Deep brain stimulation of the subthalamic nucleus improves reward-based decision-learning in Parkinson's disease.

Van Wouwe, N.C.; Ridderinkhof, K.R.; Van den Wildenberg, W.P.M.; Band, G.P.H.; Abisogun, A.; Elias, W.J.; ... ; Wylie, S.A.

Citation

Deep brain stimulation of the subthalamic nucleus improves reward-based decision-learning in Parkinson's disease

Nelleke C. van Wouwe[1]*, K. R. Ridderinkhof[2,3], W. P. M. van den Wildenberg[2], G. P. H. Band[1], A. Abisogun[4], W. J. Elias[5], R. Frysinger[5] and S. A. Wylie[4]

[1] Leiden University Institute of Psychology, Leiden, Netherlands
[2] Amsterdam Center for the Study of Adaptive Control in Brain and Behaviour (Acacia), Department of Psychology, University of Amsterdam, Amsterdam, Netherlands
[3] Cognitive Science Center Amsterdam, University of Amsterdam, Amsterdam, Netherlands
[4] Neurology Department, University of Virginia Health Systems, Charlottesville, VA, USA
[5] Neurosurgery Department, University of Virginia Health Systems, Charlottesville, VA, USA

Recently, the subthalamic nucleus (STN) has been shown to be critically involved in decision-making, action selection, and motor control. Here we investigate the effect of deep brain stimulation (DBS) of the STN on reward-based decision-learning in patients diagnosed with Parkinson's disease (PD). We determined computational measures of outcome evaluation and reward prediction from PD patients who performed a probabilistic reward-based decision-learning task. In previous work, these measures covaried with activation in the nucleus caudatus (outcome evaluation during the early phases of learning) and the putamen (reward prediction during later phases of learning). We observed that stimulation of the STN motor regions in PD patients served to improve reward-based decision-learning, probably through its effect on activity in frontostriatal motor loops (prominently involving the putamen and, hence, reward prediction). In a subset of relatively younger patients with relatively shorter disease duration, the effects of DBS appeared to spread to more cognitive regions of the STN, benefiting loops that connect the caudate to various prefrontal areas important for outcome evaluation. These results highlight positive effects of STN stimulation on cognitive functions that may benefit PD patients in daily-life association-learning situations.

Keywords: Parkinson's disease, deep brain stimulation, subthalamic nucleus, probabilistic learning

INTRODUCTION

Making appropriate choices between distinct options in daily life (for example friend or foe, food, or non-food) is vital for optimal behavior and requires learning the causal relation between events, actions and their outcomes. Decisions about how best to respond in a situation are often guided by past learning, particularly when expectations about the outcomes of those decisions are well formed. In novel situations, expectations about the favorability of a decision's outcome (i.e., leads to reward versus leads to punishment) are uncertain, and the associations between a situation, a response to it, and the outcome of that decision must be learned on the basis of trial and error.

Reward-based decision-learning paradigms enable us to measure the process of learning associations between stimuli, actions, and their related rewards. Several brain areas have been linked to key aspects of reward-based decision-learning, including prefrontal regions (e.g., the dorsolateral and orbito-frontal cortices) and the basal ganglia. Among the latter structures, the subthalamic nucleus (STN) has been implicated recently as a key structure in decision-making processes (Frank et al., 2007). The purpose of the present investigation was to determine how STN modulation affects reward-based learning in patients with Parkinson's disease (PD) who have been treated with STN deep brain stimulation (DBS).

THE ROLE OF THE SUBTHALAMIC NUCLEUS IN REWARD-BASED DECISION-LEARNING

Parkinson's disease is a neurodegenerative condition that affects dopaminergic neurons of the substantia nigra that project in a compact bundle of fibers to the dorsolateral striatum (mostly the putamen; Bjorklund and Dunnett, 2007). DA projections within the striatum are differentially affected by the progression of PD. Early in the disease, dopamine is most severely depleted in motor areas (putamen and supplementary motor areas), which produce motor deficits, such as tremor, bradykinesia, and rigidity (McAuley, 2003). As the disease progresses, dopamine depletions affecting cognitive circuits of the basal ganglia contribute to impairments in reversal learning, decision-making, working memory, response inhibition, and speed/accuracy balancing (Cooper et al., 1992; Swainson et al., 2000; Cools et al., 2001; Wylie et al., 2009a,b).

Although the most common treatment for PD consists of dopamine (DA) precursors (typically L-DOPA) and agonists, DBS of the STN has become the treatment of choice in patients whose symptoms are less well controlled by medications (Limousin et al., 1995; Lang, 2000). The remedial effects of DBS on the motor symptoms of PD are substantial (Bergman et al., 1990; Benazzouz and Hallett, 2000; Benabid, 2003; Meissner et al., 2005; Liu et al., 2006), although the specific mechanisms underlying its therapeutic effects are still debated. PD patients treated with DBS of the STN

for the purpose of alleviating motor symptoms afford the unique opportunity to investigate the effects of stimulation of this specific basal ganglia region on cognitive functions.

Substantial evidence from animal studies (Bergman et al., 1994; Baunez et al., 2001; Karachi et al., 2005) and PD patient studies (Jahanshahi et al., 2000; Schroeder et al., 2002; Witt et al., 2004; van den Wildenberg et al., 2006; Wylie et al., 2010) documents that the STN is critically involved in both motor control and action selection (Boraud et al., 2002). The role of the STN and the effects of DBS of the STN on reward-based decision-learning processes have not been studied as extensively. The effects reported in the literature appear more variable, ranging from null effects to impairments in some studies and improvements in others (Jahanshahi et al., 2000; Saint-Cyr et al., 2000; Funkiewiez et al., 2006; Frank et al., 2007). The STN receives sensorimotor, cognitive and limbic input from the external segment of the globus pallidus (GPe). Although these projections stem from functionally separate sources, the boundaries between sensorimotor, cognitive, and affective territories within the STN are not sharply defined (Karachi et al., 2005), nor is there a clear segregation between modalities in the output of the STN (Sato et al., 2000). Mallet et al. (2007) recently proposed that the STN not only regulates input from different modalities, but also integrates sensorimotor, emotional, and cognitive aspects of behavior.

Additionally, the STN may receive reward-related information (i.e., expected magnitude of reward) from medial orbito-frontal cortex OFC projections to STN similar to the input from premotor cortex, as has been shown in rats (Maurice et al., 1998) and hold the response output system (thalamus) in check until the expected reward options for a certain response are evaluated (Frank et al., 2007). Stimulating the STN may disinhibit the limbic circuits analogous to the disinhibition of motor circuits.

Animal studies also indicate that the STN plays a role in reward-based decision-making. In rats, STN lesions increase the incentive salience of reward-related stimuli (Uslaner and Robinson, 2006; Uslaner et al., 2008), which could be an indication of enhanced motivation and may affect learning by increasing reward motivation.

Studies on reversal learning, which depends mainly on learning from negative feedback, showed improved performance with STN stimulation in medication-withdrawn PD patients (Funkiewiez et al., 2006) and in animals with STN lesion (El Massioui et al., 2007). In contrast, learning based on positive and negative feedback in a reward-based probabilistic learning task remained unchanged in mildly medicated PD patients ON compared to OFF STN stimulation (Frank et al., 2007). Associative learning (stimulus–response learning not based on reward or action outcome) declined in PD patients treated by DBS of the STN, either with or without L-DOPA medication (Jahanshahi et al., 2000; Saint-Cyr et al., 2000). Given these mixed results, the present experiment sought to further examine the influence of STN stimulation on reward-based learning. Using an experimental paradigm designed by Haruno and Kawato (2006a,b), we distinguished the effects of STN stimulation on two key phases of reward-based learning, an early phase marked by the processing of reward-prediction errors (RPE) and a later phase that captures the strength of the formation of stimulus–action–reward associations. These two phases of learning were linked to dissociable regions of the striatum, both of which may be influenced by STN stimulation. Before turning to our specific predictions regarding STN stimulation, we briefly describe the role of striatal regions in reward-based learning reported by Haruno and Kawato.

THE ROLE OF THE STRIATUM IN REWARD-BASED DECISION-LEARNING

Lesion and human imaging studies demonstrate an important contribution of the striatum in reward-based decision-learning and support a functional dissociation between various aspects of the striatum (for an overview see Balleine et al., 2007). Recent fMRI studies by Haruno and Kawato (2006a,b) suggest that distinct regions within dorsal striatum may contribute to different phases of learning. These authors used a probabilistic reward learning task that required participants to learn stimulus–action associations to maximize rewards. Stimulus–action associations were rewarded probabilistically, with some associations rewarded more frequently (e.g., 90, 80, or 70%) and others rewarded less frequently (10, 20, or 30%). In the initial phase of learning, these associations were unknown, so participants engaged in trial and error decision-making. Because choices often led to losses as well as to rewards during this early phase, expectations about the outcome of a particular decision were sometimes disconfirmed and sometimes reinforced. As a result, performance depended crucially on the processing of RPE in order to adjust decision-making strategies. The RPE quantifies the mismatch (error) between the actual reward received upon a response to a stimulus and the expected reward that is based on prior outcomes associated with this response to the stimulus. The RPE is continuously updated with each response to a stimulus, thus allowing every new response choice and outcome to be weighed against the accumulative experience of reward associated with certain responses. In this way, violations of reward expectations, especially early in the course of learning, can be used to guide future decisions so as to optimize response choices that most consistently produce rewards. As learning progressed, participants were more accurate in predicting the stimulus–action associations that maximized reward (stimulus–action-dependent reward prediction; SADRP). The SADRP quantifies the strength of the evolving association between a particular stimulus, response, and reward. Thus, the increasing value of the SADRP reflects the gradual buildup of learning, reaching its highest value toward the end of the learning experience. Theoretically, as learning evolves, the value of the RPE should steadily decrease (as subjects become better at predicting reward outcomes) while the SADRP value should steadily increase (as subjects learn which responses optimize reward). See Section "Materials and methods" for computational formulas on RPE and SADRP.

