Dopamine modulates striatal response to reward and punishment in patients with Parkinson's disease: a pharmacological challenge fMRI study

Miklos Argyelan[a,b], Mohammad Herzallah[c,d], Wataru Sako[a], Ivana DeLucia[a], Deepak Sarpal[b], An Vo[a], Toni Fitzpatrick[a], Ahmed A. Moustafa[c,e], David Eidelberg[a] and Mark Gluck[c]

It is well established that Parkinson's disease leads to impaired learning from reward and enhanced learning from punishment. The administration of dopaminergic medications reverses this learning pattern. However, few studies have investigated the neural underpinnings of these cognitive processes. In this study, using fMRI, we tested a group of Parkinson's disease patients on and off dopaminergic medications and matched healthy individuals. All individuals completed an fMRI cognitive task that dissociates feedback learning from reward versus punishment. The administration of dopaminergic medications attenuated blood oxygen level dependent (BOLD) responses to punishment in the bilateral putamen, in bilateral dorsolateral prefrontal cortex and the left premotor cortex. Further, the administration of dopaminergic medications resulted in a higher ratio of BOLD activity between reward and punishment trials in these brain areas. BOLD activity in these brain areas was significantly correlated with learning from punishment, but not from reward trials. Furthermore, the administration of dopaminergic medications altered BOLD activity in the right insula and ventromedial prefrontal cortex when Parkinson's disease patients were anticipating feedback. These findings are in agreement with a large body of literature indicating that Parkinson's disease is associated with enhanced learning from punishment. However, it was surprising that dopaminergic medications modulated punishment learning as opposed to reward learning, although reward learning has been directly linked to dopaminergic function. We argue that these results might be attributed to both a change in the balance between direct and indirect pathway activation in the basal ganglia as well as the differential activity of D1 versus D2 dopamine receptors. *NeuroReport* 00:000–000 Copyright © 2018 Wolters Kluwer Health, Inc. All rights reserved.

Introduction

Converging evidence suggests that Parkinson's disease impairs sensitivity to rewards and enhances learning based on punishment feedback, whereas antiparkinsonian dopaminergic agents reverse this learning pattern [1,2]. This is thought to stem from the imbalance between direct and indirect basal ganglia pathways in Parkinson's disease, which is reversed by dopaminergic medications. Various studies have suggested a functional segregation between the direct and indirect pathways in the striatum, implicating the direct pathway in reward processing and the indirect pathway in punishment learning [1,3–7]. The direct and indirect pathways reflect a dual organization of the coritco-striato-thalamo-cortico loops, with distinct connectivity patterns and physiological properties of the striatum and basal ganglia. Dopamine stimulates the direct pathway through D1 receptors that initiates a ('Go') signal, whereas it inhibits indirect pathways through D2 receptors that results in a ('NoGo') signal. Therefore, the lack of dopamine, as in Parkinson's disease, may increase indirect pathway activity relative to direct pathway processing, leading to the observed imbalance in punishment-based and reward-based learning. Conversely, too much dopamine might cause an overly activated direct pathway resulting in higher sensitivity to reward as compared with punishment [3].

To date, the neural underpinnings of the effects of Parkinson's disease versus those of dopaminergic medications on learning have not been sufficiently studied using neuroimaging. Electrophysiological studies suggest that

dopamine neuronal firing increases in response to unexpected reward and dips after the omission of reward [6,8]. Further, dopaminergic neurons code for motivational salience during both reward and punishment [6,9,10]. Previous studies in healthy volunteers have reported striatal activation during learning from both reward and punishment feedback [11]. However, it was unclear how striatal activity during reward and punishment feedback is related to dopaminergic modulation.

In this study, we aimed to investigate the relationship between neural activation during learning from reward and punishment in Parkinson's disease patients before and during treatment with dopaminergic medications. We utilized a computer-based cognitive task, developed by Gluck and colleagues at Rutgers University-Newark, and similar to that used by Bodi et al. [2], to dissociate learning from reward and punishment. Our main hypothesis was that the dopamine-depleted striatum in medication-withdrawn Parkinson's disease patients will have increased activation in response to punishment than to reward, whereas the administration of dopaminergic medication would reverse this relationship: neural responses would be higher with reward than with punishment.

Participants and methods
Participants
We tested 11 patients with Parkinson's disease (four female, age: 63.1 ± 10.3, age range: 45–75 years) and 11 healthy volunteers (age: 57.6 ± 8.5, age range: 43–72 years; no significant age difference between groups: $t = 1.37$, $d.f. = 20$, $P = 0.19$). All Parkinson's disease patients were tested under two conditions; a dopamine OFF and ON condition, in a within-patient design. Of the 11 Parkinson's disease patients, six were on levodopa only, whereas the other five patients were on a combination of levodopa and dopamine agonist (pramipexole). The average dosage of levodopa and pramipexole were 542 ± 172 and 2.8 ± 0.7 mg, respectively (altogether levodopa equivalent dose: 822 ± 242 mg). As expected, treated patients (ON) had better Unified Parkinson's Disease Rating Scale (UPDRS) scores. For the OFF state, participating Parkinson's disease patients were withdrawn from their medications (levodopa or pramipexole or both, respectively) for a minimum of 18 h before testing. Patients experienced severe worsening in their symptoms on OFF days reflected in their UPDRS (OFF UPDRS: 18.6 ± 7.2; ON UPDRS: 10.8 ± 6.1, paired-samples t-test, $P < 0.001$). The average disease duration was 5.5 ± 2.8 years, and average H&Y was 1.9 ± 0.6.

