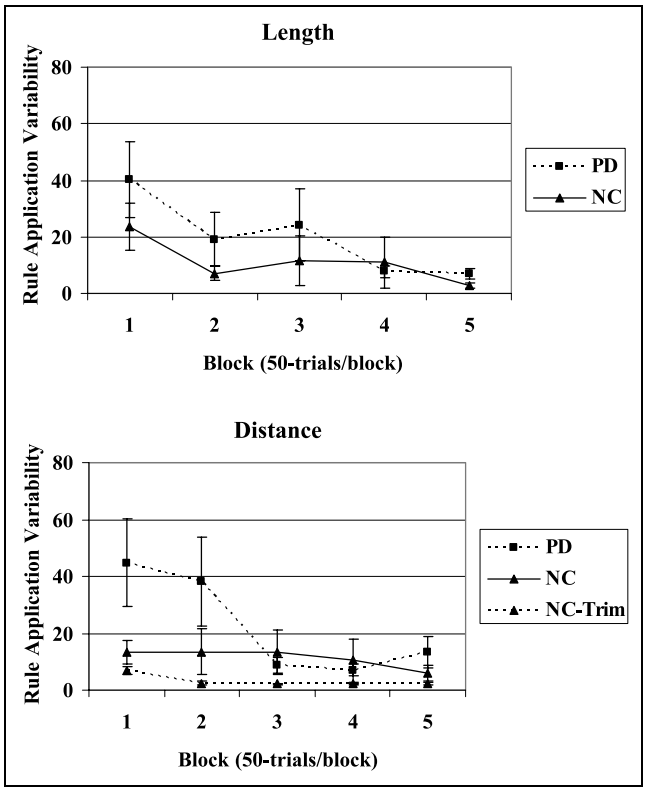

Figure 6. Rule application variability index for the length and distance category structures for the PD, NC, and NC trim (see text for details) participant groups for each block (standard error bars included).

.05, but the main effect of group was nonsignificant, $F(1,20) = 1.79, p > .10$. Post hoc analyses suggested that the interaction was due only to a significant effect in Block 1 ($p < .05$). When the patients with PD were compared with the trimmed control group on the distance task, the effect of participant group, $F(1,17) = 11.11$, $p < .01$; block, $F(4,68) = 4.84$, $p < .01$; and the interaction, $F(4,68) = 3.43$, $p < .05$, were significant. Follow-up t tests suggested the rule application variability was significantly larger for the PD participants in all but the fifth block of trials ($p < .05$ for blocks 1–4), but even in the fifth block the effect was nearly significant ($p = .054$). Rule application variability was generally large for the patients with PD early in learning and gradually declined with learning, whereas rule application variability was low and constant across learning for controls. By the final block of trials, the difference between the patients with PD and controls had been reduced.

GENERAL DISCUSSION

Recent research suggests the basal ganglia serve an important role in category learning (Maddox & Filoteo, 2001; Poldrack et al., 2001; Ashby et al., 1998; Filoteo et al., 2004; Myers et al., in press). This work follows the emerging interest in the basal ganglia's involvement in nonmotor functions, and its general role in learning. Interestingly, similar developments can be observed in the literature on cerebellar function, especially with respect to a possible role in nonmotor learning (e.g., Canavan et al., 1994; Fiez et al., 1992). The current study was designed to provide a direct neuropsychological comparison between the functions of the basal ganglia and cerebellum on rule-based category learning, using patients with either degenerative disorders affecting one of these structures or in the case of the cerebellum, patients with focal lesions. The inclusion of the cerebellar group is important not only for assessing the role of this structure on category learning but also for the evaluation of computational models that have emphasized a central role of the basal ganglia in category learning. It is, of course, essential to conduct experiments that seek confirmatory evidence (e.g., Are patients with PD impaired on such tasks?). Frequently overlooked, however, is the importance of identifying whether a particular deficit is specific to a certain type of neurological disorder (see Ravizza & Ivry, 2001). Finding similar deficits in two distinct patient groups limits the utility of neuropsychological data for constraining computational models.

