Dopamine Selectively Modulates the Outcome of Learning Unnatural Action–Valence Associations

Nelleke C. Van Wouwe, Daniel O. Claassen, Joseph S. Neimat*,
Kristen E. Kanoff, and Scott A. Wylie*

Abstract

■ Learning the contingencies between stimulus, action, and outcomes is disrupted in disorders associated with altered dopamine (DA) function in the BG, such as Parkinson disease (PD). Although the role of DA in learning to act has been extensively investigated in PD, the role of DA in "learning to withhold" (or inhibit) action to influence outcomes is not as well understood. The current study investigated the role of DA in learning to act or to withhold action to receive rewarding, or avoid punishing outcomes, in patients with PD tested "off" and "on" dopaminergic medication ($n = 19$) versus healthy controls ($n = 30$). Participants performed a reward-based learning task that orthogonalized action and outcome valence (action–reward, inaction–reward, action–punishment, inaction–punishment). We tested whether DA would bias learning toward action, toward reward, or to particular action–outcome interactions. All participants demonstrated inherent learning biases preferring action with reward and inaction to avoid punishment, and this was unaffected by medication. Instead, DA produced a complex modulation of learning less natural action–outcome associations. "Off" DA medication, patients demonstrated impairments in learning to withhold action to gain reward, suggesting a difficulty to overcome a bias toward associating inaction with punishment avoidance. On DA medication, these patterns changed, and patients showed a reduced ability to learn to act to avoid punishment, indicating a bias toward action and reward. The current findings suggest that DA in PD has a complex influence on the formation of action–outcome associations, particularly those involving less natural linkages between action and outcome valence. ■

INTRODUCTION

An emerging literature reveals that action control processes and reward/punishment learning processes, which are typically studied in isolation, are highly interactive (van Wouwe et al., 2015; Chowdhury, Guitart-Masip, Lambert, Dayan, et al., 2013; Chowdhury, Guitart-Masip, Lambert, Dolan, & Düzel, 2013; Guitart-Masip, Chowdhury, et al., 2012; Guitart-Masip, Huys, et al., 2012). Longstanding theories linking frontal BG circuitries, particularly dopamine (DA) modulation, to action selection and reinforcement learning mechanisms have elevated these networks as leading candidates for the integrative formation of stimulus–action–outcome associations (Aron et al., 2007; Bogacz & Gurney, 2007; McClure, Berns, & Montague, 2003; Schultz, 2002; Alexander, DeLong, & Strick, 1986).

Disorders impacting DA function, like Parkinson disease (PD), disrupt the ability to learn contingencies between stimuli, actions, and outcomes (Frank, Seeberger, & O'Reilly, 2004; Cools, Barker, Sahakian, & Robbins, 2001, 2003; Swainson et al., 2000). Reinforcement learning models propose that DA depletions in PD and restorative DA medication modulate action-based approach and avoidance learning in opposite directions (Bódi et al., 2009; Moustafa, Cohen, Sherman, & Frank, 2008; Cools, 2006; Frank, 2005; Frank et al., 2004; but see Rutledge et al., 2009). DA depletion in the medication-withdrawn patient with PD is thought to prevent phasic DA bursts that facilitate action-based reward learning but does not interfere with DA pauses or dips necessary for withdrawal-based punishment avoidance learning (Frank, 2011; Frank et al., 2004). Thus, patients with PD "off" DA medication should show a learning bias favoring withdrawing to avoid punishment relative to acting to acquire reward. This pattern should reverse in patients "on" DA medication; that is, action-based reward learning should be improved at the cost of diminished punishment avoidance learning. Evidence supporting these patterns has been quite mixed. However, a prevailing limitation across studies is exclusive reliance on action-based learning paradigms, which fail to directly measure the proficiency of learning to withhold action.

Learning to withhold action to produce positive outcomes or avoid negative ones is as important for adapting in novel environments as learning to act to influence outcomes. Frontal BG circuitries are linked to inhibitory action control (Forstmann et al., 2012; Forstmann, Jahfari, et al., 2008; Forstmann, van den Wildenberg, & Ridderinkhof, 2008; Aron et al., 2007; Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004; Mink, 1996;

Vanderbilt University Medical Center
*Now at University of Louisville Medical Center, Kentucky.

Alexander et al., 1986), but the mechanisms involved in learning to inhibit action to influence outcomes have received minimal attention. DA depletion in PD impairs inhibitory action control, an effect that is modifiable by DA medication (Wylie et al., 2012; Wylie, Ridderinkhof, Bashore, & van den Wildenberg, 2010; Wylie et al., 2009a, 2009b). However, how PD and DA affect the formation of stimulus–action–outcome and stimulus–inaction–outcome associations has not been investigated directly.

A recent learning paradigm orthogonalizes action and outcome valence so that all combinations of learning to act or to withhold action to gain reward or to avoid punishment can be measured (see van Wouwe et al., 2015; Wagenbreth et al., 2015; Guitart-Masip, Chowdhury, et al., 2012; Guitart-Masip, Huys, et al., 2012). This paradigm confirms strong, natural learning biases such that action is more easily associated with reward and withholding action is more easily associated with punishment avoidance (Freeman, Alvernaz, Tonnesen, Linderman, & Aron, 2015; Wagenbreth et al., 2015; Freeman, Razhas, & Aron, 2014; Guitart-Masip, Chowdhury, et al., 2012; Guitart-Masip, Huys, et al., 2012; Everitt, Dickinson, & Robbins, 2001). More importantly, the paradigm measures learning of unnatural action–outcome associations that violate these inherent biases (i.e., action to avoid punishment or withholding action to gain reward).

In the current study of PD, we investigated the effects of DA withdrawal and facilitation on learning in the orthogonalized action–valence learning paradigm. We investigated three alternative hypotheses arising out of the current literature. First, according to reinforcement learning models, a DA-depleted state (i.e., when patients with PD perform "off" DA medications) should impair reward learning but leave intact punishment avoidance learning, a pattern that should reverse on DA medications (i.e., improved reward learning but diminished punishment avoidance learning). Second, according to inhibitory control models, a DA-depleted state in PD should selectively impair learning to inhibit action, which should be improved in the "on" DA medication state. Third, based on a hybrid of the first two models, a DA-depleted state should be particularly detrimental to learning requiring the combination of inhibiting action and gaining reward (i.e., withholding action to acquire reward), which should then improve in a DA-medicated state. We note that tangential support for this latter prediction comes from a recent study by Guitart-Masip and colleagues (2014), who showed that learning to withhold action to gain reward was selectively improved in healthy adults taking levodopa versus placebo.

Finally, we explored the role of DA on learning effects involving an extension of the third hypothesis. Learning to withhold action to gain reward and, similarly, learning to act to avoid punishment actually require the formation of less natural action–valence associations. Compared with the formation of natural action–valence associations (i.e., action to gain reward, withhold action to avoid punishment), forming unnatural action–valence associations

produces distinct cortical potentials resembling conflict or error detection signals generated from medial pFCs (MPFCs; Cavanagh, Eisenberg, Guitart-Masip, Huys, & Frank, 2013). Given that DA has also been implicated in conflict and prediction error signaling (Duthoo et al., 2013; Farooqui et al., 2011; Bonnin, Houeto, Gil, & Bouquet, 2010), we examined whether the predicted effect of DA on learning to "withhold action to gain reward" actually involves a more general principle related to the learning of unnatural action–valence associations. If true, the off-DA medication state would hamper learning of both unnatural action–valence associations, but the administration of DA medication would then be expected to remediate the learning of those conditions.