All participants submitted written informed consent that was previously approved by the International Review Board of Feinstein Institute for Medical Research, Manhasset, New York, USA. Participants were recruited from the Long Island Jewish Health System, New York, USA.

Experimental task
We used a computer-based cognitive task that dissociates learning from reward and punishment. The task that was developed by Gluck and colleagues at Rutgers University, Newark was previously used in fMRI [12] and cognitive studies of Parkinson's disease patients [2]. We modified the cover story to simulate real life experiences. Further, we increased the duration of stimulus presentation. In each trial, participants were presented with a casino roulette wheel (with a different background image) and were instructed that these roulette wheels were either biased to the red or the black side of the screen as shown in Fig. 1. When the roulette wheel appeared on the screen, patients were asked to put their tokens on red or black by pressing MRI compatible buttons with their thumb or index finger, respectively. Patients were told that based on where the roulette ball landed, they would win money, lose money, or get an empty box. The task did not include any further details about the biases of the four different roulette wheels. Therefore, patients were required to learn the associations by trial-and-error. Each participant started the experiment with 500 tokens. Patients were instructed to earn as many tokens as possible.

Each trial began with the presentation of one of four roulette wheels with their respective background images and a casino table with instructions to place bets. The roulette wheel (cue) was presented for a fixed period of 3 s and was followed by a fixation cross for 8–10 s. Feedback was then presented for 3 s followed by a fixation cross for 10–12 s. Thus, an average trial was 26 s long. Feedback was either a reward with a smiley face indicating winning 25 tokens, a punishment with a sad face indicating losing 25 tokens, or an empty box indicating a neutral outcome (no feedback).

Two of the roulette wheels were biased toward red, and the other two were biased toward black. Two roulette wheels (one red-biased and one black-biased) were preassigned to result in reward if answered optimally and no feedback if answered nonoptimally, whereas the other two give punishment if answered nonoptimally and no feedback if answered optimally. In reward trials, when participants selected an optimal choice, they received reward feedback with 80% probability. They received no feedback for the remaining 20% of the trials when they selected the optimal choice. However, if the participant selected a nonoptimal choice, they would receive reward feedback with 20% probability, but no feedback for the remaining 80% of the trials. Similarly, on the punishment trials, when participants selected the optimal choice, they received no feedback 80% of trials and punishment in the remaining 20% of trials and vice versa.
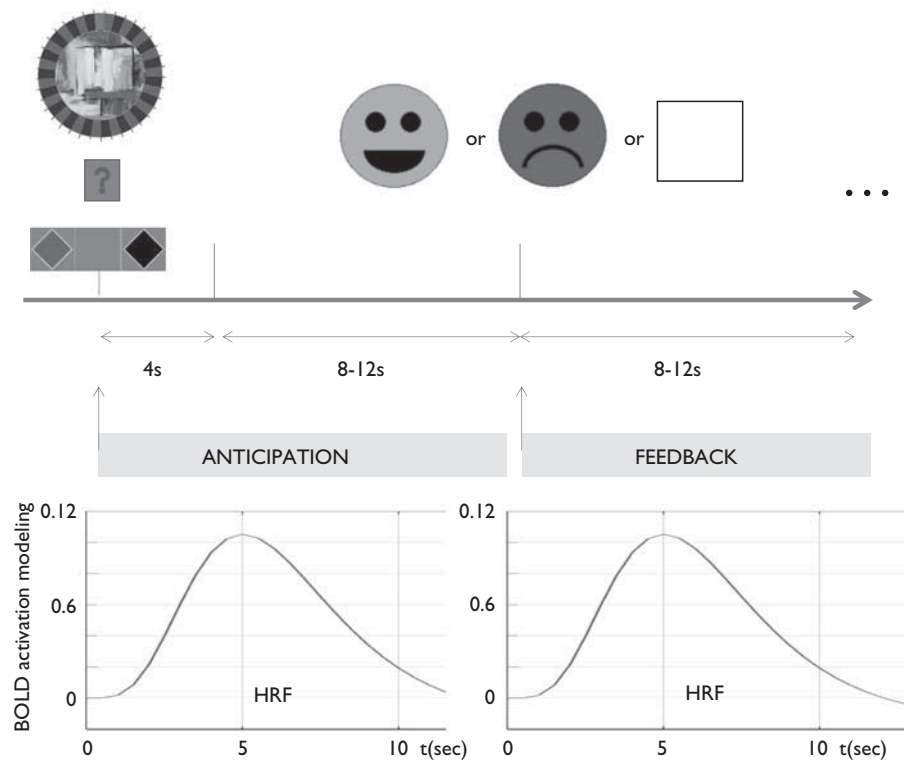
Fig. 1

Illustration of the reward and punishment learning task. First, there is the presentation of the cue. The participant has 4 s to choose between two options on the roulette wheel (red or black; light gray or black in the figure). Feedback is shown after 8–12 s (randomized) of waiting period. Similarly, feedback is followed by 8–12 s long period. These longer periods prevent hemodynamic response functions (HRFs) to interfere across trials and provide a framework to model anticipation and feedback separately. BOLD, blood oxygen level dependent.

