SUPPLEMENTARY MATERIALS

Dopaminergic medication reduces striatal sensitivity to negative outcomes in Parkinson's disease

Brónagh McCoy[1], Sara Jahfari[2,3], Gwenda Engels[4], Tomas Knapen[1,2†] and Jan Theeuwes[1†]

[1] Department of Experimental and Applied Psychology, Vrije Universiteit, Amsterdam, The Netherlands
[2] Spinoza Centre for Neuroimaging, Royal Academy of Sciences, Amsterdam, The Netherlands
[3] Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands
[4] Department of Clinical, Neuro and Developmental Psychology, Vrije Universiteit, Amsterdam, The Netherlands

† These authors share senior authorship

Supplementary Information

Supplementary Figures 1-12

Supplementary Tables 1-6

Supplementary Information

Additional participant information

Prior to participation, all subjects were screened according to the following inclusion criteria: age range 40-75 years old, normal/corrected-to-normal vision, and a prior diagnosis of PD in patients. Exclusion criteria were: current psychotropic medication usage (other than dopamine-/Parkinson-related medication in patients), major somatic disorder or psychosis, dementia diagnosis, or a history of head injury, stroke or any other neurological diseases. Patients were furthermore not included if they took selective serotonin reuptake inhibitors (SSRIs), in order to primarily examine the effects of dopamine, as serotonin has also been implicated in learning mechanisms (Daw *et al.*, 2002; den Ouden *et al.*, 2013).

All participants provided written informed consent in accordance with the Declaration of Helsinki. Payment for participation was a minimum of €100 (PD patients, three sessions) or €70 (HCs, two sessions) for participation. A reward bonus was additionally paid out based on performance during the reinforcement learning task (PD ON mean= €8.86 ± 0.99, PD OFF mean= €8.85 ± 1.00, HC mean= €9.34 ± 1.42, per learning run).

The study was set up to capture both within-subject effects of dopaminergic medication in Parkinson's disease as well as across-subject effects of disease (patients compared to healthy controls), across multiple behavioral and fMRI measurements. Since within-subject comparisons require a smaller sample size than across-subject, we aimed to collect sample sizes based on the lower-powered across-subjects (HC vs. PD) comparisons. We expected medium effect sizes of disease status, so to achieve 80% power, with a type 1 error of .05 (two-tailed), this required a sample size of 24 participants in each group.

Computational model estimation

The model was estimated using Markov Chain Monte Carlo (MCMC) inference. Models were implemented using the Stan programming language (Stan Development Team, 2015; Hoffman and Gelman, 2016). We ran three chains of 5000 samples each (discarding the first 2500 of each chain for burn-in), and ensured convergence using manual examination of the trace plots (hairy caterpillars, easily moving around the parameter space) and evaluation of $\hat{r}$ statistics, which were all <1.1 (Gelman and Rubin, 1992). We additionally generated a number of quantities of interest from the model, including individual subjects' trial-by-trial RPE for inclusion in fMRI whole-brain and deconvolution analyses. Simulations displayed in Supplementary Fig. 2 show adequate data recovery, along with the fitted mean group-level posterior distributions. Further assessment of model suitability demonstrates that the model is identified, as indicated by only weakly correlated parameter estimates

(Supplementary Fig. 11) and satisfactory parameter recovery of the modes of the fitted parameter distributions for each participant (Supplementary Fig. 12).