This study was also motivated by the fact that the current literature has only begun to ask how the basal ganglia contribute to rule-based categorization tasks. Although previous studies have reported that patients with PD are impaired on such tasks (Ashby et al., 2003), there have been conditions in which patients with PD performed as well as controls (Maddox & Filoteo, 2001). Moreover, the focus of the work to date has been to contrast either rule-based and information integration tasks or various types of decision bounds. In the present study, we focused on rule-based tasks and manipulated the degree to which the decision process required the use of an internal referent. This manipulation was included given that the motor control literature has suggested variation in the relative contribution of the basal ganglia and cerebellum to internally and externally guided movements (Goldberg, 1985). We asked whether a similar distinction might apply in the cognitive domain.

The results showed a clear dissociation between the performance of the PD and CB patients. Whereas the patients with PD were impaired on both of the category-learning tasks, the patients with CB performed as well as the control participants. The null result for the latter group was observed in both patients with unilateral focal lesions as well as in those with bilateral degeneration. These results indicate that the cerebellum is not essential for this form of nonmotor learning. Interestingly, cerebellar degeneration has been linked to impaired performance on a task in which the participants had to learn arbitrary stimulus–response associations, even when the movement requirements were minimal (Canavan et al., 1994; Bracke-Tolkmitt et al., 1989; see also Fiez et al., 1992). It would be useful in future studies to

directly compare categorical and noncategorical nonmotor learning with this patient group. Bischoff-Grethe, Ivry, and Grafton (2002) proposed that cerebellar deficits on certain nonmotor tasks may reflect impairment in the maintenance of stimulus–response associations, a process that would be taxed when the response rule cannot be defined categorically.

The patients with PD had lower overall accuracy scores and the learning function for the group tended to asymptote at a lower level than that found for the control participants. In addition, using a learning criterion of 55% correct on the last block of trials indicates that some of the patients with PD failed to learn over the course of the session. In terms of measures of overall performance, we did not observe any clear difference between the two categorization tasks in terms of the overall accuracy data. Contrary to our predictions, the PD deficit was similar on the task requiring an internal referent (distance condition) as it was on the task requiring an external referent (length condition). The only indication that the internal referent condition may have been more compromised in the PD group was in the classification of participants as learners or nonlearners. Using the strict 55% criterion, an equal number (n = 3) of the PD and controls failed to learn on the length condition. On the distance condition, the number of nonlearners for the PD and controls was 4 and 1, respectively.

The lack of a difference between the internal and external criterion conditions should be interpreted with caution for three reasons. First, as with any null result, it is not possible to draw strong conclusions and future work may provide alternative ways to test the internal/external distinction. Second, only 13 patients with PD were tested in the current study, a number on the low end with respect to recent studies of cognitive performance in this population. This sample size was similar to our cerebellar group, thus providing a more balanced test for comparisons involving the two patient groups. However, the small sample size reduces our power for identifying within-group dissociations such as that between the two categorization tasks. Third, due to an oversight on our part, we failed to counterbalance the test order for the two conditions, with all participants first tested on the length condition and then the distance condition. This confound is problematic for comparisons between the two conditions, although in an analysis of order effects from our previous work, we have found no substantive carryover effects after the first 50 or so trials (J. V. Filoteo, W. T. Maddox, A. D. Ing, & D. D. Song, unpublished data). Given these concerns, our emphasis here is on the fact that the patients with PD were impaired on these categorization tasks, whereas the patients with CB performed similar to the controls.

The model-based analyses allowed us to explore the source of the PD category learning deficit in more detail. Even when limited to those individuals who learned the task (using the 55% criterion) this approach produced a number of interesting insights. First, patients with PD were as adept as the controls in deriving an appropriate rule-based strategy to solve each task. This suggests that the patients with PD had no problem in ignoring irrelevant information. Second, the estimates of the decision criterion were farther from the optimal location for the patients with PD. We assume that these shifts reflect response biases given that there was no consistent pattern in the direction of the biases. Third, the patients with PD were less consistent in their application of the dimension-based rules. Fourth, the increase in response bias and rule application variability for the patients with PD was only statistically reliable in the distance condition. When coupled with the number of nonlearners, three dependent variables suggest that the patients with PD were disproportionately impaired in the condition that required the use of an internal referent.