All participants completed four scanning runs per session except one Parkinson's disease patient who had only two runs in the OFF condition. The order of stimulus presentation was randomly determined. A run consisted of 20 trials, with each run lasting 8 min and 40 s. We calculated optimal responses and reaction time (RT) in each run for both reward and punishment. We also measured learning scores, which was the difference between the average of the optimal responses in the third and fourth runs from the average optimal responses in the first run. Using the same approach, we also calculated RT for reward and punishment trials.

MRI acquisition

MRI scans were conducted at the North Shore University Hospital on a single 3T scanner (GE Signa HDx; General Electric, Milwaukee, Wisconsin, USA). For image registration, we acquired anatomical scans in the coronal plane using an inversion-recovery prepared three-dimensional fast spoiled gradient sequence (repetition time = 7.5 ms, echo time = 3 ms, inversion time = 650 ms, matrix = 256 × 256, field of view = 240 mm) producing 216 contiguous images (slice thickness = 1 mm) through the whole brain. All scans were reviewed by a radiologist and a member of the research team. Any scan with significant artifacts was repeated. We also acquired functional scans comprising of 264 echo-planner imaging (EPI) volumes per run with the following parameters: repetition time = 2000 ms, echo time = 30 ms, matrix = 64 × 64, FOV = 240 mm, slice thickness = 3 mm, 40 continuous axial oblique slices (one voxel = 3.75 × 3.75 × 3 mm). During data acquisition, the behavioral task was presented on an MR compatible screen.

Image processing and statistical analysis

We used FSL (*http://www.fmrib.ox.ac.uk*) FEAT for preprocessing and for the statistical analysis of our images. First level analysis consisted of standard preprocessing, modeling, and the calculation of the transformation matrix of the registration. Standard preprocessing and modeling included removal of the first four 'dummy' scans, motion correction, and spatial smoothing (6-mm full width at half maximum Gaussian kernel) followed by general linear model (GLM) estimation of parameters associated with the design matrix variables. We modeled the hemodynamic response functions (HRFs) in two different ways. First, we modeled the HRF associated with the feedback types (reward, punishment, and no

feedback). Second, we analyzed our data in the anticipation phase, modeling our HRF locked to the different type of cues (reward or punishment). In both type of analysis (feedback or anticipation) a custom boxcar function associated with the 3 s presentation of feedback and cue was convoluted with standard HRF. In addition to these variables, we added to the GLM the first time derivative of these variables and the six movement parameters to regress out large effects of movement. Further, standard registration and normalization to MNI152 space were also carried out, and registration matrices were calculated for the second level analysis. During the first level analysis, we also contrasted punishment versus reward in both feedback and anticipation. This was followed by a group level analysis – multiway analysis of variance, with factors of feedback type (reward or punishment), medication condition (ON, OFF), repetitions (four runs). Post-hoc calculations were carried out with featquery and lab-based R scripts. All the reported results were cluster corrected for the whole brain with P less than 0.05 (with standard FSL $z = 2.3$ initial threshold). No mask was used at any point during the analysis.

Results
Behavioral results
Parkinson's disease patients were unable to learn to categorize the roulette wheels correctly. Learning curves did not increase significantly in any of the treatment conditions ($F_{1,164} = 0.3$, $P = 0.56$) with the average correct response of 61.5 ± 25.3 and $51.8 \pm 23.9\%$ across trial types for OFF and ON conditions, respectively. Learning score did not show differences between reward and punishment trials in any of the treatment conditions (condition: $F_{1,38} = 1.68$, $P = 0.2$; trial type: $F_{1,38} = 1.76$, $P = 0.19$, interaction: $F_{1,38} = 0.13$, $P = 0.72$). However, Parkinson's disease patients' RT suggested that they learned whether a roulette wheel represents a reward or a punishment cue. Patients responded to reward cues progressively faster during the course of the runs, in both ON and OFF condition (average RT changes: reward OFF and ON: -370 and -302 ms, punishment OFF and ON: 53 and 54 ms, respectively, repeated analysis of variance, condition: $F_{1,38} = 0.06$, $P = 0.81$; trial type: $F_{1,38} = 6.3$, $P = 0.02$, interaction: $F_{1,38} = 0.05$, $P = 0.82$; Fig. 2). This significant difference indicates that Parkinson's disease patients differentiated trials on the basis of their associated feedback. The lack of difference in RTs in general (1914 ± 487 and 1911 ± 466 ms, OFF and ON, respectively) between treatment conditions suggests that our behavioral and neuroimaging results are not merely the result of Parkinson's disease motor symptoms and dopaminergic medications.

Age-matched healthy volunteers showed very similar results: $61.5 \pm 25.4\%$ correct response across both trial types. Healthy controls showed moderate improvement

in learning scores only in punishment, but not reward, trials across runs [8.1% ($t = 1.56$, $d.f. = 10$, $P = 0.15$) and 12.3% ($t = 2.46$, $d.f. = 10$, $P = 0.03$) in reward and punishment trials, respectively]. RTs were similar to those of the Parkinson's disease cohort (RT $= 1515 \pm 359$ ms) and became progressively faster for reward trials as compared with punishment trials (reward: -204 ms, punishment: $+1$ ms; $F_{1,19} = 6.19$, $P = 0.02$; Supplementary Fig. S1, Supplemental digital content 1, *http://links.lww.com/WNR/A454*).

