EJN fens European Journal of Neuroscience

European Journal of Neuroscience, Vol. 35, pp. 1144–1151, 2012 doi:10.1111/j.1460-9568.2012.08043.x

Decomposing effects of dopaminergic medication in Parkinson's disease on probabilistic action selection – learning or performance?

P. Smittenaar,[1] H. W. Chase,[2] E. Aarts,[3] B. Nusselein,[4] B. R. Bloem[5] and R. Cools[4,6]
[1]The Wellcome Trust Centre for Neuroimaging, University College London, London, UK
[2]Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA
[3]Helen Wills Neuroscience Institute, University of California, Berkeley, CA, USA
[4]Donders Institute for Brain, Cognition and Behavior, Centre for Cognitive Neuroimaging, Radboud University Nijmegen, Nijmegen, The Netherlands
[5]Department of Neurology, Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands
[6]Department of Psychiatry, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands

Keywords: decision-making, dopamine, reinforcement learning, valence, value-based performance

Abstract

Dopamine has long been implicated in reward-based learning and the expression of such learned associations on performance. Robust evidence supports its effects on learning and performance, but teasing these apart has proved challenging. Here we have adapted a classic test of value-based learning, the probabilistic selection task, to disentangle effects of dopamine on value-based performance from effects on value-based learning. Valence-specific effects of dopamine on this specific task cannot be accounted for by modulation of learning, and therefore must reflect modulation of performance. We found that dopaminergic medication, consisting of levodopa and/or dopamine agonists taken at own dose, in 18 patients with mild Parkinson's disease (Hoehn and Yahr < 2.5) potentiated reward-based approach in terms of both accuracy and reaction times, while leaving punishment-based avoidance unaffected. These data demonstrate that the effects of dopamine on probabilistic action selection are at least partly mediated by effects on the expression of learned associations rather than on learning itself, and help refine current models of dopamine's role in reward.

Introduction

Current models of the role of midbrain dopamine in reward highlight its contribution to behavior through effects on learning, motivation, overcoming effort or, more generally, behavioral activation (Berridge & Robinson, 1998; Salamone *et al.*, 2005; Robbins & Everitt, 2007). The predominant view in computational and systems neuroscience is that dopamine serves to promote reinforcement learning, i.e. trial-and-error instrumental learning to choose rewarding actions (Houk & Wise, 1995; Montague *et al.*, 1996; Schultz *et al.*, 1997; Samejima *et al.*, 2005; Paton *et al.*, 2006). This idea is derived from the insight that the phasic firing of midbrain dopaminergic neurons of primates quantitatively resembles a 'reward prediction error' signal used in computational algorithms for reinforcement learning (Ljungberg *et al.*, 1992; Montague *et al.*, 1996, 2004; Hollerman & Schultz, 1998; Sutton & Barto, 1998; Waelti *et al.*, 2001; Bayer & Glimcher, 2005; Frank, 2005).

Further evidence for a role of dopamine in human reinforcement learning comes from controlled medication withdrawal studies in Parkinson's disease (PD). PD is associated with dopamine cell loss in the substantia nigra pars compacta (Hassler, 1938), which projects to the basal ganglia through mesolimbic and nigrostriatal projections. Dopamine depletion in PD is alleviated through dopamine-enhancing drugs and the role of dopamine in human cognition can be analysed by examining the effects of dopaminergic medication in PD.

Studies employing this approach have revealed effects of dopamine on reinforcement learning in value-based action selection paradigms (e.g. Frank *et al.*, 2004). For example, Frank *et al.* (2004, 2007) have demonstrated that dopaminergic medication in PD increased the likelihood of choosing reward-associated stimuli at the expense of avoiding punishment-associated stimuli. Based on these data it has been argued that dopaminergic medication in PD potentiates the relative tendency to learn from reward versus punishment. One problem with this argument is that success on these tasks depends not only on the gradual learning of stimulus–response associations based on reinforcement, but also on the expression of such learning during performance. Accordingly, it is pertinent to dissociate effects of dopamine on value-based performance, defined as processes affecting the selection and execution of responses independent of learning, from effects on reinforcement

Correspondence: P. Smittenaar, as above.
E-mail: peter.smittenaar.10@ucl.ac.uk

Received 6 October 2011, revised 10 January 2012, accepted 18 January 2012

learning. Indeed, dopamine is increasingly recognized to be involved not just in the acquisition but also in the performance of motivated behavior (Ikemoto & Panksepp, 1999; Berridge, 2007; Robbins & Everitt, 2007; Salamone *et al.*, 2007).

To disentangle the effects of dopamine on learning from those on performance, we present a novel task based on the probabilistic selection task by Frank *et al.* (2004). With this new approach, subjects learn a series of associations between affectively neutral cues and affectively neutral outcomes (i.e. stimulus–stimulus learning), rather than a series of associations between affectively neutral stimuli and reward or punishment (i.e. reinforcement learning). Outcome values were assigned only *after* learning and prior to testing. In the absence of reinforcement learning, any valence-specific effect of medication on approaching reward-associated stimuli versus avoiding punishment-associated stimuli in this paradigm must be due to an effect of medication on value-based performance rather than on the learning of stimulus–response–reinforcement associations. This does not imply all previously reported effects of medication on reinforcement learning must also reflect performance instead of learning effects. Rather, our results help refine current models of dopamine by showing that dopamine can also alter behavioral control independent of its established role in learning and plasticity.

Methods

Subjects

Eighteen patients with PD and 14 age- and education-matched healthy controls participated. Clinical and demographic characteristics are shown in Table 1; inclusion and exclusion criteria are given in Table S1.

All patients were diagnosed by an experienced movement disorders neurologist (B.R.B. or Dr R. Esselink) at the Parkinson centre of the Department of Neurology of Radboud University Nijmegen Medical Centre. All patients had idiopathic PD, according to the UK Brain Bank criteria. Clinical disease severity was assessed at the start of each session using the motor subscale (part III) of the Unified Parkinson's Disease Rating Scale (UPDRS; Fahn & Elton, 1987). All patients were taking dopaminergic medication (levodopa, dopamine receptor agonists, or both; details of each individual patient's medication (Wenzelburger *et al.*, 2002) are summarized in Table S2).