The authors discovered that caudate nucleus activity was most strongly associated with the processing of outcome errors (i.e., RPE) early in the course of learning. In contrast, activity in the putamen was highly correlated with the prediction of rewarded stimulus–action associations (i.e., SADRP) during the latest phase of the learning task. To explain these patterns, the authors proposed that the caudate (embedded in the cortical striatal loop which includes the orbito-frontal cortex, OFC, and dorsolateral prefrontal cortex, DL-PFC) is involved in generating and testing hypotheses regarding reward optimization. Global reward-related features of the stimulus–action–reward associations are propagated from the caudate to motor loops (which include the putamen and premotor

areas) by means of a dopamine signal that is subserved by reciprocal projections between the striatum and the substantia nigra (Haruno and Kawato, 2006b). During later stages of learning, putamen activity increases with reward predictions (i.e., with learning SADRPs). Activity in the putamen increases to incorporate more specific motor information with the associated stimuli and the expected reward; that is, the reward associated with a specific stimulus and a specific action becomes more predictable and learning is gradually fine-tuned (Haruno and Kawato, 2006b). As these SADRP values increase, the RPE is reduced as subjects more accurately anticipate the rewards associated with their actions.

PRESENT STUDY
Patients with PD performed the Haruno and Kawato (2006a) task with and without stimulation of their STN. We determined the effect of STN stimulation on RPE during the early phase of learning and on formation of SADRP during the last phase of learning. DBS is targeted primarily at the motor regions of the STN, and such stimulation clearly enhances motor functions (Kleiner-Fisman et al., 2003), which are supported primarily by regions of the dorsal putamen and its associated circuitry in the motor loop. Moreover, an extensive literature indicates that deficient motor learning processes are a prominent feature of PD (McAuley, 2003; Faure et al., 2005). Therefore, our strongest prediction was that that STN stimulation would produce beneficial effects on the formation of stimulus–action–reward associations (i.e., SADRP values) during later stages of learning, indirectly implicating stimulation effects on putamen activity. Stimulating the STN may affect cognitive and limbic loops as well, because STN output is not sharply segregated (Mallet et al., 2007). Hence, we also assessed whether STN stimulation impacts the processing of RPE early in the course of new learning, which might indirectly implicate alteration in caudate activity.

MATERIALS AND METHODS
PATIENTS AND SURGERY DETAILS
Our study included 12 PD patients (three females; mean age = 61.1 years) who were treated successfully with bilateral DBS of the STN and who were concurrently treated with dopaminergic medications (with the exception of one patient). Patients remained on their prescribed doses of dopamine medications during their testing and were studied during their optimal therapeutic window. **Table 1** shows participant information.

All participants were free of dementia and did not demonstrate clinical levels of depression at the time of testing. Participants with PD were recruited from the Movement Disorders Clinic at the University of Virginia Medical Center and diagnosed with PD by a neurologist specializing in movement disorders. All PD patients ambulated independently and were rated a Hoehn–Yahr Stage III or less when their STN were being stimulated.

The surgical procedure for STN DBS utilized standard stereotactic techniques with microelectrode recordings for electrophysiological localization and has been described previously (Elias et al., 2007). Briefly, macroelectrodes (Medtronic Model 3389) consisting of four platinum–iridium cylindrical surfaces, each with a diameter of 1.27 mm, length of 1.5 mm, and edge-to-edge separation of 0.5 mm, were guided into the STN using MRI-guided stereotaxy and intraoperative microelectrode recordings. The planned

coordinates for macroelectrode placement was based on direct visualization of the STN on T2-weighted magnetic resonance images. Final electrode position was based on microelectrode recordings and confirmed intraoperatively with macrostimulation after implantation of the DBS electrode. Selection of final bipolar contacts and stimulation settings were determined on an individual basis to optimize control over clinically manifest motor symptoms.

The PD patients completed participation under two conditions, once when the STN was stimulated and once without STN stimulation. Patients were tested at a minimum of 3 months post-surgery. Exclusion criteria were: dementia; history of neurological condition other than PD; untreated or unstable mood disorder; history of bipolar affective disorder, schizophrenia, or other psychiatric condition known to compromise executive cognitive functioning; and untreated or unstable medical condition known to interfere with cognitive functioning (e.g., diabetes, pulmonary disease). All participants had normal or corrected-to-normal vision. All subjects participated voluntarily and gave their written informed consent prior to participation, as part of procedures that complied fully with relevant laws and with standards of ethical conduct in human research as regulated by the University of Virginia human investigation committee.

TASK AND APPARATUS
Questionnaires
The mini-mental status examination (MMSE, Folstein et al., 1975) assessed the global cognitive state of each patient to verify the absence of dementia (i.e., MMSE score higher than 25). To capture the effects of DBS of the STN on fine motor dexterity and speed, we administered the Purdue Pegboard task (Lezak, 1995) and a finger tapping test during each condition. The latter required participants to use the index finger of each hand to tap a tapping board as fast as possible during a period of 10 s. The tapping task alternated between each hand until three attempts were completed with each hand.
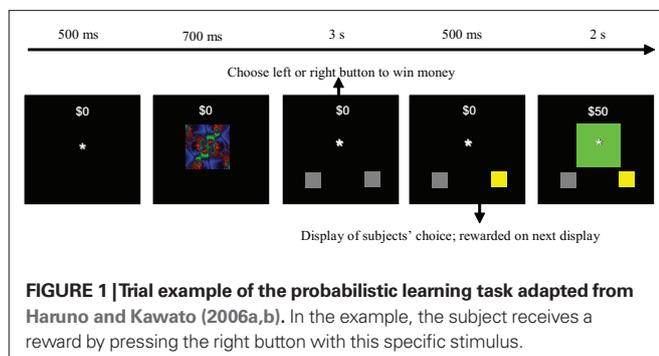
Experimental paradigm
A probabilistic learning task, adapted from Haruno and Kawato (2006a,b), was implemented on an IBM-compatible computer with a 17-inch digital display monitor. The computer screen, placed at a distance of 91 cm, was positioned such that stimuli appeared at eye level. Stimuli consisted of colored pictures against a dark background. Responses to stimuli were right or left thumb button presses registered by comfortable handheld grips.

The probabilistic learning task was designed to estimate RPE and measure the learning of SADRP, which have been linked to caudate nucleus and putamen activity, respectively. Subjects were instructed that the goal of the task was to make as much money as possible by pressing a left or a right button press to each picture stimulus that appeared on the computer screen. Each response provided the chance to either win or lose $50 in game money (note: participants were not remunerated for their participation). **Figure 1** depicts the sequence of a trial from the task. Each trial began with the presentation of a fixation point (an asterisk) in the center of the screen, and subjects were instructed to focus on this point in anticipation of the presentation of a picture stimulus. After a duration of 500 ms, the fixation point was extinguished and one of three picture stimuli (colored

Table 1 | Patient information.

Variable	Sample (N = 12)	
	Mean	**SE**
Age (years)	**61.1**	**2.3**
Sex (male/female)	9/3	
MMSE	29.1	0.3
Finger tapping ON (# taps)	42	2
Finger tapping OFF (# taps)	34	3
Pegboard ON (s)	31.4	2.5
Pegboard OFF (s)	34.6	3.2
Years since disease onset	11.4	1.8
l-DOPA (daily dose mg)	425.0	81.53
STIMULATION DETAILS		
Left STN		
Voltage (V)	3.2	0.2
Rate (Hz)	138.2	4.2
Pulse width (μs)	68.2	4.2
Right STN		
Voltage (V)	3.1	0.3
Rate (Hz)	138	4.7
Pulse width (μs)	74	4.2

FIGURE 1 | Trial example of the probabilistic learning task adapted from Haruno and Kawato (2006a,b). In the example, the subject receives a reward by pressing the right button with this specific stimulus.

fractals) appeared in the same location as the fixation point. The picture stimulus subtended visual angles of $5.67° \times 4.41°$ (9×7 cm). The picture stimulus remained on the screen for 700 ms. Participants were instructed to view the picture stimulus, but not to respond until the picture stimulus disappeared and was replaced by a response screen. The response screen consisted of the fixation point and two gray boxes displayed at the bottom left and bottom right portions of the screen, respectively (see **Figure 1**). Upon the presentation of the response screen, the participant was instructed to make a left or a right button press, which would then be indicated on the screen by a change in color (from gray to yellow) of the box that corresponded to the response side that was chosen (left button press = left box turns yellow). The participant was given 3 s to issue a response. After the button press was indicated on the screen, a large box with feedback appeared in the center of the screen. If the participant chose the correct response, the large box appeared in green, indicating that $50 had been won. If the incorrect response was chosen, the box appeared in red, indicating that the participant had lost $50. Throughout the entire trial, a running tab of the total amount of money won by the

participant was depicted in the upper center portion of the screen. Thus, if the participant won or lost $50 on a particular trial, the running total was immediately updated.