Feedback related BOLD activations
Feedback, in general, strongly activated the dorsal attention network, the salience network, and the occipital areas across trial types (Supplementary Fig. S2, regions listed in Table S1, cluster corrected $P < 0.05$, Supplemental digital content 1, *http://links.lww.com/WNR/A454*). There was no deactivation at this same level of significance of the main effect.

When we compared the effect of medication status without differentiating trial types, no effect was detected. Although BOLD activity was not modulated by medication status in reward trials, we found that medication status significantly attenuated response to punishment trials (OFF > ON) in several brain regions [left and right dorsolateral prefrontal cortex (DLPFC), left and right putamen and left premotor cortex; cluster corrected $P < 0.05$, Table 1]. When we explicitly tested whole brain voxel level interaction between feedback type (reward vs. punishment) and medication status (OFF vs. ON) the same brain regions remained significantly engaged (Fig. 3a, cluster corrected $P < 0.05$).

Post-hoc analysis of these regions revealed the same pattern across all the regions (condition: $F_{1,792} = 11.5$, $P < 0.001$, feedback: $F_{1,792} = 1.1$, $P = 0.3$, region: $F_{1,792} = 1.9$, $P = 0.17$, feedback-condition interaction: $F_{1,792} = 39.0$, $P < 0.001$): significantly higher BOLD signal during punishment trials than in the reward trials in the OFF condition. Activation during punishment trials significantly decreased in ON condition (Fig. 3b). Side by side comparison of the activations of the healthy controls and Parkinson's disease patients in different treatment conditions indicated that activation patterns in healthy controls looked similar to OFF conditions in all regions (Supplementary Fig. S3, Supplemental digital content 1, *http://links.lww.com/WNR/A454*, Fig. 3b).

By correlating BOLD responses with behavioral results, we found that the average difference in activation between punishment and reward trials correlated with learning scores. This was true for all five regions that were activated (left and right DLPFC, left and right putamen and left premotor cortex). The higher the BOLD activity in these brain regions during punishment trials in reference to reward trials the better the individual learned to categorize punishment cues. Post-hoc analyses

Fig. 2
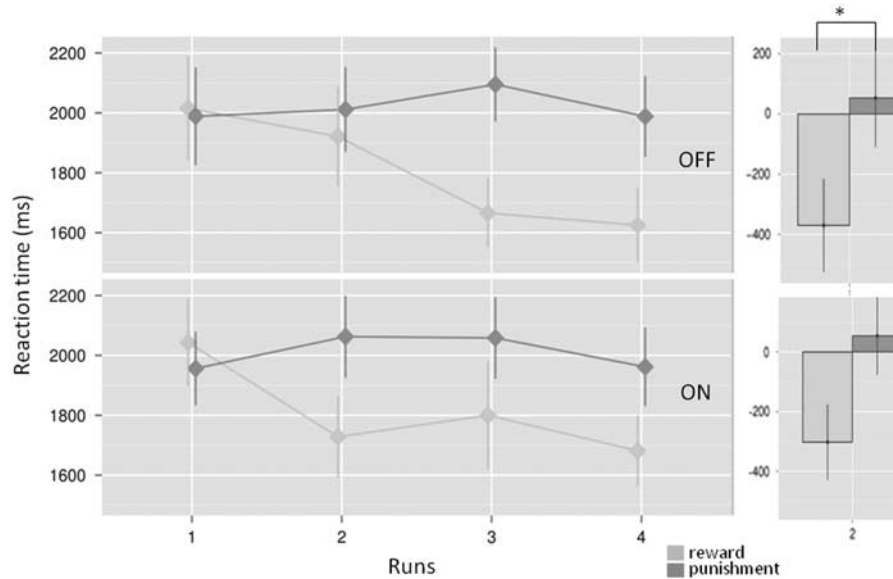
Reaction time change. The figure illustrates that the reaction times decrease selectively in reward cue-related trials. The chart indicates that participant learned to distinguish between the cues (reward or punishment). There was no difference across treatment conditions. *$P < 0.05$.

Table 1 Regions modulated by dopamine administration

Regions	x	y	z	Peak z
Feedback modulated by dopamine				
Premotor cortex, BA 6 L	−32	−2	52	3.7
Putamen R	20	4	14	3.4
Putamen L	−22	12	4	3.4
DLPFC R	50	43	21	3.4
DLPFC L	−43	43	22	3.2
Anticipation modulated by dopamine				
vmPFC, BA 11m and 14r	4	39	−18	4.8
Insula R	38	20	2	3.9

x,y,z, MNI coordinates; peak coordinates reported $z > 2.3$ and cluster corrected $P < 0.05$.
DLPFC, dorso-lateral prefrontal cortex; L, left; R, right; vmPFC, ventro-medial prefrontal cortex.

showed that activity in these brain regions showed no connection to learning to categorize reward cues (GLM: punishment–reward BOLD signal: $F_{1,60} = 15.8$, $P < 0.001$, feedback type: $F_{1,60} = 9.6$, $P = 0.003$; medication: $F_{1,60} = 3.3$, $P = 0.08$, and region: $F_{4,60} = 0.16$, $P = 0.96$ (Fig. 4, only first two areas showed here, for all the values see Supplementary Table S2, Supplemental digital content 1, *http://links.lww.com/WNR/A454*).

