Differential influence of levodopa on reward-based learning in Parkinson's disease

Susanne Graef[1,2]*, **Guido Biele[1,3]**, **Lea K. Krugel[1,4]**, **Frank Marzinzik[4]**, **Michael Wahl[4]**, **Johann Wotka[4]**, **Fabian Klostermann[4]** and **Hauke R. Heekeren[1]***

[1] Max Planck Institute for Human Development, Berlin, Germany
[2] Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
[3] Department of Psychology, University of Oslo, Oslo, Norway
[4] Department of Neurology, Charité, Campus Benjamin Franklin, Berlin, Germany

The mesocorticolimbic dopamine (DA) system linking the dopaminergic midbrain to the prefrontal cortex and subcortical striatum has been shown to be sensitive to reinforcement in animals and humans. Within this system, coexistent segregated striato-frontal circuits have been linked to different functions. In the present study, we tested patients with Parkinson's disease (PD), a neurodegenerative disorder characterized by dopaminergic cell loss, on two reward-based learning tasks assumed to differentially involve dorsal and ventral striato-frontal circuits. 15 non-depressed and non-demented PD patients on levodopa monotherapy were tested both on and off medication. Levodopa had beneficial effects on the performance on an instrumental learning task with constant stimulus-reward associations, hypothesized to rely on dorsal striato-frontal circuits. In contrast, performance on a reversal learning task with changing reward contingencies, relying on ventral striato-frontal structures, was better in the unmedicated state. These results are in line with the "overdose hypothesis" which assumes detrimental effects of dopaminergic medication on functions relying upon less affected regions in PD. This study demonstrates, in a within-subject design, a double dissociation of dopaminergic medication and performance on two reward-based learning tasks differing in regard to whether reward contingencies are constant or dynamic. There was no evidence for a dose effect of levodopa on reward-based behavior with the patients' actual levodopa dose being uncorrelated to their performance on the reward-based learning tasks.

Keywords: levodopa, decision-making, reinforcement learning, reversal learning, overdose hypothesis, PD, reward contingencies

INTRODUCTION

Therapy with the dopamine (DA) precursor levodopa or DA agonists has proved successful in the treatment of motor symptoms associated with Parkinson's disease (PD). However, the progressive degeneration of dopaminergic neurons in PD not only affects motor abilities. Instead, consistent with the wide distribution of dopaminergic projection targets throughout the brain, dopaminergic cell loss in PD patients is associated with a number of non-motor symptoms, e.g., symptoms relating to cognition, mood and motivation. The effects of dopaminergic drugs on cognitive symptoms are complex. While improvement has been reported in some domains, deteriorating effects or no changes have been found in others (for an overview, see Nieoullon, 2002; Cools, 2006).

Numerous cognitive and motivational processes have been linked to the brain reward system, which relies on the neurotransmitter DA. For instance, dopaminergic neurons projecting to the prefrontal cortex and the subcortical striatum have been shown repeatedly to be sensitive to reinforcement in animals and humans (Schultz and Dickinson, 2000; Bayer and Glimcher, 2005). Since PD is characterized by a loss of dopaminergic neurons, reward-based learning is one of the domains affected in PD (Knowlton et al., 1996; Shohamy et al., 2004).

Due to a characteristic progression of dopaminergic cell loss in the substantia nigra, DA depletion in PD progresses from dorsolateral to ventromedial striatal structures and their associated striato-frontal circuits (see Cools, 2006). These dorsal and ventral striato-frontal circuits have been implicated in different subfunctions within the processing of rewards (Cools et al., 2001a, 2002; Packard and Knowlton, 2002).

Reward-based performance in PD patients depends on the properties of the task used, and therewith on its underlying neural substrates. Perretta et al. (2005) showed that medicated PD patients in an early stage of the disease perform differently on two reward-based learning tasks, thus providing support for the assumption that those tasks rely on different striato-frontal circuits. The patients performed worse on the Iowa Gambling Task, a task that has been associated with the prefrontal cortex (e.g., Manes et al., 2002; Fellows and Farah, 2005), as compared to the Weather Prediction Task (WPT), which has been related to the striatum, especially the dorsal parts of the striatum (Knowlton et al., 1996). The performance on reward-based learning tasks also seems to be subject to the medication state PD patients are tested in (e.g., Swainson et al., 2000; Cools et al., 2001a,b, 2003), with Czernecki et al. (2002) however reporting no overall effect of dopaminergic treatment on their reward-based learning paradigms.

A common finding when testing PD patients in different medication states on reward-based tasks is the impaired performance of patients on medication as opposed to PD patients off medication on reversal learning tasks (Swainson et al., 2000; Cools et al., 2001a,

2007; Mehta et al., 2001). Reversal learning, as opposed to simple instrumental learning, is characterized by the challenge to switch behavior according to experimentally introduced changes in reward contingencies. The processing of such a behavioral adaptation during reversal tasks has been shown to rely on ventral striato-frontal circuitry including the ventral striatum (e.g., Cools et al., 2002; Heekeren et al., 2007) and the orbitofrontal cortex (Rolls, 2000).

Medication-dependent performance differences have been explained with the "overdose hypothesis" (Gotham et al., 1988). Due to the progression of PD from dorsal to ventral structures (Cools, 2006), the poorer performance on certain tasks while patients are on levodopa might arise from the fact that doses necessary to replete DA levels in the more severely affected regions (e.g., dorsal striatum) are too high for less affected regions (e.g., orbitofrontal cortex, ventral striatum), thus "overdosing" those regions of the brain and disrupting their function. Therefore, patients on dopaminergic medication would perform worse on tests that require the orbitofrontal cortex or ventral striato-frontal circuits in general. Other studies generalized the "overdose hypothesis" to an "inverted U" shape relationship between DA levels and DA-depending function (Williams and Goldman-Rakic, 1995; Cools et al., 2001a; Mehta et al., 2001), with the optimal level depending on the nature of the function. This is consistent with reports that both insufficient and excessive DA levels can have adverse effects on cognition in experimental animals (Arnsten, 1998).

More recently, Frank and colleagues provided a more detailed mechanistic account of the DA "overdose hypothesis" (e.g., Frank et al., 2004). According to their theory, the critical prerequisite for basal ganglia-dependent learning is the availability of a large dynamic range in DA release: the discrimination between outcome values of different responses is only possible if the DA signal is able to increase or decrease substantially from its baseline levels. Therewith, they tie in with previous reports about rewards being connected with phasic DA bursts, whereas omission of rewards or punishments are connected with dips in DA (Schultz, 2006).

To investigate such an "inverted U" shape relationship which assumes both beneficial and detrimental (overdosing) dopaminergic effects, within-subject designs are particularly useful. More recently, a few studies have applied within-subject designs to allow for group comparisons that are not influenced by differences between the groups other than the dopaminergic state (e.g., Czernecki et al., 2002; Moustafa et al., 2008a,b; Rutledge et al., 2009). However, in addition to studying the effects of dopaminergic medication on reward-based performance, we were also interested in investigating to what extent the nature of the stimulus-reward associations within a task would have an effect on the performance of PD patients. In the present study, we therefore examined the combined effects of levodopa withdrawal and an experimentally imposed difference in reward contingencies by testing PD patients on two distinct reward-based learning tasks: a simple instrumental learning task characterized by constant reward contingencies and a reversal learning task characterized by dynamic reward contingencies. PD patients on levodopa monotherapy were assessed twice, once on and once off their medication. We deliberately studied only patients on levodopa monotherapy who were not taking any other antiparkinsonian medication (e.g., DA agonists). Due to the differing pharmacological properties of levodopa and

DA agonists, we deem it important to separately investigate the influence of those medication groups on feedback-based tasks. To our knowledge, comparably few studies of reward-based behavior have examined patients treated with levodopa only. Therefore, with this study, we hope to shed some more light on the influence of levodopa on two reward-based tasks that differ in their stimulus-reward contingencies. We hypothesized that PD patients off levodopa (as compared to their respective performance on medication) would perform worse on the simple instrumental learning task characterized by constant reward contingencies and assumed to involve mainly dorsal striato-frontal circuits (for studies that have linked instrumental learning to the dorsal striatum, see Yin et al., 2004; Kimchi et al., 2009), but that they would perform better on a reversal learning task with changing reward contingencies, which is assumed to rely on ventral striato-frontal circuits (e.g., Rolls, 2000; Cools et al., 2002; Heekeren et al., 2007). In line with our hypothesis, we found a double dissociation between task performance and medication state suggesting that performance differences relate to a differential influence of levodopa on reinforcement learning depending on task-specific reward contingencies and underlying striato-frontal circuits.