General procedure

Patients were tested both on and off their dopaminergic medication, at least 7 days apart (nine patients were tested on medication first). In the ON condition, patients took their regular second dose of the day (around noon) approximately 45 min before the start of the experiment. Prolonged release medication was taken at the regular time, typically early in the morning. In the OFF condition, all dopaminergic medication was withheld for at least 21 h (or 51 h for prolonged release medication) before start of the experiment. Healthy controls were tested twice to estimate test–retest (e.g. practice) effects.

All subjects provided informed written consent prior to their participation. All procedures were approved by the Committee for the Protection of Human Subjects (CMO region Arnhem Nijmegen; protocol number 2008/159) in accordance with the declaration of Helsinki.

The experiment was performed as part of a larger study for which patients and controls were scanned using functional magnetic resonance imaging (these results will be reported elsewhere). The experimental paradigm of interest for the current paper was administered outside the scanner, approximately 60 min after subjects were scanned.

Background neuropsychological tests

A battery of tests were used to probe a range of neuropsychological functions: Beck Depression Inventory (BDI) (Beck *et al.*, 1961), Mini-Mental State Examination (Folstein *et al.*, 1975), Frontal Assessment Battery (Dubois *et al.*, 2000), premorbid intelligence levels using the Dutch version of the National Adult Reading Test (DART) (Nelson & O'Connell, 1978; Schmand *et al.*, 1991), cognitive impulsiveness, motor impulsiveness and non-planning impulsiveness using the Barratt Impulsiveness Scale-11 (Barratt, 1985), and cognitive processing speed using the box completion task (Salthouse, 1994). A number cancellation task was used to assess sustained attention and concentration. The data were analysed using independent and paired-samples *t*-tests.

Task

The probabilistic selection task used previously by Frank *et al.* (2004) was modified for our purposes (Fig. 1). The task was programmed using Presentation 14.1 (Neurobs, Inc., Albany, CA, USA; http://www.neurobs.com). The instructions as given to the subjects are described in Appendix S1.

Like the original task, the current task consisted of two phases: an initial learning phase and a subsequent test phase. Unlike the original learning phase, our learning phase required subjects to learn a series of associations between affectively neutral cues and affectively neutral outcomes rather than a series of associations between affectively neutral stimuli and reward or punishment. Here, reward and punishment values were assigned to the outcomes only after learning and prior to testing (Fig. 1).

In the learning phase, three different pairs of Hiragana characters (AB, CD, EF) were presented in a random order, with the assignment of Hiragana characters to elements A–F randomized between subjects. Each character represented a cue and was associated, stochastically, with one of two affectively neutral outcomes, represented by two colored shapes (Table 2). Cues were presented in white (8 cm height

TABLE 1. Demographic and clinical data

	n	Men	Age (years)	UPDRS ON	UPDRS OFF	L-Dopa equivalent dose (mg)	Disease duration (years)
Control subjects	14	9	58.8 (3.3)	–	–	–	–
PD patients	18	12	55.4 (2.2)	20.8 (1.7)	29.9 (1.9)	555.8 (119.3)	4.8 (0.7)

Values represent mean (SEM). UPDRS, Unified Parkinson's Disease Rating Scale.

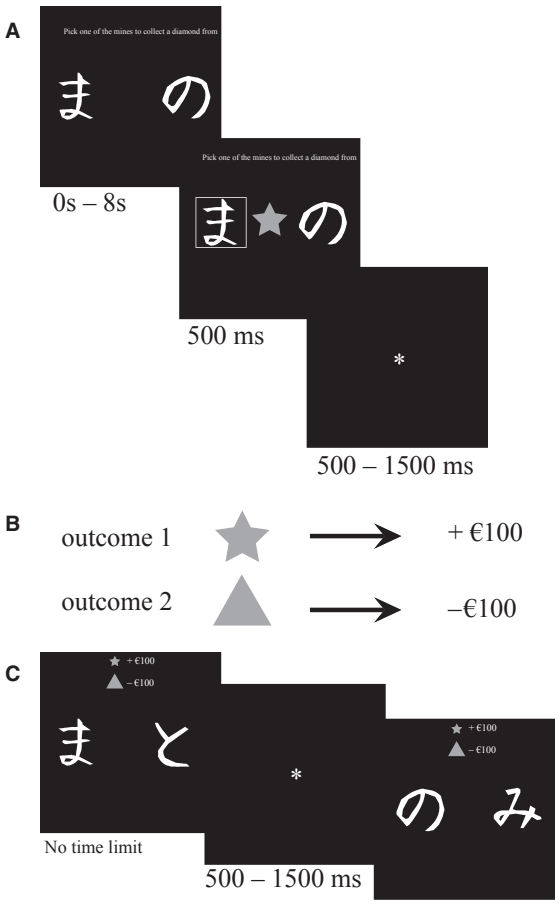

FIG. 1. Description of the probabilistic task with delayed valuation. (A) During the learning phase subjects gradually acquire an association between cues, represented by Hiragana characters, and two possible affectively neutral outcomes, represented by colored shapes. During each trial, subjects are shown one of the pairs, which are presented in pseudorandom order. After choosing one of the cues, one of two possible outcomes is shown, chosen randomly based on the contingencies shown in Table 2. Over the course of acquisition, subjects learn to associate cues A, C and E with neutral outcome one, and cues B, D and F with neutral outcome two. Note that acquisition of these associations does not involve reinforcement learning, as the subjects have not yet been told the value of the outcomes. (B) After 60 trials with each pair, subjects are then explicitly told what the value of the outcome is: outcome one yields money, whereas outcome two will lead to loss of money. This money is not actually paid, but subjects are informed they will see their score at the end of the experiment. (C) During the test phase, all 15 possible combinations of cues are shown six times and participants are instructed to maximize their gains, using their experience from the learning phase. No feedback is given to prevent ongoing learning, but a reminder of outcome values is shown at the top of the screen.

on-screen) and outcomes in color on a black background (5 cm height on-screen), ~80 cm from the subject.