Anticipation: cue-related BOLD activity
The main effect of anticipation (Supplementary Fig. S4, regions listed in Table S1, Supplemental digital content 1, *http://links.lww.com/WNR/A454*) involved similar areas as the main effect of feedback (Supplementary Fig. S2, Supplemental digital content 1, *http://links.lww.com/WNR/A454*). The dorsal attention network, the salience network, and the occipital areas were similarly activated

across trial types and medication conditions after the presentation of cues. BOLD activity in the anticipation phase was more robust than it was after feedback, as it is apparent from comparing Supplementary Figs S2 and S4 (Supplemental digital content 1, *http://links.lww.com/WNR/A454*) on the same scale. Anticipation, unlike feedback, elicited significant 'negative activation' in the ventromedial prefrontal cortex (vmPFC) (area 11m and 14r [13]) and the parieto-temporal juncture (Supplementary Fig. S4, Supplementary digital content 1, *http://links.lww.com/WNR/A454*). Although the interpretation of the negative peaks is often problematic, these regions likely represent true deactivations given the spatial overlap with the default mode network. Voxel wised paired *t*-test between OFF and ON condition revealed that vmPFC/gyrus rectus had significantly lower BOLD activity during the ON condition in both cue types (Fig. 5a, Table 1). Moreover, the right anterior insula was significantly activated in the ON condition (Fig. 5a and b, Table 1). There was no interaction between medication condition and cue types. Anticipatory activity did not correlate with the behavioral findings. In the right insula, the BOLD activation during OFF condition showed closer resemblance with healthy controls (Supplementary Fig. S5, Supplemental digital content 1, *http://links.lww.com/WNR/A454*; Fig. 5b). This was similar to what we reported above regarding the brain regions activated by feedback. However, the pattern in the vmPFC/gyrus rectus was similar between healthy volunteers and ON condition (Supplementary Fig. S5,

Fig. 3
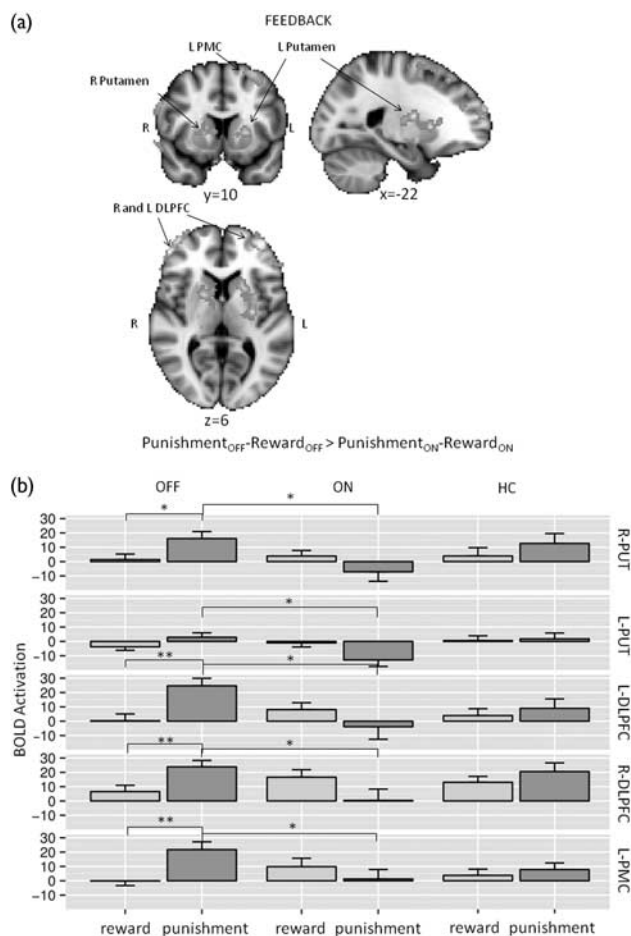
(a) Neuroimaging results: interaction between trial types and treatment condition. (b) Post-hoc presentation of the significant regions of the interaction analysis in (a). The first two columns are OFF and ON conditions in Parkinson's disease, whereas the last column represents healthy control (HC). Each row represents the average values of the respective regions [right and left putamen (R-PUT and L-PUT), right and left dorsolateral prefrontal cortex (R-DLPFC and L-DLPFC), left premotor cortex (L-PMC)] from the primary analysis. The most significant differences were found between OFF and ON in the punishment trials. *$P < 0.05$, **$P < 0.01$.