MATERIALS AND METHODS
SUBJECTS
Twenty patients with mild idiopathic PD were recruited from the neurological outpatients' clinic of the Charité University Hospital, Campus Benjamin Franklin, Berlin, Germany. All patients were diagnosed by a neurologist and met the UK Brain Bank criteria. Inclusion criteria were levodopa monotherapy (i.e., no concurrent therapy with DA agonists) and the ability to be tested both on medication (under the therapeutic effect of levodopa) and off medication (after abstaining from levodopa for at least 12 h). Patients were excluded if they had a significant history of any other neurological or psychiatric illness, if their modified Hoehn and Yahr stage (Hoehn and Yahr, 1967) was higher than three and if they showed signs of a clinical depression (as measured with Beck's Depression Inventory (Beck et al., 1961), BDI scores > 16) or dementia (as measured with the Mini Mental State Exam (Folstein et al., 1975), MMSE scores < 25) in the medicated state. The Parkinson Neuropsychometric Dementia Assessment (PANDA, Kessler et al., 2006), an additional PD-specific dementia screening, was applied as well. On average, all groups achieved scores in the cognitively unimpaired range. Two patients achieved a score below 15 on the PANDA, in each case however only on one of their two assessments. We did not exclude those two patients from the study because the scores were obtained in opposed medication states (on medication for one patient, off medication for the other) and MMSE scores were in the normal range on both assessments. Additionally, patients were tested on the Unified Parkinson's Disease Rating Scale (UPDRS, Fahn et al., 1987), on the Fatigue Severity Scale (FSS, Krupp et al., 1989) and on a questionnaire assessing patients' quality of life (PDQ-39, Peto et al., 1995). The patients had been diagnosed with PD around 1.7 years prior (±1.6) and reported noticing the first symptoms around 4.3 years (±4.0) ago. The average dose of levodopa taken by PD patients was 553.3 mg (SD: 227.8 mg, range: 200–1100 mg/day). The only significant difference between the two PD medication

states was the Unified Parkinson's Disease Rating Scale (UPDRS, Fahn et al., 1987) score, which was significantly higher for patients off levodopa ($t = 7.56$, df $= 12$, $p = 0.000$). Against the background of the criteria used, five patients had to be excluded since they were not testable in the OFF state. Consequently, a total of 15 PD patients could be included in the study, six of whom were female. All of these patients came in for two assessments, once on and once off medication, with a 4-week interval in between.

Sixteen healthy elderly subjects (six female) without any history of neurological or psychiatric illness served as a control group. They were recruited from a database at the Max-Planck-Institute for Human Development, Berlin. This study was approved by the Research Ethics Committee of the Charité University Medicine Berlin and all participants gave written informed consent.

The patient and control group were matched in regard to age ($t = -0.058$, df $= 29$, $p = 0.299$) and years of education ($t = -1.146$, df $= 27$, $p = 0.262$). See **Table 1** for subjects' details. The patient groups differed significantly from the control subjects only in regard to their Hoehn and Yahr (NC vs PD ON: $t = 9.885$, df $= 14$, $p = 0.000$; NC vs PD OFF: $t = 10.269$, df $= 14$, $p = 0.000$) and UPDRS scores (NC vs PD ON: $t = 8.042$, df $= 17.8$, $p = 0.000$; NC vs PD OFF: $t = 9.801$, df $= 16.1$, $p = 0.000$), but not on any of the other clinical measures (BDI, MMSE, PANDA, FSS: $p > 0.0167$).

Within the patient sample, seven patients were tested on medication first while eight patients were tested off medication first. They were therefore divided into two subgroups: the "ON first" subgroup ($N = 7$) and the "OFF first" subgroup ($N = 8$). **Table 2**

Table 1 | Descriptive statistics for normal control subjects (NC) and PD patients.

	NC ($N = 16$) Mean $\pm$ SD (range)	PD total ($N = 15$) Mean $\pm$ SD (range)	PD ON ($N = 15$) Mean $\pm$ SD (range)	PD OFF ($N = 15$) Mean $\pm$ SD (range)
Age	67.75 ± 4.55 (62–78)	65.27 ± 8.14 (52–83)		
Years of education	15.50 ± 3.00 (11–22)	14.00 ± 3.95 (8–20)		
Hoehn and Yahr stage	0 ± 0.0		1.68 ± 0.64* (1–3)	1.96 ± 0.78* (1–3)
UPDRS score	2.00 ± 1.51 (0–6)		19.64 ± 7.70* (10–33)	30.93 ± 11.01* (13–51)
BDI score	5.31 ± 2.98 (0–11)		6.93 ± 5.55 (0–16)	8.50 ± 7.77 (0–32)
MMSE score	29.00 ± 1.00 (27–30)		28.64 ± 1.15 (27–30)	28.57 ± 0.94 (27–30)
PANDA score	25.43 ± 3.2 (18–29)		22.14 ± 4.9 (13–30)	21.71 ± 5.1 (14–28)
FSS score	3.15 ± 1.4 (1.0–4.8)		3.60 ± 2.0 (1.0–7.0)	3.60 ± 1.9 (1.11–6.67)
PDQ score			29.93 ± 24.7 (2–88)	35.21 ± 28.5 (4–97)

Significantly different from control subjects ($p < 0.0167$). BDI, Beck depression inventory; FSS, fatigue severity scale; MMSE, mini mental state examination; PANDA, Parkinson neuropsychometric dementia assessment; PDQ, Parkinson's disease questionnaire; UPDRS, unified Parkinson's disease rating scale; SD, standard deviation.

Table 2 | Group characteristics for ON-beginner and OFF-beginner subgroups.

	ON first Mean $\pm$ SD (range)	OFF first Mean $\pm$ SD (range)	t	df	p
Age	69.1 ± 7.9 (59–83)	61.9 ± 7.1 (52–69)	−1.874	13	0.084
Education in years	13.7 ± 5.0 (8.0–20.0)	14.3 ± 3.1 (11–20)	0.253	13	0.804
Years since diagnosis	1.9 ± 1.6 (0–4.0)	1.5 ± 1.7 (0–5.0)	−0.514	13	0.616
Years since first symptoms	5.6 ± 5.4 (0–14.0)	3.1 ± 2.0 (1–6.5)	−1.133[a]	7.515	0.292
Daily levodopa dose in mg	607.1 ± 300.6 (200–1100)	506.3 ± 156.8 (200–700)	−0.832	13	0.421
Hoehn and Yahr stage ON	2.1 ± 0.6 (1.0–3.0)	1.3 ± 0.4 (1.0–2.0)	−3.192	13	**0.007[b]**
Hoehn and Yahr stage OFF	2.7 ± 0.4 (2.0–3.0)	1.4 ± 0.4 (1.0–2.0)	−6.070	13	**0.000**
UPDRS score ON	23.4 ± 7.7 (12–33)	15.4 ± 5.7 (10–26)	−2.322	13	**0.037**
UPDRS score OFF	37.1 ± 10.2 (23–51)	24.5 ± 7.7 (13–40)	−2.733	13	**0.017**
BDI score ON	10.1 ± 4.9 (2–16)	4.1 ± 4.7 (0–14)	−2.438	13	**0.030**
BDI score OFF	13.3 ± 9.5 (6–32)	4.9 ± 3.6 (0–9)	−2.334	12	**0.038**
MMSE score ON	28.5 ± 0.8 (28–30)	28.8 ± 1.4 (27–30)	0.418[a]	11.608	0.684
MMSE score OFF	28.3 ± 0.8 (27–29)	28.8 ± 1.0 (27–30)	0.812	12	0.433
PDQ score ON	46.5 ± 26.7 (5–88)	17.5 ± 14.5 (2–37)	−2.618	12	**0.022**
PDQ score OFF	57.3 ± 27.7 (12–97)	18.6 ± 15.0 (4–45)	−3.380	12	**0.005**
FSS score ON	5.1 ± 2.0 (1.9–7.0)	2.5 ± 1.2 (1.0–3.9)	−2.996	12	**0.011**
FSS score OFF	5.0 ± 1.8 (1.9–6.7)	2.6 ± 1.2 (1.1–4.9)	−2.491	11	**0.030**

[a]Numbers in italics represent tests, in which variances could not be considered as equal. [b]Bold numbers represent significant differences ($p < 0.05$). BDI, Beck depression inventory; FSS, fatigue severity scale; MMSE, mini mental state examination; PDQ, Parkinson's disease questionnaire; UPDRS, unified Parkinson's disease rating scale; SD, standard deviation.

shows group characteristics for the two patient subgroups. Patients for this study were recruited in a consecutive manner. Therefore, the patients in the two subgroups were not matched for parameters other than disease duration and years of education. Consequently, significant differences were found between the two subgroups with regard to Hoehn and Yahr stage, UPDRS, BDI, PDQ and FSS scores.

PROCEDURE

Patients were tested at the time of day when they subjectively felt best, testing sessions for control subjects started at 9 am. After filling in screening tests for depression (BDI) as well as for fatigue (FSS) and patient's quality of life in the patient sample, participants were seated at a 1.5-m distance from the presentation computer. They carried out four different tasks, two of which are reported here. To avoid sequence effects, task sequence was randomized across all subjects. There were regular breaks between the tasks during which instructions for the task to come were given. When participants had no further questions regarding the task, a test run was started. For both the instrumental and reversal task, picture stimuli in these test runs were different from the stimuli used in the actual experiment. For the task with constant reward contingencies, a simple probabilistic instrumental learning task, participants had to complete at least one block of 30 test trials until one picture pair was recognized correctly. For the reversal learning task with changing reward contingencies, subjects had to pursue until successfully achieving two reversals under probabilistic conditions to ensure that the concept of switching had been understood. Some subjects needed a little more practice than others before they grasped the switching concept. However, we believe that minor differences in practice time are justified in order to ensure that the instruction had been understood correctly and that performance differences in the actual task would not be based on misinterpreting instructions. Participants made their decision by pressing the according button on a specially designed button box with four buttons that were arranged in a rhomboid shape (one at the top, one at the bottom, one on the right and one on the left side). For the probabilistic instrumental learning task, the upper and lower buttons were masked so that only the right and left buttons were available. When only two buttons were needed, participants were asked to use one hand only (their preferred hand). For the reversal learning task, however, participants were allowed to use both hands. To control for differences in motor skills due to the medication state, patients and control subjects had to carry out a finger-tapping task before the actual experiment started. During that task, participants had to press a specific button on the button box repeatedly with their right and left index finger, respectively, during 30 s each.