Behavioural analysis of AB trials

We carried out an additional analysis on AB trials to examine medication- or disease-related differences in AB choice behaviour associated with receiving either positive or negative feedback. Data were separated into trials in which positive feedback was received on the previous trial of that pair (for regression 1) and trials in which negative feedback was received on the previous trial (for regression 2). These were divided into separate subsets since including a regressor for each of them in the same GLM would lead to rank deficiency (i.e. one is (necessarily) a linear combination of the other). We carried out a mixed-effects logistic regression analysis on each subset. In regression 1, the dependent variable (*DV*) was a binary indicator for choosing A (*DV*=1) or choosing B (*DV*=0) on the current trial. The first independent variable (*IV1*) coded for whether the A (optimal) or B (suboptimal) stimulus was chosen on the previous trial (*IV1* = 1 or -1, respectively). In this way, a large positive beta estimate from the regression represented choosing A on the previous trial, receiving positive feedback, and choosing A again on the next trial. Medication and disease status were included as covariates, as in the other mixed-effects regressions, along with a varying subject intercept to account for individual differences. In regression 2 (trials following negative feedback), we used the same DV and covariates, but with *IV1* coding the opposite sign to that in regression 1, i.e., choosing optimal A (*IV1*= -1) and choosing suboptimal B (*IV1*=1) on the previous trial. This inversion of sign means that a large positive beta estimate from the regression captured choosing B on the previous trial, receiving negative feedback, and switching to the optimal A stimulus on the next trial. For trials following positive feedback (regression 1), we found a main effect of *IV1* (β (SE) = 0.84 (0.05), $z = 17.33$, $p < .001$), i.e. in general, participants stuck with the optimal A stimulus if rewarded for that choice on the previous trial. For trials following negative feedback, we found a large effect of medication (β (SE) = 0.52 (0.13), $z = 3.96$, $p < .001$), with a trend effect of disease (β (SE) = 0.64 (0.33), $z = 1.95$, $p = .051$) and *IV_1* (β (SE) = −0.17 (0.09), $z = 1.92$, $p = .055$). The main effect of medication shows that in trials following negative feedback, PD ON chose the optimal A stimulus more often than PD OFF, regardless of what they chose in the previous trial. This suggests that, for the AB pair at least, PD ON were less inclined to use negative feedback to influence subsequent choice behaviour.

Session order effects

Experimental sessions were performed twice on PD patients and once on HCs. PD ON/OFF medication status was counterbalanced across first/second session in patients, so any session order effects on performance would be expected only in relation to HCs. We therefore carried out additional analyses on the behavioral data of the learning and transfer phases, to include session ("*Sess*") as a separate binary covariate in the mixed effects regression analyses reported in the Methods and

materials section. We coded this as *Sess*=0 for HC since they only had one session, with *Sess*=0 for a patient's first session and *Sess*=1 for a patient's second session. The interaction between this variable and stimulus pair was also included, with the rest of the regression set up as in Equation 5 in manuscript.

For the learning phase, the effects already reported in the manuscript were also significant after carrying out this analysis; there was a main effect of stimulus pair (β (SE) = 0.37 (0.04), z = 8.75, $p << .001$), medication (β (SE) = 0.14 (0.04), $z = 3.41$, $p < .001$), an interaction between medication and stimulus pair (β (SE) = 0.21 (0.05), $z = 4.27$, $p < .001$), and an interaction between disease and stimulus pair (β (SE) = 0.18 (0.06), $z = 3.11$, p=.002) (see Supplementary Fig. 3A). This analysis revealed an additional interaction between session and stimulus pair (β (SE) = −0.11 (0.05), z = 2.30, p=.02), which suggests that when patients were in their second session, the difference in accuracy between the AB and EF pair was smaller than when participants were in their first session. There were no interactions between session and medication or disease.

For the transfer phase, the previously reported significant effects were again present; there was an interaction between medication (PD ON or OFF) and Approach A/Avoid B trial type (β (SE) = 0.35 (0.06), $z = 5.78$, $p < .001$) and an interaction between disease (HC or PD OFF) and Approach A/Avoid B trial type (β (SE) = 0.30 (0.07), $z = 4.25$, $p < .001$) (see Supplementary Fig. 3B). There was also a main effect of session (β (SE) = -0.13 (0.06), $z = 2.15$, $p = .03$), with patients in their second session performing worse in general than participants in their first session. There were no interactions between session and medication or disease.