METHODS

Participants

Participants with PD ($n = 19$) were recruited from the Movement Disorders Clinic at Vanderbilt University Medical Center, and healthy controls (HCs, $n = 30$) were recruited from community advertisement or as qualifying family members of participants with PD. All participants met the following exclusion criteria: no history of (i) neurological condition (besides PD); (ii) bipolar affective disorder, schizophrenia, or other psychiatric condition known to compromise executive cognitive functions; (iii) moderate to severe depression; or (iv) medical condition known to interfere with cognition (e.g., diabetes, pulmonary disease). Patients met Brain Bank criteria for PD and were diagnosed by a movement disorder neurologist (D. O. C.), and all patients were treated currently with levodopa monotherapy ($n = 8$), DA agonist monotherapy ($n = 6$), or levodopa plus agonist dual therapy ($n = 4$). PD motor symptoms were graded using the Unified Parkinson's Disease Rating Scale motor subscore (Part III); in addition, they all received a rating of Stage III or less using the Hoehn and Yahr (1967) scale. On the basis of these criteria, each participant with PD was experiencing mild to early moderate symptoms. Dosages for the DA medications were converted to levodopa equivalent daily dose values by the method described in Weintraub et al. (2006).

All patients with PD performed at a level on the Montreal Cognitive Assessment (Nasreddine et al., 2005) that ruled out dementia but permitted very mild to minimal gross cognitive difficulties (mean = 25, $SD = 2.6$). HCs all scored greater than 27 (mean = 30, $SD = 0.8$) on the Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975). All participants reported stable mood functioning and absence of major depression and did not meet clinical criteria for MCI or dementia based on a neurological examination. The mean Center for Epidemiological Studies-Depression Scale scores (PD) or Beck Depression Inventory scores (HC) were below the standard cutoff scores of 16 and 10, respectively, suggesting the absence of depressive symptoms. One patient with PD was recruited but excluded

from the analyses because the participant did not understand the task instructions (all learning conditions < 50% accurate). Patient clinical and demographic information is presented in Table 1.

All participants had corrected-to-normal vision and provided informed consent before participating in the study in full compliance with the standards of ethical conduct in human investigation as regulated by the Vanderbilt University Institutional Review Board.

Design and Procedure

HC participants performed one session of the task. Participants with PD completed two sessions, once while taking all of their prescribed dopaminergic medications and in their optimal "on" phase of their medication cycle and a second time after a 36- to 48-hr withdrawal from their dopaminergic medication (levodopa: 36 hr, agonist: 48 hr). The order of visits was counterbalanced across participants with PD and completed at approximately the same time of day. No changes in medication dosages or addition or discontinuation of either drug for clinical purposes were made at any time during study participation. Participants were exposed to new stimuli in each

Table 1. Demographic Data for the Patients with PD and HCs

	HC	PD
Sample size (*n*)	30	18
Age (years)	62.0 (8.8)	66.3 (6.0)
Sex (M:F)*	14:16	14:4
Education (years)	16.3 (2.0)	15.9 (2.3)
MoCA	–	25.0 (2.6)
MMSE	29.6 (.8)	–
AMNART	–	115.8 (9.2)
CES-D	–	12.1 (6.6)
BDI	4.1 (4.0)	–
QUIP-ICD	–	8.4 (6.8)
QUIP-total	–	17.6 (12.5)
LEDD	–	506 (397)
Disease duration (years)	–	3.0 (2.8)
UPDRS motor	–	23.8 (11.9)

Standard deviation in parentheses. MoCA = Montreal Cognitive Assessment; AMNART = American modification of the National Adult Reading Test; CES-D = Center for Epidemiological Studies-Depression Scale; BDI = Beck Depression Inventory; QUIP = Questionnaire for Impulsive-Compulsive Disorders (ICD includes only the following behaviors: gambling, sexual behavior, buying, and eating); LEDD = Levodopa Equivalent Daily Dosage; MMSE = Mini-Mental State Examination; UPDRS = Unified Parkinson's Disease Rating Scale.
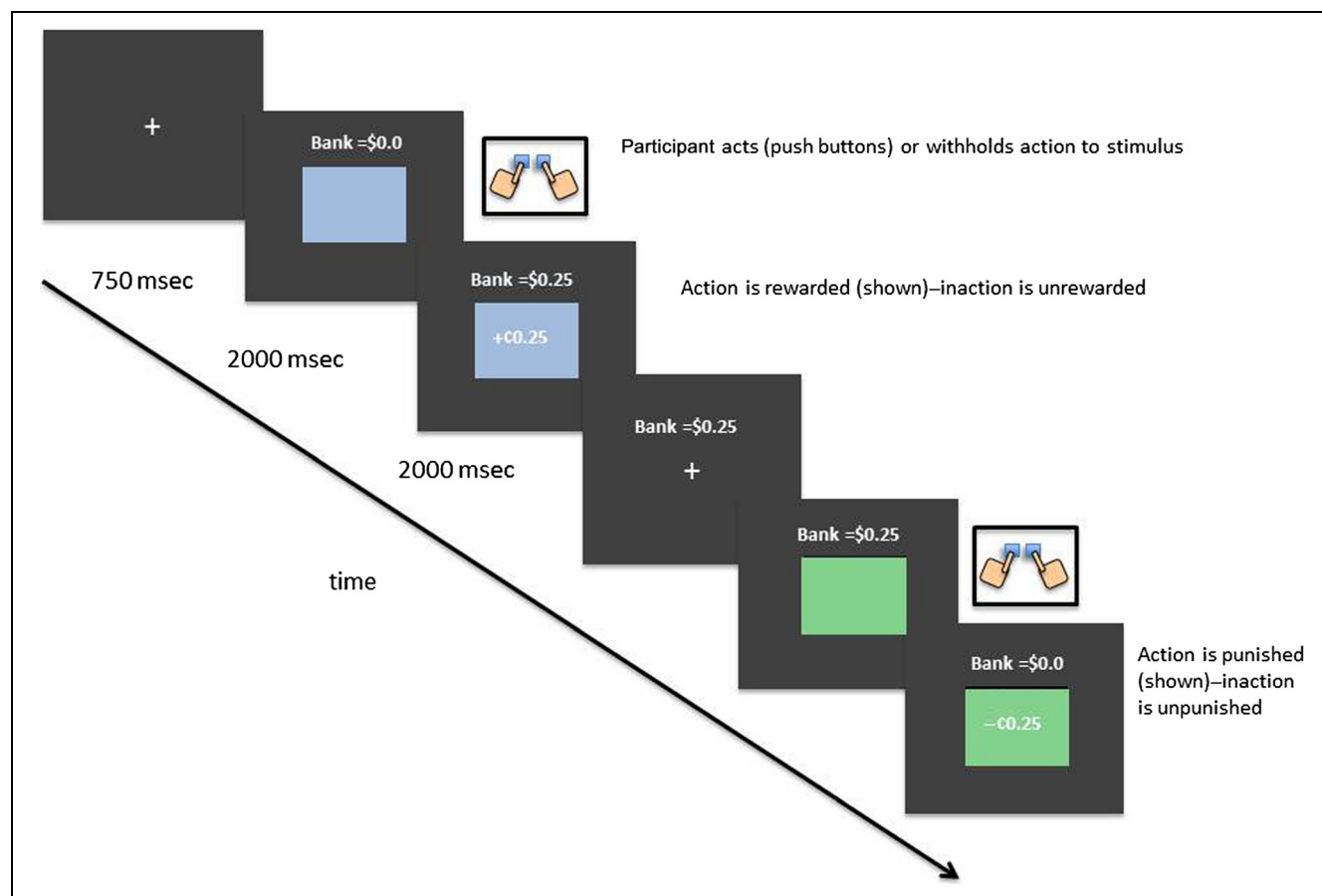
*$p < .05$, comparing HCs with all patients with PD.