It should be noted that whereas the distance condition required the use of an internal referent, the length condition could be solved with either an external referent or an internal referent. Although speculative, it is possible that a subgroup of patients with PD adopted a strategy that was based on an internal referent when categorizing on the basis of target line length, and that the group impairment was dominated by those using this strategy. Interviewing the participants at the end of testing for each condition might allow us to identify their (explicit) strategies. We did not include this manipulation in the current study because we were worried about carryover effects from one condition to the other. Nonetheless, it would be useful to include an appropriate interview in future studies to assess whether the participants favor internal or external referents and see if such strategies have differential effects on performance. At present, the results are suggestive that the categorization deficit is more pronounced when the task requires the use of an internal referent. However, we do want to emphasize that on almost all measures, the performance of the patients with PD was similar on the two tasks, even if the effects on some measures only reached statistical significance for the distance condition.

Perhaps the most striking aspect of this study is the finding that the decrease in overall accuracy for the patients with PD on the categorization tasks was primarily due to their inability to identify the appropriate decision criterion. We assume that this reflects some form of bias, and the analysis of the distribution of the responses for all participants, learners and nonlearners, is consistent with this interpretation. A bias could reflect a systematic distortion of particular stimulus features (e.g., bias to perceive the target line as longer than controls), a preference to respond with one hand or on one response key, or a preference to perseverate with an initial response. The current experiments were not designed to dissociate these factors given that we did not anticipate the importance of bias, an idea that is

not explicitly incorporated in current neurobiological models of category learning such as COVIS (Ashby et al., 1998).

However, the exaggeration of a response bias in PD is consistent with other ideas about basal ganglia function. The development of habitual responses involves a many-to-few mapping in which widely varied stimulus conditions become associated with a small set of actions. For example, a stereotypic facial grooming response is produced by rats after various forms of somatosensory stimulation (Berridge, Fentress, & Parr, 1987). The basal ganglia have been hypothesized to play a critical role in these mapping operations (e.g., Graybiel & Kimura, 1995). Moreover, dysfunction of the basal ganglia is associated with pathological biases: the patient with obsessive–compulsive disorder or Huntington's disease is compelled to repeatedly produce the same action pattern. Similarly, bradykinesia in patients with PD has been conceptualized as a problem in switching from one action pattern to another (Wing, 1988) or, more generally, effect changes in either motor or mental set (Hayes, Davidson, Keele, & Rafal, 1998). This form of inertia could be considered a response bias or a preference to persist with an action plan even if it is no longer appropriate. The current results suggest that susceptibility to response biases may also be a prominent contributor to category learning deficits in PD. The patients identify the appropriate stimulus information, but their ability to employ the optimal decision bound is impeded by a bias to prefer one response over the other.

METHODS

Participants

Thirteen patients with PD, 14 patients with CB, and 14 control participants were tested. Patients were either referred to the study by neurologists at an outpatient clinic at the VA Medical Center in Martinez, CA, or recruited at meetings of Parkinson and ataxia support groups in the San Francisco Bay Area. The CB group included 6 patients with unilateral CB, either from stroke (n = 4) or tumor (n = 2). Reconstructions of their lesions are presented in Figure 7. The remaining 8 patients in this group had a diagnosis of cerebellar atrophy, confirmed by clinical exam, radiological records, and, in 3 cases, genetic testing. Medical histories were obtained from all participants to exclude individuals with secondary neurological conditions, significant psychiatric disorders, or current substance abuse.

The control participants were recruited through advertisements at local senior centers. These participants were chosen to be comparable, as a group, to the two patient groups on variables of age and education.