fMRI preprocessing

The following information was generated from FMRIPREP based on the preprocessing pipeline used in this study. Each T1w (T1-weighted) volume was corrected for INU (intensity non-uniformity) using N4BiasFieldCorrection v2.1.0 (Tustison *et al.*, 2010) and skull-stripped using antsBrainExtraction.sh v2.1.0 (using the OASIS template). Brain surfaces were reconstructed using recon-all from FreeSurfer v6.0.1 (Dale *et al.*, 1999), and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray-matter of Mindboggle (Klein *et al.*, 2017). Spatial normalization to the ICBM 152 Nonlinear Asymmetrical template version 2009c (Fonov *et al.*, 2009) was performed through nonlinear registration with the antsRegistration tool of ANTs v2.1.0 (Avants *et al.*, 2008), using brain-extracted versions of both T1w volume and template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using fast (Zhang *et al.*, 2001) (FSL version 5.0.9). Functional data was motion corrected using *mcflirt* (FSL version 5.0.9) (Jenkinson *et al.*, 2002)."Fieldmap-less" distortion correction was performed by co-registering the functional image to the same-subject T1w image with intensity inverted (Huntenburg *et al.*, 2012; Wang *et al.*, 2017) constrained with an average fieldmap

template (Treiber *et al.*, 2016), implemented with *antsRegistration* (ANTs). This was followed by co-registration to the corresponding T1w using boundary-based registration (Greve and Fischl, 2009) with 9 degrees of freedom, using *bbregister* (FreeSurfer version 6.0.1). Slice timing correction was not performed on the data. Motion correcting transformations, field distortion correcting warp, BOLD-to-T1w transformation and T1w-to-template (MNI) warp were concatenated and applied in a single step using *antsApplyTransforms* (ANTs version 2.1.0) using Lanczos interpolation. Physiological noise regressors were extracted applying CompCor (Behzadi *et al.*, 2007). Principal components were estimated for the two CompCor variants: temporal (tCompCor) and anatomical (aCompCor). A mask to exclude signal with cortical origin was obtained by eroding the brain mask, ensuring it only contained subcortical structures. Six tCompCor components were then calculated including only the top 5% variable voxels within that subcortical mask. For aCompCor, six components were calculated within the intersection of the subcortical mask and the union of CSF and WM masks calculated in T1w space, after their projection to the native space of each functional run. Frame-wise displacement (Power *et al.*, 2014) was calculated for each functional run using the implementation of Nipype. Many internal operations of FMRIPREP use Nilearn (Abraham *et al.*, 2014), principally within the BOLD-processing workflow. For more pipeline details, please refer to https://fmriprep.readthedocs.io/en/latest/workflows.html.

Single-trial whole-brain analysis

We were interested in estimating trial-by-trial representations of learning (e.g. RPEs) in the brain. For this, we carried out a single-trial whole-brain analysis to capitalize on the variability in BOLD signal across trials, using Nipype's FSL interface (Gorgolewski *et al.*, 2011, 2017). A Least Squares All (LSA) GLM was fit to each subject's brain data, per learning run (see Mumford *et al.*, 2012). The feedback onset of each trial was included as a separate regressor, all in one model. We included 13 confound regressors to remove nuisance effects that might have contributed to the brain signal: Framewise Displacement (FD), 6 rigid-body transform motion parameters (3 translational, 3 rotational), and 6 aCompCor physiological noise regressors (to help exclude physiological noise in the CSF and WM). Spatial smoothing was performed using a Gaussian kernel with a full width at half maximum of 4 mm. A Savitsky-Golay filter was used for high-pass filtering, with a window length of 120 seconds and polynomial order of 3. The first-level design was set up using the Nipype interface to FMRI Expert Analysis Tool (FEAT) from FSL (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl). Delta functions of all regressors in the model were convolved with the canonical hemodynamic response function (HRF) and regressed against each subject's fMRI data, using Nipype's FSL FILMGLS interface. From this, a contrast of parameter estimates (COPE) was obtained for each trial of every subject. Next, we performed a second-stage analysis on the single-trial copes, to model per trial feedback valence and RPE. The feedback regressor was coded as +1 or -1 for positive and negative feedback trials respectively, to model brain activity that co-varied with valence. A single