During the learning phase, the subject's task was to learn the associations between cues and affectively neutral outcomes through observational learning. On each trial, a pair of cues was presented (right/left location randomized). Subjects had 8 s to respond with the left or right key to choose between the two cues. Upon choosing a cue, a white box was shown around that cue and the (neutral) outcome was shown in between the two cues for 0.5 s. After a jittered fixation period of 500–1500 ms, a new trial started. Subjects learned to associate cues with outcomes by keeping track of the outcomes of their choices. The learning phase consisted of four blocks of 60 trials. The order of presentation of the three pairs was randomized over three

TABLE 2. Cue–outcome contingencies during the learning phase

Pair	1		2		3	
Cue	A	B	C	D	E	F
Outcome 1 (%)	80	20	70	30	60	40
Outcome 2 (%)	20	80	30	70	40	60

trials, and for every 10 trials of each pair the observed outcome was controlled such that it corresponded to the intended contingency.

Accuracy during the learning phase could not be directly assessed because there was no 'right' or 'wrong' choice. Therefore, learning was assessed using a visual analogue scale (VAS) questionnaire after every 60 acquisition trials. Using the arrow keys, subjects first estimated for each cue the relative abundance of outcomes one and two on an 11-step VAS. The learning error was calculated as abs(estimate – actual), i.e. as the absolute difference between the subject's estimate and the correct values for each cue. For example, an estimate of two for cue A would be an error of one, because the correct estimate for cue A was three. This was not an arbitrary unit: one step on the VAS represented a 10% step in the ratio of the two cues (e.g. 60/40 to 70/30). Over- and under-estimations of the outcomes were treated the same. Subjects then indicated how certain they were of this estimate on a scale from one (uncertain) to 11 (certain) (for an illustration, see Fig. S1). The cues were shown one by one, in a random order. After the VAS questionnaire, subjects had a 15-s break.

After completion of the learning phase, the outcomes were assigned value; subjects were instructed that outcome 1, most strongly associated with cue A, would yield €100, while outcome 2, most strongly associated with cue B, would yield a loss of €100 (for literal instructions, see Appendix S2). During the test phase, which was identical to that used by Frank *et al.* (2004), subjects were shown all 15 possible novel combinations of the previously learned cues six times, with a reminder of the outcome values at the top of the screen. They were instructed to maximize their profits by choosing the best cue in each pair, based on the outcomes associated with them and their instructed values. Response time was unlimited, and testing was conducted in extinction to prevent any reinforcement learning. Subjects were informed that their final score would be shown to them after task completion. There was a 500- to 1500-ms jitter between response and presentation of the next pair.

All subjects, with the exception of two control subjects, were tested on two separate occasions. Both sessions used a unique set of stimuli for the cues and outcomes. To minimize test–retest effects, subjects on each session practised on a shorter, but otherwise identical version of the task with a unique set of stimuli. One of the authors was present during practice to ensure subjects understood the task.

Data analysis

All statistical tests were performed in SPSS for Windows (SPSS Inc., Chicago IL). The first set of analyses was designed to examine any test–retest effects on performance. Reaction times (RTs) < 200 ms were excluded, and RTs were transformed using the formula ln(RT) (RT in ms) to correct for a positive skew common in RT data. Proportional scores of accuracy, for which the variance is proportional to the mean, were transformed using the formula $2 \times \text{asin}\sqrt{x}$ (Howell, 1997, p. 324). Group (PD versus controls) was included as a between-subject factor, and testing session (first versus second) as a within-subject factor. These analyses revealed that there were no significant

test–retest effects, nor differences in test–retest effects between groups (see Results). This enabled us to conduct the second set of analyses, aimed at elucidating effects of dopaminergic medication on performance, irrespective of session order.

We performed two repeated-measures ANOVAs on the patient data. In the first model we examined the effect of the within-subject factors medication status (ON versus OFF) and valence (approach-A versus avoid-B) on test phase accuracy. In this model, the critical dependent measure was the number of times the reward-associated cue A was chosen (approach-A) when any pair was shown during the test phase that included A (but not B). Conversely, avoid-B represented the number of times B was avoided (i.e. not chosen) when any pair was shown that included loss-associated cue B (but not A). In the second model, we examined the effect of within-subject factors medication status and valence on RTs of successful approach-A and avoid-B trials.

Both models included a covariate that captured a bias for choosing cue A or B during the learning phase. We reasoned that if subjects happened to choose one of the two cues more often, then this might lead to improved learning of that specific cue–outcome association. This would add variance of no interest to the data, which in turn would reduce power to detect effects of interest. The covariate was calculated by taking the proportion of trials on which cue A was chosen in AB-pairs during the learning phase of each session (a value between zero and one). The covariate represented the difference in bias between the ON and OFF session (a value between −1 and 1).

Simple effects analyses of accuracy and RT measures were conducted to compare controls with PD patients ON and OFF medication, respectively. For these analyses, data were averaged across both sessions for 12 controls or duplicated for two controls who participated only once, such that all comparisons between controls and patients involved identical control data. The learning bias covariate (for either the ON or the OFF session) was also included.

Results

The controlled medication procedure was successful – motor signs were significantly reduced by medication, as shown by lower UPDRS scores in patients ON relative to OFF medication (Table 1; $F_{1,17} = 35.79$, $P < 0.0001$.

Background data

Background neuropsychological tests (Table 3) revealed that verbal IQ (as measured with the DART) was higher in controls than in patients ($t_{30} = 3.12$, $P = 0.004$); patients were more depressed than controls, as measured with the BDI ($t_{30} = 3.41$, $P = 0.002$); patients responded more slowly than controls on the box completion task both ON ($t_{30} = 2.82$, $P = 0.008$) and OFF ($t_{30} = 2.92$, $P = 0.007$) medication; and patients responded more slowly than controls on the number cancellation task both ON ($t_{30} = 2.59$, $P = 0.02$) and OFF ($t_{30} = 2.93$, $P = 0.01$) medication. All other scores were matched between patients and controls, including educational background measured as part of the DART (Fischer's exact test, $P = 0.26$).

Probabilistic action selection

Learning phase

Learning was assessed in terms of VAS error (see Methods) (Fig. 2). There was no main effect of cue type (A or B) on learning in patients ($F_{1,16} < 1$) and no cue type × medication interaction, either when taken across all blocks, or at the end of learning (both $F_{1,16} < 1$; Fig. 2B). This allowed us to collapse across cue A and B estimates for further analyses concerning learning (Fig. 2A). There was a main effect of block (four levels, one for each acquisition block; $F_{3,48} = 3.80$, $P = 0.02$) due to a reduction of error over blocks, indicating that the subject group as a whole learned the cue–outcome associations. There was also a main effect of medication ($F_{1,16} = 10.56$, $P = 0.01$): patients exhibited reduced error in their VAS estimates when ON compared to OFF medication. However, the medication × block interaction ($F_{3,48} = 1.13$, $P = 0.35$) was not significant. There was no main effect of control versus patients ON ($F_{1,29} < 1$) or OFF ($F_{1,29} < 1$) medication. Control subjects also did not show a difference in learning across blocks compared with patients ON (group × block $F_{3,87} < 1$) or OFF (group × block $F_{3,87} = 1.32$, $P = 0.27$) medication.