Supplemental digital content 1, *http://links.lww.com/WNR/A454*; Fig. 5b).
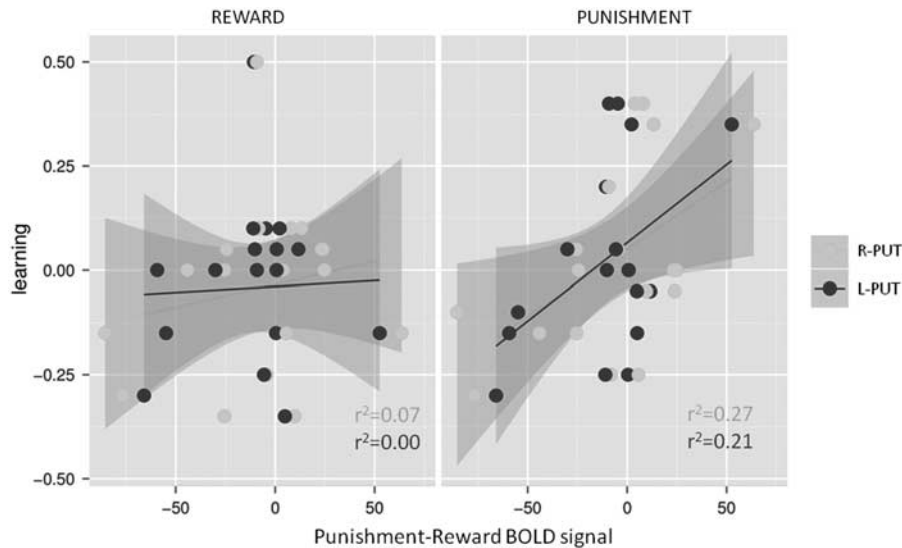
Discussion

This study demonstrated that the administration of dopaminergic medications to Parkinson's disease patients was associated with modulation of the neural response to both feedback and anticipation. First, dopaminergic medications attenuated the neural response to punishment feedback in the left and right putamen, the left and right DLPFC, and the left premotor cortex. The same regions exhibited a higher ratio of activity between reward and punishment during the ON condition as compared with the OFF condition. Neural activity

correlated with the behavioral output in punishment trials only, implicating the role of these regions in processing punishment feedback. Second, we detected enhanced right insula activation and attenuated vmPFC/gyrus rectus activation during the anticipation phase in the ON-medication condition. This effect was similar for both reward and punishment trials.

These data are consistent with previous behavioral studies that found opposite patterns of sensitivity to reward and punishment between dopamine depleted (OFF) and dopamine supplemented (ON) conditions in Parkinson's disease [1,2]. In agreement with these previous results, we have found higher BOLD activity in response to punishment as compared with reward during the OFF condition. Conversely, during the ON condition, punishment-related BOLD activity was significantly decreased, whereas reward-related activity remained unaffected, such that the ratio of activity between reward and punishment was higher during the ON condition as compared with the OFF condition. Our analysis revealed that this pattern of BOLD activity was not only limited to the dorsal striatum but was also observed in the DLPFC bilaterally and in the left premotor cortex. These findings indicate that the observed BOLD activity reflects modulation along the fronto-striatal circuits, not only within the striatum. The feedback-related pattern of BOLD activity correlated strongly among the observed regions (putamen, DLPFC, and premotor area) indicating that these regions operate within the same neural network to execute this behavior. These observations are also congruent with the proposed model of functional segregation between direct and indirect pathways within the basal ganglia, attributing changes in BOLD activity during the ON versus OFF conditions to differential engagement of the indirect pathway [1,3–7]. An alternative explanation is offered by Keeler *et al*. [14] who argue that these cognitive differences can be explained by the activation of D1 versus D2 receptors in the direct and indirect pathways, respectively, where the direct pathway codes for habitual responses, whereas indirect pathway codes for goal-directed behavior. This theory predicts that shortening of reaction time is an indicator that task is becoming more habitual, thus mediated by higher D1 receptor activation. In our behavioral results, RTs became shorter only during reward-related trials, during both OFF and ON conditions, indicating that both of these theories are potentially consistent with our results. Most critically, according to both theories, our data suggest that the administration of dopaminergic medications interfered mainly with the indirect D2 pathway, not with the direct D1 pathway. These results are surprising as levodopa is thought to change phasic dopamine signaling and enhance activity in the direct D1 pathway. D2 receptors, on the other hand, are thought to be driven by tonic dopamine levels [15]. Previous computational modeling of phasic and tonic

Fig. 4

REWARD   PUNISHMENT

$r^2=0.07$
$r^2=0.00$

$r^2=0.27$
$r^2=0.21$

R-PUT
L-PUT

learning

Punishment-Reward BOLD signal

Neuroimaging findings correlated with learning. Average activation differences during the feedback period between punishment and reward trials predicted learning from punishment, but not from reward. L-PUT, left putamen; R-PUT, right putamen.

characteristics of dopamine release in the dorsal striatum have suggested, however, that the relationship between D1 and D2 receptor pathways is much more complex than initially thought [16], and D2 receptor postsynaptic signaling can be significantly activated through phasic dopamine modulation.

Inferring a dopamine-related mechanism-based BOLD signals can be problematic as BOLD responses are not a direct measure of dopamine activation. Recent evidence, however, suggests that dopamine release can increase BOLD activity in the dorsal striatum in a dose dependent manner [17]. Our healthy control data provide additional insights into the interpretation of the BOLD signal by establishing a baseline for acquired BOLD signal from Parkinson's disease patients. The decaying nature of dopamine signals during cognitive tasks can be explained by phasic dopamine prediction error signaling in response to unexpected feedback in the first runs. This suggests that feedback-related BOLD activity during the first half of the session (run 1 and 2) might be attributed to variations in dopaminergic signaling (Supplementary Fig. S4, Supplemental digital content 1, *http://links.lww.com/WNR/A454*). The decrease in BOLD activity in response to reward versus punishment feedback further strengthens the case that the measured BOLD signals are related to reward prediction error [8]. Because of the nature of our task, reward trials elicited high reward prediction error initially. As patients learned that a cue was rewarding (manifested in decreasing reaction time), reward prediction error decreased. However, irrespective of the patient's knowledge of the punishment-related cues, it always elicited high punishment prediction error

during punishment trials indicating that processing of punishment feedback is not a mirror image of reward feedback, but could have different neural underpinnings [3,4]. However, it is important to note that although punishment prediction error signaling can decrease dopaminergic activation, it can manifest as higher BOLD activity because of active disinhibition of the indirect pathway.