As part of our recruitment procedures, patients are given a battery of standardized neurological and neuropsychological tests (Table 2). Patients with PD are assessed with two instruments, the Hoehn and Yahr Staging (Hoehn & Yahr, 1967) and the UPDRS (Fahn, Elton, & Members of the UPDRS Development Committee, 1987). The patients were under their normal medication regimen when the assessments were taken and when tested on the categorization tasks. The PD patients' motor impairments ranged from mild unilateral involvement only (1.0) to moderate to severe bilateral involvement (3.5) on the Hoehn and Yahr. This range was also reflected in the UPDRS scores, although the correlation is not perfect given that the latter also assesses a wider range of cognitive abilities and problems associated with daily living activities. Patients with CB were evaluated with International Cooperative Ataxia Rating Scale (ICARS; Trouillas et al., 1997). Based on the 100-point scale, the CB patients' impairments ranged from 4.5 (essentially no clinical signs of ataxia) to 49.75 (moderate to severe ataxia).

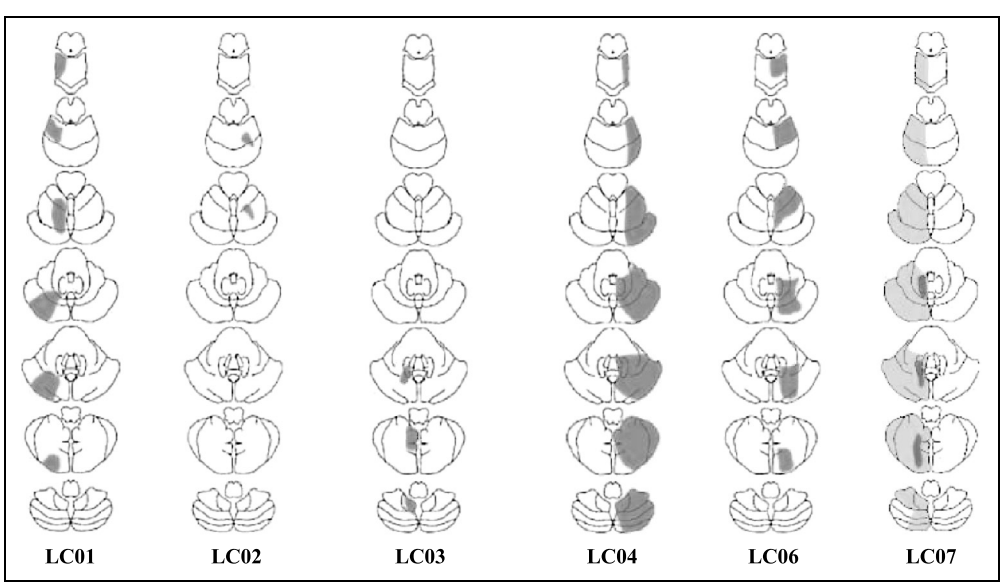

Figure 7. Reconstruction for focal cerebellar patients.

The neuropsychological battery included tests designed to measure different aspects of cognitive function. The Mini-Mental State Examination (MMSE) evaluates general cognitive function and orientation and is used as a screening tool for dementia. The National Adult Reading Test (NART) is designed to assess premorbid intellectual functioning. We also administered two subtests of the Wechsler Adult Intelligence Scale—Third Edition (WAIS-III), Digit Span and Matrix Reasoning. The former is taken to provide a measure of verbal working memory; the latter assesses perceptual organization. Our choice of tests was motivated to provide an overview of various cognitive abilities that might be relevant to categorization while focusing on those that do not require speeded responses given the motor impairments of many of the patients. The scores for the patients with PD on the four cognitive assessments all fell within the normal range and were not significantly different from the scores for the control participants. The patients with CB were significantly impaired on the NART subtest compared with the controls, $t(23) = 2.39$, $p < .05$, and reached near significance on the Digit Span as well, $t(22) = 2.05$, $p = .053$. Although interesting, CB patients showed normal category learning.