Medication was associated with increased confidence ratings in the cue estimates ($F_{1,16} = 8.50$, $P = 0.01$; see Fig. 2C), in accordance with a beneficial effect of medication on learning as reported above. However, there was no medication × cue interaction either when taken across all blocks or at the end of learning (both $F_{1,16} < 1$). This indicates that medication did not affect the relative confidence patients had in their A versus B estimates. Taken together, these results indicate there was no incidental bias in cue A versus cue B learning or confidence between medication conditions that might explain any valence-specific effect in the test phase. Note that overall differences in the learning phase between groups or medication conditions cannot account for any valence-specific effects (i.e. differences between approach and avoid conditions) in the test phase.

Test phase

All patients and 12/14 controls were tested twice on the same task. There was no effect of session on accuracy ($F_{1,28} < 1$), and there was no session × valence interaction effect on accuracy ($F_{1,28} < 1$). These findings allowed us to (1) average control data across both sessions

TABLE 3. Background neuropsychological data

					FAB		Digit span		Block completion		Number cancellation	
	DART*	MMSE	BIS total	BDI*	ON	OFF	ON	OFF	ON*	OFF*	ON*	OFF*
Patients	99.6 (4.1)	28.6 (0.4)	62.7 (1.5)	8.5 (1.0)	17.4 (0.2)	16.5 (0.3)	6.0 (0.3)	5.7 (0.2)	108 (6)	112 (7)	315 (7)	314 (10)
Controls	118 (4.1)	28.6 (0.3)	60.9 (1.7)	3.7 (0.9)	17.2 (0.2)		6.2 (0.3)		81 (8)		260 (16)	

Values represent mean (SEM). Asterisks indicate significance at $P < 0.05$ in two-tailed control versus patient Student's *t*-tests, uncorrected for multiple comparisons. DART, Dutch Adult Reading Test; MMSE, Mini-Mental State Examination; BIS, Barratt Impulsivity Scale; BDI, Beck Depression Inventory; FAB, Frontal Assessment Battery.

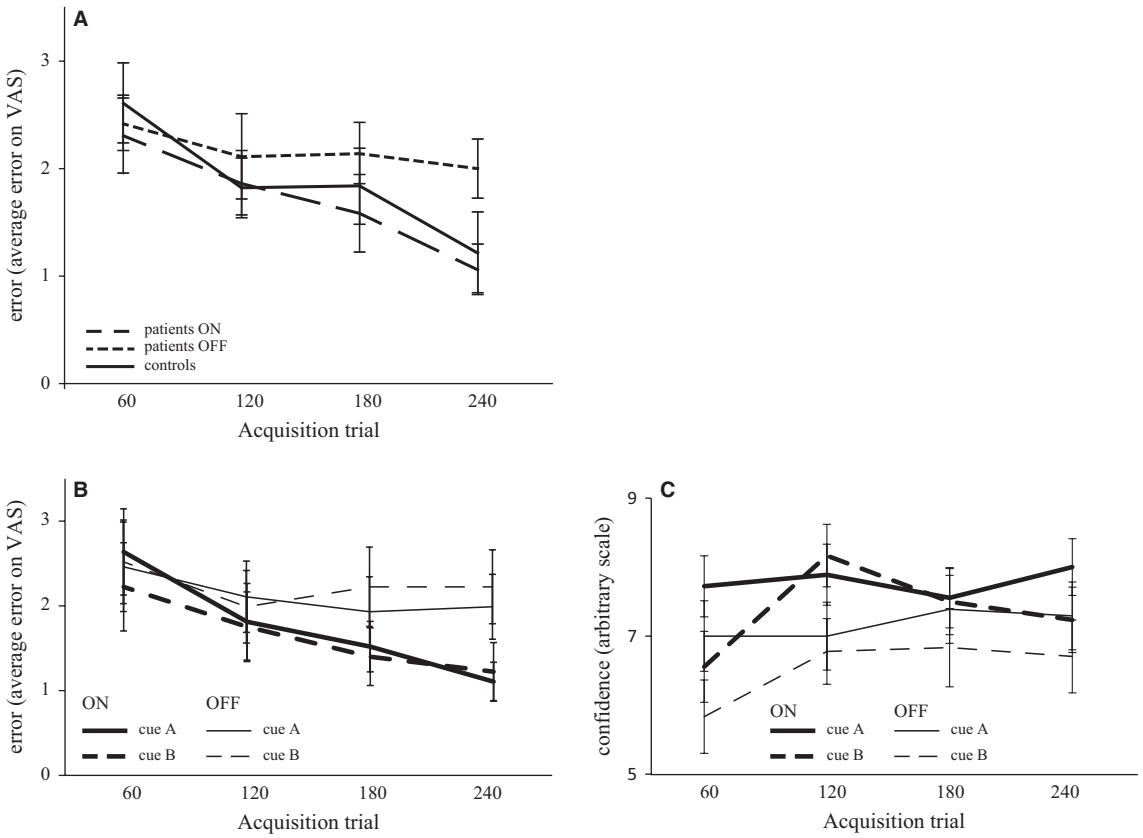

FIG. 2. Assessment of learning through a visual analogue scale questionnaire administered during the learning phase. (A) Errors on cue A and B estimates. Lower scores on the *y*-axis indicate a smaller error. Patients performed better ON than OFF medication (see text for details). Note that any difference in learning could lead to a general but not a valence-specific change in performance on the test phase. (B) Patient data presented separately as a function of medication and cue type to confirm there was no accidental cue bias in either medication condition. (C) Confidence ratings. Error bars indicate twice the standard error of the mean.

and (2) examine medication effects regardless of whether patients were tested ON or OFF their medication first.