Subjects completed a block of 48 trials for each of the three conditions. For each condition, a novel set of three picture stimuli were used. The reward outcome of each response to a picture stimulus was determined in the following way: (1) for each picture, one response hand was assigned as the optimal choice and the other response hand was designated as the non-optimal choice; (2) in the first condition, selecting the optimal response hand resulted in a 90% probability of winning $50 and a 10% probability of losing $50; (3) in a second condition, selecting the optimal response hand resulted in an 80% probability of winning $50 and a 20% probability of losing $50; (4) in a third condition, selecting the optimal response hand resulted in an 70% probability of winning $50 and a 30% probability of losing $50. In all conditions, the probabilities of winning versus losing were reversed for the non-optimal relative to the optimal response hand. As an example, in the 90/10 condition a left response to fractal stimulus 1 (FS1) yielded a 50 dollar reward with a probability of 0.9 (90%) and a 50 dollar loss with a probability of 0.1 (10%). A right response to FS1 yielded a 50 dollar loss with a probability of 0.9 and a 50 reward with a probability of 0.1. Therefore, the optimal behavior for FS1 in the 90/10 condition was to press the left button, which participants had to learn by trial and error. The dominant probabilities for optimal behavior regarding the other fractal stimuli (FS2 and FS3) in the 90/10 condition were also 0.9. The optimal response for each fractal was pseudorandomized over left and right hands, for example optimal behavior could be FS1: right, FS2: left, FS3: right, which means that these responses were rewarded with positive feedback 90% of the time. Similarly, a response pattern could consist of two fractals that were rewarded (most of the time) with a left hand response and one with a right hand response. For each condition, the specific response options were randomly attached to each of the fractals. Across conditions, sessions and subjects, left and right hand dominant response patterns occurred equally often. Additionally, the fractal stimuli were presented randomly and with equal frequency within a condition. Condition order was also randomized.

PROCEDURE

Participants completed two similar versions of the task on the same day. The versions used different picture stimuli. The task was completed while the patient's STN was stimulated (ON condition), and again with stimulation turned off (OFF condition). The order of testing with respect to the status of STN stimulation was counterbalanced and randomly determined among patients. Prior to completing the task, each participant signed the consent form and completed the MMSE.

Also, each participant completed the pegboard and finger tapping tasks with and without STN stimulation. After STN stimulation was turned on or off, patients waited 30 min before commencing the cognitive task.

DATA ANALYSIS
Computational model to estimate SADRP and RPE

A reinforcement model (Q-learning, Sutton and Barto, 1998) was used to estimate each participant's SADRP and RPE during learning. Q-learning is an implementation of a temporal difference

model which assumes that stimulus–action–reward associations are acquired as a single representation during learning. The SADRP value (Q) consists of the predicted amount of reward for a certain decision (left or right response, r) made for a specific stimulus (one of three fractal stimuli, FS). It thus relates reward to sensory input and actions. Individual predicted reward values (SADRPs) for each action (two response) and each fractal stimulus (three different fractal stimuli) will be calculated at time t, $Q_t(FS, r)$ which adds up to six SADRP values per condition. The RPE represents the actual reward received (rt) minus the expected reward, $RPE = rt − (Q_t(FS, r)$. For the next occurrence of the same stimulus and action, SADRP and RPE values are updated according to the "Q-learning algorithm" to maximize reward (Sutton and Barto, 1998), $Q_{t+1}(FS, r) = Q_t(FS, r) + a_t^{FS}(rt − (Q_t(FS, r)))$.

The learning rate is updated separately for each FS according to the following rule: $a_t^{FS} = (a_{t-1}^{FS}) /(1 + a_{t-1}^{FS})$.

The formula of this learning rate is often used in reinforcement learning studies or studies on adaptive control (Young, 1984; Bertsekas and Tsitsiklis, 1996; Dayan et al., 2000; Haruno and Kawato, 2006a,b). It provides an estimation of a learning parameter which is updated recurrently with the presentation of a stimulus. In the current study, a_t^{FS} reduces with the presentation of each fractal stimulus, but remains equal if a specific FS is not presented.

The learning rate (a_t^{FS}) decreases towards the end of the learning stage (when SADRP becomes reliable). This is an important feature of a_t^{FS} because it means that, at the end of learning, the SADRP is less affected by an unexpected RPE (due to the probabilistic nature of the task).

The RPE is large at the beginning of learning (i.e., first 24 trials), while the SADRP value is small. Major changes in SADRP are especially expected in the first stage of learning. In a later stage of learning (i.e., last 24 trials) SADRP becomes accurate and does not show large changes (converges to an asymptotic value). Additionally, RPEs are expected to be small at the end of learning.

STATISTICAL ANALYSES

Motor performance on the finger tapping and pegboard tests was analyzed separately by a one-tailed paired samples t-test. We expected motor performance to improve with STN stimulation compared to without it. A one-sample t-test was used to test whether MMSE scores (OFF stimulation) were significantly larger than 25.

To analyze performance on the reward learning task, we first calculated a learning rate (cumulative percentage correct over trials) for each condition as a function of DBS state to investigate whether STN stimulation affected learning in general. Cumulative accuracy (at the last trial) was analyzed by a repeated-measures analysis of variance (RM-ANOVA) including the within subject variables Condition (90/10, 80/20, 70/30) and Stimulation (OFF, ON).

Second, SADRP and RPE values were analyzed separately by RM-ANOVA, including the within-subjects variables of Stimulation (OFF, ON) and Time (First, Second half of each block) and Condition (90/10, 80/20, 70/30) to investigate whether the patients show learning on RPE and SADRP from the first to the second half of the experiment. The analyses were based on the mean RPE value from the first half of the task (calculated on first 24 trials) and the second half of the task (based on the second

24 trials) and the mean SADRP value from the first and second half of the task. We predicted a beneficial effect of STN stimulation on SADRP during the last stage of learning and on RPE during the first stage of learning.

Subsequently, SADRP and RPE values of the theoretically relevant learning phases were more elaborately analyzed. The RPE analyses were based on the mean RPE value calculated on the first half, and the SADRP analyses on the mean SADRP value based on the last half of the experiment. SADRP and RPE values were analyzed separately by RM-ANOVA, including the within-subjects variables of Stimulation (OFF, ON) and Condition (90/10, 80/20, 70/30). Condition types are represented as the dominant versus non-dominant probability. Specific predictions were tested by using a simple contrast test; that is, the 90/10 condition was compared with the 80/20 condition and the 70/30 condition.

Since individual disease characteristics of PD patients may affect cognitive performance, like disease duration and age (Kaasinen and Rinne, 2002), we also took these variables into account. Disease duration and age were correlated with the dependent variables (improvement in RPE and SADRP comparing ON and OFF stimulation) to identify which individual characteristics could be predictive for performance in the learning task. The variables that turned out to correlate significantly with the dependent variables were used as predictors in the subsequent regression analysis.

First, we correlated change in RPE (RPE ON minus OFF = ΔRPE) and change in SADRP (SADRP ON minus OFF = ΔSADRP), separately for each condition, with individual characteristics (disease duration and age). Negative ΔRPE indicates that participants improved, whereas positive ΔRPE indicates that they were impaired ON compared to OFF stimulation. SADRP values are expected to increase ON versus OFF stimulation; therefore high ΔSADRP indicates improved performance.

A stepwise regression analysis was then performed with the variables that turned out to significantly correlate with RPE and SADRP, that is, disease duration and age with ΔRPE in the 90/10 condition. ΔSADRP did not significantly correlate with any of the individual characteristics. Thus, dependent variables in the regression analysis were change in RPE ON compared to OFF (ΔRPE in condition 90/10) and independent variables consisted of disease duration and age.

RESULTS

MOTOR PERFORMANCE AND MENTAL STATUS

Consistent with improved fine motor control associated previously with DBS of the STN, turning stimulation ON (compared to turning the device OFF) increased finger tapping speed ($t(11) = 3.5$, $p < 0.01$) and nearly significantly sped pegboard performance ($t(11) = −1.7$, $p = 0.06$; one-sided). The mean MMSE score was significantly larger than 25, $M = 29.1$ ($t(11) = 12.2$, $p < 0.001$), indicating that our participants were not demented.