RPE covariate regressor was constructed by taking the per-trial RPE (retaining its positive or negative sign), de-meaning this distribution across trials, and convolving these events with the HRF. Since feedback valence and RPE are correlated, i.e. positive feedback is accompanied by a positive RPE, this allowed us to assign brain activity co-varying specifically with valence or RPE. This was run as a fixed effects multiple regression model using FLAMEO on a per subject basis. Fixed effects multiple regression models for collapsing across runs and deriving within-patient medication difference COPEs were carried out in a similar way. Medication difference COPEs of feedback and RPE were then brought to the group level in a random effects model, using FSL's FLAME 1+2 and outlier detection procedures. All group level Z (Gaussianized T) statistic images were thresholded using clusters of $z > 2.3$ and a cluster-corrected significance threshold of $p < 0.01$. Group-level analyses were carried out in this way for each regressor (feedback valence and RPE) on each separate group (HC, PD ON, PD OFF), on the within-subject medication differences ON > OFF and OFF > ON, and on across-subject disease differences HC vs PD (OFF or ON). All group-level z-statistic contrasts can be viewed at https://doi.org/10.6084/m9.figshare.6989024.v2. The MNI152_T1_1mm_brain standard brain was converted to functional space using FSL FLIRT, eroded by 1 voxel, and used as a brain mask for all of the analyses described above.

ROIs

We obtained high-resolution probabilistic atlas masks from a recent open-source dataset (Pauli *et al.*, 2018). These sub-cortical ROIs have been well-established in playing an important role in reinforcement learning (Schultz *et al.*, 1997; Brown *et al.*, 1999; Hazy *et al.*, 2010; O'Doherty *et al.*, 2017). We focused on striatal ROIs (caudate nucleus, putamen, and nucleus accumbens) for the learning phase deconvolution analysis, as these have been most extensively studied in the past (Frank, 2005; Cools *et al.*, 2007; Cox *et al.*, 2015; Jahfari *et al.*, 2018). FSL FLIRT was used to register the masks to FMRIPREP output space. BOLD percent signal change during the transfer phase was extracted from ROIs informed by the learning phase of the task. We took the cluster-corrected RPE medication difference (OFF-ON) z-statistic COPE from the learning phase and multiplied it by independent striatal ROIs from the Pauli et al. (2018) dataset. These masks were thresholded to exclude the lowest 25% of voxels and then binarized. The masks are available on figshare, at: https://doi.org/10.6084/m9.figshare.6989024.v2.

Deconvolution analysis of feedback interval and RPEs

Deconvolution analyses were carried out on BOLD timeseries from striatal ROIs (see *ROIs* for mask information). We extracted BOLD time courses and teased apart the covariation with BOLD signal of positive and negative RPEs separately, using the fMRI timeseries data as preprocessed by FMRIPREP. A Savitsky-Golay filter was used for high-pass filtering, with a window length of 120 seconds and polynomial order of 3. These timeseries were then converted to percent signal change

(PSC). PSC was calculated by dividing the timeseries by the mean of the entire timeseries, multiplying by 100, and then subtracting 100, to get a mean-centered output timeseries. Data from each subject were weighted per voxel according to the probability of belonging to a particular striatal ROI, and then averaged across voxels of that ROI. We set up a model with three regressors: stimulus onsets (with RT duration), positive feedback onsets and negative feedback onsets. Positive and negative RPEs were z-scored separately and included as covariates of their respective positive or negative feedback event type. The deconvolution was implemented using the Python-based *nideconv* package (de Hollander and Knapen, 2017). Events and covariates were deconvolved with a Fourier basis set, which uses a combination of sines and cosines to model the data. This was implemented instead of the standard finite impulse response (FIR) function as it substantially reduces the number of regressors, thereby improving the robustness of parameter estimates. Five Fourier regressors (1 intercept, 2 sine waves and 2 cosine waves) were used for each of the positive and negative feedback events and positive and negative RPE covariates. We also included several confound regressors in the model: FD, 6 rigid-body transform motion parameters (3 translational, 3 rotational), WM, stdDVARS (standardized derivative of RMS variance over voxels), and 6 aCompCor physiological noise regressors. Time courses were then estimated simultaneously using a least-squares fit, for the time window -2 to 13.05 seconds (7 TRs) around feedback onsets. Up-sampling during the fitting procedure was implemented 20-fold as part of *nideconv* functionality. The resulting time courses were then brought to the group level, and within-patient medication differences were calculated per up-sampled timepoint of each fit. Clusters of significant intervals were identified using permutation-based one-sample t-tests (t-threshold set at $p<.05$, n=5000 permutations) as implemented in *mne-python* (version 0.15.2; (Gramfort *et al.*, 2013, 2014). Shaded regions in Fig. 4 and Supplementary Fig. 8 represent 68% confidence intervals (±1 SEM; bootstrapped using n=5000 permutations). Shaded regions in Supplementary Fig. 7 represent 95% confidence intervals (±1 SEM; bootstrapped using n=5000 permutations), to make group differences clearly visible.