All participants received monetary compensation for their participation. The study protocol was approved by the institutional review boards at the VA Medical Center in Martinez and at the University of California, Berkeley.

Stimuli and Stimulus Generation

The experiment used the randomization technique introduced by Ashby and Gott (1988) in which optimal accuracy, within-category scatter, and category coherence are identical in both conditions. The stimuli for the distance condition were generated by sampling randomly from two bivariate normal distributions (see Table 4). Given this and that the covariance was set to zero, the two categories within each condition differed only in the location of the means. Under these constraints, the optimal decision bound is linear and, given the selected stimulus sets, optimal performance would correspond to 90% correct for each condition. The stimuli for the length and distance conditions are displayed in Figure 2.

Table 4. Category Distribution Parameters and the Experimental Conditions

	Category A					*Category B*				
Condition	μ_l	μ_v	σ_l	σ_v	cov_{lv}	μ_l	μ_v	σ_l	σ_v	cov_{flv}
Length	118	150	25	75	0	182	150	25	75	0
Distance	150	118	25	75	0	150	182	25	75	0

The same set of 50 stimuli per condition were used for all participants.

The stimuli were computer generated and displayed as black lines on a white background with 640 × 400 screen resolution. The target line was always above the reference line and varied in length across trials. The reference line was double the thickness of the target line and was always 400 pixels in length. The vertical position of the reference line was randomly selected with a range between 120 and 140 pixels from the bottom of the screen.

Procedure

Each observer was tested individually in a room with normal ambient lighting. A scripted tutorial was used to explain categorization as a process that is commonly used in everyday life that can involve rules used to assign groups of objects to distinct categories. Participants were then informed that in the experimental task, they would see two lines on the computer screen. We (the experimenters) had created two categories and their task was to discover, through trial and error, the rule that defined the two categories. The participants were told that each category was equally likely and that perfect performance was not possible. They were told to be as accurate as possible and not to worry about the speed of their responses.

A trial consisted of the presentation on the computer screen of a reference and target line. They were instructed to press the ''Z'' key on the keyboard if they thought the stimulus was a member of Category A and the ''/'' key if they thought the stimulus was a member of Category B. Immediately after responding, the feedback message ''CORRECT'' or ''ERROR'' appeared at the top of the screen. The feedback message was displayed for 1000 msec and the stimulus remained visible throughout this period. After this, the screen was blank for an intertrial interval of 1000 msec. The instructions clearly emphasized that we were only measuring accuracy. No emphasis was given to the speed of the participants' responses and, in fact, we did not collect reaction time data.

For each condition, participants completed five blocks of 50 trials each. The order of the 50 stimuli was randomized separately for each observer for each block. Each participant was first tested on the length condition and, after a short break, the distance condition. The entire test session lasted approximately 1 hr. The neurological and neuropsychological assessments were generally completed at a different session.

Acknowledgments

This research was supported in part by National Institutes of Health grants MH59196, NS30256, NS33504, and NS40813. We thank Shawn Ell and Roshan Cools for several excellent

comments on an earlier version of this manuscript, Greg Ashby and Vince Filoteo for discussions regarding these data, and two anonymous reviewers.

Reprint requests should be sent to W. Todd Maddox, Department of Psychology, 1 University Station A8000, University of Texas, Austin, TX 78712, or via e-mail: maddox@psy.utexas.edu, or to Richard B. Ivry, Department of Psychology, 3210 Tolman Hall #1650, University of California, Berkeley, CA 94618, or via e-mail: ivry@socrates.berkeley.edu.