In the case of the patients, the routine described above was repeated in exactly the same way approximately 4 weeks later: task sequence and time of testing was the same as on the first assessment; participants were given the same instructions and had to accomplish test runs for the tasks again. Standardization of the study procedure was given by means of standardized task instructions, obligatory test runs on both testing sessions and identical time of testing and task sequence for both sessions.

TASK DESIGN

Probabilistic reversal learning task with changing reward contingencies

This task is a modification of a task used in previous studies (see also Reischies, 1999; Marschner et al., 2005; Mell et al., 2005; Krugel et al., 2009). For a total of 228 trials, participants chose one out of four options presented simultaneously on a computer screen. Non-monetary feedback (5, 15, or 30 points) was presented after each choice. The participant's objective was to collect as many points as possible. At a time, only one option was the maximally rewarded one. To make the task less predictable, the choice of this option led to the maximum number of points in 75% of trials only. After reaching a specified learning criterion (choosing the highest-rewarded option in six out of eight consecutive trials), the reward contingencies changed so that another option became the highest-rewarded one. In this task, the main measure of interest was the number of switches ("reversals") achieved. The stimuli used did not elicit any one-sided preferences and were considered easily recognizable in pre-tests with healthy controls. Additionally, they stayed at the same position throughout the experiment to keep working memory requirements low. The decision was carried out by pressing one of four buttons on the button box. The buttons were arranged in the same order as the stimuli on the screen. Each trial lasted a maximum of 4250 ms: the options were displayed for 2750 ms first, the subject then had to make a decision within the next 750 ms, and finally, the feedback was shown for 750 ms. If participants did not press one of the assigned buttons during that time, a hint appeared on the screen telling the subject that the response was given too slowly ("Leider zu langsam" – "Unfortunately too slow").

Probabilistic instrumental learning task with constant reward contingencies

In each trial, two geometric figures were presented on a computer screen. For 30 trials in a row, the same two options were displayed and participants repeatedly chose from this pair of options to find out which option was the higher rewarded one. The positive reinforcement for which participants had to aim was a yellow smiley face. To make the task less predictable, the two options had different pre-defined stochastic probabilities with which they led to this desired outcome: the target effectuated the smiley face in either 70, 60, or 50% of all trials, whereas the distractor was always rewarded in only 40% of all trials. In those trials that were not rewarded with a smiley face, a red-colored sad face was shown. Participants knew in advance that each of altogether eight pairs of stimuli would be presented repeatedly over 30 trials, during which they would have to find out which option leads to the correct outcome (the smiley face) more frequently. The geometric figures used in this task were taken from the Rey Visual Design Learning Test (Rey, 1964) and modified at the Cognitive Neuroscience Lab at the University of Michigan. They were found to elicit no one-sided preferences when evaluated by healthy control subjects. All of the symbols were of the same color and varied only with respect to shape. To avoid side effects, the symbols changed sides randomly across the 30 trials. Participants chose by pressing either the left or right button on a specially designed button box. Each trial lasted a maximum of 3450 ms: the options were displayed for 2250 ms first, then the

subject had to make a decision within the next 500 ms, and finally the feedback was shown for 700 ms. If participants did not press one of the assigned buttons during the decision time, a hint appeared on the screen telling the subject that the response was given too slowly ("Leider zu langsam" – "Unfortunately too slow"). Choices, reaction time, and number of missed trials were recorded. The main measure of interest in this task was the percentage of correct choices.

Both tasks were programmed with Presentation® software, version 11.1 (Neurobehavioral Systems, Inc; Albany, CA/USA) and administered on a 17" computer screen at a standard distance of 1.5 m from the subject's eyes. Pre-tests were conducted with 25 healthy elderly controls to adjust experiment parameters according to the intended setup. We chose to administrate two separate tasks instead of just one task with several distinct phases of stimulus-reward learning versus reversal learning in order to clearly separate the effects of constant versus dynamic reward contingencies. While the instrumental learning task allowed us to study the performance on a task with constant reward contingencies over a total of 240 trials (8 blocks of 30 trials), the reversal learning task allowed repeated switches in the reward contingencies with the final number of switches depending on the participant's performance. See **Figure 1** for the setup of the two reward-based learning tasks used in the study.

DATA ANALYSIS

Since we were mainly interested in the influence of dopaminergic medication on reward-based learning, we focused the statistical analysis on comparing the performances across the two medication states. We calculated repeated measures analyses of covariance (ANCOVA) with session as repeated measures factor to investigate interaction effects between the subgroups and testing sessions. Patients were divided into the subgroups "ON first" and "OFF first" for patients beginning on medication and patients beginning off medication, respectively. As half of the patients started on medication and the other half off medication, regression effects to or from the mean are unlikely to have had a significant impact on the results and therefore, they can be neglected as potential biasing factors.

Since patients were recruited in a consecutive manner and because significant differences with regard to Hoehn and Yahr stage, UPDRS, BDI, PDQ and FSS scores were found between the two subgroups as a consequence (see **Table 2**), a comparison between the subgroups may make it necessary to include covariates in the analyses. When including covariates into the analysis, only the scores assessed while on medication were used since the scores are correlated between the ON and OFF state. In addition, change scores between the OFF and ON state were compared to account for group differences in the extent of change from one state to another. The extent of change from the medicated to the unmedicated state was significantly greater for the "ON first" subgroup for UPDRS and PDQ scores. Therefore, the following covariates were included when calculating ANCOVA comparing the "ON first" and "OFF first" subgroups: Hoehn and Yahr stage (ON medication), UPDRS score (ON medication), BDI score (ON medication), PDQ score (ON medication), FSS score (ON medication), UPDRS change score (OFF–ON), PDQ change score (OFF–ON). Consequently, estimated marginal means were calculated as is usually the case when accounting for covariates.

Before calculating ANCOVAs, boxplots were generated to display distribution of the data. As suggested by Frigge et al. (1989), data values outside 1.5 IQR (inter-quartile range) were considered as outliers.

For exploratory comparisons between healthy control subjects and the patients in their respective medication states (ON and OFF), multiple independent-sample t-tests between the control group and the patient group as a whole ($N = 15$) in both its medication states (ON and OFF) were calculated. Differences between the patients on versus off medication (the whole group of patients on levodopa vs. the whole group off levodopa) were generally tested with paired t-tests (the prerequisite of normal distribution was fulfilled for all variables). All statistics were calculated using the Statistical Package for Social Sciences (SPSS® 15.0, SPSS Inc.) and MATLAB® programming software (The MathWorks™, version R2007a). When calculating t-tests with unspecific hypotheses, generally, an α-level of 5% was applied. However, because

FIGURE 1 | Setup of the two reward-based learning tasks.
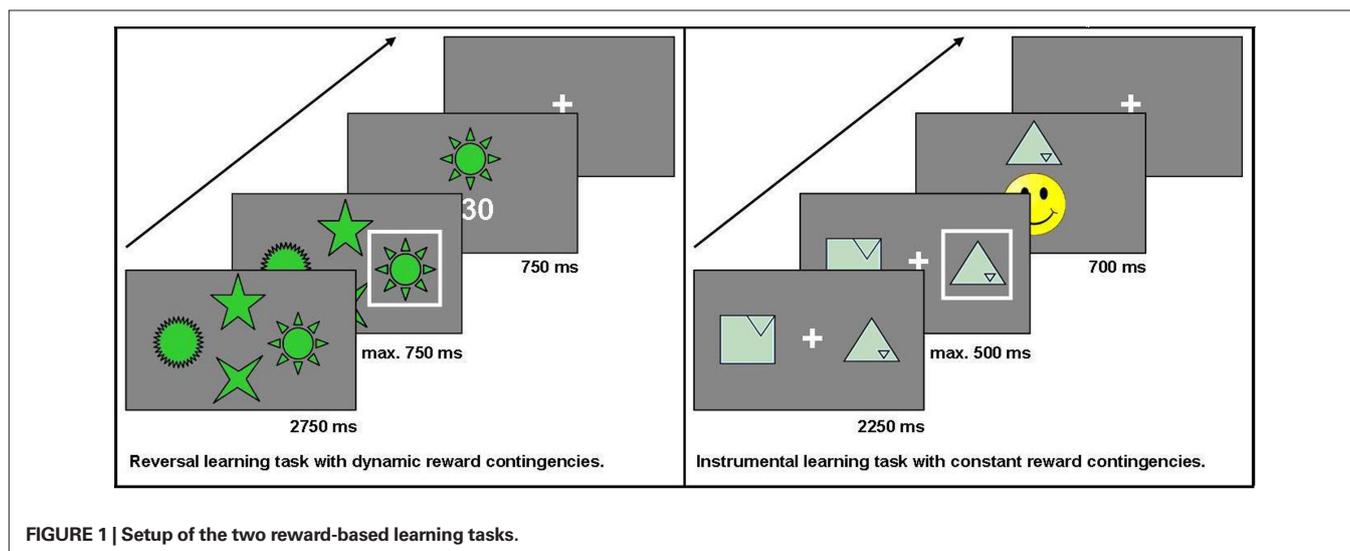
of the problem of α-inflation when calculating multiple *t*-tests, a Bonferroni-adjusted α of 1.67 was used when comparing three groups. Due to technical problems during acquisition of the data for the probabilistic instrumental learning task, data from 2 of the 16 control subjects had to be discarded. Therefore, 14 controls and 15 patients entered the analysis for this task.
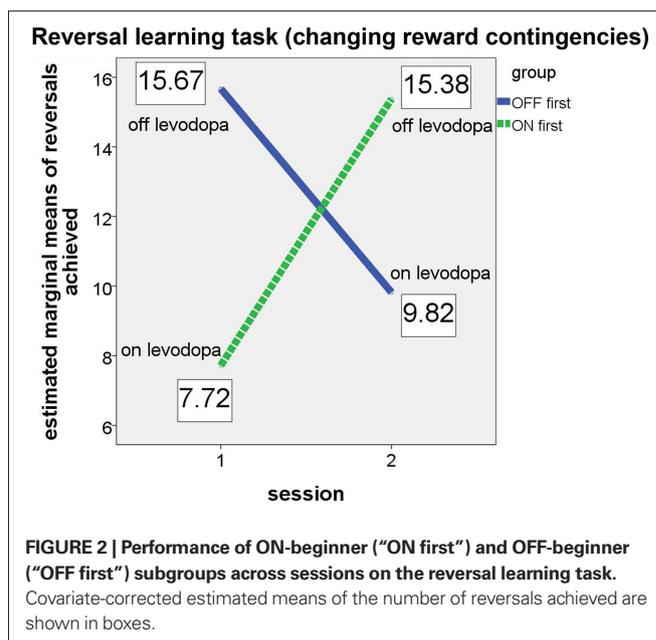
RESULTS