Figure 2 presents the cumulative accuracy values separate for each condition. **Figure 3** shows the mean RPE separately for the first and second half of the experiment. **Figure 4** displays the mean SADRP separately for the first and second half of the experiment.
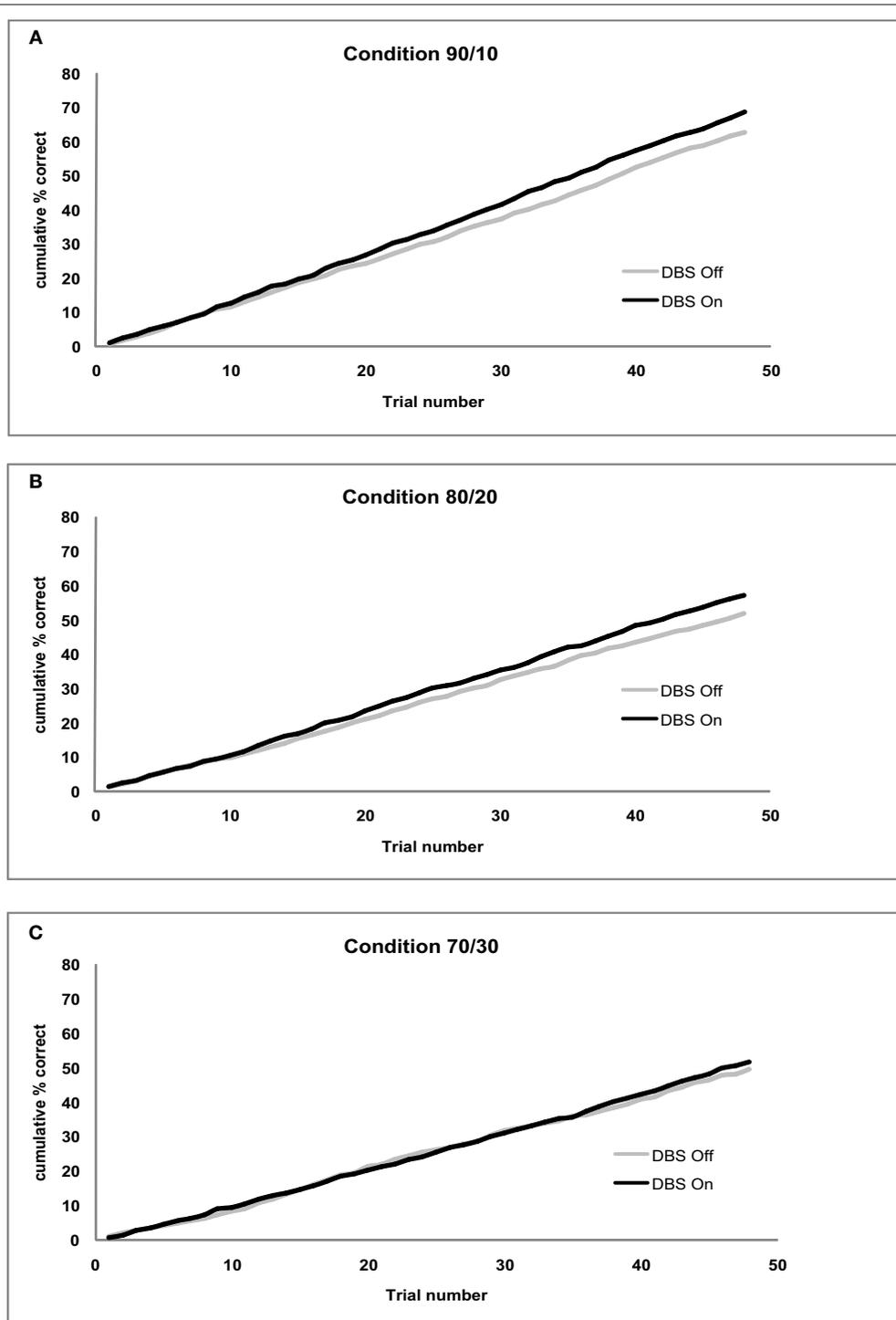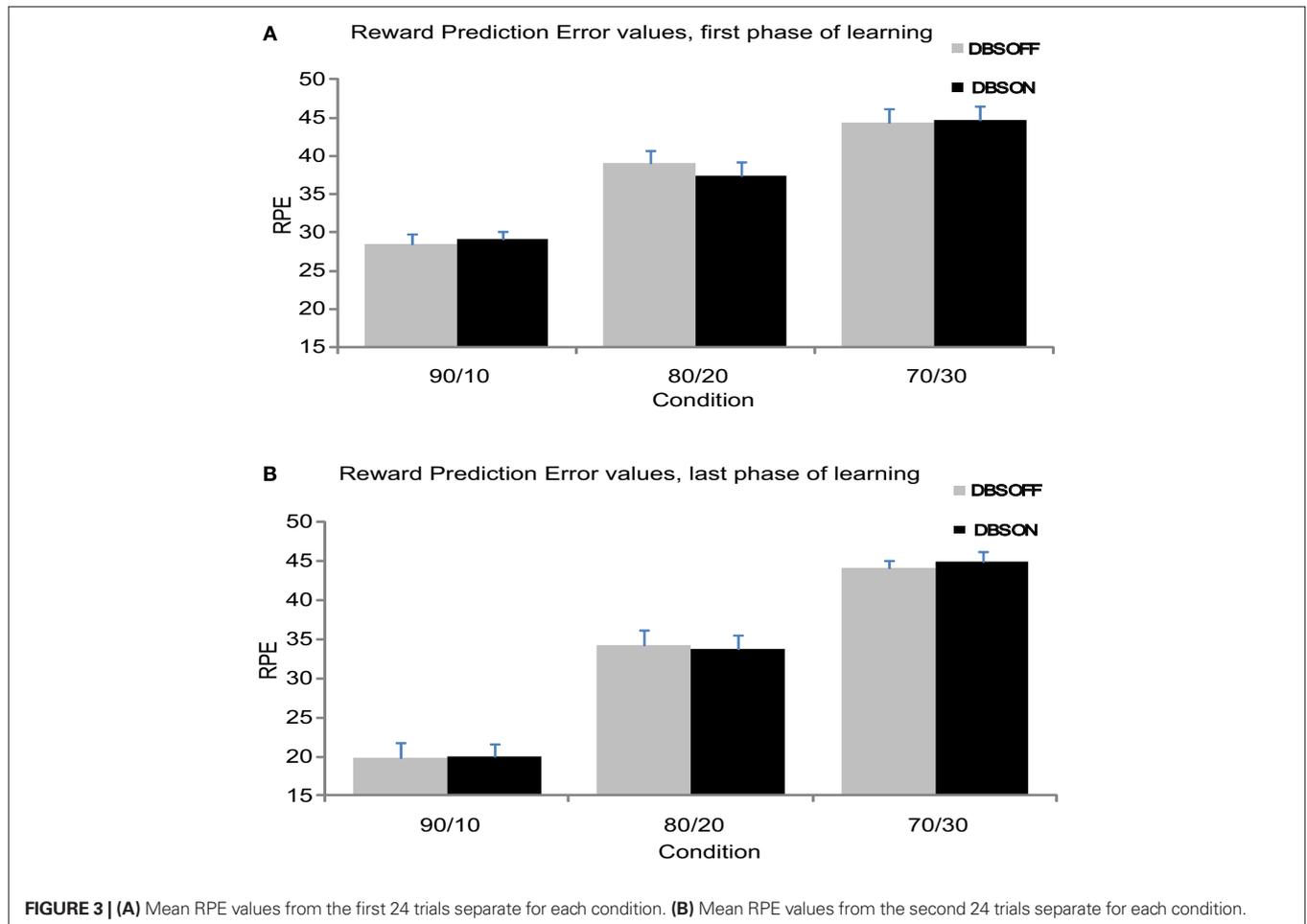
FIGURE 2 | (A) Cumulative accuracy for the 90/10 Condition by Stimulation (On/Off). **(B)** Cumulative accuracy for the 80/20 Condition by Stimulation (On/Off). **(C)** Cumulative accuracy for the 70/30 Condition by Stimulation (On/Off).

CUMULATIVE ACCURACY

Subthalamic nucleus stimulation produced a marginally significant effect on cumulative accuracy ($F(1,11) = 4.24$, $p = 0.06$). Patients showed higher accuracy values with their stimulation on ($M_{ON} = 58.97\%$, $M_{OFF} = 54.57\%$) compared to when the stimulation was off. Cumulative accuracy differed across Conditions ($F(1,11) = 16.98$, $p < 0.001$) showing larger accuracy values in the 90/10 ($M_{90/10} = 65.54\%$) compared to the 80/20 ($M_{80/20} = 54.34$) and 70/30 condition ($M_{70/30} = 50.43\%$). The Condition effect did not interact with Stimulation ($F(2,22) = 0.20$, $p > 0.8$). See **Figure 2** for cumulative accuracy values plotted trial-by-trial separate for each Condition.

FIGURE 3 | (A) Mean RPE values from the first 24 trials separate for each condition. **(B)** Mean RPE values from the second 24 trials separate for each condition.