Deconvolution analysis of choice interval and Q-values

We established a separate deconvolution analysis to assess modulations of the BOLD signal by Q-values during the choice period, i.e. while the stimulus options were presented. We set up this model with the same three regressors as in the previous analysis: stimulus onsets (with RT duration), positive feedback onsets and negative feedback onsets, all with a Fourier basis set. However, here the Q-values of the chosen stimulus were z-scored and entered as a covariate on the stimulus onset events. RPEs were not included in the model. All other aspects of the analysis were the same as those described previously.

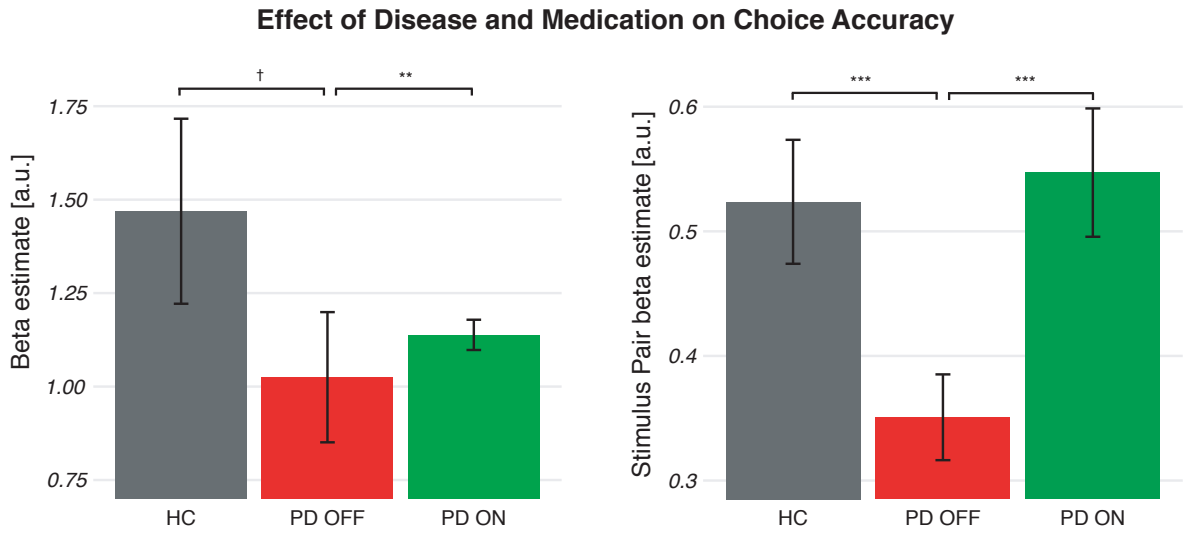
Applying the analysis to striatal ROIs, we did not find a significant medication-related difference in the caudate nucleus or nucleus accumbens. However, the analysis did reveal a significant

PD ON > OFF difference in modulation of the BOLD signal by Q-values in the putamen (see Supplementary Fig. 9).