PROBABILISTIC REVERSAL LEARNING TASK WITH CHANGING REWARD CONTINGENCIES

For the reversal learning task, two subjects (one PD patient and one control subject) were identified as statistical outliers (data values outside 1.5 IQR; Frigge et al., 1989) and had to be excluded from the analysis since the sample size was quite small and thus susceptible to deviations. As a consequence, the following results are based on 14 PD patients and 15 control subjects.

Our main focus lay on the comparison of patients on and off dopaminergic medication. Therefore, we compared our two patient subgroups, the "ON first" group beginning on levodopa and the "OFF first" group starting off levodopa. We expected that the "OFF first" subgroup would achieve more reversals on their first session (i.e., while the patients were off medication) than on the second session (while they were on medication). The opposite was expected for the "ON first" subgroup, with less reversals expected on the first assessment (on medication) and more reversals on the second assessment (off medication). It is important to note that, in terms of the ANCOVA, the crucial result would be a significant interaction between group and session and not a significant main effect of the group. Since each of the two subgroups (ON first, OFF first) included patients that were tested both on and off levodopa, a main effect of the group would not be relevant as a sign of differences in the medication states. The interaction of group and session, however, makes it possible to draw conclusions about the effect of medication state on performance.

The results confirmed our hypotheses (see **Figure 2**). The significant interaction ($F(1, 4) = 302.483$, $p < .001$, partial eta^2 = 0.987) and the associated estimated marginal means show that performance on the probabilistic reversal task was indeed better for patients off medication.

Both subgroups achieved a similar number of reversals in the OFF state, irrespective of the session, in which they were assessed without medication. In the ON state, the "OFF first" subgroup achieved approximately two reversals more than the "ON first" subgroup. Since for this subgroup, the ON session was the second assessment, it seems likely that a learning effect might have occurred across the two sessions on top of the effect of medication. For the ANCOVA, the main effect "session" was significant as well ($F(1, 4) = 104.149$, $p = 0.001$, partial eta^2 = 0.963), suggesting that there was indeed a learning effect over sessions for the reversal learning task. Homogeneity of the covariance matrices (Box-M-Test = 2.859, $F(3, 433459) = 0.763$, $p = 0.515$) and equality of error variances was given (Levene's test for reversals [session 1]: $F(1, 11) = 1.579$, $p = 0.235$, and for reversals [session 2]: $F(1, 11) = 1.488$, $p = 0.248$). There was no significant main effect of group ($F(1, 4) = 0.049$, $p = 0.835$). Also, simple effects of group tested for each session separately did not elicit any significant differences between the two subgroups on either session (session 1: $F(1, 4) = 2.163$, $p = .215$; session 2: $F(1, 4) = 1.060$, $p = 0.361$).

FIGURE 2 | Performance of ON-beginner ("ON first") and OFF-beginner ("OFF first") subgroups across sessions on the reversal learning task. Covariate-corrected estimated means of the number of reversals achieved are shown in boxes.

Since our two patient subgroups, the OFF-beginners and the ON-beginners, differed significantly in several clinical parameters (see **Table 2**) due to the consecutive recruitment, we calculated an ANCOVA to ensure that the clinical differences would not mask a potential interaction effect of (sub)group × session. Unfortunately however, the number of degrees of freedom decreased considerably due to the inclusion of covariates. Because the relevant outcome of the ANCOVA was the interaction effect (which looks at the differential session effects *within* two separate subgroups) and not the main effect of group (which would be a *between*-subject comparison), one could argue that a comparison even without the inclusion of covariates may be justifiable. However, it should be noted that the comparison of two heterogeneous patient subgroups without controlling for covariates might mask potential effects of dopaminergic medication on reward-based performance. Therefore, calculation of an ANOVA without controlling for covariates will result in a higher number of degrees of freedom, but will also make it more likely that a potential effect might be masked. We calculated an additional ANOVA without covariates comparing the "OFF first" and "ON first" subgroups to allow for a higher number of degrees of freedom and to confirm whether our results would also hold true when differences between the two subgroups other than the dopaminergic state would not be controlled for in the form of covariates. Again, the relevant interaction group × session was significant ($F(1, 12) = 8.268$, $p = 0.014$, partial eta^2 = 0.408). There were no main effects of group ($F(1, 12) = 0.449$, $p = 0.515$) or session ($F(1, 12) = 1.435$, $p = 0.254$). Homogeneity of the covariance matrices (Box-M-Test = 5.153, $F(3, 25920) = 1.407$, $p = 0.238$) and equality of error variances (Levene's test for reversals [session 1]: $F(1, 12) = 0.740$, $p = 0.406$, and for reversals [session 2]: $F(1, 12) = 0.057$, $p = 0.815$) was given for the ANOVA as well.

Table 3 shows group performances and simple group comparisons between the control group and the PD patients (on and off levodopa). The control group did not differ significantly from the PD

patients in any of the two medication states. Interestingly though, patients achieved a significantly higher number of reversals off medication compared to their performance on medication. This confirms the results of the analyses of (co)variance, which showed subgroups in their respective OFF states (first session for "OFF first" subgroup, second session for "ON first" subgroup) to perform better than in their respective ON states.

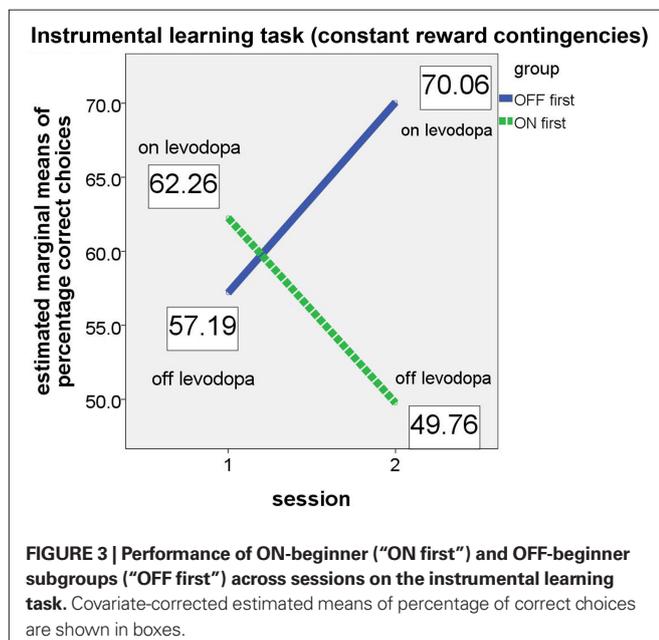### PROBABILISTIC INSTRUMENTAL LEARNING TASK WITH CONSTANT REWARD CONTINGENCIES

For the instrumental learning task with constant reward contingencies, we expected that the "ON first" subgroup would perform better on their first session (i.e., while the patients were on medication) than on the second session (while they were off medication). The opposite was expected for the "OFF first" subgroup, with a poorer performance expected on the first assessment (off medication) and better performance on the second assessment (on medication). As in the analysis of the reversal learning task, this effect should result in a significant group × session interaction in the ANCOVA. Again, the results confirmed our hypotheses (see **Figure 3**).

Table 3 | Number of reversals achieved on the reversal learning task.

	N	Mean	SD	*t*-test for equality of means			
		Group statistics			*t*	df	Sig. (two-tailed)
NC	15	10.73	2.865	NC vs. ON	−0.191	27	0.850
PD off	14	13.14	3.695	NC vs. OFF	1.899	27	0.068
PD on	14	10.50	3.920	ON vs. OFF[a]	3.137	13	0.008[b]

[a]*Paired samples t-test.*
[b]*Sig. difference, p < 0.0167.*

Instrumental learning task (constant reward contingencies)

group
OFF first
ON first

70.06

on levodopa

62.26

on levodopa

57.19

off levodopa

49.76

off levodopa

FIGURE 3 | Performance of ON-beginner ("ON first") and OFF-beginner subgroups ("OFF first") across sessions on the instrumental learning task. Covariate-corrected estimated means of percentage of correct choices are shown in boxes.