Notes

1. Rule-based category-learning tasks are often contrasted with information-integration category-learning tasks. Information-integration category-learning tasks are those in which accuracy is maximized only if information from two or more stimulus components (or dimensions) is integrated at some 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. On information-integration category-learning tasks, patients with basal ganglia degeneration have been found to show deficits under some but not other conditions (e.g., Filoteo, Maddox, Salmon, & Song, 2004; Ashby, Noble, Filoteo, Waldron, & Ell, 2003; Maddox & Filoteo, 2001).
2. Several different methods are available for model testing. The three models examined in this article were nested in the sense that a simpler model could always be derived from a more complex model by setting some of the parameters from the more complex model to constants. For example, the optimal rule-based model is a special case of the suboptimal, rule-based model where the criterion value is set to the optimal value. With nested models, likelihood ratio tests can be used to determine whether the additional free parameters of a more general model provide a significant improvement in fit over the simpler model. This was the approach taken in this article. Another popular approach is to compare the Bayesian Information Criteria (BIC) statistic across models, choosing the model with the smallest BIC value. The BIC statistic is more conservative and thus is biased toward simpler models. Using this approach, we found that 97% and 98% of the patients with PD used a rule-based strategy in the length and distance conditions, respectively, and 96% and 100% of the control participants used a rule-based strategy in the same conditions.

REFERENCES

Ashby, F. G. (1992). Multidimensional models of categorization. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition.* Hillsdale, NJ: Erlbaum.

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 neurobiological basis of category learning. *Trends in Cognitive Sciences, 5,* 204–210.

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., Noble, S., Filoteo, J. V., Waldron, E. M., & Ell, S. W. (2003). Category learning deficits in Parkinson's Disease. *Neuropsychology, 17,* 115–124.

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review, 93,* 154–179.

Berridge, K. C., Fentress, J. C., & Parr, H. (1987). Natural syntax rules control action sequence of rats. *Behavioural Brain Research, 23,* 59–68.

Bischoff-Grethe, A., Ivry, R. B., & Grafton, S. T. (2002). Cerebellar involvement in response reassignment rather than attention. *Journal of Neuroscience, 22,* 546–553.

Bracke-Tolkmitt, R., Linden, A., Canavan, A. G. M., Rochstroh, B., Scholz, E., Wessel, K., & Diener, H. C. (1989). The cerebellum contributes to mental skills. *Behavioral Neuroscience, 103,* 442–446.

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

Canavan, A. G., Sprengelmeyer, R., Diener, H. C., & Homberg, V. (1994). Conditional associative learning is impaired in cerebellar disease in humans. *Behavioral Neuroscience, 108,* 475–485.

Debaere, F., Wenderoth, N., Sunaert, S., Van Hecke, P., & Swinnen, S. P. (2003). Internal vs external generation of movements: Differential neural pathways involved in bimanual coordination performed in the presence or absence of augmented visual feedback. *Neuroimage, 19,* 764–776.

Desmond, J. E., Gabrieli, J. D., Wagner, A. D., Ginier, B. L., & Glover, G. H. (1997). Lobular patterns of cerebellar activation in verbal working-memory and finger-tapping tasks as revealed by functional MRI. *Journal of Neuroscience, 17,* 9675–9685.

Diedrichsen, J., Ivry, R. B., & Pressing, J. (2003). Cerebellar and basal ganglia contributions to interval timing. In W. H. Meck (Ed.), *Functional and neural mechanisms of interval timing* (pp. 457–481). Boca Raton, FL: CRC Press.

Doyon, J., Penhune, V., & Ungerleider, L. G. (2003). Distinct contribution of the cortico-striatal and cortico-cerebellar systems to motor skill learning. *Neuropsychologia 41,* 252–262.

Estes, W. K. (1994). *Classification and cognition*. New York: Oxord University Press.

Fahn, S., Elton, R., & Members of the UPDRS Development Committee. (1987). Unified Parkinson's Disease Rating Scale. In S. Fahn, C. D. Marsden, D. B. Calne, & M. Goldstein (Eds.), *Recent developments in Parkinson's disease* (Vol. 2, pp. 153–63, 293–304). Florham Park, NJ: Macmillan Health Care Information.

Fiez, J. A., Petersen, S. E., Cheney, M. K., & Raichle, M. E. (1992). Impaired non-motor learning and error detection associated with cerebellar damage. A single case study. *Brain, 115,* 155-178.