visit, so both versions of the task were different and required new learning. Presentation of the stimuli within each session was randomized.

Figure 1 depicts two example trials of the action–valence learning task (a similar design was used in van Wouwe et al., 2015). Participants were instructed that the goal of this task was to learn to act or withhold action to each of four color patches to maximize monetary earnings by gaining rewards and avoiding losses (i.e., punishments). Specifically, participants viewed a series of color patches that were presented one at a time in the center of a computer screen. A trial began with the presentation of a centered fixation point for 750 msec. The fixation point was then replaced by the appearance of one of four colored patches at fixation that remained on the screen for 2000 msec (see Figure 1). Upon the presentation of a color patch, participants were instructed that they had 2 sec to either act (i.e., make a two-handed button press) or withhold action. After the 2000-msec window expired, feedback was displayed for 2000 msec in the center of the color patch indicating that the action decision led to monetary reward (+25 cents), monetary punishment (−25 cents), or no monetary outcome (0 cents). The feedback and color patch were then extinguished, and the next trial began. A running total of earnings was presented in the upper center of the screen throughout the task. The four color patches appeared in random order and with equal probability (10 times for each color) across four blocks of 40 learning trials. Thus, each color appeared for a total of 40 exposures across the four learning blocks. Each block of trials took around 3 min to complete, with a brief 1-min break between blocks.

Unbeknown to the participant, two of the color patches provided outcomes that were either rewarded or unrewarded, and the remaining two colors provided outcomes that were either punished or unpunished. Thus, the former colors were associated with reward learning, whereas the latter colors were associated with punishment avoidance learning. Also unknown to the participant, one color from each set produced the optimal outcome (i.e., either gain reward or avoid punishment) by acting, but the other color from each set produced the optimal outcome by withholding action. This design completed the 2 × 2 factor design that orthogonalized both valence and action (see Table 2).

To make the learning challenging, we designed the task so that feedback was partly probabilistic rather than fully deterministic (i.e., rewards or punishments did not occur with 100% certainty for a particular action choice). This semiprobabilistic design was applied to each color patch as outlined below. Although participants were not aware of the exact probabilities of each action–outcome association, they were instructed that each action associated with a particular stimulus would lead to a particular outcome most of the time (but not always). They also received 15 practice trials during which they experienced

Figure 1. Example trials of the action–valence learning task.

the probabilistic nature of the task. Table 2 provides a summary with the optimal response for each stimulus condition (1–4):

1. Stimulus A: Learning to act to gain reward. Selecting action to this stimulus is rewarded 80% of the time (unrewarded: 20%), but withholding action to it is unrewarded 100% of the time; only action yields reward.
2. Stimulus B: Learning to suppress action to gain reward. Suppressing (withholding) action to this stimulus is rewarded 80% of the time (unrewarded: 20%), but selecting action to it is unrewarded 100% of the time; only withholding action yields reward.
3. Stimulus C: Learning to act to avoid punishment. Selecting action to this stimulus avoids punishment 80% of time (punished: 20%), but withholding action to it is punished 100% of the time; only acting yields punishment avoidance.

4. Stimulus D: Learning to suppress action to avoid punishment. Suppressing (withholding) action to this stimulus avoids punishment 80% of time (punished: 20%), but selecting action to it is punished 100% of time; only withholding action yields punishment avoidance.

Data Analysis for Action–Valence Learning Task

Across blocks, participants completed 160 learning trials, including 40 trials for each color patch. Accuracy was defined by the percentage of trials in which the participant selected the optimal response per block.

To confirm that participants learned throughout the task, we first analyzed performance across the four learning blocks and across conditions (average accuracy per learning block), separate for PD on and off medication

Table 2. Optimal Response for Each of the Four Action–Valence Combinations (Stimuli A–D)

	Reward	*Punishment Avoidance*
Action	Act to gain reward (A)	Act to avoid punishment (C)
Inaction	Withhold to gain reward (B)	Withhold to avoid punishment (D)

and HCs. Previous studies have shown DA effects on the asymptote of learning (Guitart-Masip et al., 2014); we therefore applied our more specific analyses on the final learning block once it was confirmed that participants learned across blocks.[1]

To test the hypothesis that DA depletion in PD impairs reward learning and inhibition learning, we analyzed performance from the final block of the learning task using a mixed ANOVA to distinguish effects related to the within-participant factors of Action (action, inaction) and Valence (reward acquisition, punishment avoidance) and the between-participant factor Group (PD, HC).

Subsequently, to test the effect of DA depletion in PD on learning natural versus unnatural action–valence associations, we analyzed performance from the final block of the learning task using a mixed ANOVA to distinguish effects related to the within-participant factors of Condition (inaction–reward, inaction–punishment avoidance, action–reward, action–punishment avoidance) and the between-participant factor Group (PD, HC).

Similar analyses (within-participant ANOVA) were applied to test the medication effect for each of the three hypotheses; however, instead of a between-participant Group factor, there was an additional within-participant Medication (off, on) factor. In addition, to exclude that changes in the unnatural conditions were due to stronger learning on the natural action–valence associations, we calculated a Pavlovian bias (Cavanagh et al., 2013). This measure averages the bias to "act" when there is a reward at stake (reward invigoration, i.e., the number of action responses on the conditions "act to gain reward" plus action responses on "withhold to gain reward" divided by total action responses) with the bias to refrain from action when it concerns avoiding punishment (punishment-based suppression; number of withhold responses on conditions "act to avoid punishment" plus withhold responses on "withhold to avoid punishment" divided by total withhold responses).

RESULTS

Overall Learning Performance and Decision Time

Figure 2 displays the accuracy for each condition across learning blocks, separate for HC and PD off and on medications. All groups showed an increase in learning across blocks and across action–valence conditions, (accuracy: $F(3, 51)_{\text{OFFmeds}} = 4.86$, $p < .05$; $F(3, 51)_{\text{ONmeds}} = 9.41$, $p < .01$; $F(3, 87)_{\text{HC}} = 22.51$, $p < .01$; Figure 2). Performance in the final block of trials provides a direct measure of how well participants learned each condition; thus, subsequent analyses will focus on performance from the final block.