Whereas both main effects (for session and group) were not significant (session: $F_{(1, 5)} = 1.347$, $p = 0.298$; group: $F_{(1,5)} = 0.961$, $p = 0.372$), the relevant interaction session × group yielded significance ($F_{(1,5)} = 8.094$, $p = 0.036$, partial eta$^2 = 0.618$), thus supporting the hypothesis that patients on levodopa medication perform better on a simple instrumental learning task than patients off medication. Homogeneity of the covariance matrices (Box-M-Test = 4.095, $F_{(3, 9784)} = 1.109$, $p = 0.344$) and equality of error variances was given (Levene's test for percentage correct [session 1]: $F_{(1, 12)} = .068$, $p = 0.799$, and for percentage correct [session 2]: $F_{(1, 12)} = 2.095$, $p = 0.173$). At first glance, it might seem as if the significant group × session interaction is driven by an effect in the second session since the difference between the two subgroups appears very small in the first session. However, simple group effects tested for each session separately did not show significant differences between the two subgroups in either session (session 1: $F_{(1, 5)} = 0.404$, $p = 0.553$; session 2: $F_{(1, 5)} = 4.259$, $p = 0.094$).

For the instrumental learning task with constant reward contingencies, an ANOVA (no covariates, more degrees of freedom) did not show a significant group × session interaction ($F_{(1,13)} = 0.104$, $p = 0.752$). Main effects for group ($F_{(1,13)} = 1.998$, $p = 0.181$) or session ($F_{(1, 13)} = 0.385$, $p = 0.545$) were also not significant.

Table 4 shows simple group comparisons between control subjects and patients on and off levodopa. No significant differences between the performance of healthy controls and the performance of the PD group (medicated or unmedicated) were found.

REACTION TIMES, MISSED TRIALS AND FINGER-TAPPING TASK

Mean reaction times (see **Table 5**) and the percentage of missed trials (see **Table 6**) did not differ significantly between groups for any task. There were no significant differences between groups with regard to the finger-tapping task either (see **Table 7**), suggesting no marked impairment in elementary motor control.

Unexpectedly, the mean reaction times tended to be largest for the control group. On both tasks, controls took longer to decide by button press than PD patients both on and off medication. On average, PD patients on medication were the quickest to respond. However, none of the differences in reaction time between all three groups was significant.

CORRELATION ANALYSIS OF LEVODOPA DOSAGE WITH TASK PERFORMANCE

The patients differed with regard to their prescribed levodopa dose. The lowest dosage taken was 200 mg/day whereas the highest was 1100 mg/day. However, the daily levodopa dosage (in mg) of PD patients was not found to correlate with either the percentage of correct choices in the simple probabilistic learning task ($N = 15$; ON: $r = 0.158$, $p = 0.574$; OFF: $r = -0.008$, $p = 0.978$) or the number of reversals achieved on the reversal task ($N = 14$; ON: $r = -.254$, $p = 0.381$; OFF: $r = -0.131$, $p = 0.655$; see **Figure 4**).

DISCUSSION

This study aimed at investigating the role of DA in reward-based learning depending on whether reward contingencies are constant or dynamic. Learning can be considered as an ongoing process of trying to achieve

Table 4 | Percentage of correct choices on the instrumental learning task with constant reward contingencies.

	Group statistics			t-test for equality of means			
	N	Mean	SD		t	df	Sig. (two-tailed)
NC	14	68.493	14.017	NC vs. ON	1.802	27	0.083
PD on	15	60.068	11.081	NC vs. OFF	−2.185	27	0.038
PD off	15	59.048	8.856	ON vs. OFF[a]	−0.374	14	0.714

[a]Paired samples t-test.

Table 5 | Mean reaction times and t-test results for the reward-based learning tasks in healthy controls (NC) and PD patients on levodopa (ON) and off levodopa (OFF).

	Group characteristics			t-test for equality of means			
	N	Mean	SD		t	df	Sig. (two-tailed)
INSTRUMENTAL LEARNING TASK							
NC	14	586.6	75.9	NC vs. ON	−1.953	27	0.061
PD ON	15	516.4	112.7	NC vs. OFF	−1.225	27	0.231
PD OFF	15	545.8	100.7	ON vs. OFF	1.031	14	0.320
REVERSAL LEARNING TASK							
NC	15	512.9	72.2	NC vs. ON	−1.085	27	0.287
PD ON	14	479.0	95.6	NC vs. OFF	−0.419	27	0.679
PD OFF	14	500.3	89.6	ON vs. OFF	0.945	13	0.362

Table 6 | Mean percentage of missed trials across both reward-based learning tasks.

	Group characteristics			t-test for equality of means			
	N	Mean	SD		t	df	Sig. (two-tailed)
INSTRUMENTAL LEARNING TASK							
NC	14	3.60	3.20	NC vs. ON	1.033	27	0.311
PD ON	15	6.31	9.29	NC vs. OFF	1.141	17.894	0.269
PD OFF	15	6.36	8.76	ON vs. OFF	−0.046	14	0.964
REVERSAL LEARNING TASK							
NC	15	2.89	3.23	NC vs. ON	−0.116	27	0.909
PD ON	14	2.76	3.17	NC vs. OFF	0.300	27	0.767
PD OFF	14	3.29	3.85	ON vs. OFF	0.854	13	0.409

an optimal adaptation to the environment, in our case the task environment or the properties of a feedback-based task. Within reward-based learning, an important property of a task is the way in which stimuli are associated with their respective feedback. These stimulus-reward associations can be either constant over time or dynamic. Dynamic reward contingencies are much more exemplary for a "real world" envi-

ronment and can be illustrated in a reversal learning task, a task that requires participants to adapt their behavior according to switches in the rewarding properties of the choices at hand.

It is well known that feedback-based learning is affected by PD (e.g., Knowlton et al., 1996). However, different accounts have been given depending on the nature of the feedback-based task used (e.g., Perretta et al., 2005) and the medication status of the patients studied (e.g., Cools et al., 2001a). Tying in with those findings, we set out to evaluate whether the medication state in PD patients has a differential effect depending on the stimulus-reward properties of two feedback-based tasks.
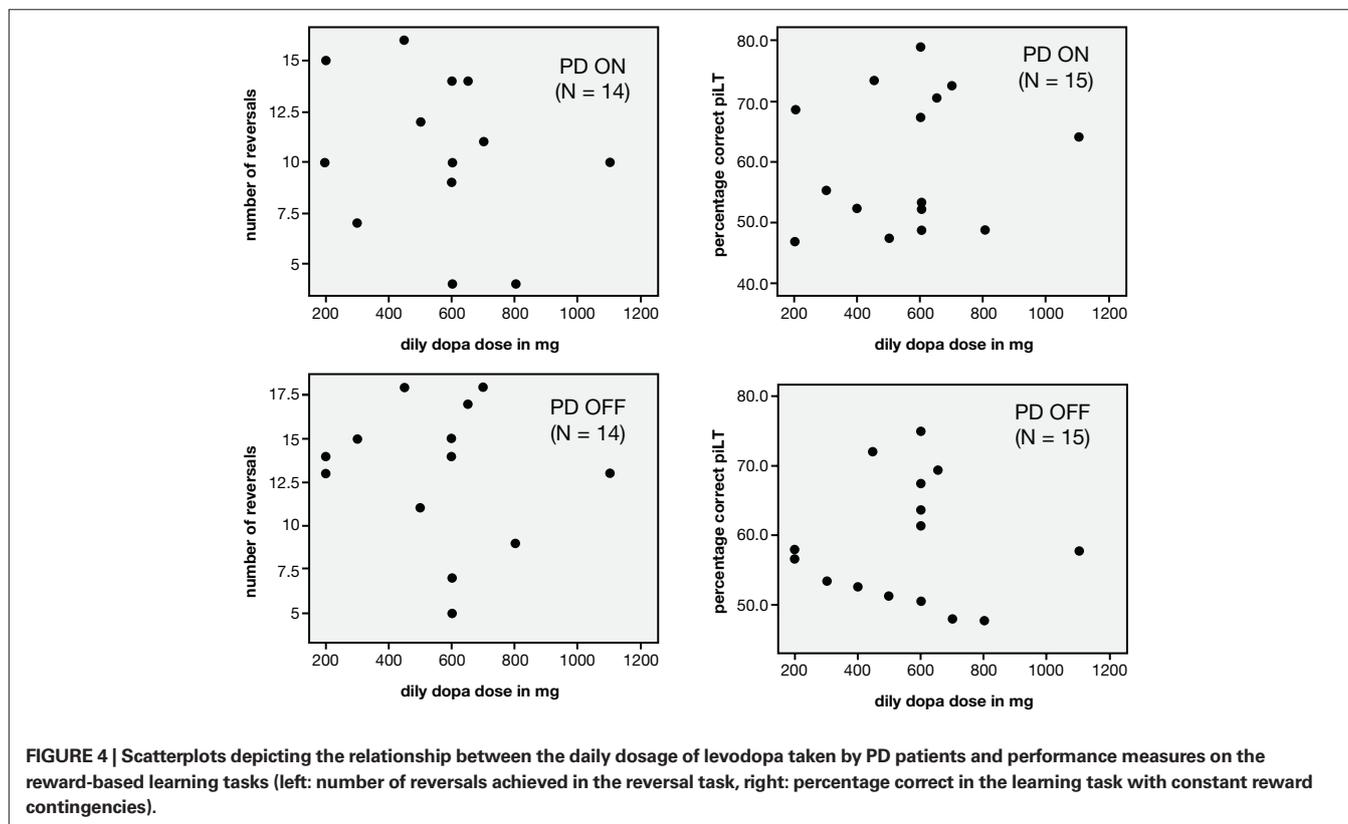
We compared the performance of PD patients on and off levodopa on two reward-based learning tasks that differ in their stimulus-reward contingencies. Whereas the instrumental learning task required participants to learn which of two options elicited higher rewards, the reversal learning task did not primarily examine option-outcome association learning but mainly the unlearning of those associations against the background of changing reward contingencies. We found a double dissociation between patients' medication state and performance on the reward-based learning tasks: Patients on medication achieved a higher percentage of correct choices on the simple option-outcome association task characterized by constant reward contingencies, which, much like habit learning in animals, would be expected to primarily rely on dorsal striatal structures (Yin et al., 2004; Seger, 2006; but see De Wit et al., 2010). Conversely, when patients were off medication, they performed better on a reversal learning task with dynamic reward contingencies, i.e., they achieved more reversals than on medication. Reversal learning, unlike simple instrumental learning, requires unlearning of previously installed stimulus-reward associations, which crucially depends on ventral striato-frontal circuits including the ventral striatum and the orbitofrontal cortex (Rolls, 2000; Cools et al., 2007).