REWARD-PREDICTION ERROR

Reward-prediction error values were significantly larger at the beginning of the experiment compared to the end of the experiment ($F(1,11) = 25.91$, $p < 0.001$), which indicates that the patients reduced their RPE over time (see **Figure 3**). Additionally, RPE values varied across Conditions ($F(1,11) = 124.37$, $p < 0.001$) revealing larger RPE values in the 70/30 and 80/20 compared to the 90/10 condition. The Condition effect interacted with Time ($F(2,22) = 12.35$, $p < 0.001$); the 90/10 condition showed a larger reduction in RPE from beginning to end of learning compared to the other conditions. STN stimulation did not influence RPE ($F < 1$), nor did Stimulation interact with Time or Condition ($F < 1$).

Haruno and Kawato reported that the first phase of learning correlated with activity in the caudate nucleus. While the above analysis showed no effect of STN stimulation on RPE across the entire block of trials, we focused the second analysis on the first phase of learning within each block to link our study to previous findings and provide better clarification about the effects of STN stimulation on RPE values linked to caudate nucleus activity (Haruno and Kawato, 2006a). The second RM-ANOVA, performed on the first phase of learning, showed that RPE values varied across Conditions ($F(2,22) = 42.17$, $p < 0.001$). Planned contrasts revealed a smaller RPE in the 90/10 Condition ($M_{90/10} = 28.76$) compared to the 80/20 and 70/30 Condition $F_{90/10-80/20}(1,11) = 25.96$, $p < 0.001$ $M_{80/20} = 38.26$; $F_{90/10-70/30}(1,11) = 79.60$, $p < 0.001$ $M_{70/30} = 44.54$

(see **Figure 2A**). STN stimulation did not influence RPE ($F < 1$). Moreover, RPE did not vary as a function of the interaction between Condition and Stimulation ($F(2,22) < 1$). These findings suggest that stimulating the STN did not influence RPE processing linked in previous studies to caudate activity during the early phase of learning.

STIMULUS–ACTION-DEPENDENT REWARD PREDICTION

The second half of the experiment yielded significantly larger SADRP values than the first half ($F(1,11) = 10.74$, $p < 0.01$), thus patients strengthened the association between stimulus, action and reward prediction over time (see **Figure 4**). Additionally, SADRP values differed across Conditions ($F(1,11) = 9.11$, $p < 0.001$) showing larger SADRP values in the 90/10 compared to the 80/20 and 70/30 condition. The Condition effect interacted with Time ($F(2,22) = 6.07$, $p < 0.01$); the 90/10 condition showed a larger increase in SADRP from beginning to end of learning compared to the other conditions. Overall, STN stimulation produced a nearly significant effect on SADRP values ($F(1,11) = 3.83$, $p = 0.08$) that was modulated by Time ($F(1,11) = 4.54$, $p = 0.06$) but not by Condition ($F < 1$).

Haruno and Kawato (2006a) reported that the last phase of learning correlated with activity in the putamen. The above analysis showed a tendency for STN stimulation to impact learning across the entire block of trials, but especially in the second half of each
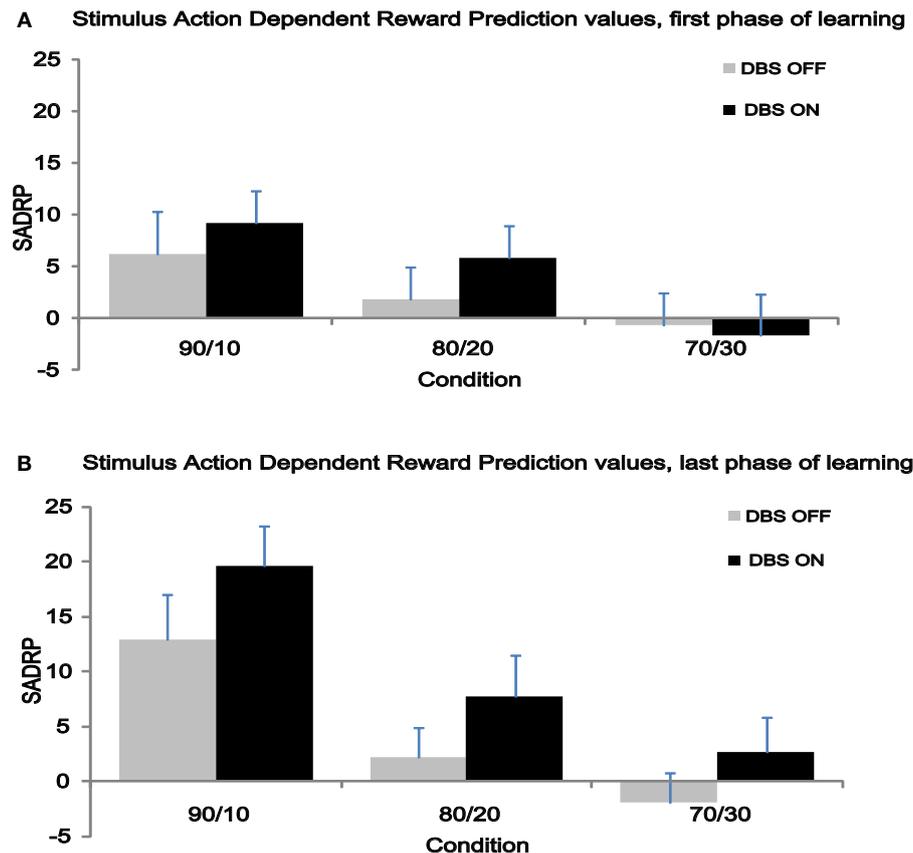
FIGURE 4 | (A) Mean SADRP values from the first 24 trials separate for each condition. **(B)** Mean SADRP values from the second 24 trials separate for each condition.

block. An additional analysis therefore focused on the last phase of learning to link our study to previous findings and provide better clarification about the effects of STN stimulation on SADRP values linked to putamen activity.

This analysis showed that learning of associations depended on the probability of being rewarded for a correct response (i.e., the Condition effect; $F(2,22) = 17.36$, $p < 0.001$). SADRP values in the 90/10 Condition ($M_{90/10} = 16.23$) were significantly larger than values obtained in the 80/20 and 70/30 Condition ($F_{90/10-80/20}(1,11) = 23.64$, $p < 0.01$) $M_{80/20} = 5.00$; ($F_{90/10-70/30}(1,11) = 32.76$, $p < 0.001$) $M_{70/30} = 0.40$ (see **Figure 3B**). Thus, patients learned better when the correct action was more likely to be rewarded. In contrast to the analysis of RPE, STN stimulation benefited the learning of SADRPs in the last phase ($F(1,11) = 8.11$, $p < 0.05$). Specifically, participants showed significantly larger SADRP values when their STN were stimulated ($M_{ON} = 10.00$) than when they were not ($M_{OFF} = 4.41$). No significant interaction between Stimulation and Condition was found ($F(2,22) < 1$) suggesting that STN stimulation improved performance in all conditions equally.

RELATION ΔSADRP AND ΔRPE ON–OFF WITH DISEASE DURATION AND AGE
Figure 5 shows ΔRPE ON–OFF in the 90/10 Condition plotted as a function of disease duration and age.

Although the ANOVA did not show a general effect of STN stimulation on RPE values, correlation analyses revealed that the change in RPE between Stimulation conditions was sensitive to individual characteristics. ΔRPE within the 90/10 Condition correlated significantly with disease duration ($r = 0.68$, $p < 0.05$) and showed a large correlation with age, $r = 0.57$, which was marginally significant ($p = 0.05$).
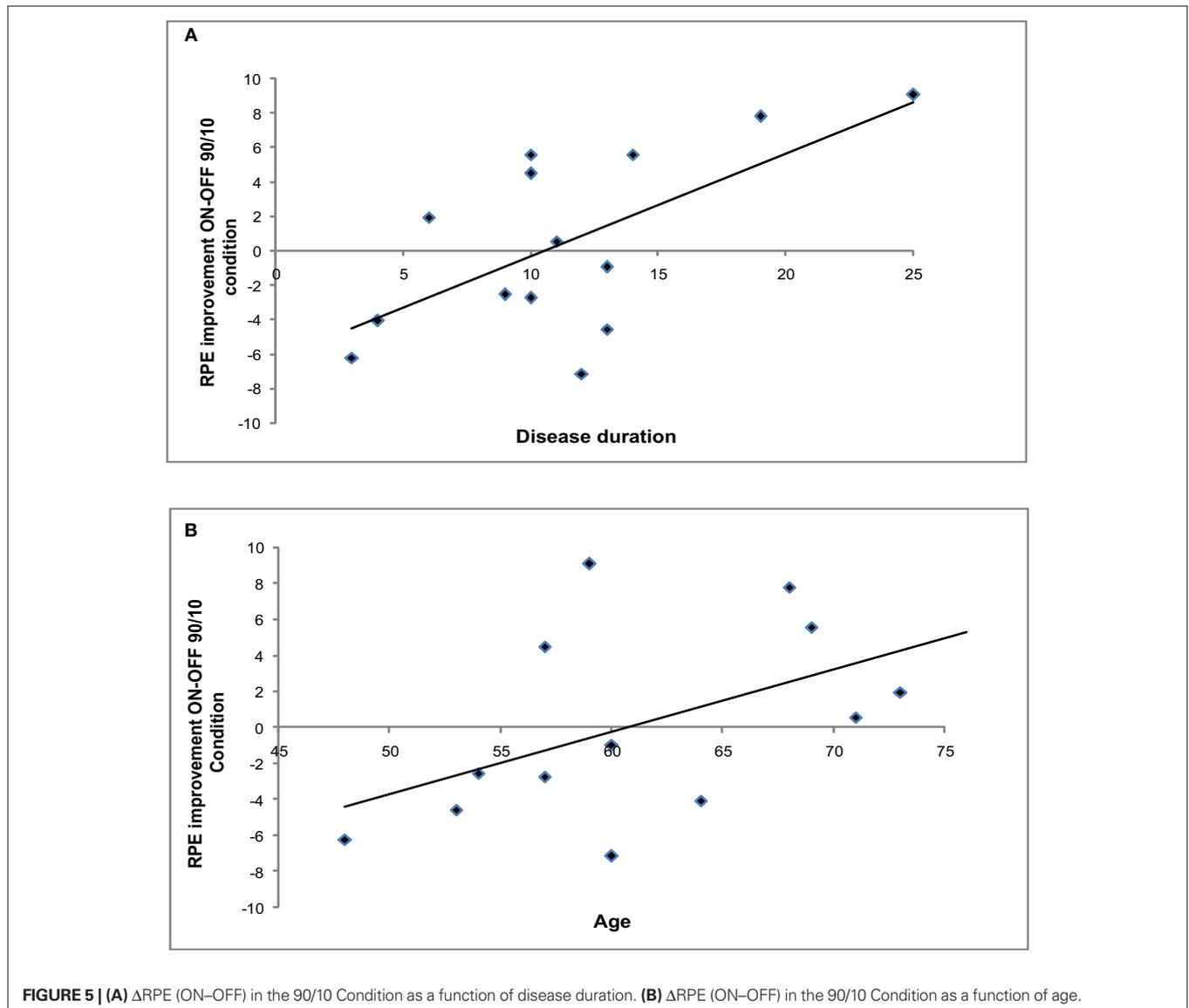
Thus, patients who were younger and earlier in the course of the disease showed the largest improvement in the RPE during the initial stages of learning when their STN were being stimulated.