Speed of responding was not emphasized during task instructions, and speed could only be measured on trials where participants chose to act. However, to rule out any group effects on decision speed or possible decision speed–accuracy trade-off effects, we analyzed the RTs across action conditions by group and by medication state. Across action conditions, there was no difference between HCs and patients with PD in their response speed (group: $F(1, 46) = 0.76, p = .39$). In addition, there was no effect of medication on the RTs (medication: $F(1, 17) = 0.68$, $p = .42$). All participants showed faster RTs for learning to act to gain reward (HC = 822 msec, PD "off" = 807 msec, PD "on" = 809 msec) than for learning to act to
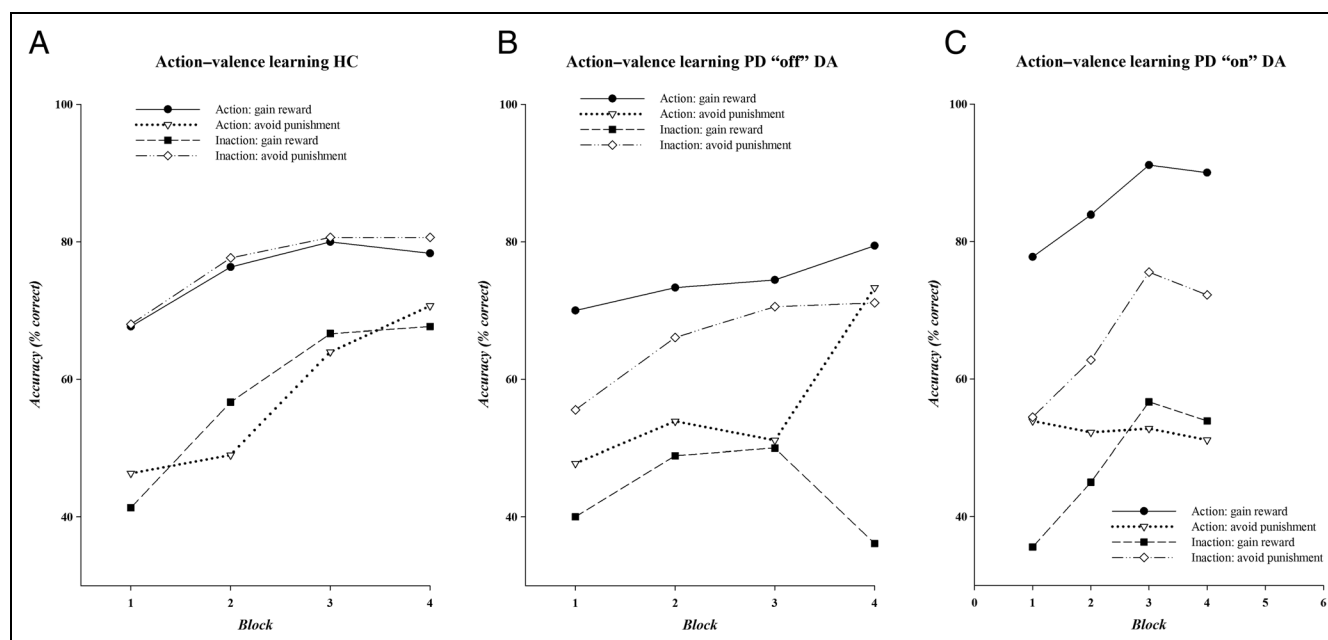

Figure 2. Accuracy rates on each of the action–valence combinations across four learning blocks, separate for (A) HCs, (B) patients with PD in their "off" medication state, and (C) patients with PD in their "on" DA medication state.

avoid punishment (HC = 1047 msec, PD "off" = 971 msec, PD "on" = 823 msec; valence$_{HC-PD}$: $F(1, 46) = 36.54, p < .001$; valence$_{PD\_Off-PD\_On}$: $F(1, 17) = 23.26, p < .001$).

PD "Off" Dopaminergic Medication versus HC

Hypotheses 1 and 2: PD Effects on Reward Learning and Inhibition Learning

Figure 3 displays the accuracy on each of the action–valence associations for the final learning block, separately for HCs and patients with PD in each medication state.

Overall, all participants showed higher learning for action (75% accuracy) than for withholding action (64% accuracy; action: $F(1, 46) = 6.13, p < .05$) and for learning to avoid punishment (74% accuracy) than for learning to gain reward (65% accuracy; valence: $F(1, 46) = 6.86, p < .05$). Action and Valence factors produced the expected
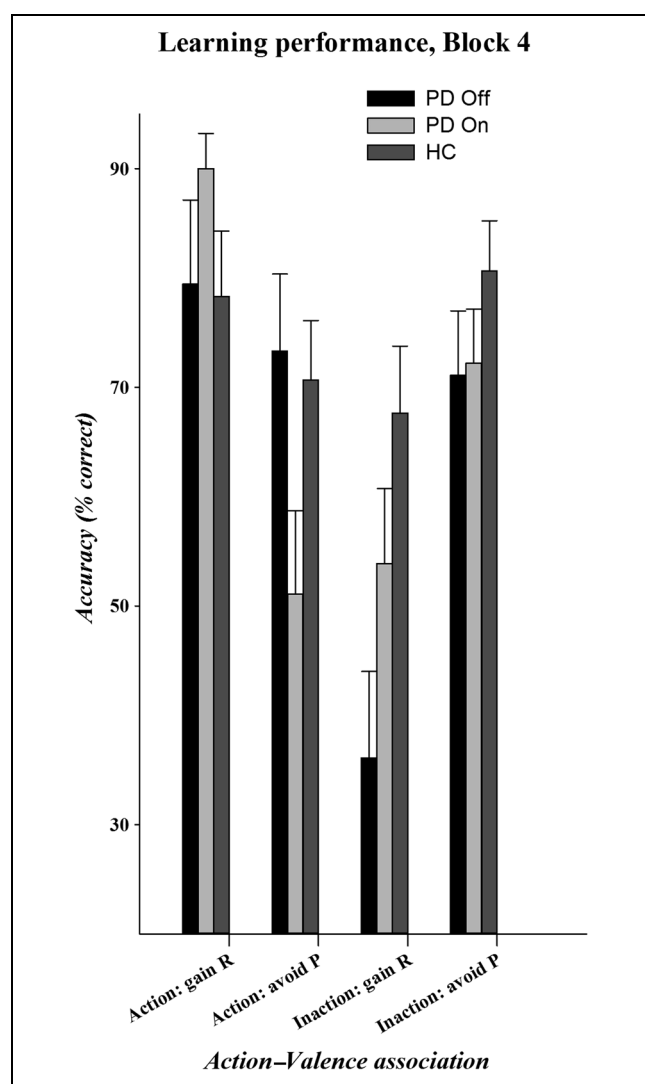

Figure 3. Accuracy rates on each of the action–valence combinations in the final learning block for HCs, patients with PD in their "off" medication state, and patients with PD in their "on" DA medication state.

interactive effect on learning performance (Action × Valence: $F(1, 46) = 10.17, p < .01$). Specifically, learning to act to obtain reward (79% accuracy) or avoid punishment (72% accuracy) was associated with similar high learning performance ($t(47) = 1.06, p = .3$). In contrast, participants learned to withhold action to avoid punishment at a similar high learning performance (76% accuracy), but learning to withhold action to obtain reward was associated with a significant reduction in accuracy (51% accuracy; $t(47) = 4.7, p < .001$). Notably, these general performance patterns were similar across groups (Action × Valence × Group: $F(1, 46) = 1.11, p = .30$).