Filoteo, J. V., Maddox, W. T., Salmon, D. P., & Song, D. D. (2004). Information-integration category learning in patients with striatal dysfunction. *Neuropsychology,* in press.

Flowers, K. A. (1976). Visual ''closed-loop'' and ''open-loop'' characteristics of voluntary movement in patients with Parkinsonism and intention tremor. *Brain, 99,* 269–310.

Goldberg, G. (1985). Supplementary motor area structure and function: review and hypotheses. *Behavioral and Brain Sciences, 8,* 567–616.

Grafton, S. T., Mazziotta, J. C., Presty, S., Friston, K. J., Frackowiak, R. S., & Phelps, M. E. (1992). Functional anatomy of human procedural learning determined with regional cerebral blood flow and PET. *Journal of Neuroscience, 12,* 2542–2548.

Graybiel, A. M., & Kimura, M. (1995). Adaptive neural networks in the basal ganglia. In J. C. Houk, J. L. Davis, & D. G. Beiser (Eds.), *Models of information processing in the basal ganglia* (pp. 103–116). Cambridge: MIT Press.

Green, D. M., & Swets, J. A. (1967). *Signal detection theory and psychophysics.* New York: Wiley.
Haaland, K. Y., Harrington, D. L., O'Brien, S., & Hermanowicz, N. (1997). Cognitive-motor learning in Parkinson's disease. *Neuropsychology, 11,* 180–186.
Hayes, A. E., Davidson, M. C., Keele, S. W., & Rafal, R. D. (1998). Toward a functional analysis of the basal ganglia. *Journal of Cognitive Neuroscience, 10,* 178–198.
Helmuth, L., Ivry, R., & Shimizu, N. (1997). Preserved performance by cerebellar patients on tests of word generation, discrimination learning, and attention. *Learning and Memory, 3,* 456–474.
Helmuth, L. L., Mayr, U., & Daum, I. (2000). Sequence learning in Parkinson's disease: A comparison of spatial-attention and number-response sequences. *Neuropsychologia, 38,* 1443–1451.
Hoehn, M. M., & Yahr, M. D. (1967). Parkinsonism: Onset, progression, and mortality. *Neurology, 17,* 427–442.
Jackson, G. M., Jackson, S. R., Harrison, J., Henderson, L., & Kennard, C. (1995). Serial reaction time learning and Parkinson's disease: Evidence for a procedural learning deficit. *Neuropsychologia, 33,* 577–593.
Keri, S. (2003). The cognitive neuroscience of category learning. *Brain Research Reviews, 43,* 85–109.
Krebs, H. I., Hogan, N., Hening, W., Adamovich, S. V., & Poizner, H. (2001). Procedural motor learning in Parkinson's disease. *Experimental Brain Research, 141,* 425–437.
Maddox, W. T., & Ashby, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception & Psychophysics, 53,* 49–70.
Maddox, W. T., Ashby, F. G., Ing, A. D., & Pickering, A. D. (2004). Disrupting feedback processing interferes with rule-based but not information integration category learning. *Memory & Cognition, 32,* 582–591.
Maddox, W. T., & Filoteo, J. V. (2001). Striatal contributions to category learning: Quantitative modeling of simple linear and complex nonlinear rule learning in patients with Parkinson's Disease. *Journal of the International Neuropsychological Society, 7,* 710–727.
Martin, T. A., Keating, J. G., Goodkin, H. P., Bastian, A. J., & Thach, W. T. (1996). Throwing while looking through prisms. I. Focal olivocerebellar lesions impair adaptation. *Brain, 119,* 1183–1198.
Middleton, F. A., & Strick, P. L. (2000). Basal ganglia output and cognition: Evidence from anatomical, behavioral, and clinical studies. *Brain and Cognition, 42,* 183–200.
Middleton, F. A., & Strick, P. L. (2001). Cerebellar projections to the prefrontal cortex of the primate. *Journal of Neuroscience, 21,* 700–712.