Reaction times

One patient did not make any successful avoid-B trials when ON medication and was excluded from RT analyses involving avoid-B (Fig. 3). There was a main effect of valence on RTs ($F_{1,15} = 12.83$, $P = 0.003$) due to faster RTs on approach-A than on avoid-B trials, but no main effect of medication ($F_{1,15} = 3.20$, $P = 0.09$). A significant medication × valence interaction for RTs was found ($F_{1,15} = 5.48$, $P = 0.03$). This interaction was driven by faster responses of patients ON medication compared with OFF medication in approach-A trials ($F_{1,16} = 10.68$, $P = 0.01$). There was no such difference between avoid-B RTs ON versus OFF medication ($F_{1,15} < 1$). There was a valence × disease interaction between controls and patients ON medication ($F_{1,28} = 10.76$, $P = 0.003$), but not between controls and patients OFF medication ($F_{1,29} < 1$). Again, the significant interaction was driven by patients ON medication responding significantly faster than controls on approach-A trials ($F_{1,29} = 4.86$, $P = 0.04$), but there was no such difference on avoid-B trials ($F_{1,28} < 1$).

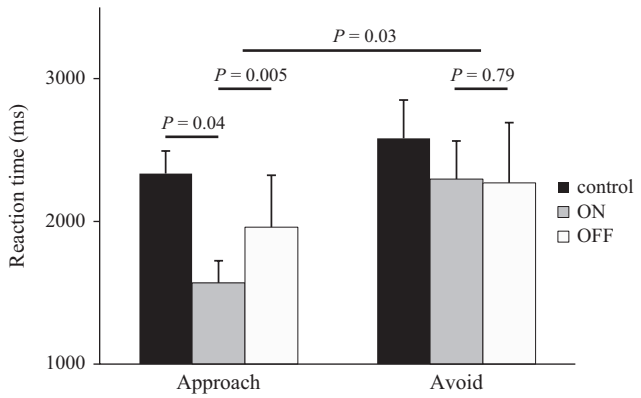

FIG. 3. Reaction times on approach-A and avoid-B trials. Errors bars indicate standard error of the mean.

Supplemental analyses were performed to correct for variability in the confidence in cue estimates after learning. To this end, the reported confidence ratings in cue estimates on the VAS were included as a covariate (see Appendix S3). The same valence-specific effect of dopaminergic medication was observed. Together, these results show that dopaminergic medication in PD patients speeds responding to obtain reward without affecting RTs on trials in which punishment must be avoided.

Accuracy

Control subjects had similar accuracy on approach-A and avoid-B ($F_{1,12} = 2.08$, $P = 0.18$) (Fig. 4). There was a main effect of valence (approach versus avoid) on accuracy in PD patients ($F_{1,16} = 5.75$, $P = 0.03$), due to better approach than avoid performance, and no main effect of medication ($F_{1,16} = 2.67$, $P = 0.12$). The crucial

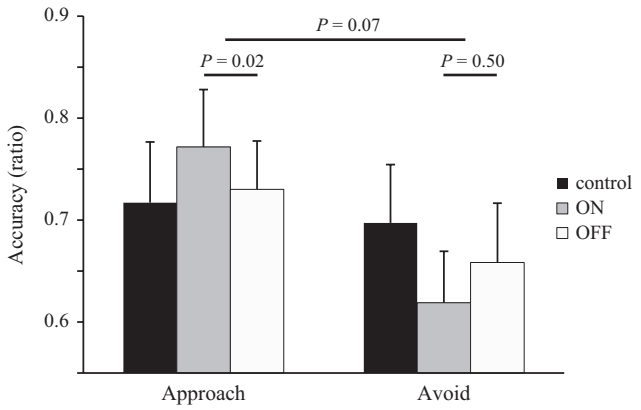

FIG. 4. Accuracy on approach-A and avoid-B trials. Error bars indicate standard error of the mean.

interaction, medication × valence, approached significance ($F_{1,16}$ = 3.78, P = 0.07). The disease × valence interaction between PD patients ON medication and controls also approached significance ($F_{1,29}$ = 3.58, P = 0.07). Simple effects analyses revealed that medication in PD patients significantly improved approach-A accuracy ($F_{1,16}$ = 7.19, P = 0.02), but had no effect on avoid-B ($F_{1,16}$ < 1).

As for RTs, supplemental analyses with confidence ratings at the end of learning inserted as a covariate confirmed the presence of a valence-specific effect of dopaminergic medication, even after correcting for variability in these confidence ratings (see Appendix S3).

Discussion

The present study was designed to investigate the hypothesis that effects of dopaminergic medication in PD on probabilistic action selection (Frank *et al.*, 2004) reflect, at least in part, modulation of value-based performance rather than modulation of reinforcement learning. Consistent with this hypothesis, dopaminergic medication in PD patients had a valence-specific effect that could not be attributed to modulation of reinforcement learning, and therefore must reflect modulation of value-based performance. Thus, dopaminergic medication decreased RTs to obtain reward-associated stimuli, whereas RTs to avoid punishment-associated stimuli were unaffected. This speeding was not associated with a speed–accuracy trade-off. On the contrary, the results indicated that dopaminergic medication not only speeded reward-driven choice but also improved the accuracy of such reward-driven choices.

Our findings generally concur with previous results, obtained using the original probabilistic selection task, showing that dopaminergic medication improved reward-based choice (Frank *et al.*, 2004). This previous result was interpreted to reflect a medication-induced bias towards learning from reward versus punishment. The learning bias was argued to transfer to the subsequent test phase, during which the learned associations were probed and the learning bias revealed. Our study indicates that these prior results may require reinterpretation. We demonstrate an effect on reward-based speeding and choice, even though subjects were informed about the reward and punishment values of the outcomes only *after* learning. This indicates that at least part of the effect of dopaminergic medication on our task, and probably on the original task, might not be due to modulation of reinforcement learning, as is commonly thought, but due to modulation of value-based performance (Frank, 2005). Specifically, dopaminergic medication altered the ability to apply instructions concerning which outcomes were rewarding to existing stimulus–outcome associations. This suggests that dopaminergic medication in PD might affect reward-based choice not only via habit-like (model-free) reinforcement learning, but also via goal-directed (model-based) control (Belin *et al.*, 2009; Bornstein & Daw, 2011; Cools, 2011; Frank, 2011; Jocham *et al.*, 2011).