There were important differences between the group with PD and the HC group in their patterns of performance by the end of learning. Overall, HCs (74% accuracy) tended to show higher overall performance than patients with PD (65% accuracy; group: $F(1, 46) = 3.31, p = .08$). However, group differences emerged as a function of action learning (Action × Group: $F(1, 46) = 5.78, p < .05$). Whereas the HC group and the group with PD performed equally well at learning to act to influence outcomes (PD = 76%, HC = 75% accuracy; $F(1, 46) = 0.09, p = .80$), patients with PD showed a drastic reduction in performance on withholding action (54% accuracy) compared with the HC group (74% accuracy; $F(1, 47) = 7.25, p < .05$). Performance on learning to obtain reward versus avoid punishment, irrespective of the action required, also tended to differ across groups (Valence × Group: $F(1, 46) = 3.25, p = .08$). Patients with PD showed reduced learning to obtain reward compared with HCs (PD = 58%, HC = 73% accuracy; $F(1, 46) = 5.75, p < .05$), but the groups showed equivalent learning to avoid punishment (PD = 72%, HC = 76% accuracy; $F(1, 46) = 0.55, p = .6$).

Hypotheses 3: PD Effects on Learning Unnatural Action–Valence Associations

HCs tended to show higher overall learning proficiency compared with patients with PD (74% vs. 65%; group: $F(1, 46) = 3.31, p = .08$). Performance varied across action–valence conditions (condition: $F(1, 46) = 7.96, p < .001$), and this effect varied by group (Condition × Group: $F(1, 46) = 3.34, p < .05$). Inspection of Figure 3 shows the source of this interaction. Specifically, natural associations are learned more effectively than the unnatural associations, but patients with PD have a selective difficulty learning to "withhold action to obtain reward." To further examine these patterns, we deconstructed learning for natural and unnatural conditions. Across natural conditions, patients with PD and HCs showed similar learning proficiency (Group × Condition: $F(1, 46) = 0.78, p = .38$; group: $F(1, 46) = 0.46, p = .5$). In contrast, learning across the unnatural conditions showed a significant difference between groups (Group × Condition: $F(1, 46) = 10.23, p < .005$; group: $F(1, 46) = 3.43, p = .07$). Specifically, patients with PD were significantly

poorer at learning to "withhold action to obtain reward" (36.1%; note that patients with PD never exceed chance-based performance (50%) on this condition across blocks (see Figure 2B) compared with HCs (67.7%), $F(1, 46) = 9.99, p < .005$, whereas the groups showed similar learning to "act to avoid punishment," $F(1, 46) = 0.09, p = .77$.

Medication Effect: PD "Off" versus "On" Dopaminergic Medication

Hypotheses 1 and 2: DA Effects on Reward Learning and Inhibition

Overall, patients with PD demonstrated better performance when learning to act (74% accuracy) versus to withhold (58% accuracy; action: $F(1, 17) = 8.23, p < .05$) but no overall performance differences between reward acquisition and punishment avoidance learning (valence: $F(1, 17) = 0.26, p = .62$). Again, Action and Valence factors produced the expected interactive effect on learning performance (Action × Valence: $F(1, 17) = 33.0, p < .001$). Withholding action was more easily associated with avoiding punishment (72% accuracy) than with obtaining reward (45% accuracy; $t(17) = 4.72, p < .01$), whereas learning to act was more easily associated with reward acquisition (85% accuracy) than with punishment avoidance (62% accuracy; $t(17) = 3.63, p < .01$). Notably, these general learning patterns were not altered by medication state (Action × Valence × Medication: $F(1, 17) = 1.04, p = .30$).

Although DA medication state did not influence overall learning performance (PD "off": 65% accuracy, PD "on": 67% accuracy; medication, $F(1, 17) = 0.32, p = .58$), DA exerted specific effects on performance patterns. Learning to act or to withhold action varied by DA medication state (Action × Medication: $F(1, 17) = 5.2, p < .05$). As reported above, patients "off" DA medication showed poorer learning to withhold action (54% accuracy) compared with learning to act (76% accuracy; $F(1, 17) = 10.79, p < .001$), but on DA medication, learning to act and to withhold action were performed with similar accuracy (learning to withhold = 63% accuracy, learning to act = 71% accuracy; $F(1, 17) = 1.89, p = .19$).

DA medication state also modulated performance on learning to gain reward versus avoid punishment (Valence × Medication: $F(1, 17) = 12.29, p < .01$). Reduced reward learning in the "off" DA state (58% accuracy) was improved significantly in the "on" DA state (72% accuracy; $F(1, 17) = 8.29, p < .05$). In contrast, higher punishment avoidance learning in the "off" DA state (72% accuracy) was reduced in the "on" DA state (62% accuracy; $F(1, 17) = 5.29, p < .05$).

Hypotheses 3: DA Medication Effects on Learning Unnatural Action–Valence Associations

Overall, there was no main effect of medication state on learning proficiency (Medication: $F(1, 17) = 0.32, p =$

.58). However, learning proficiency varied across the four conditions (Condition: $F(3, 51) = 13.32, p < .001$), and this effect further varied across medication states (Condition × Medication: $F(3, 51) = 5.80, p < .01$). Inspection of Figure 3 again points to PD and DA effects on the unnatural compared with natural action–valence conditions.

To investigate these patterns more directly, we analyzed performance across the groups separately for natural and unnatural conditions. Medication state had no overall or selective effect on learning across the natural conditions (Medication: $F(1, 17) = 1.43, p = .25$; Medication × Condition: $F(1, 17) = 1.09, p = .31$). In contrast, medication state had a direct and opposite influence on learning proficiency across the unnatural action–valence conditions (Medication × Condition: $F(1, 17) = 14.93, p < .01$). Specifically, DA medication improved learning proficiency in the "withhold action to gain reward" condition by 18% but produced a 22% reduction in learning proficiency in the "act to avoid punishment" condition ($t(17) = 3.86, p < .01$).

To rule out that the DA medication effects on learning of the unnatural conditions were not due to changes in bias toward the more natural associations, we computed a bias measure (i.e., a Pavlovian bias; Cavanagh et al., 2013). The Pavlovian bias was similar for PD "off" (0.65) versus "on" medication (0.62; $t(17) = 0.68, p = .51$), which indicates that medication did not increase a learning bias for the natural action–valence associations. This suggests that DA state directly modulated the formation of unnatural action–valence associations.