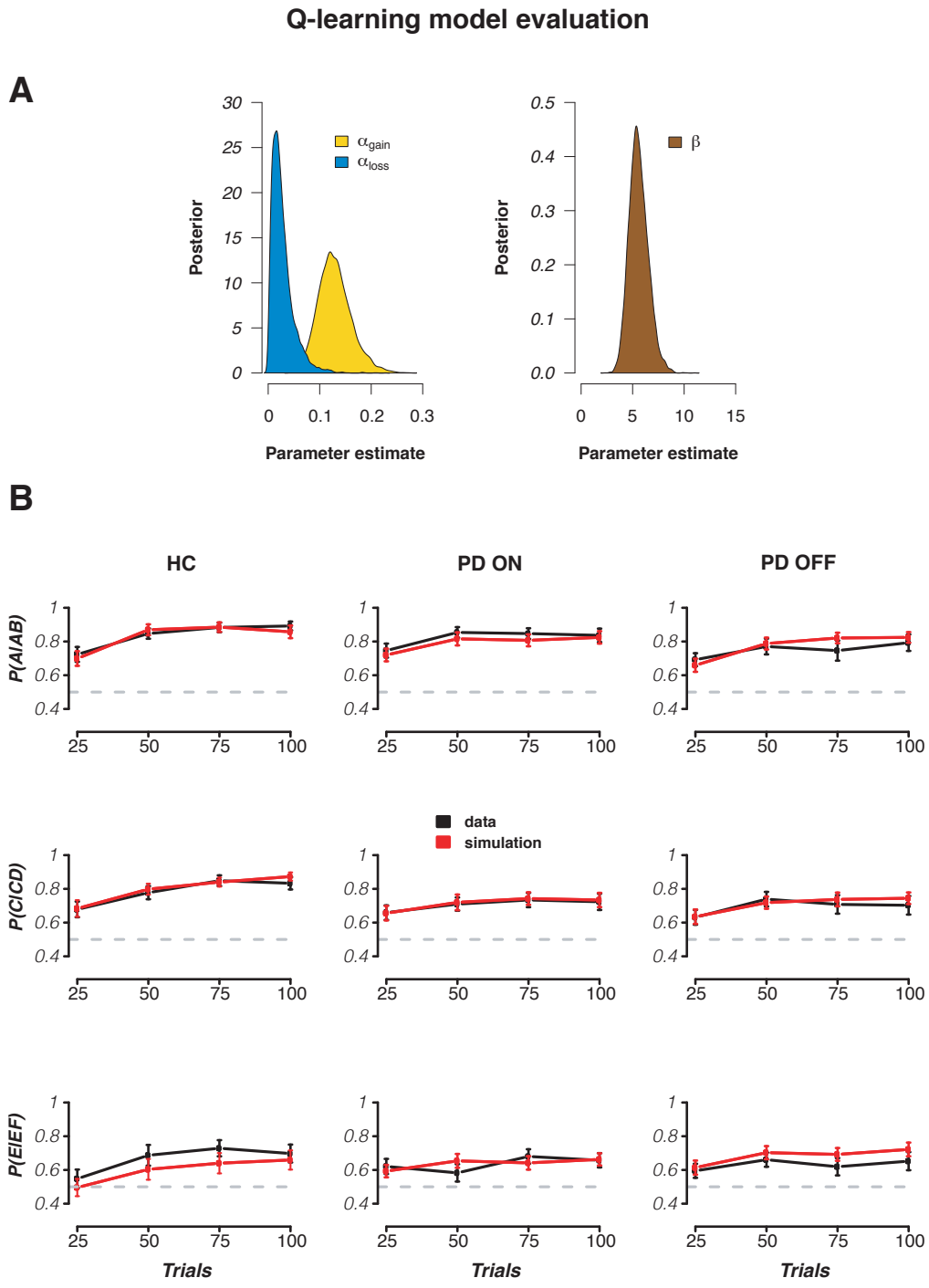
Transfer phase BOLD percent signal change