In contrast, ΔSADRP within the 90/10 Condition did not correlate with age or disease duration (see **Table 2** for correlations). RPE and SADRP values of the 80/20 and 70/30 Condition did not reveal any significant correlations with Disease duration, or Age and were thus not included in the regression analysis.

Stepwise regression, with Disease duration and Age entered sequentially as predictors, showed disease duration to be a significant predictor of ΔRPE ($F(1,11) = 14.06$, $p < 0.01$), with age also explaining additional variance ($F(2,11) = 13.39$, $p < 0.01$). These effects are presented in **Table 3** for R and ΔR^2 for each of the predictive variables.

DISCUSSION
The present study investigated the effect of STN stimulation on separate components of reward-based learning: outcome evaluation (the processing of RPE to update hypotheses) and reward

FIGURE 5 | (A) ΔRPE (ON–OFF) in the 90/10 Condition as a function of disease duration. **(B)** ΔRPE (ON–OFF) in the 90/10 Condition as a function of age.

anticipation (the formation of SADRP) that have been tied to distinct regions in the striatum and their associated circuitries. The probabilistic reward-based decision-learning task used here successfully reproduced the behavioral findings first reported by Haruno and Kawato (2006a,b); participants' learning improved from the beginning to the end of the task, that is, prediction errors (i.e., RPE) reduced whereas the formation of predictive stimulus–action–reward associations (i.e., SADRP values) increased over time. Specifically, participants were able to adapt to RPE during early stages of learning and showed increased learning of SADRP across trials, especially in the condition with the highest degree of reward-related predictability. Similar to Haruno and Kawato (2006a), the conditions with a lower degree of reward-predictability turned out to be more difficult; performance dropped dramatically in patients with and without their STN stimulated.

Parkinson's disease patients completed the task twice, once with and once without STN stimulation. We predicted that STN stimulation would improve the formation of stimulus–action–reward

associations as reflected in higher SADRP values at the end of the task. Consistent with this prediction, SADRP at the late stages of learning was larger when the STN was stimulated compared to when it was not. Because SADRP values have been linked to activity in the putamen, this finding provides indirect support for the idea that DBS of the STN may benefit the action-oriented learning functions of the severely dopamine-depleted putamen in PD patients. This finding fits well with studies of PD showing that DBS of the STN enhances motor performance and control (Benabid, 2003; Kleiner-Fisman et al., 2003; van den Wildenberg et al., 2006; Wylie et al., 2010) and improves reward-based learning (Funkiewiez et al., 2006). It is also consistent with studies of STN lesions in rats (note that stimulation of the hyperactive STN in human PD patients is assumed to lead to roughly comparable inactivation effects as lesioning the STN in animals). Such studies indicate that STN lesions increase "wanting" and thereby facilitate reward-based learning (Uslaner and Robinson, 2006; Uslaner et al., 2008), particularly when the probability of receiving positive

Table 2 | Correlations between ΔSADRP and ΔRPE (ON compared to OFF STN stimulation) and disease duration and age.

Variables	Disease duration	Age
Disease duration (years)	–	0.15
Age (years)	0.23	–
ΔRPE 90/10 condition	0.68*	0.57
ΔRPE 80/20 condition	−0.03	0.26
ΔRPE 70/30 condition	0.16	0.01
ΔSADRP 90/10 condition	−0.52	-0.16
ΔSADRP 80/20 condition	0.31	0.19
ΔSADRP 70/30 condition	−0.01	0.05

*$p < 0.05$.

Table 3 | Linear stepwise regression on ΔRPE ON–OFF in the 90/10 condition as a function of disease duration and age.

Variables	β	R	ΔR^2
Step 1 Disease duration	0.70	0.76	0.58
Step 2 Age	0.41	0.87	0.16

reward is high. In our study, the modulating effect of STN stimulation was indeed most salient in the highly predictive condition (dominant probability 0.9).

In addition to determining the effects of DBS of the STN on learning proficiency, we also considered its effects on the processing of RPEs that occur in the early stages of learning. For learning to be successful, subjects must evaluate discrepancies between expected (or predicted) reward associated with a particular decision and the actual outcome of that decision. When an error occurs (i.e., predicted reward does not match the actual outcome), expectancies about possible outcomes associated with a decision can be updated to increase the likelihood of selecting a more optimal (i.e., reward-yielding) response in the future. As expectancies about the outcomes of particular decisions become more accurate, subjects are less swayed by the occasional violation of these reward expectancies and learn to optimize their selection of the most advantageous response to a stimulus. Thus, the processing of RPE, especially early in the course of the learning experience, is a fundamental aspect of effective learning. Overall, DBS of the STN did not influence RPE values, not even when zooming in on the initial learning phase. This suggests that the processing of RPEs, which has been linked to caudate nucleus activity (part of the cognitive corticostriatal loop), was insensitive to STN modulation.

Based on the current finding of a selective stimulation effect on SADRP, it can be inferred that the motor learning functions attributed to putamen function are relatively more impacted by STN stimulation than the reward prediction processing functions linked to caudate. The latter argues against the notion that stimulating the STN might also affect cognitive and limbic loops (Mallet et al., 2007). Of course, this reasoning assumes that motor, cognitive, and limbic loops are relatively closed and segregated, an assumption that is not universally made in other basal ganglia models (Joel, 2001), which predict that modulation of the motor loop affects the associative and limbic loops concurrently by means of open circuits connecting these processing functions. While the current study cannot exclude the possibility that circuits devoted to associative and limbic functions were impacted by STN stimulation, the dissociation of stimulation effects on RPE and SADRP is most consistent with a relatively greater impact on motor control and motor learning functions. More work is clearly needed to clarify the nature of segregation and overlap among motor, cognitive, and limbic circuits within the STN.

Although we are unable to verify the precise locations of the final STN electrodes (as is the case in a majority of published studies), we suspect that the sensitivity of the SADRP, but not the RPE, to STN stimulation may reflect the fact that the most effective contacts from a clinical standpoint are situated in a relatively more dorsal region of the STN. This region is more commonly linked to motor control functions as opposed to ventral regions that are speculated to more closely align with higher cognitive and reward processing circuits Hershey et al., 2010). Future studies that directly compare the effects of stimulation across relatively more dorsal and ventral regions of the STN would be highly informative.

Notably, an interesting finding concerning RPE values emerged when we took into consideration individual differences within the PD group. Specifically, younger patients with relatively short disease duration showed improvement in RPE values when their STN was being stimulated compared to when it was not. The reason for this association is unclear, although interestingly another study reported beneficial effects of DBS of the STN on aspects of learning among PD patients who were younger (mean age 54.5 years, SD = 7.5) and who had shorter disease duration (mean disease duration 10.7 years, SD = 3.9; Funkiewiez et al., 2006). Similarly, several clinical studies reported that younger patients and patients with a relatively short disease duration benefit more from stimulation of the STN in terms of general motor performance than older patients and patients who had a longer disease duration (Charles et al., 2002; Pahwa et al., 2006; Schupbach et al., 2007). It is tempting to speculate that perhaps compensatory mechanisms might be more effective in younger patients with a relatively short disease duration.