DISCUSSION

The current study investigated the influences of DA loss in PD and DA restorative medication on learning to act and learning to withhold (inhibit) action to gain rewarding outcomes and avoid punishing outcomes. The action–valence learning paradigm produced the expected outcome patterns. At the end of the learning task, the HC group and the group with PD ("off" and "on" DA) were better at acting to gain reward than to withhold action to gain reward. In addition, both groups learned to withhold action to avoid punishment more effectively than to withhold action to gain reward. These patterns replicate previous studies using a similar paradigm in healthy adults (van Wouwe et al., 2015; Wagenbreth et al., 2015; Chowdhury, Guitart-Masip, Lambert, Dayan, et al., 2013; Chowdhury, Guitart-Masip, Lambert, Dolan, et al., 2013; Guitart-Masip, Chowdhury, et al., 2012; Guitart-Masip, Huys, et al., 2012). Our findings also replicate improved outcomes on learning to withhold to gain reward with DA as found in a levodopa study in HCs by Guitart-Masip et al. (2014).

Our findings expand this work by showing that DA depletion and pharmacological facilitation of DA in PD had no impact on learning of natural action–valence associations but instead produced a complex modulation of

learning involving the formation of unnatural action and outcome valence associations. Compared with HCs, patients with PD "off" DA showed deficient reward learning and learning to withhold actions, supporting predictions from reinforcement learning and inhibitory control theories of DA. However, the orthogonalized learning paradigm disclosed a novel source of these effects related to deficient learning of one of the unnatural action–valence associations. Specifically, in the DA-withdrawn state, patients with PD were deficient at learning to inhibit action to produce rewarding outcomes. This suggests that patients with PD had more difficulty overcoming the more natural bias linking inaction to punishment avoidance compared with HCs.

DA medication shifted these learning patterns and biases. DA improved both inaction learning and reward outcome learning compared with the "off" DA state. These patterns also conform to predictions about DA effects on inhibitory control and reward learning. However, the underlying source of these effects emerged again from an interaction between action and valence involving unnatural associations. The deficit learning to associate inaction with reward found in the "off" DA state was improved on DA. Instead, patients on DA experienced difficulty learning to act to avoid punishment.

Although the direction of the DA effect was not similar across the two unnatural conditions, these patterns suggest that DA might play a specific role in modulating less natural associations between action and outcomes. DA may be critical to overriding strong, natural learning biases so that less natural links between action and outcomes can be established.

Putative Neural Mechanisms

The opponent system model of reinforcement learning (Frank, Samanta, Moustafa, & Sherman, 2007; O'Reilly & Frank, 2006; Frank, 2005) proposes that DA modulates reward and punishment learning via two separate BG pathways embedded in the cortico-striato-thalamocortical circuitry, respectively, the direct or go pathway, which is facilitated by DA acting on D1 receptors, and the indirect or no-go pathway, which is suppressed by DA acting on D2 receptors. Reward-based action learning is promoted by DA bursts that enhance D1 facilitation of the go pathway and D2 suppression of the no-go pathway, whereas punishment-based avoidance learning is promoted by pauses in DA activity that attenuate D1 facilitation of the go pathway as well as D2 suppression of the no-go pathway, leading to inaction. According to this model, a decrease in DA (patients with PD "off" DA medication) should bias toward punishment avoidance learning and reduce reward action learning. Conversely, DA medication should produce the opposite bias toward enhanced reward action learning with a concomitant reduction in punishment avoidance learning. Contrary to these predictions, learning to associate action to reward and inaction to

punishment avoidance was unaffected by PD in "off" or "on" DA states. Rather, the shifts in learning performance were mostly driven by dopaminergic effects on the two less natural learning conditions, learning to act to avoid punishment and learning to inhibit action to gain reward (although the direction of the effect of DA was not similar across the two unnatural conditions). This invites an alternative framework for understanding DA effects on action–valence learning.

A recent study by Cavanagh and colleagues (2013) demonstrated unique neural signatures in event-related brain potentials when processing the unnatural action–valence conditions. During learning, the presentation of unnatural associations produced enhanced midline prefrontal signals commonly linked to MPFC activity and signaling related to conflict detection and mobilization of control after detected conflict (Cavanagh, Zambrano-Vazquez, & Allen, 2012; Cohen, Elger, & Fell, 2009; Kerns, 2006; Yeung, Botvinick, & Cohen, 2004). Recent computational models based on single-cell recordings in animals posit that MPFC detects positive (unexpected reward or omission of punishment) or negative (unexpected punishment or omission of reward) reward prediction errors (RPEs; Silvetti, Alexander, Verguts, & Brown, 2014), which then projects to the DA system to adjust behavior; that is, the prediction error drives DA activity to train and adapt learning signals. According to this view, DA activity would be most critical when prediction errors were highest, that is, when learning unnatural associations. DA depletions would interfere with the generation of positive RPE (i.e., phasic DA bursts that signal unexpected reward; Cohen & Frank, 2009; Shen, Flajolet, Greengard, & Surmeier, 2008) that would be most critical for associating inaction with reward. Conversely, increasing DA with medications would restore positive RPE signaling but interfere with negative RPE signaling (i.e., preventing phasic DA dips or pauses that signal unexpected punishment; Cohen & Frank, 2009; Shen et al., 2008) most critical for linking action with punishment avoidance.

Future imaging and computational modeling studies with patients with PD on the current paradigm could study whether the loss of DA selectively impacts the processing of prediction errors critical for overriding inaction–punishment avoidance biases to associate withholding of action to produce rewarding outcomes. Adding DA medication should interfere with the processing of conflict or prediction errors critical for overriding the more natural action–reward learning biases.

Our findings suggest that DA is most critical when learning requires overriding natural action–valence tendencies in favor of unnatural associations between action and valence. This role of DA fits squarely with ideas linking DA to action–outcome prediction errors while preserving ideas related to shifts in biases between reward and punishment avoidance learning. How could this be instantiated in the framework of BG direct and indirect pathways? One potential mechanism was postulated

recently by Kravitz and Kreitzer (2012). They proposed that the integration of action and valence factors occurs through the separate processes of long-term potentiation (LTP) and long-term depression (LTD) operating on direct and indirect BG pathways. Specifically, like the opponent system reinforcement learning model, DA bursts facilitate direct pathway LTP associating action to reward, whereas DA dips induce indirect pathway LTP promoting inaction to avoid punishment. The novel addition to the proposed model was the inclusion of LTD processes to account for unnatural action–valence associations. Specifically, LTD of the direct pathway could serve as a mechanism to associate reward with inaction, whereas LTD of the indirect pathway could link action with punishment avoidance. Although speculative, this model promotes consideration of the orthogonalized learning conditions represented in the current paradigm. Notably, animal models of PD have emphasized how PD changes LTD and LTP (Dupuis et al., 2013) and that this modulates learning performance (Beeler et al., 2012). Human behavioral and imaging studies supporting these ideas await future investigation.

Limitations and Potential Applications

The current study reveals that DA states have complex, value-dependent influences on the interaction between instrumental and Pavlovian learning processes. Beyond the analysis of behavioral effects, greater precision in disclosing the dissociable effect of DA on the unnatural conditions, DA's relation to prediction errors during learning, and the dynamics of the learning process could be achieved by applying computational models (e.g., Guitart-Masip, Huys et al., 2012; Rutledge et al., 2009).