This result is important because it emphasizes a role for dopamine in reward-driven behavioral control *beyond* that already established in reinforcement learning and plasticity. At the same time, it should be noted that the present finding does not imply that previously observed effects on probabilistic action selection (Frank *et al.*, 2004, 2007; Bodi *et al.*, 2009; Palminteri *et al.*, 2009) cannot be driven by a combined effect on learning plus performance, or even on learning alone. Indeed, extensive evidence from controlled manipulations of dopamine implicate it in learning (e.g. Tsai *et al.*, 2009; for a review see Maia & Frank, 2011). This hypothesis concurs with a general behavioral activational account that assigns both a performance-based energetic component as well as reinforcement-related functions to dopamine (Robbins & Everitt, 1982, 1992, 2007; Wise, 2004). Early experimental work by Gallistel (1974) highlighted both reinforcing and activational effects of (putatively dopaminergic) brain stimulation reward. Consistent with this hypothesis is our observation that effects on accuracy on our task, in which reward could influence performance but not learning, were less pronounced than those seen in previous studies that allowed reward to influence both learning and performance (Frank *et al.*, 2004, 2007; Cools *et al.*, 2006; Bodi *et al.*, 2009; Palminteri *et al.*, 2009; Voon *et al.*, 2010).

We found that valence-specific medication effects on performance were driven by beneficial effects on reward-based approach, but not by detrimental effects on punishment-based avoidance (but see Appendix S3; selective beneficial effects on approach were more robust for RTs than for accuracy). Contrasting with this, prior studies with PD patients have often reported that valence-specific effects on learning reflect both a beneficial effect on reward-based learning *and* a detrimental effect on punishment-based learning (Frank *et al.*, 2004, 2007; Moustafa *et al.*, 2008; Bodi *et al.*, 2009; Palminteri *et al.*, 2009; Voon *et al.*, 2010; but see Rutledge *et al.*, 2009). If anything, some learning studies have found greater effects on punishment-based learning than on reward-based learning (Frank *et al.*, 2004; Cools *et al.*, 2006). This apparent inconsistency with our results suggests that the distinction between learning and performance should be considered when isolating neuromodulatory effects on value-based action selection. Thus, dopaminergic medication might impair *learning* from punishment without affecting punishment-based *performance*.

Despite this divergence from previous data, our results concur with recent models, according to which dopamine specifically mediates interactions between affect and action in the appetitive domain (Niv *et al.*, 2006). In the aversive domain, interactions between affect (punishment) and action have instead been ascribed to serotonin (Crockett *et al.*, 2009; Boureau & Dayan, 2011; Cools *et al.*, 2011). If so, then punishment-based avoidance on our task might be more sensitive to central serotonin manipulations (for example, acute tryptophan depletion) than to dopamine manipulations.

Medication effects were most pronounced in terms of RTs. Although previous studies have not reported RT effects on the (original) probabilistic action selection task (Frank *et al.*, 2004, 2007), we are not the first to report dopaminergic medication effects on RTs as a function of reward. For example, Moustafa *et al.* (2008) have shown that PD patients were better at learning to speed up to maximize reward when they were medicated than when they were not

medicated. The present data extend this finding by demonstrating that this reward-based speeding by medication can reflect modulation of performance only, possibly through direct modulation of the values associated with the cues, or of the precision of these value estimates (Daw *et al.*, 2005).

It might be noted that our conclusion that the effect of dopamine on action selection reflects modulation of both learning and performance is consistent with the model by Frank (2005), where increases in dopamine potentiate Go activity (via excitatory effects of dopamine on D1 receptors), while inhibiting NoGo activity (via D2 receptors). This effect of dopamine on Go and NoGo activations during the response selection phase directly affects choice and RTs, because previously established Go associations are amplified by added D1 excitatory current, and relative differences between Go and NoGo activity determine the probability that the response is gated and the speed with which it is gated (Maia & Frank, 2011; M. J. Frank, pers. comm.).

The present finding that dopaminergic medication increases reward-based speeding and choice concurs with renewed interest in the last decade in dopamine's role in motivation, vigor and effort, and more generally behavioral activation (Berridge & Robinson, 1998; Satoh *et al.*, 2003; Robinson *et al.*, 2005; Salamone *et al.*, 2005; Niv *et al.*, 2006, 2007; Berridge, 2007). Our finding concords with data demonstrating that amphetamine potentiates behavioral control by neutral stimuli previously associated with reward in a dopamine-dependent way (Robbins *et al.*, 1989), even in the absence of reinforcement contingency (Wyvell & Berridge, 2000). The observation that this D-amphetamine-induced potentiation of behavioral control by previously rewarded stimuli is abolished by lesions of the nucleus accumbens (Parkinson *et al.*, 1999) suggests an important role for the nucleus accumbens in the aberrant potentiation of behavioral control by rewarding stimuli. The hypothesis that the nucleus accumbens mediates this effect of medication on reward-based performance also concurs with the dopamine overdose hypothesis (Cools *et al.*, 2001; Cools, 2006), which predicts that the medication-induced potentiation of control by reward is beyond normal; indeed, patients differed from controls only when they were ON medication and not when they were OFF medication.

In addition to showing an effect on reward-based performance, the present study also revealed that PD patients exhibited a dopamine-dependent deficit during the observational learning phase. During the learning phase, patients provided more accurate estimates of stimulus–stimulus associations when they were ON medication than when they were OFF medication. Caution is warranted when interpreting this effect in terms of a learning deficit, because there was no statistical evidence that medication improved the learning rate. Although inconclusive, these results are remarkable in light of previous conclusions that PD patients exhibit only a feedback-based, and not an observational learning impairment (Shohamy *et al.*, 2004). The present results suggest that an observational learning impairment may be found when the task is sufficiently difficult, perhaps reflecting indirect effects via interactions between the medial temporal lobe and the basal ganglia (Poldrack *et al.*, 2001; Wilkinson *et al.*, 2009; Sadeh *et al.*, 2010).

In conclusion, the present study disentangled distinct learning and performance components of dopaminergic medication effects on probabilistic action selection. To this end, a task was employed in which reward and punishment values were assigned to outcomes after learning. Results showed that dopaminergic medication potentiated reward- relative to punishment-based choice and performance, an effect that could not be attributed to modulation of learning. These data highlight a role for dopamine in reward-driven behavioral control *beyond* that already established in learning and plasticity.