In contrast to the feedback-related BOLD signals, the anticipation BOLD signals can be explained by dopaminergic signaling associated with the salience of stimuli given the lack of change in these signals over time (Supplementary Fig. S5, Supplemental digital content 1, *http://links.lww.com/WNR/A454*). Further, the anticipatory BOLD signals were mainly observed in salience-related brain regions [18]. It is well established that salient stimuli elicit BOLD activity in the anterior insula as a part of the salience network, whereas the vmPFC tends to deactivate as a part of default mode network. According to this model, administration of dopaminergic medications made participants more sensitive to external stimuli. We acknowledge, however, that our task was not designed to dissociate motivational salience-related and motivational value-related BOLD activations [6]. Hence, the BOLD signals could be driven by both salience and value.

In contrast to other studies by Gluck and colleagues using variations on this task, patients in our study were unable to learn the optimal cue-outcome associations. This was likely because of the current task design where cue and feedback were separated by 8–12 s, in contrast to previous studies where feedback was directly displayed

Fig. 5
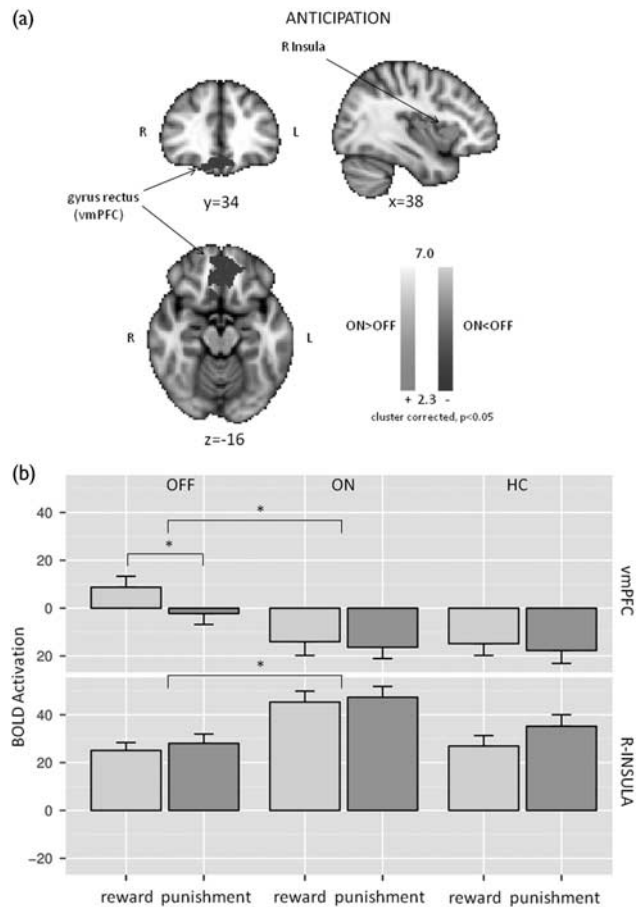
(a) Neuroimaging results: the main effect of treatment condition during the anticipation of the feedback. Right insula (R-insula) had increased activation during ON condition, whereas left ventro-medial prefrontal cortex (vmPFC) had lower activation during ON condition. (b) Post-hoc presentation of the significant regions of the treatment condition analysis in (a). The first two columns are OFF and ON conditions in Parkinson's disease, whereas the last column represents healthy control (HC). Each row represents the average values of the respective regions (vmPFC, R-insula) from the primary analysis. The most significant differences were found between OFF and ON conditions. *P < 0.05.

after the patient selected a choice [2,12]. The longer interstimuli interval was chosen to make a clearer distinction between feedback-related and anticipation-related BOLD signals. However, this design made processing cue–feedback associations much more difficult. Also, because of the increased interstimuli intervals, we could only deliver half as many trials as in the previous studies. Although patients did not learn the optimal associations, the trial type (reward or punishment) was increasingly identified across the runs as implicated by the decreasing reaction time after reward-only associated cues. Thus, these RT changes verify that patients were learning and performing reinforcement learning.

Another likely consequence of the current task design was that although a strong interaction between dopaminergic

medication and the neural response to punishment was evident, the effect on reward learning was unclear. Furthermore, as the anticipation phase (part of the current design) did not distinguish between reward-related and punishment-related BOLD signals, future work might benefit from examining the neural correlates of motivational bias with the original task design [2]. Using high resolution fMRI, Mattfeld *et al.* [12] found that in healthy young participants, unique subregions of the striatum, separated along both a dorsal/ventral and anterior/posterior axis, differentially participate in the learning of associations through reward and punishment. This double dissociation within the striatum can help differentiate between learning from reward versus punishment and how variability in dopaminergic signaling individually modulates them.