One limitation is that PD certainly involves brain changes beyond DA, including reported alterations in other neuromodulators (e.g., serotonin), which could affect learning processes measured in the current study (Guitart-Masip et al., 2014; Geurts, Huys, den Ouden, & Cools, 2013). Moreover, we did not quantify DA system or structural changes in frontostriatal brain areas, which have also been associated with reductions in stimulus–action–outcome learning in PD (O'Callaghan et al., 2013). These factors will be important issues to control and study in future investigations. Another limitation is that we treated dopaminergic medications uniformly, although different classes of medications (e.g., agonists vs. levodopa) have partly dissociable effects on D1 and D2 receptor activity expressed along the direct and indirect pathways. Future work aimed at comparing D2 versus D1 DA effects pharmacologically and with PET on the current experimental task will be insightful. An additional limitation is that we did not match sex across groups (men were overrepresented in the PD group). On the basis of the absence of any sex effects in the HC sex ($F(1, 28) = 1.8$, $p = .19$), we do not expect that sex plays a major contribution in explaining our findings,

but a possible dissociable medication effect across sex should be taken into account in future studies.

A range of neurologic (e.g., Huntington's disease) and neuropsychiatric (e.g., obsessive–compulsive disorder, Tourette's syndrome) disorders alter frontal BG circuitry, impacting action learning and inhibition control mechanisms (Holl, Wilkinson, Tabrizi, Painold, & Jahanshahi, 2013; Gillan et al., 2011; Worbe et al., 2011). The current learning paradigm could offer more novel insights into the effects of these conditions on action–outcome learning processes. The action–valence paradigm also offers a novel tool for investigating the effects of pharmacologic and surgical interventions (e.g., deep brain stimulation) aimed at modulating BG activity selectively. We recently demonstrated that implicit processing of already formed natural and unnatural action–valence associations, even when irrelevant to a current task, produces dissociable effects on cognitive control mechanisms (van Wouwe et al., 2015). How established action–valence learning and biases interact with cognitive control mechanisms that are also linked to frontal BG circuits could provide new insight into a range of behavioral deficits involving poor control over strong urges and impulses.

Acknowledgments

Note

1. To validate our findings across task performance, we included learning block (1 and 4) as an additional factor in all analyses that were applied to test our three hypotheses proposed in the Introduction (between and within groups). The details of these analyses and results are described in a supplementary section and can be found on www.researchgate.net/profile/Nelleke_Van_Wouwe/contributions and are similar to our findings with the final learning block. Importantly, they replicate PD effects on learning unnatural action–valence associations across learning blocks, Group × Condition × Block, $F(1, 46) = 4.92$, $p < .05$, and DA medication effects on learning unnatural action–valence associations across learning blocks, Condition × Medication × Block, $F(3, 51) = 4.85$, $p < .01$.

REFERENCES

Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience, 9,* 357–381.

Aron, A. R., Durston, S., Eagle, D. M., Logan, G. D., Stinear, C. M., & Stuphorn, V. (2007). Converging evidence for a fronto-basal-ganglia network for inhibitory control of action and cognition. *Journal of Neuroscience, 27,* 11860–11864.

Beeler, J. A., Frank, M. J., McDaid, J., Alexander, E., Turkson, S., Bernardez Sarria, M. S., et al. (2012). A role for dopamine-mediated learning in the pathophysiology and treatment of Parkinson's disease. *Cell Reports, 2,* 1747–1761.

Bódi, N., Kéri, S., Nagy, H., Moustafa, A., Myers, C. E., Daw, N., et al. (2009). Reward-learning and the novelty-seeking personality: A between- and within-subjects study of the effects of dopamine agonists on young Parkinson's patients. *Brain, 132,* 2385–2395.

Bogacz, R., & Gurney, K. (2007). The basal ganglia and cortex implement optimal decision making between alternative actions. *Neural Computation, 19,* 442–477.

Bonnin, C. A., Houeto, J. L., Gil, R., & Bouquet, C. A. (2010). Adjustments of conflict monitoring in Parkinson's disease. *Neuropsychology, 24,* 542–546.

Cavanagh, J. F., Eisenberg, I., Guitart-Masip, M., Huys, Q., & Frank, M. J. (2013). Frontal theta overrides Pavlovian learning biases. *Journal of Neuroscience, 33,* 8541–8548.

Cavanagh, J. F., Zambrano-Vazquez, L., & Allen, J. J. (2012). Theta lingua franca: A common mid-frontal substrate for action monitoring processes. *Psychophysiology, 49,* 220–238.

Chowdhury, R., Guitart-Masip, M., Lambert, C., Dayan, P., Huys, Q., Düzel, E., et al. (2013). Dopamine restores reward prediction errors in old age. *Nature Neuroscience, 16,* 648–653.

Chowdhury, R., Guitart-Masip, M., Lambert, C., Dolan, R. J., & Düzel, E. (2013). Structural integrity of the substantia nigra and subthalamic nucleus predicts flexibility of instrumental learning in older-age individuals. *Neurobiology of Aging, 34,* 2261–2270.

Cohen, M. X., Elger, C. E., & Fell, J. (2009). Oscillatory activity and phase-amplitude coupling in the human medial frontal cortex during decision making. *Journal of Cognitive Neuroscience, 21,* 390–402.

Cohen, M. X., & Frank, M. J. (2009). Neurocomputational models of basal ganglia function in learning, memory and choice. *Behavioural Brain Research, 199,* 141–156.

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

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

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

Dupuis, J. P., Feyder, M., Miguelez, C., Garcia, L., Morin, S., Choquet, D., et al. (2013). Dopamine-dependent long-term depression at subthalamo-nigral synapses is lost in experimental Parkinsonism. *Journal of Neuroscience, 33,* 14331–14341.

Duthoo, W., Braem, S., Houtman, F., Schouppe, N., Santens, P., & Notebaert, W. (2013). Dopaminergic medication counteracts conflict adaptation in patients with Parkinson's disease. *Neuropsychology, 27,* 556–561.

Everitt, B. J., Dickinson, A., & Robbins, T. W. (2001). The neuropsychological basis of addictive behaviour. *Brain Research. Brain Research Reviews, 36,* 129–138.

Farooqui, A. A., Bhutani, N., Kulashekhar, S., Behari, M., Goel, V., & Murthy, A. (2011). Impaired conflict monitoring in Parkinson's disease patients during an oculomotor redirect task. *Experimental Brain Research, 208,* 1–10.

Folstein, M., Folstein, S., & McHugh, P. (1975). "Mini-mental state." A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189–198.

Forstmann, B. U., Jahfari, S., Scholte, H. S., Wolfensteller, U., van den Wildenberg, W. P. M., & Ridderinkhof, K. R. (2008). Function and structure of the right inferior frontal cortex predict individual differences in response inhibition: A model-based approach. *Journal of Neuroscience, 28,* 9790–9796.

Forstmann, B. U., Keuken, M. C., Jahfari, S., Bazin, P. L., Neumann, J., Schafer, A., et al. (2012). Cortico-subthalamic white matter tract strength predicts interindividual efficacy in stopping a motor response. *Neuroimage, 60,* 370–375.