Our analysis showed a very strong effect of medication on reversal learning, which was not only detected by means of a paired t-test between all patients on levodopa and all patients off levodopa, but also by means of a repeated measures analysis of variance across two (clinically heterogeneous) patient subgroups. The relevant interaction between subgroup (OFF-beginners vs. ON-beginners) and time of testing (first vs. second session) was significant, both when controlling for differences between the patient subgroups by including covariates and also across the heterogeneous subgroups (without covariates). For the instrumental learning task with constant reward contingencies, the effect of levodopa was less distinct. For this task, to show a significant interaction between subgroup and time of testing the covariates needed to be included in the analysis of (co)variance. This suggests that the differences between the two patient subgroups (ON-beginners vs. OFF-beginners) on clinical scales like the UPDRS or PDQ might have masked the medication effect in the ANOVA without covariates, and that the effect of levodopa on instrumental learning can only be seen when other possible differences between the subgroups are controlled for. One possible explanation for the weaker effect in the instrumental learning task is that it was relatively simple; participants also considered it to be the easier task. Given that there was no ceiling effect, such an explanation seems unlikely though. An alternative explanation is that participants searched – despite differ-

Table 7 | Means and standard deviations for the number of taps in the finger-tapping task for normal controls (NC) and PD patients.

	NC	PD ON	PD OFF		t	df	Sig. (two-tailed)
Right hand	118.6 ± 34.5	115.7 ± 38.7	117.5 ± 39.4	Tapping right:			
				NC vs. ON	0.218	27	0.829
				NC vs. OFF	0.085	27	0.933
				ON vs. OFF	0.331	14	0.745
				Tapping left:			
Left hand	115.4 ± 23.6	112.7 ± 38.8	112.5 ± 44.1	NC vs. ON	0.224	27	0.824
				NC vs. OFF	0.223	27	0.825
				ON vs. OFF	−0.035	14	0.973

Independent and paired-samples t-tests were calculated to compare mean number of taps achieved within 30 s.

FIGURE 4 | Scatterplots depicting the relationship between the daily dosage of levodopa taken by PD patients and performance measures on the reward-based learning tasks (left: number of reversals achieved in the reversal task, right: percentage correct in the learning task with constant reward contingencies).

ing instructions – for patterns in the reward sequences. When no pattern is present, such behavior introduces noise in performance scores (Wolford et al., 2004). Due to the weaker effects achieved on the instrumental learning task, we suggest that future studies should further investigate the effect of levodopa monotherapy on learning of constant stimulus-reward associations. Ideally, those studies should match ON-beginners and OFF-beginners according to clinical features such as UPDRS, Hoehn and Yahr, fatigue, depression and quality of life scores. Unfortunately, our patients had to be recruited in a consecutive manner, which resulted in subgroups that differed on those clinical measures and which in turn led us to control for those differences by means of covariates.

The inclusion of covariates, however, results in a decreased number of degrees of freedom in the statistical analysis. Any future investigation of the effect of levodopa on simple instrumental learning would benefit from studying a larger sample of patients with better comparable patient subgroups. From the results achieved in the present study, we can reason that there is a very strong (detrimental) effect of levodopa on feedback-based learning characterized by changing reward contingencies. The effect of levodopa on feedback-based learning characterized by constant reward contingencies is less clear. Nevertheless, there is some evidence for a beneficial effect of levodopa on this kind of learning, which should be investigated further.

We tested a group of healthy volunteers in addition to the patient sample to explore how patients on and off levodopa compare to normal control subjects on our two reward-based tasks. When pooled together, PD patients did not differ significantly from the control group on the reward-based tasks. In the case of the reversal learning task, this finding is not surprising as performance of controls lay in between that of patients on and patients off medication. This might suggest that the PD patients' performance is in the same range as that of healthy elderly subjects, a finding that is in line with previous studies which did not report a difference between the performance of PD patients and elderly control subjects on feedback-based tasks (e.g., Schmitt-Eliassen et al., 2007). With regard to the simple instrumental learning task, we suppose that different factors have contributed to the lacking group differences. Despite a tendency of controls to perform best and PD patients off medication to perform worst, none of the group comparisons were significant. One reason might be that the instrumental learning task was comparatively easy. Furthermore, our sample size was relatively small and our PD patients were only mildly to moderately affected by PD (Hoehn and Yahr score < 4). In addition, most patients had relatively short disease and symptom durations. This is reflected in the fact that we did not find any group differences with regard to reaction times or trial omissions in either task nor did we find significant differences on the finger-tapping task conducted to control for performance differences due to motor impairments. More pronounced performance differences might be detected in later stages of the disease.

Similarly, we found no significant correlations between the levodopa dose and any of the reward-based learning tasks, neither in the "ON" nor in the "OFF" condition. This is in accord with a study by Jahanshahi et al. (2010) who reported that the levodopa equivalent dose in their PD patients and the performance on a probabilistic classification paradigm were not significantly correlated.

We tested the control subjects only once. Since our patients were tested twice (once on, once off medication), a better comparison between patients and control subjects would have been possible if we had tested the controls twice as well. Because our focus lay on the comparison of the different medication states of the patients, we attempted to minimize the influence of a potential practice effect by dividing the patient sample into two subgroups – one starting off levodopa, the other one on levodopa – and thereby evening out the session effect across medication states. The fact that we found significant group × session interactions for the "ON first" and "OFF first" subgroups, despite potential learning effects that might have occurred on the second assessment, argues for the strength of our findings. If a subgroup performs worse on the second session (which was the case for the OFF-beginners on the reversal learning task and for the ON-beginners on the instrumental learning task) even though one might expect an advantage with increasing practice, it emphasizes the strong influence of the dopaminergic medication on the task performance. Nevertheless, it is advisable for future studies with a similar design to test both the patients and the control group twice to allow for a better comparison between those groups.

The double dissociation between medication state and performance in two reward-based learning tasks is compatible with the "overdose hypothesis", a theory first expressed by Gotham et al. (1988). This theory predicts that PD patients on dopaminergic medication are better at tasks that require dorsal striato-frontal structures (which are affected more severely in PD), but perform more poorly on tasks requiring the ventral striatum or orbitofrontal cortex, parts of the brain that are less severely affected. This is due to the "overdosing" effect of dopaminergic medication with DA doses necessary to restore DA levels in the dorsal striatum being too high for the less affected ventral structures. Evidence for this theory comes from a number of studies (e.g., Cools et al., 2001a; Mehta et al., 2001). Instrumental learning of constant stimulus-reward associations has been linked repeatedly to dorsal striatal structures (Yin et al., 2004; Kimchi et al., 2009; Shiflett et al., 2010), whereas reversal learning has been found to rely on ventral striatal (Cools et al., 2002; Heekeren et al., 2007) and ventral prefrontal/orbitofrontal structures (Rolls, 2000; Xue et al., 2008). By applying these findings to our tasks (i.e., dorsal structures underlying the instrumental learning task and ventral structures underlying the reversal learning task), our results confirm the predictions of the "overdose hypothesis".

On the basis of the "overdose hypothesis", the framework provided by Frank and colleagues allows a more detailed interpretation of our results (Frank et al., 2004; Frank, 2005). Starting from the finding that PD patients off medication learn better from negative feedback, whereas medicated patients are better able to learn from positive feedback, the theory emphasizes the relevance of a large dynamic range of the DA signal for feedback learning. The ability to produce DA dips is important for learning from negative feedback, whereas the production of DA bursts is crucial for learning from positive feedback. Frank et al. (2004) suggest that these abilities are impaired in PD patients. Following their argumentation, performance differences on our reversal learning task can be explained in terms of learning from negative feedback: Since the aim of the reversal task was to find out the highest-rewarded option at any moment in the course of the experiment, choices for the formerly "good" option resulted in bad outcomes after the occurrence of a reversal, which, in turn, signaled the need to switch choice behavior. Even though feedback was in the positive range only (i.e., there were no negative points or "losses"), outcomes lower than outcome expectations result in negative prediction errors (Hollerman and Schultz, 1998; Bayer and Glimcher, 2005) and are perceived as negative feedback. According to the framework of Frank et al. (2004), DA overdosing of task-relevant ventral structures hinders the effectiveness of phasic DA dips in bringing DA levels below the threshold necessary for learning from negative feedback. In contrast, unmedicated patients with near-normal DA levels in ventral striato-frontal structures are able to produce effective DA dips in response to negative prediction errors. The higher DA levels of medicated patients would be advantageous in the simpler instrumental learning task, where subjects have to rely on positive prediction errors to learn which of two alternatives is more beneficial. Accordingly, unmedicated patients perform worse in this task due to their decreased DA levels in the task-relevant dorsal striatal structures, which make DA bursts less effective.