Middleton, F. A., & Strick, P. L. (2002). Basal-ganglia 'projections' to the prefrontal cortex of the primate. *Cerebral Cortex, 12,* 926–935.
Molinari, M., Leggio, M. G., Solida, A., Ciorra, R., Misciagna, S., Silveri, M. C., & Petrosini, L. (1997). Cerebellum and procedural learning: Evidence from focal cerebellar lesions, *Brain, 120,* 1753–1762.
Myers, C., Shohamy, D., Gluck, M., Grossman, S., Kluger, A., Ferris, S., Golomb, J., Schnirman, G., & Schwartz, R. (in press). Dissociating hippocampal vs. basal ganglia contributions to learning and transfer. *Journal of Cognitive Neuroscience.*
Packard, M. G., & Knowlton, B. J. (2002). Learning and memory functions of the basal ganglia. *Annual Review of Neuroscience, 25,* 563–693.
Pascual-Leone, A., Grafman, J., Clark, K., Stewart, M., Massaquoi, S., Lou, J. S., & Hallett, M. (1993). Procedural learning in Parkinson's disease and cerebellar degeneration. *Annals of Neurology, 34,* 594–602.
Poldrack, R. A., Prabhakaran, V., Seger, C. A., & Gabrieli, J. D. E. (1999). Striatal activation during acquisition of a cognitive skill. *Neuropsychology, 13,* 564–574.
Poldrack, R., Clark, J., Pare-Blagoev, J., Shohamy, D., Creso-Moyano, J., Myers, C., & Gluck, M. (2001). Interactive memory systems in the brain. *Nature, 414,* 546–550.
Ravizza, S., & Ivry, R. B. (2001). Comparison of the basal ganglia and cerebellum in shifting attention. *Journal of Cognitive Neuroscience, 13,* 285–297.
Sanes, J. N., Dimitrov, B., & Hallett, M. (1990). Motor learning in patients with cerebellar dysfunction. *Brain, 113,* 103–120.
Schmahmann, J. D. (Ed.). (1997). *The cerebellum and cognition.* San Diego, CA: Academic Press.
Selemon, L. D., & Goldman-Rakic, P. (1985). Longitudinal topography and interdigitation of corticostriatal projections in the rhesus monkeys. *Journal of Neuroscience, 5,* 776–794.
Shin, J. C., & Ivry, R. B. (2003). Spatial and temporal sequence learning in patients with Parkinson's disease or cerebellar lesions. *Journal of Cognitive Neuroscience, 15,* 1232–1243.
Shohamy, D., Myers, C. E., Grossman, S., Sage, J., Gluck, M. A., & Poldrack, R. A. (2004). Cortico-striatal contributions to feedback-based learning: Converging data from neuroimaging and neuropsychology. *Brain, 127,* 851–859.
Smith, E. E., & Medin, D. L. (1981). *Categories and concepts.* Cambridge: Harvard University Press.
Smith, M. A., Brandt, J., & Shadmehr, R. (2000). Motor disorder in Huntington's disease begins as a dysfunction in error feedback control. *Nature, 403,* 544–549.
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.
Trouillas, P., Takayanagi, T., Hallett, M., Currier, R. D., Subramony, S. H., Wessel, K., Bryer, A., Diener, H. C., Massaquoi, S., Gomez, C. M., Coutinho, P., Ben Hamida, M., Campanella, G., Filla, A., Schut, L., Timann, D., Honnorat, J., Nighoghossian, N., & Manyam, B. (1997). International Cooperative Ataxia Rating Scale for pharmacological assessment of the cerebellar syndrome. *Journal of Neurological Science, 145,* 205–211.
Ullman, M. T. (2001). A neurocognitive perspective on language: The declarative/procedural model. *Nature Neuroscience Review, 2,* 717–726.
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.
Wing, A. M. (1998). A comparison of the rate of pinch grip force increases and decreases in parkinsonian bradykinesia. *Neuropsychologia, 26,* 479–482.