Forstmann, B. U., van den Wildenberg, W. P. M., & Ridderinkhof, K. R. (2008). Neural mechanisms, temporal dynamics, and individual differences in interference control. *Journal of Cognitive Neuroscience, 20,* 1854–1865.

Frank, M., Samanta, J., Moustafa, A., & Sherman, S. (2007). Hold your horses: Impulsivity, deep brain stimulation, and medication in Parkinsonism. *Science, 318,* 1309–1312.

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

Frank, M. J. (2011). Computational models of motivated action selection in corticostriatal circuits. *Current Opinion in Neurobiology, 21,* 381–386.

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

Freeman, S. M., Alvernaz, D., Tonnesen, A., Linderman, D., & Aron, A. R. (2015). Suppressing a motivationally-triggered action tendency engages a response control mechanism that prevents future provocation. *Neuropsychologia, 68,* 218–231.

Freeman, S. M., Razhas, I., & Aron, A. R. (2014). Top–down response suppression mitigates action tendencies triggered by a motivating stimulus. *Current Biology, 24,* 212–216.

Geurts, D. E., Huys, Q. J., den Ouden, H. E., & Cools, R. (2013). Serotonin and aversive Pavlovian control of instrumental behavior in humans. *Journal of Neuroscience, 33,* 18932–18939.

Gillan, C. M., Papmeyer, M., Morein-Zamir, S., Sahakian, B. J., Fineberg, N. A., Robbins, T. W., et al. (2011). Disruption in the balance between goal-directed behavior and habit learning in obsessive–compulsive disorder. *American Journal of Psychiatry, 168,* 718–726.

Guitart-Masip, M., Chowdhury, R., Sharot, T., Dayan, P., Duzel, E., & Dolan, R. J. (2012). Action controls dopaminergic enhancement of reward representations. *Proceedings of the National Academy of Sciences, U.S.A., 109,* 7511–7516.

Guitart-Masip, M., Economides, M., Huys, Q. J., Frank, M. J., Chowdhury, R., Duzel, E., et al. (2014). Differential, but not opponent, effects of L-DOPA and citalopram on action learning with reward and punishment. *Psychopharmacology, 231,* 955–966.

Guitart-Masip, M., Huys, Q. J., Fuentemilla, L., Dayan, P., Duzel, E., & Dolan, R. J. (2012). Go and no-go learning in reward and punishment: Interactions between affect and effect. *Neuroimage, 62,* 154–166.

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

Holl, A. K., Wilkinson, L., Tabrizi, S. J., Painold, A., & Jahanshahi, M. (2013). Selective executive dysfunction but intact risky decision-making in early Huntington's disease. *Movement Disorders, 28,* 1104–1109.

Kerns, J. G. (2006). Anterior cingulate and prefrontal cortex activity in an fMRI study of trial-to-trial adjustments on the Simon task. *Neuroimage, 33,* 399–405.

Kravitz, A. V., & Kreitzer, A. C. (2012). Striatal mechanisms underlying movement, reinforcement, and punishment. *Physiology (Bethesda), 27,* 167–177.

McClure, S. M., Berns, G. S., & Montague, P. R. (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron, 38,* 339–346.

Mink, J. W. (1996). The basal ganglia: Focused selection and inhibition of competing motor programs. *Progress in Neurobiology, 50,* 381–425.

Moustafa, A. A., Cohen, M. X., Sherman, S. J., & Frank, M. J. (2008). A role for dopamine in temporal decision making and reward maximization in Parkinsonism. *Journal of Neuroscience, 28,* 12294–12304.

Nasreddine, Z. S., Phillips, N. A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., et al. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society, 53,* 695–699.

O'Callaghan, C., Moustafa, A. A., de Wit, S., Shine, J. M., Robbins, T. W., Lewis, S. J. G., et al. (2013). Fronto-striatal gray matter contributions to discrimination learning in Parkinson's disease. *Frontiers in Computational Neuroscience, 7,* 180.

O'Reilly, R., & Frank, M. (2006). Making working memory work: A computational model of learning in the prefrontal cortex and basal ganglia. *Neural Computation, 18,* 283–328.

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science, 306,* 443–447.

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

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron, 36,* 241–263.

Shen, W., Flajolet, M., Greengard, P., & Surmeier, D. J. (2008). Dichotomous dopaminergic control of striatal synaptic plasticity. *Science, 321,* 848–851.

Silvetti, M., Alexander, W., Verguts, T., & Brown, J. W. (2014). From conflict management to reward-based decision making: Actors and critics in primate medial frontal cortex. *Neuroscience & Biobehavioral Reviews, 46,* 44–57.

Swainson, R., Rogers, R. D., Sahakian, B. J., Summers, B. A., Polkey, C. E., & 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.

van Wouwe, N. C., van den Wildenberg, W. P., Ridderinkhof, K. R., Claassen, D. O., Neimat, J. S., & Wylie, S. A. (2015). Easy to learn, hard to suppress: The impact of learned stimulus-outcome associations on subsequent action control. *Brain and Cognition, 101,* 17–34.

Wagenbreth, C., Zaehle, T., Galazky, I., Voges, J., Guitart-Masip, M., Heinze, H. J., et al. (2015). Deep brain stimulation of the subthalamic nucleus modulates reward processing and action selection in Parkinson patients. *Journal of Neurology, 262,* 1541–1547.

Weintraub, D., Siderowf, A. D., Potenza, M. N., Goveas, J., Morales, K. H., Duda, J. E., et al. (2006). Association of dopamine agonist use with impulse control disorders in Parkinson disease. *Archives of Neurology, 63,* 969–973.

Worbe, Y., Palminteri, S., Hartmann, A., Vidailhet, M., Lehericy, S., & Pessiglione, M. (2011). Reinforcement learning and Gilles de la Tourette syndrome dissociation of clinical phenotypes and pharmacological treatments. *Archives of General Psychiatry, 68,* 1257–1266.

Wylie, S. A., Claassen, D. O., Huizenga, H. M., Schewel, K. D., Ridderinkhof, K. R., Bashore, T. R., et al. (2012). Dopamine agonists and the suppression of impulsive motor actions in Parkinson disease. *Journal of Cognitive Neuroscience, 24,* 1709–1724.

Wylie, S. A., Ridderinkhof, K. R., Bashore, T. R., & van den Wildenberg, W. P. (2010). The effect of Parkinson's disease on the dynamics of on-line and proactive cognitive control during action selection. *Journal of Cognitive Neuroscience, 22,* 2058–2073.

Wylie, S. A., van den Wildenberg, W. P., Ridderinkhof, K. R., Bashore, T. R., Powell, V. D., Manning, C. A., et al. (2009a). The effect of Parkinson's disease on interference control during action selection. *Neuropsychologia, 47,* 145–157.

Wylie, S. A., van den Wildenberg, W. P., Ridderinkhof, K. R., Bashore, T. R., Powell, V. D., Manning, C. A., et al. (2009b). The effect of speed-accuracy strategy on response interference control in Parkinson's disease. *Neuropsychologia, 47,* 1844–1853.

Yeung, N., Botvinick, M., & Cohen, J. (2004). The neural basis of error detection: Conflict monitoring and the error-related negativity. *Psychological Review, 111,* 931–959.