Supporting Information

Additional supporting information may be found in the online version of this article:

Appendix S1 Instructions for the adapted probabilistic selection task.
Appendix S2 Instructions of valuation.
Appendix S3 Confidence ratings as covariate.
Fig. S1. Visual analogue scale: assessment of learning. First, subjects estimate the contingency for the specific cue (shown on top). Second, they indicate how sure they are of this estimate.
Table S1. In- and exclusion criteria for Parkinson's disease patients.
Table S2. Dopaminergic medication of all Parkinson's disease patients individually.

Please note: As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer reviewed and may be re-organised for online delivery, but are not copy-edited or typeset by Wiley-Blackwell. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

Acknowledgements

This study was supported by an NWO VIDI grant to R.C., a fellowship from the Hersenstichting to R.C., and a Human Frontiers Science Program grant to Kae Nakamura, R.C. and Nathaniel Daw. B.R.B. is supported by an NWO VIDI grant.

Abbreviations

BDI, Beck's depression inventory; DART, Dutch adult reading test; MMSE, mini-mental state examination; PD, Parkinson's disease; RT, reaction time; UPDRS, Unified Parkinson's Disease Rating Scale; VAS, visual analogue scale.

References

Barratt, E.S. (1985) Impulsiveness subtraits: Arousal and information processing. In Spence, J.T. & Itard, C.E. (Eds), *Motivation, emotion, and personality*. Amsterdam, Elsevier, pp. 137–146.

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

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

Belin, D., Jonkman, S., Dickinson, A., Robbins, T.W. & Everitt, B.J. (2009) Parallel and interactive learning processes within the basal ganglia: relevance for the understanding of addiction. *Behav. Brain Res.*, **199**, 89–102.

Berridge, K.C. (2007) The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology*, **191**, 391–431.

Berridge, K.C. & Robinson, T.E. (1998) What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res. Brain Res. Rev.*, **28**, 309–369.

Bodi, N., Keri, S., Nagy, H., Moustafa, A., Myers, C.E., Daw, N., Dibo, G., Takats, A., Bereczki, D. & Gluck, M.A. (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.

Bornstein, A.M. & Daw, N.D. (2011) Multiplicity of control in the basal ganglia: computational roles of striatal subregions. *Curr. Opin. Neurobiol.*, **21**, 374–380.

Boureau, Y.L. & Dayan, P. (2011) Opponency revisited: competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology*, **36**, 74–97.

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

Cools, R. (2011) Dopaminergic control of the striatum for high-level cognition. *Curr. Opin. Neurobiol.*, **21**, 402–407.

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. *Cereb. Cortex*, **11**, 1136–1143.