The DA "denervation hypothesis" (e.g., Kulisevsky, 2000) has been brought up as an alternative explanation of medication-dependent differences in task performance. According to this hypothesis, levodopa-induced cognitive impairment depends on

the extent of DA cell loss in particular brain sites. The theory states that "frontal" disturbances occur only after a massive degeneration of the DA system. Consequently, "overdose hypothesis" and "denervation hypothesis" make different assumptions in regard to the role of disease severity in DA-related cognitive impairment. Whereas the "overdose hypothesis" predicts greater impairment upon DA administration in less affected patients, the "denervation hypothesis" opts for a more marked impairment in the severely affected. Since we tested patients with relatively short symptom and disease durations that suffered from mild to moderate PD, our data do not allow drawing conclusions about reward-based learning impairments in severely affected patients. However, since a large dynamic range of the DA signal has been emphasized to be needed for the appropriate weighting of feedback (Frank et al., 2004), it is conceivable that along with progressive cell loss in ventral striato-frontal loops, DA medication might not be able to restore firing patterns needed to effectively encode learning signals in reversal learning. Therefore, the "overdose hypothesis" and the "denervation hypothesis" might be less contradictory than it seems at first glance, especially in later stages of the disease. Further research is needed to evaluate to what extent dopaminergic medication exerts beneficial or detrimental effects on functions relying on ventral striato-frontal circuits in advanced stages of the disease.

One advantage of the present study is the within-subject design that was used to compare the different medication states in PD. Due to the within-subject design, it is coherent that our results can be ascribed to DA availability and not to other pre-existing group differences that might have existed between groups in between-subject designs. That our patients received levodopa monotherapy has to be mentioned as an additional advantage. We included only patients on levodopa to avoid heterogeneity in the dopaminergic treatment. Some of the studies concerned with reward-based performance in PD have investigated patients that were not exclusively treated with levodopa (e.g., Shohamy et al., 2006; Jahanshahi et al., 2010), thus bearing the risk of confounds. DA agonists have different pharmacological properties than levodopa. Since levodopa is a precursor to DA, it influences both tonic and phasic DA transmission (see Breitenstein et al., 2006). It is taken up by dopaminergic neurons and, after conversion to DA, can be released into the synaptic cleft in a stimulation-dependent manner. DA agonists are not taken up by DA neurons and thus do not increase presynaptic DA availability and cannot be released phasically. Instead, they exert their DA-mimicking function through binding to postsynaptic DA receptors and thereby influence tonic DA levels. This change in tonic DA is believed to reduce the effect of phasic, stimulation-dependent endogenous DA release (Breitenstein et al., 2006). Consequently, while levodopa allows or even boosts phasic DA bursts and dips while simultaneously raising tonic DA levels, DA agonists only increase tonic DA levels. DA agonists act tonically on postsynaptic receptors of the D1 or D2 type, but also on presynaptic D2-like autoreceptors. For that reason, different mechanisms of action are possible: either, occupation of postsynaptic DA receptors may reduce the neurotransmission of phasic postsynaptic DA signals, or stimulation of sensitive presynaptic autoreceptors may reduce DA release and thus phasic DA signaling (Breitenstein et al., 2006; Cools, 2006), e.g., the phasic bursts believed to encode the reward

prediction error (Schultz and Dickinson, 2000). This should result in a narrower range of DA firing and a consecutive impairment to reach the thresholds necessary for learning both from positive and negative feedback in PD patients treated with a DA agonist. Accordingly, even in healthy subjects, administration of DA agonists impairs performance both in simple instrumental learning tasks with constant stimulus-reward associations and reversal learning (e.g., Mehta et al., 2001; Pizzagalli et al., 2008; Boulougouris et al., 2009; Haluk and Floresco, 2009; Santesso et al., 2009). As a consequence, in contrast to levodopa, DA agonists would be expected to impair performance on our instrumental learning task, a task relying on positive feedback, i.e., phasic DA bursts.

The differing medication might explain some inconsistencies between beneficial medication (levodopa monotherapy) effects on our instrumental learning task and detrimental medication (levodopa plus other antiparkinsonian drugs) effects on instrumental paradigms in PD patients (Shohamy et al., 2006; Jahanshahi et al., 2010). However, a closer look at those studies reveals that the tasks conducted are not quite comparable to our tasks. For instance, a recent study by Jahanshahi et al. (2010) found PD patients on medication to be impaired on the WPT. Knowlton et al. (1996) had associated the WPT with the dorsal striatum. Therefore, the findings by Jahanshahi et al. (2010) seem to contradict our result that the performance on a task linked to dorsal striato-frontal circuitry is improved on medication. However, the authors cite evidence from an imaging task that found ventral striatal activations during a parallel version of the WPT (Rodriguez et al., 2006). Considering this evidence, the results are in agreement with our finding that medicated patients are impaired on a task associated with ventral striato-frontal circuitry. Another study seems to contradict our finding of enhanced instrumental learning on levodopa. In this study, PD patients on levodopa were found to be impaired on a two-phase learning and transfer task (Shohamy et al., 2006). In the first phase, the subjects learned a concurrent discrimination. For eight discrimination pairs, they had to find out which of the two objects within each pair was the rewarded one. When compared to patients off levodopa, medicated patients performed worse. This seems to argue against our findings. Yet, the task used in that study is also unlike our instrumental learning task. The main difference is that the task used by Shohamy et al. (2006) was characterized by a higher memory load because the discrimination pairs were not presented consecutively but rather in a concurrent fashion. When tested on a version of the task with reduced memory load, i.e., when the discrimination pairs were presented successively, as was the case in our instrumental learning task, PD patients on levodopa were not impaired anymore compared to healthy control subjects. This suggests that the working memory requirements of a task influence performance on reward-based tasks as well. Unfortunately, patients off levodopa were not tested on this reduced memory load version of the task.

Similar relationships between subcortical DA action and performance in reward-based learning tasks have been shown for both genetically determined interindividual differences in healthy subjects and differences imposed by pharmacological challenge of healthy subjects. Pessiglione et al. (2006) found enhanced reward-based performance when administering levodopa, but impaired performance in a simple probabilistic instrumental learning

paradigm when administering the DA antagonist haloperidole to healthy adults. This relationship is consistent with our finding that PD patients on medication perform better on the simple instrumental learning task than PD patients off medication. Furthermore, in a reversal learning task similar to the one applied here, healthy carriers of a polymorphism that presumably leads to elevated prefrontal and subcortical DA tones achieved less reversals than individuals without the mutation (Krugel et al., 2009). Thus, a detrimental association of elevated DA tone on reversal learning performance seems to be observable already in a normal range of fluctuation in healthy subjects.

By showing a detrimental effect of levodopa on the performance in a reward-based learning task characterized by dynamic reward contingencies (and associated with ventral striato-frontal circuitry) and a beneficial effect on the performance in a task characterized by constant reward contingencies (associated with dorsal striato-frontal circuitry), the present study demonstrates in a within-subject design that reward-based task performance differs depending on the availability of levodopa, on specific features of the task (stimulus-reward contingencies) as well as on underlying striato-frontal circuits.

ACKNOWLEDGMENTS

REFERENCES

Arnsten, A. F. T. (1998). Catecholamine modulation of prefrontal cortical cognitive function. *Trends Cogn. Sci.* 2, 436–447.

Bayer, H. M., and Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47, 129–141.

Beck, A. T., Erbaugh, J., Ward, C. H., Mock, J., and Mendelsohn, M. (1961). An inventory for measuring depression. *Arch. Gen. Psychiatry* 4, 561–571.

Boulougouris, V., Castane, A., and Robbins, T. W. (2009). Dopamine D2/D3 receptor agonist quinpirole impairs spatial reversal learning in rats: investigation of D3 receptor involvement in persistent behavior. *Psychopharmacology (Berl.)* 202, 611–620.

Breitenstein, C., Korsukewitz, C., Floel, A., Kretzschmar, T., Diederich, K., and Knecht, S. (2006). Tonic dopaminergic stimulation impairs associative learning in healthy subjects. *Neuropsychopharmacology* 31, 2552–2564.

Cools, R. (2006). Dopaminergic modulation of cognitive function-implications for L-DOPA treatment in Parkinson's disease. *Neurosci. Biobehav. Rev.* 30, 1–23.

Cools, R., Barker, R. A., Sahakian, B. J., and Robbins, T. W. (2001a). Enhanced or impaired cognitive function in Parkinson's disease as a function of dopaminergic medication and task demands. *Cereb. Cortex* 11, 1136–1143.

Cools, R., Barker, R. A., Sahakian, B. J., and Robbins, T. W. (2001b). Mechanisms of cognitive set flexibility in Parkinson's disease. *Brain* 124, 2503–2512.

Cools, R., Barker, R. A., Sahakian, B. J., and Robbins, T. W. (2003). L-Dopa medication remediates cognitive inflexibility, but increases impulsivity in patients with Parkinson's disease. *Neuropsychologia* 41, 1431–1441.

Cools, R., Clark, L., Owen, A. M., and Robbins, T. W. (2002). Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *J. Neurosci.* 22, 4563–4567.

Cools, R., Lewis, S. J. G., Clark, L., Barker, R. A., and Robbins, T. W. (2007). L-DOPA disrupts activity in the nucleus accumbens during reversal learning in Parkinson's disease. *Neuropsychopharmacology* 32, 180–189.

Czernecki, V., Pillon, B., Houeto, J. L., Pochon, J. B., Levy, R., and Dubois, B. (2002). Motivation, reward, and Parkinson's disease: influence of dopatherapy. *Neuropsychologia* 40, 2257–2267.