Using a different probabilistic reward learning task, Frank et al. (2007) failed to observe effects of STN stimulation on either positive or negative feedback learning in PD patients. According to a neurocomputational model developed to simulate behavior on their task (Frank, 2005), the STN provides a global NoGo signal because projections from the STN to GPi are diffuse and not response specific. Thus STN stimulation was predicted to have little effect on learning specific stimulus–response associations, but to lead to more impulsive decisions on high conflict trials. However, a comparison between the results of our study with Frank's study is not so straightforward since there are several differences in the study designs. To begin with, the task required speeded responses. Also, the most appropriate comparison is probably to contrast the effects of DBS between the learning phase of Frank's task and the learning phase of the Haruno task. In Appendix, Frank reported the absence of a DBS effect on learning across the probability conditions (i.e., 80–20, 70–30, 60–40). There is also another important distinction between the Haruno and Frank tasks. In the latter, reward is associated directly to a stimulus, whereas in the Haruno task, reward is associated with a stimulus–response ensemble. This added dimension of stimulus–reward learning may render the comparison untenable as well.

In contrast with Frank (2005) model, other BG models (e.g., Albin et al., 1989) would have predicted that stimulating the STN in PD patients would impair NoGo learning but improve Go learning. Stimulating the STN might have reduced the excessive activity in the NoGo pathway in PD patients in our study and thereby improved SADRP learning.

The beneficial effect of STN stimulation on the putamen may have been established through STN influences on multiple sites within the motor loop. STN stimulation may have modulated the processing of motor input information from GPe (entering the GPe via the putamen). Moreover, STN is directly activated by projections from the motor cortex (hyperdirect pathway, Nambu et al., 2000). Thus, if several competing responses are active in the motor cortex, the STN becomes increasingly activated, which leads to a global NoGo signal. Stimulating the STN may change the way these signals are processed, for example, if an already overactive STN in PD is excited by the motor cortex this leads to oscillatory activity and tremor, whereas extraneous stimulation or lesioning of the STN normalizes this activity (Bergman et al., 1990). Recently, a combined ERP and rTMS study in PD patients treated with DBS STN (Balaz et al., 2010) showed that modulating the inferior frontal cortex (IFC) by rTMS directly fastens ERPs in the STN via the hyperdirect pathway, thus reflecting speeded cognitive processing. Likewise, STN stimulation in the current study could have improved cortical processing via this hyperdirect pathway. Parametric modulation of STN stimulation in different functional STN areas might shed further light on the modulating role of STN in reward-based learning.

To summarize, our data suggest that the STN plays a modulatory role in reward-based learning, particularly in the formation of stimulus–action–reward associations. STN stimulation modulated S–R learning and was associated with more efficient reward processing when clinical characteristics were taken into account.

RELATION TO OTHER STUDIES

In the current study, feedback-based response selection was improved by STN stimulation. This is in line with the finding that action selection improves with STN stimulation (van den Wildenberg et al., 2006), but in contrast with findings from rat studies indicating that STN lesions induce impulsive responding (Baunez and Robbins, 1997; Baunez et al., 2001). Notably, STN stimulation may exert dissociable effects on impulsive behavior and cognitive control processes in ways that appear contradictory but reflect separable and temporally dissociable processes. That is, PD patients with STN stimulation might be more prone to both fast response capture (due to enhanced sensitivity of STN to inputs of the pre-SMA) as well as improved selective suppression (because the STN is more response to rIFC inputs) which develops more slowly (see Wylie et al., 2010). In the probabilistic learning task used in the current study, impulsive behavior would have led to less effective feedback processing and more random choices, which we did not find.

LIMITATIONS

There are some limitations related to the experimental paradigm and thereby the interpretation of the results. Although SADRP and RPE have been linked to the role of DA bursts at different time points and in different stages of learning, and have been shown to correlate with different striatal structures, at the behavioral level they are not entirely independent. That is, a decrease in RPE values increases SADRP values (according to the computational model). Thus, a null result of stimulation status on RPE values at the beginning of the task but an effect on SADRP at the end of learning does not entirely exclude that the caudate is modulated by STN stimulation. Rather, it suggests that the STN stimulation affects the putamen relatively more than the caudate, and affects late stages of learning more than early stages.

Currently it is unknown how many trials (and how much feedback) are needed to activate the caudate and putamen in PD patients and in what way this is modulated by STN stimulation, although there is some evidence that PD patients need more trials to learn (Shohamy et al., 2008). Future studies should test the critical time course of caudate and putamen involvement in probabilistic reward-based decision-learning in PD by means of an fMRI study.

Additionally, the PD patients in our study remained on their regular DA medication, although these dosages are smaller than in medicated-only PD patients. Nevertheless, DBS of the STN affects reward-based decision-learning above and beyond a DA effect. Future studies that consider the medication and DBS effects separately as well as their interaction will be important.

Finally, it is currently unclear to what extent variations in stimulation parameters, electrode location, and contact selection influence cognitive processes (Voon et al., 2006), and the current data cannot exclude the possibility that the electrical current may have spread to structures adjacent to the STN. Thus, there is a need for parametric investigation of stimulation settings on processes of reward-based stimulus–action learning.

CONCLUSION

In conclusion, DBS of the STN for treatment of PD motor symptoms also has a beneficial effect on learning stimulus–action–reward associations, a process shown previously to be associated with putamen activity. Thus, with DBS of the STN, PD patients were more effective at using feedback from their decisions to guide learning how to respond optimally to a stimulus situation. Moreover, relatively young patients with shorter disease duration particularly improved by DBS of the STN in their processing of reward errors early in the course of learning, which is essential for guiding new learning.

REFERENCES

Albin, R. L., Young, A. B., and Penney, J. B. (1989). The functional-anatomy of basal ganglia disorders. *Trends Neurosci.* 12, 366–375.

Balaz, M., Srovnalova, H., Rektorova, I., and Rektor, I. (2010). The effect of cortical repetitive transcranial magnetic stimulation on cognitive event-related potentials recorded in the subthalamic nucleus. *Exp. Brain Res.* 203, 317–327.

Balleine, B. W., Delgado, M. R., and Hikosaka, O. (2007). The role of the dorsal striatum in reward and decision-making. *J. Neurosci.* 27, 8161–8165.

Baunez, C., Humby, T., Eagle, D., Ryan, L., Dunnett, S., and Robbins, T. (2001). Effects of STN lesions on simple vs choice reaction time tasks in the rat: preserved motor readiness, but impaired response selection. *Eur. J. Neurosci.* 13, 1609–1616.

Baunez, C., and Robbins, T. (1997). Bilateral lesions of the subthalamic nucleus induce multiple deficits in an attentional task in rats. *Eur. J. Neurosci.* 9, 2086–2099.

Benabid, A. L. (2003). Deep brain stimulation for Parkinson's disease. *Curr. Opin. Neurobiol.* 13, 696–706.

Benazzouz, A., and Hallett, M. (2000). Mechanism of action of deep brain stimulation. *Neurology* 55, S13–S16.

Bergman, H., Wichmann, T., and DeLong, M. (1990). Reversal of experimental parkinsonism by lesions of the subthalamic nucleus. *Science* 249, 1436–1438.

Bergman, H., Wichmann, T., Karmon, B., and DeLong, M. (1994). The primate subthalamic nucleus. II. Neuronal activity in the MPTP model of parkinsonism. *J. Neurophysiol.* 72, 507–520.

Bertsekas, D. P., and Tsitsiklis, J. N. (1996). *Neuro-Dynamic Programming.* Belmont, MA: Athena Scientific.

Bjorklund, A., and Dunnett, S. B. (2007). Dopamine neuron systems in the brain: an update. *Trends Neurosci.* 30, 194–202.

Boraud, T., Bezard, E., Bioulac, B., and Gross, C. (2002). From single extracellular unit recording in experimental and human Parkinsonism to the development of a functional concept of the role played by the basal ganglia in motor control. *Prog. Neurobiol.* 66, 265–283.

Charles, P. D., Van Blercom, N., Krack, P., Lee, S. L., Xie, J., Besson, G., Benabid, A.-L., and Pollak, P. (2002). Predictors of effective bilateral subthalamic nucleus stimulation for PD. *Neurology* 59, 932–934.

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

Cooper, J. A., Sagar, H. J., Doherty, S. M., Jordan, N., Ridswell, P., and Sullivan, E. V. (1992). Different effects of dopaminergic and anticholineric therapies on cognitive and motor function in Parkinson's disease. *Brain* 115, 1701–1725.

Dayan, P., Kakade, S., and Montague, P. (2000). Learning and selective attention. *Nat. Neurosci.* 3, S1218–S1223.

El Massioui, N., Cheruel, F., Faure, A., and Conde, F. (2007). Learning and memory dissociation in rats with lesions to the subthalamic nucleus or to the dorsal striatum. *Neuroscience* 147, 906–918.

Elias, W., Fu, K., and Frysinger, R. (2007). Cortical and subcortical brain shift during stereotactic procedures. *J. Neurosurg.* 107, 983–988.

Faure, A., Haberland, U., Condé, F., and El Massioui, N. (2005). Lesion to the nigrostriatal dopamine system disrupts stimulus-response habit formation. *J. Neurosci.* 25, 2771–2780.