Cools, R., Altamirano, L. & D'Esposito, M. (2006) Reversal learning in Parkinson's disease depends on medication status and outcome valence. *Neuropsychologia*, **44**, 1663–1673.
Cools, R., Nakamura, K. & Daw, N.D. (2011) Serotonin and dopamine: unifying affective, activational, and decision functions. *Neuropsychopharmacology*, **36**, 98–113.
Crockett, M.J., Clark, L. & Robbins, T.W. (2009) Reconciling the role of serotonin in behavioral inhibition and aversion: acute tryptophan depletion abolishes punishment-induced inhibition in humans. *J. Neurosci.*, **29**, 11993–11999.
Daw, N.D., Niv, Y. & Dayan, P. (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.*, **8**, 1704–1711.
Dubois, B., Slachevsky, A., Litvan, I. & Pillon, B. (2000) The FAB: a Frontal Assessment Battery at bedside. *Neurology*, **55**, 1621–1626.
Fahn, S., Elton, R.L. & Members of the UPDRS Development Committee. (1987) Unified Parkinson's disease rating scale. In: Fahn, S., Marsden, C.D., Calne, D.B. & Goldstein, M. (Eds), *Recent Developments in Parkinson's Disease, Vol 2*. Macmillan Health Care Information 1987, Florham Park, NJ, pp. 153–163.
Folstein, M.F., Folstein, S.E. & McHugh, P.R. (1975) Mini-mental state A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.*, **12**, 189–198.
Frank, M.J. (2005) Dynamic Dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated parkinsonism. *J. Cogn. Neurosci.*, **17**, 51–72.
Frank, M.J. (2011) Computational models of motivated action selection in corticostriatal circuits. *Curr. Opin. Neurobiol.*, **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.
Frank, M.J., Moustafa, A., Haughey, H.M., Curran, T. & Hutchison, K.E. (2007) Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proc. Nat. Acad. Sci. USA*, **104**, 16311–16316.
Gallistel, C.R. (1974) Note on temporal summation in the reward system. *J Comp Physiol Psychol*, **87**, 870–875.
Hassler, R. (1938) Zur Pathologie der Paralysis agitans und des postenzephalitischen Parkinsonismus. *J Psychol Neurol*, **48**, 387–476.
Hollerman, J.R. & Schultz, W. (1998) Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat. Neurosci.*, **1**, 304–309.
Houk, J.C. & Wise, S.P. (1995) Distributed modular architectures linking basal ganglia, cerebellum, and cerebral cortex: their role in planning and controlling action. *Cereb. Cortex*, **5**, 95–110.
Howell, D.C. (1997). *Statistical Methods for Psychology*. Duxbury Press, Belmont, CA.
Ikemoto, S. & Panksepp, J. (1999) The role of nucleus accumbens dopamine in motivated behavior: a unifying interpretation with special reference to reward-seeking. *Brain Res. Rev.*, **31**, 6–41.
Jocham, G., Klein, T.A. & Ullsperger, M. (2011) Dopamine-mediated reinforcement learning signals in the striatum and ventromedial prefrontal cortex underlie value-based choices. *J. Neurosci.*, **31**, 1606–1613.
Ljungberg, T., Apicella, P. & Schultz, W. (1992) Responses of monkey dopamine neurons during learning of behavioral reactions. *J. Neurophysiol.*, **67**, 145–163.
Maia, T.V. & Frank, M.J. (2011) From reinforcement learning models to psychiatric and neurological disorders. *Nat. Neurosci.*, **14**, 154–162.
Montague, P.R., Dayan, P. & Sejnowski, T.J. (1996) A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.*, **16**, 1936–1947.
Montague, P.R., Hyman, S.E. & Cohen, J.D. (2004) Computational roles for dopamine in behavioural control. *Nature*, **431**, 760–767.
Moustafa, A., Cohen, M.X., Sherman, S.J. & Frank, M.J. (2008) A role for dopamine in temporal decision making and reward maximization in parkinsonism. *J. Neurosci.*, **28**, 12294–12304.
Nelson, H.E. & O'Connell, A. (1978) Dementia: the estimation of premorbid intelligence levels using the New Adult Reading Test. *Cortex*, **14**, 234–244.
Niv, Y., Joel, D. & Dayan, P. (2006) A normative perspective on motivation. *Trends Cogn. Sci.*, **10**, 375–381.
Niv, Y., Daw, N.D., Joel, D. & Dayan, P. (2007) Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology*, **191**, 507–520.
Palminteri, S., Lebreton, M., Worbe, Y., Grabli, D., Hartmann, A. & Pessiglione, M. (2009) Pharmacological modulation of subliminal learning in Parkinson's and Tourette's syndromes. *Proc. Natl. Acad. Sci. USA*, **106**, 19179–19184.
Parkinson, J.A., Olmstead, M.C., Burns, L.H., Robbins, T.W. & Everitt, B.J. (1999) Dissociation in effects of lesions of the nucleus accumbens core and shell on appetitive pavlovian approach behavior and the potentiation of conditioned reinforcement and locomotor activity by D-amphetamine. *J. Neurosci.*, **19**, 2401–2411.
Paton, J.J., Belova, M.A., Morrison, S.E. & Salzman, C.D. (2006) The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature*, **439**, 865–870.
Poldrack, R.A., Clark, J., Paré-Blagoev, E.J., Shohamy, D., Myers, C. & Gluck, M.A. (2001) Interactive memory systems in the human brain. *Nature*, **414**, 546–550.
Robbins, T.W. & Everitt, B.J. (1982) Functional studies of the central catecholamines. *Int. Rev. Neurobiol.*, **23**, 303–365.
Robbins, T.W. & Everitt, B.J. (1992) Functions of dopamine in the dorsal and ventral striatum. *Semin. Neurosci.*, **4**, 119–127.
Robbins, T.W. & Everitt, B.J. (2007) A role for mesencephalic dopamine in activation: commentary on Berridge (2006). *Psychopharmacology*, **191**, 433–437.
Robbins, T.W., Cador, M., Taylor, J.R. & Everitt, B.J. (1989) Limbic-striatal interactions in reward-related processes. *Neurosci. Biobehav. Rev.*, **13**, 155–162.
Robinson, S., Sandstrom, S.M., Denenberg, V.H. & Palmiter, R.D. (2005) Distinguishing whether dopamine regulates liking, wanting, and/or learning about rewards. *Behav. Neurosci.*, **119**, 5–15.
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. *J. Neurosci.*, **29**, 15104–15114.
Sadeh, T., Shohamy, D., Levy, D.R., Reggev, N. & Maril, A. (2011) Cooperation between the hippocampus and the striatum during episodic encoding. *J. Cogn. Neurosci.*, **23**, 1597–1608.
Salamone, J.D., Correa, M., Mingote, S.M. & Weber, S.M. (2005) Beyond the reward hypothesis: alternative functions of nucleus accumbens dopamine. *Curr. Opin. Pharmacol.*, **5**, 34–41.
Salamone, J.D., Correa, M., Farrar, A. & Mingote, S.M. (2007) Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology*, **191**, 461–482.
Salthouse, T.A. (1994) The nature of the influence of speed on adult age differences in cognition. *Dev. Psychol.*, **30**, 240–259.
Samejima, K., Ueda, Y., Doya, K. & Kimura, M. (2005) Representation of action-specific reward values in the striatum. *Science*, **310**, 1337–1340.
Satoh, T., Nakai, S., Sato, T. & Kimura, M. (2003) Correlated coding of motivation and outcome of decision by dopamine neurons. *J. Neurosci.*, **23**, 9913–9923.
Schmand, B., Bakker, D., Saan, R. & Louman, J. (1991) [The Dutch Reading Test for Adults: a measure of premorbid intelligence level]. *Tijdschr. Gerontol. Geriatr.*, **22**, 15–19.
Schultz, W., Dayan, P. & Montague, P.R. (1997) A neural substrate of prediction and reward. *Science*, **275**, 1593–1599.
Shohamy, D., Myers, C.E., Grossman, S., Sage, J., Gluck, M.A. & Poldrack, R.A. (2004) Cortico-striatal contributions to feedback-based learning: converging data from neuroimaging and neuropsychology. *Brain*, **127**, 851–859.
Sutton, R.S. & Barto, A.G. (1998). *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA.
Tsai, H.C., Zhang, F., Adamantidis, A., Stuber, G.D., Bonci, A., de Lecea, L. & Deisseroth, K. (2009) Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science*, **324**, 1080–1084.
Voon, V., Pessiglione, M., Brezing, C., Gallea, C., Fernandez, H., Dolan, R.J. & Hallett, M. (2010) Mechanisms underlying dopamine-mediated reward bias in compulsive behaviors. *Neuron*, **65**, 135–142.
Waelti, P., Dickinson, A. & Schultz, W. (2001) Dopamine responses comply with basic assumptions of formal learning theory. *Nature*, **412**, 43–48.
Wenzelburger, R., Zhang, B., Pohle, S., Klebe, S., Lorenz, D., Herzog, J., Wilms, H., Deuschl, G. & Krack, P. (2002) Force overflow and levodopa-induced dyskinesias in Parkinson's disease. *Brain*, **125**, 871–879.
Wilkinson, L., Khan, Z. & Jahanshahi, M. (2009) The role of the basal ganglia and its cortical connections in sequence learning: evidence from implicit and explicit sequence learning in Parkinson's disease. *Neuropsychologia*, **47**, 2564–2573.
Wise, R.A. (2004) Dopamine, learning and motivation. *Nat. Rev. Neurosci.*, **5**, 483–494.
Wyvell, C.L. & Berridge, K.C. (2000) Intra-accumbens amphetamine increases the conditioned incentive salience of sucrose reward: enhancement of reward ''wanting'' without enhanced ''liking'' or response reinforcement. *J. Neurosci.*, **20**, 8122–8130.