Understanding the role of dopamine in reward and punishment processing could provide valuable insight into the understanding of impulse control disorders, amotivational syndromes, and major depressive disorder. In impulse control disorders, patients underestimate risks and overestimate potential rewards possibly because of the imbalance in learning from rewarding and punishing stimuli [19]. In sharp contrast, depressed patients are thought to be overly sensitive to punishment as compared with reward [20–22].

Conclusion
We demonstrated the modulation of the neural encoding of feedbacks across trial types (reward or punishment) and treatment condition (OFF and ON) involving the fronto-striatal circuitry (DLPFC, premotor, and putamen). These results indicate that during the ON condition there is a relative hypersensitivity toward rewards, which reverses during the OFF condition. This flip has the potential to make patients vulnerable to cognitive distortions, which could be the underlying neural mechanism through which dopamine depleted, and dopamine enhanced conditions can lead to behavioral manifestations such as clinical depression and compulsive gambling.

References
1 Frank MJ, Seeberger LC, O'Reilly RC. By Carrot or by Stick: Cognitive Reinforcement Learning in Parkinsonism. *Science* 2004; **306**:1940–1943.
2 Bodi N, Keri S, Nagy H, Moustafa A, Myers CE, Daw N, *et al.* 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* 2009; **132**:2385–2395.
3 Hikida T, Kimura K, Wada N, Funabiki K, Nakanishi S. Distinct roles of synaptic transmission in direct and indirect striatal pathways to reward and aversive behavior. *Neuron* 2010; **66**:896–907.
4 Kravitz AV, Tye LD, Kreitzer AC. Distinct roles for direct and indirect pathway striatal neurons in reinforcement. *Nat Neurosci* 2012; **15**:816–818.
5 Lobo MK, Covington HE, Chaudhury D, Friedman AK, Sun H, Damez-Werno D, *et al.* Cell type-specific loss of BDNF signaling mimics optogenetic control of cocaine reward. *Science* 2010; **330**:385–390.

6 Bromberg-Martin ES, Matsumoto M, Hikosaka O. Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron* 2010; **68**:815–834.

7 Ferguson SM, Eskenazi D, Ishikawa M, Wanat MJ, Phillips PEM, Dong Y, et al. Transient neuronal inhibition reveals opposing roles of indirect and direct pathways in sensitization. *Nat Neurosci* 2011; **14**:22–24.

8 Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science* 1997; **275**:1593–1599.

9 Barr GA, Moriceau S, Shionoya K, Muzny K, Gao P, Wang S, Sullivan RM. Transitions in infant learning are modulated by dopamine in the amygdala. *Nat Neurosci* 2009; **12**:1367–1369.

10 Fadok JP, Dickerson TMK, Palmiter RD. Dopamine is necessary for cue-dependent fear conditioning. *J Neurosci* 2009; **29**:11089–11097.

11 Garrison J, Erdeniz B, Done J. Prediction error in reinforcement learning: a meta-analysis of neuroimaging studies. *Neurosci Biobehav Rev* 2013; **37**:1297–1310.

12 Mattfeld AT, Gluck MA, Stark CEL. Functional specialization within the striatum along both the dorsal/ventral and anterior/posterior axes during associative learning via reward and punishment. *Learn Mem* 2011; **18**:703–711.

13 Mackey S, Petrides M. Architecture and morphology of the human ventromedial prefrontal cortex. *Eur J Neurosci* 2014; **40**:2777–2796.

14 Keeler JF, Pretsell DO, Robbins TW. Functional implications of dopamine D1 vs. D2 receptors:a 'prepare and select' model of the striatal direct vs. indirect pathways. *Neuroscience* 2014; **282**:156–175.

15 Moustafa AA, Herzallah MM, Gluck MA. Dissociating the cognitive effects of levodopa versus dopamine agonists in a neurocomputational model of learning in Parkinson's disease. *Neurodegener Dis* 2013; **11**:102–111.

16 Dreyer JK, Herrik KF, Berg RW, Hounsgaard JD. Influence of phasic and tonic dopamine release on receptor activation. *J Neurosci* 2010; **30**:14273–14283.

17 Ferenczi EA, Zalocusky KA, Liston C, Grosenick L, Warden MR, Amatya D, et al. Prefrontal cortical regulation of brainwide circuit dynamics and reward-related behavior. *Science* 2016; **351**:aac9698.

18 Seeley WW, Menon V, Schatzberg AF, Keller J, Glover GH, Kenna H, et al. dissociable intrinsic connectivity networks for salience processing and executive control. *J Neurosci* 2007; **27**:2349–2356.

19 De Ruiter MB, Veltman DJ, Goudriaan AE, Oosterlaan J, Sjoerds Z, van den Brink W. Response perseveration and ventral prefrontal sensitivity to reward and punishment in male problem gamblers and smokers. *Neuropsychopharmacology* 2008; **34**:1027–1038.

20 Herzallah MM, Moustafa AA, Natsheh JY, Abdellatif SM, Taha MB, Tayem YI, et al. Learning from negative feedback in patients with major depressive disorder is attenuated by SSRI antidepressants. *Front Integr Neurosci* 2013; **7**:67.

21 Must A, Szabó Z, Bódi N, Szász A, Janka Z, Kéri S. Sensitivity to reward and punishment and the prefrontal cortex in major depression. *J Affect Disord* 2006; **90**:209–215.

22 Eshel N, Roiser JP. Reward and punishment processing in depression. *Biol Psychiatry* 2010; **68**:118–124.