De Wit, S., Barker, R. A., Dickinson, T., and Cools, R. (2010). Habitual versus goal-directed action control in Parkinson's disease. *J. Cogn. Neurosci.* doi: 10.1162/jocn.2010.21514. [Epub ahead of print].

Fahn, S., Elton, R. L., Members of the UPDRS Development Committee (1987). "Unified Parkinson's disease rating scale," in *Recent Developments in Parkinson's Disease,* Vol. 2, eds S. Fahn, C. D. Marsden, M. D. Goldstein, and D. B. Calne (Florham Park, NJ: Macmillan Healthcare Information), 153–163.

Fellows, L. K., and Farah, M. J. (2005). Different underlying impairments in decision-making following ventromedial and dorsolateral frontal lobe damage in humans. *Cereb. Cortex* 15, 58–63.

Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). Mini-mental state: a practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* 12, 189–198.

Frank, M. J. (2005). Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. *J. Cogn. Neurosci.* 17, 51–72.

Frank, M. J., Seeberger, L. C., and O'Reilly, R. C. (2004). By carrot or by stick: cognitive reinforcement learning in Parkinsonism. *Science* 306, 1940–1943.

Frigge, M., Hoaglin, D. C., and Iglewicz, B. (1989). Some implementations of the boxplot. *Am. Stat.* 43, 50–54.

Gotham, A. M., Brown, R. G., and Marsden, C. D. (1988). Frontal cognitive function in patients with parkinsons-disease on and off levodopa. *Brain* 111, 299–321.

Haluk, D. M., and Floresco, S. B. (2009). Ventral striatal dopamine modulation of different forms of behavioral flexibility. *Neuropsychopharmacology* 34, 2041–2052.

Heekeren, H. R., Wartenburger, I., Marschner, A., Mell, T., Villringer, A., and Reischies, F. M. (2007). Role of ventral striatum in reward-based decision making. *Neuroreport* 18, 951–955.

Hoehn, M. M., and Yahr, M. D. (1967). Parkinsonism: onset, progression, and mortality. *Neurology* 17, 427–442.

Hollerman, J. R., and Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat. Neurosci.* 1, 304–309.

Jahanshahi, M., Wilkinson, L., Gahir, H., Dharminda, A., and Lagnado, D. A. (2010). Medication impairs probabilistic classification learning in Parkinson's disease. *Neuropsychologia* 48, 1096–1103.

Kessler, J., Calabrese, P., and Kalbe, E. (2006). Dementia assessment with PANDA: a new screening instrument to assess cognitive and effective dysfunction in Parkinson's disease. *Eur. J. Neurol.* 13, 66.

Kimchi, E. Y., Torregrossa, M. M., Taylor, J. R., and Laubach, M. (2009). Neuronal correlates of instrumental learning in the dorsal striatum. *J. Neurophysiol.* 102, 475–489.

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

Krugel, L. K., Biele, G., Mohr, P. N., Li, S. C., and Heekeren, H. R. (2009). Genetic variation in dopaminergic neuromodulation influences the ability to rapidly and flexibly adapt decisions. *Proc. Natl. Acad. Sci. U.S.A.* 106, 17951–17956.

Krupp, L. B., LaRocca, N. G., Muir-Nash, J., and Steinberg, A. D. (1989). The fatigue severity scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. *Arch. Neurol.* 46, 1121–1123.

Kulisevsky, J. (2000). Role of dopamine in learning and memory–Implications for the treatment of cognitive dysfunction in patients with Parkinson's disease. *Drugs Aging* 16, 365–379.

Manes, F., Sahakian, B., Clark, L., Rogers, R., Antoun, N., Aitken, M., and Robbins, T. (2002). Decision-making processes following damage to the prefrontal cortex. *Brain* 125, 624–639.

Marschner, A., Mell, T., Wartenburger, I., Villringer, A., Reischies, F. M., and Heekeren, H. R. (2005). Reward-based decision-making and aging. *Brain Res. Bull.* 67, 382–390.

Mehta, M. A., Swainson, R., Ogilvie, A. D., Sahakian, B. J., and Robbins, T. W. (2001). Improved short-term spatial memory but impaired reversal learning following the dopamine D2 agonist bromocriptine in human volunteers. *Psychopharmacology (Berl.)* 159, 10–20.

Mell, T., Heekeren, H. R., Marschner, A., Wartenburger, I., Villringer, A., and Reischies, F. M. (2005). Effect of aging on stimulus-reward association learning. *Neuropsychologia* 43, 554–563.

Moustafa, A. A., Cohen, M. X., Sherman, S. J., and Frank, M. J. (2008a). A role

for dopamine in temporal decision making and reward maximization in Parkinsonism. *J. Neurosci.* 28, 12294–12304.

Moustafa, A. A., Sherman, S. J., and Frank, M. J. (2008b). A dopaminergic basis for working memory, learning and attentional shifting in Parkinsonism. *Neuropsychologia* 46, 3144–3156.

Nieoullon, A. (2002). Dopamine and the regulation of cognition and attention. *Prog. Neurobiol.* 67, 53–83.

Packard, M. G., and Knowlton, B. J. (2002). Learning and memory functions of the Basal Ganglia. *Annu. Rev. Neurosci.* 25, 563–593.

Perretta, J. G., Pari, G., and Beninger, R. J. (2005). Effects of Parkinson disease on two putative nondeclarative learning tasks: probabilistic classification and gambling. *Cogn. Behav. Neurol.* 18, 185–192.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., and Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* 442, 1042–1045.

Peto, V., Jenkinson, C., Fitzpatrick, R., and Greenhall, R. (1995). The development and validation of a short measure of functioning and well being for individuals with Parkinson's disease. *Qual. Life Res.* 4, 241–248.

Pizzagalli, D. A., Evins, A. E., Schetter, E. C., Frank, M. J., Pajtas, P. E., Santesso, D. L., and Culhane, M. (2008). Single dose of a dopamine agonist impairs reinforcement learning in humans: behavioral evidence from a laboratory-based measure of reward responsiveness. *Psychopharmacology (Berl.)* 196, 221–232.

Reischies, F. M. (1999). Pattern of disturbance of different ventral frontal functions in organic depression. *Ann. N. Y. Acad. Sci.* 877, 775–780.

Rey, A. (1964). *L'Examen Clinique en Psychologie*. Paris: Press Universitaire de France.

Rodriguez, P. F., Aron, A. R., and Poldrack, R. A. (2006). Ventral-striatal/nucleus-accumbens sensitivity to prediction errors during classification learning. *Hum. Brain Mapp.* 27, 306–313.

Rolls, E. T. (2000). The orbitofrontal cortex and reward. *Cereb. Cortex* 10, 284–294.

Rutledge, R. B., Lazzaro, S. C., Lau, B., Myers, C. E., Gluck, M. A., and Glimcher, P. W. (2009). Dopaminergic drugs modulate learning rates and perseveration in Parkinson's patients in a dynamic foraging task. *J. Neurosci.* 29, 15104–15114.

Santesso, D. L., Evins, A. E., Frank, M. J., Schetter, E. C., Bogdan, R., and Pizzagalli, D. A. (2009). Single dose of a dopamine agonist impairs reinforcement learning in humans: evidence from event-related potentials and computational modeling of striatal-cortical function. *Hum. Brain Mapp.* 30, 1963–1976.

Schmitt-Eliassen, J., Ferstl, R., Wiesner, C., Deuschl, G., and Witt, K. (2007). Feedback-based versus observational classification learning in healthy aging and Parkinson's disease. *Brain Res.* 1142, 178–188.

Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annu. Rev. Psychol.* 57, 87–115.

Schultz, W., and Dickinson, A. (2000). Neuronal coding of prediction errors. *Annu. Rev. Neurosci.* 23, 473–500.

Seger, C. A. (2006). The basal ganglia in human learning. *Neuroscientist.* 12, 285–290.

Shiflett, M. W., Brown, R. A., and Balleine, B. W. (2010). Acquisition and performance of goal-directed instrumental actions depends on ERK signaling in distinct regions of dorsal striatum in rats. *J. Neurosci.* 30, 2951–2959.

Shohamy, D., Myers, C. E., Geghman, K. D., Sage, J., and Gluck, M. A. (2006). L-dopa impairs learning, but spares generalization, in Parkinson's disease. *Neuropsychologia* 44, 774–784.

Shohamy, D., Myers, C. E., Grossman, S., Sage, J., Gluck, M. A., and Poldrack, R. A. (2004). Cortico-striatal contributions to feedback-based learning: converging data from neuroimaging and neuropsychology. *Brain* 127, 851–859.

Swainson, R., Rogers, R. D., Sahakian, B. J., Summers, B. A., Polkey, C. E., and Robbins, T. W. (2000). Probabilistic learning and reversal deficits in patients with Parkinson's disease or frontal or temporal lobe lesions: possible adverse effects of dopaminergic medication. *Neuropsychologia* 38, 596–612.

Williams, G. V., and Goldman-Rakic, P. S. (1995). Modulation of memory fields by dopamine D1 receptors in prefrontal cortex. *Nature* 376, 572–575.

Wolford, G., Newman, S. E., Miller, M. B., and Wig, G. S. (2004). Searching for patterns in random sequences. *Can. J. Exp. Psychol.* 58, 221–228.

Xue, G., Ghahremani, D. G., and Poldrack, R. A. (2008). Neural substrates for reversing stimulus-outcome and stimulus-response associations. *J. Neurosci.* 28, 11196–11204.

Yin, H. H., Knowlton, B. J., and Balleine, B. W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur. J. Neurosci.* 19, 181–189.