Folstein, M., Folstein, S., and McHugh, P. (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., Samanta, J., Moustafa, A., and 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. *J. Cogn. Neurosci.* 17, 51–72.

Funkiewiez, A., Ardouin, C., Cools, R., Krack, P., Fraix, V., Bair, A., Chabardès, S., Benabid, A. L., Robbins, T. W., and Pollak, P. (2006). Effects of levodopa and subthalamic nucleus stimulation on cognitive and affective functioning in Parkinson's disease. *Mov. Disord.* 21, 1656–1662.

Haruno, M., and Kawato, M. (2006a). Different neural correlates of reward expectation and reward expectation error in the putamen and caudate nucleus during stimulus-action reward association learning. *J. Neurophysiol.* 95, 948–959.

Haruno, M., and Kawato, M. (2006b). Heterarchical reinforcement-learning model for integration of multiple cortico-striatal loops: fMRI examination in stimulus-action-reward association learning. *Neural Networks* 19, 1242–1254.

Hershey, T., Campbell, M. C., Videen, T. O., Lugar, H. M., Weaver, P. M., Hartlein, J., Karimi, M., Tabbal, S. D., and Perlmutter, J. S. (2010). Mapping go-no-go performance within the subthalamic nucleus region. *Brain* 133, 3625–3634.

Jahanshahi, M., Ardouin, C., Brown, R., Rothwell, J., Obeso, J., Albanese, A., Rodriguez-Oroz, M. C., Moro, E., Benabid, A. L., Pollak, P., and Limousin-Dowsey, P. (2000). The impact of deep brain stimulation on executive function in Parkinson's disease. *Brain* 123, 1142–1154.

Joel, D. (2001). Open interconnected model of basal ganglia-thalamocortical circuitry and its relevance to the clinical syndrome of Huntington's disease. *Mov. Disord.* 16, 407–423.

Kaasinen, V., and Rinne, J. O. (2002). Functional imaging studies of dopamine system and cognition in normal aging and Parkinson's disease. *Neurosci. Biobehav. Rev.* 26, 785–793.

Karachi, C., Yelnik, J., Tande, D., Tremblay, L., Hirsch, E. C., and Francois, C. (2005). The pallidosubthalamic projection: an anatomical substrate for nonmotor functions of the subthalamic nucleus in primates. *Mov. Disord.* 20, 172–180.

Kleiner-Fisman, G., Fisman, D. N., Sime, E., Saint-Cyr, J. A., Lozano, A. M., and Lang, A. E. (2003). Long-term follow up of bilateral deep brain stimulation of the subthalamic nucleus in patients with advanced Parkinson disease. *J. Neurosurg.* 99, 489–495.

Lang, A. E. (2000). Surgery for Parkinson disease. *Arch. Neurol.* 57, 1118–1125.

Lezak, M. D. (1995). *Neuropsychological Assessment,* 3rd Edn. New York: Oxford University Press.

Limousin, P., Pollak, P., Benazzouz, A., Hoffmann, D., Le Bas, J. F., Broussolle, E., Perret, J. E., and Benabid, A. L. (1995). Effect of parkinsonian signs and symptoms of bilateral subthalamic nucleus stimulation. *Lancet* 345, 91–95.

Liu, Y., Postupna, N., Falkenberg, J., and Anderson, M. E. (2006). High frequency deep brain stimulation: what are the therapeutic mechanisms? *Neurosci. Biobehav. Rev.* 32, 343–351.

Mallet, L., Schupbach, M., N'Diaye, K., Remy, P., Bardinet, E., Czernecki, V., Welter, M. L., Pelissolo, A., Ruberg, M., Agid, Y., and Yelnik, J. (2007). Stimulation of subterritories of the subthalamic nucleus reveals its role in the integration of the emotional and motor aspects of behavior. *Proc. Natl. Acad. Sci. U.S.A.* 104, 10661–10666.

Maurice, N., Deniau, J. M., Glowinski, J., and Thierry, A. M. (1998). Relationships between the prefrontal cortex and the basal ganglia in the rat: physiology of the corticosubthalamic circuits. *J. Neurosci.* 18, 9539–9546.

McAuley, J. H. (2003). The physiological basis of clinical deficits in Parkinson's disease. *Prog. Neurobiol.* 69, 27–48.

Meissner, W., Leblois, A., Hansel, D., Bioulac, B., Gross, C., Benazzouz, A., and Boraud, T. (2005). Subthalamic high frequency stimulation resets subthalamic firing and reduces abnormal oscillations. *Brain* 128, 2372–2382.

Nambu, A., Tokuno, H., Hamada, I., Kita, H., Imanishi, M., Akazawa, T., Ikeuchi, Y., and Hasegawa, N. (2000). Excitatory cortical inputs to pallidal neurons via the subthalamic nucleus in the monkey. *J. Neurophysiol.* 84, 289–300.

Pahwa, R., Factor, S. A., Lyons, K. E., Ondo, W. G., Gronseth, G., Bronte-Stewart, H., Hallett, M., Miyasaki, J., Stevens, J., and Weiner, W. J. (2006). Practice parameter: treatment of Parkinson disease with motor fluctuations and dyskinesia (an evidence-based review) Report of the Quality Standards Subcommittee of the American Academy of Neurology. *Neurology* 66, 983–995.

Saint-Cyr, J., Trépanier, L., Kumar, R., Lozano, A., and Lang, A. (2000). Neuropsychological consequences of chronic bilateral stimulation of the subthalamic nucleus in Parkinson's disease. *Brain* 123, 2091–2108.

Sato, F., Parent, M., Levesque, M., and Parent, A. (2000). Axonal branching pattern of neurons of the subthalamic nucleus in primates. *J. Comp. Neurol.* 424, 142–152.

Schroeder, U., Kuehler, A., Haslinger, B., Erhard, P., Fogel, W., Tronnier, V. M., Lange, K. W., Boecker, H., and Ceballos-Baumann, A. O. (2002). Subthalamic nucleus stimulation affects striato-anterior cingulate cortex circuit in a response conflict task: a PET study. *Brain* 125, 1995–2004.

Schupbach, W. M. M., Maltete, D., Houeto, J. L., du Montcel, S. T., Mallet, L., Welter, M. L., Gargiulo, M., Béhar, C., Bonnet, A. M., Czernecki, V., Pidoux, B., Navarro, S., Dormont, D., Cornu, P., and Agid, Y. (2007). Neurosurgery at an earlier stage of Parkinson disease: a randomized, controlled trial. *Neurology* 68, 267–271.

Shohamy, D., Myers, C. E., Kalanithi, J., and Gluck, M. A. (2008). Basal ganglia and dopamine contributions to probabilistic category learning. *Neurosci. Biobehav. Rev.* 32, 219–236.

Sutton, R. S., and Barto, A. G. (1998). *Reinforcement Learning: An Introduction.* Cambridge, MA: MIT Press.

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.

Uslaner, J. M., Dell'Orco, J. M., Pevzner, A., and Robinson, T. E. (2008). The influence of subthalamic nucleus lesions on sign-tracking to stimuli paired with food and drug rewards: facilitation of incentive salience attribution? *Neuropsychopharmacology* 33, 2352–2361.

Uslaner, J. M., and Robinson, T. (2006). Subthalamic nucleus lesions increase impulsive action and decrease impulsive choice – mediation by enhanced incentive motivation? *Eur. J. Neurosci.* 24, 2345–2354.

van den Wildenberg, W. P. M., van Boxtel, G. J. M., van der Molen, M. W., Bosch, D. A., Speelman, J. D., and Brunia, C.

H. M. (2006). Stimulation of the subthalamic region facilitates the selection and inhibition of motor responses in Parkinson's disease. *J. Cogn. Neurosci.* 18, 626–636.

Voon, V., Kubu, C., Krack, P., Houeto, J. L., and Troster, A. I. (2006). Deep brain stimulation: neuropsychological and neuropsychiatric issues. *Mov. Disord.* 21, S305–S326.

Witt, K., Pulkowski, U., Herzog, J., Lorenz, D., Hamel, W., Deuschl, G., and Krack, P. (2004). Deep brain stimulation of the subthalamic nucleus improves cognitive flexibility but impairs response inhibition in Parkinson's disease. *Arch. Neurol.* 61, 697–700.

Wylie, S. A., Ridderinkhof, K. R., Elias, W. J., Frysinger, R. C., Bashore, T. R., Downs, K. E., van Wouwe, N. C., and van den Wildenberg, W. P. M. (2010). Subthalamic nucleus stimulation influences expression and suppression

of impulsive behavior in Parkinson's disease. *Brain* 133, 3611–3624.

Wylie, S. A., van den Wildenberg, W. P., Ridderinkhof, K. R., Bashore, T., Powell, V., Manning, C. A., and Wooten, G. F. (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., Manning, C. A., and Wooten, G. F. (2009b). The effect of speed-accuracy strategy on response interference control in Parkinson's disease. *Neuropsychologia* 47, 1853–1944.

Young, P. (1984). *Recursive Estimation and Time Series*. New York: Springer-Verlag.