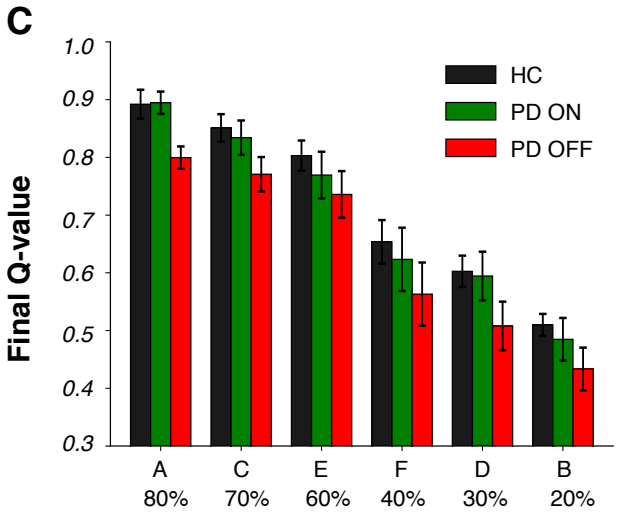
A standard GLM was set up to model BOLD responses to events in the transfer phase. Stimulus onsets and durations for three regressors were included: Approach A trials, Approach B trials, and all other trials (of no interest). Similar to the steps carried out in the learning phase whole-brain analysis, we performed 4mm smoothing, Savitsky-Golay high-pass filtering with a window length of 120 seconds and polynomial order of 3, and included the same 13 confound regressors in the design and convolved with a canonical HRF. Fixed effects analyses were performed across runs and for medication differences within patients. We then took the resulting two COPES for Approach A and Avoid B trials per subject and used FSL's *featquery* to calculate the mean percent signal change in the striatal ROIs that showed a significant learning phase medication difference in RPE (described in *ROIs*). In a similar way to the behavioral correlation analysis between learning rate and transfer accuracy (see Materials and Methods and Fig. 5B), we included learning rate EVs to explain the PSC (OFF>ON) medication difference in Avoid B>Approach A trials in striatal ROIs using robust (multiple) regression, with EVs as either: both positive and negative learning rate medication differences ($k\alpha_{gain}$ and $k\alpha_{loss}$), $k\alpha_{gain}$ only, or $k\alpha_{loss}$ only. Models were compared based on calculated BIC values (see Supplementary Table 4), and the learning to transfer PSC correlation *p*-value was obtained from the winning α_{loss}-only model. Individual medication differences were quantified as the modes of the within-subject medication difference parameter distributions.

Supplementary Figures

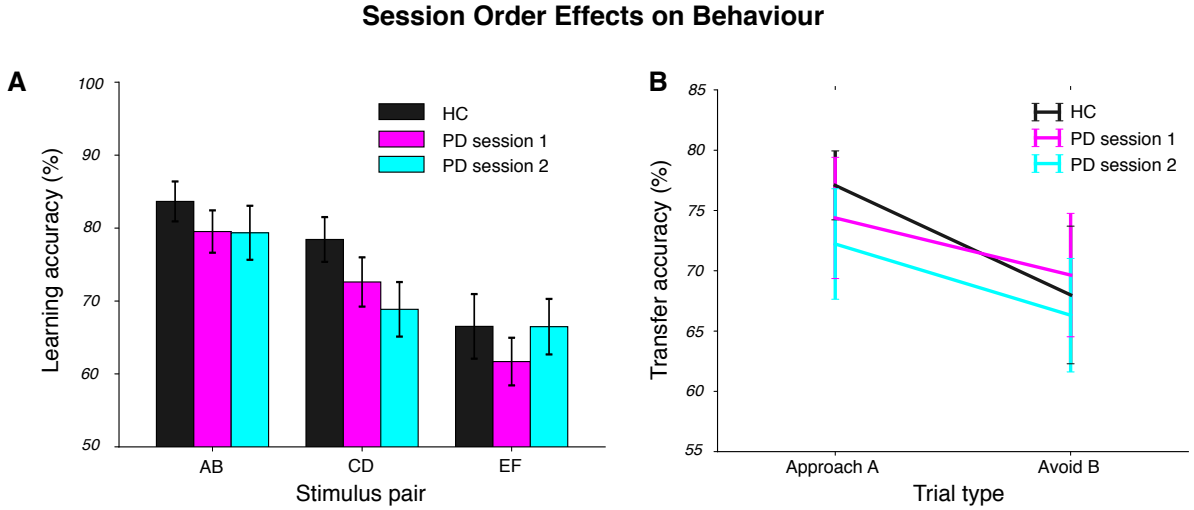


Supplementary Figure 1 | Beta parameter estimates in learning phase mixed-effects logistic regression model, with choice accuracy as the dependent variable. PD OFF was considered as 'baseline', with any relative increase in beta parameters for PD ON or HC representing the effect of medication and disease status, respectively. Here, the main effects of disease and medication on choice accuracy are presented (left), as well as interaction effects of stimulus pair and disease, and stimulus pair and medication, on choice accuracy (right).

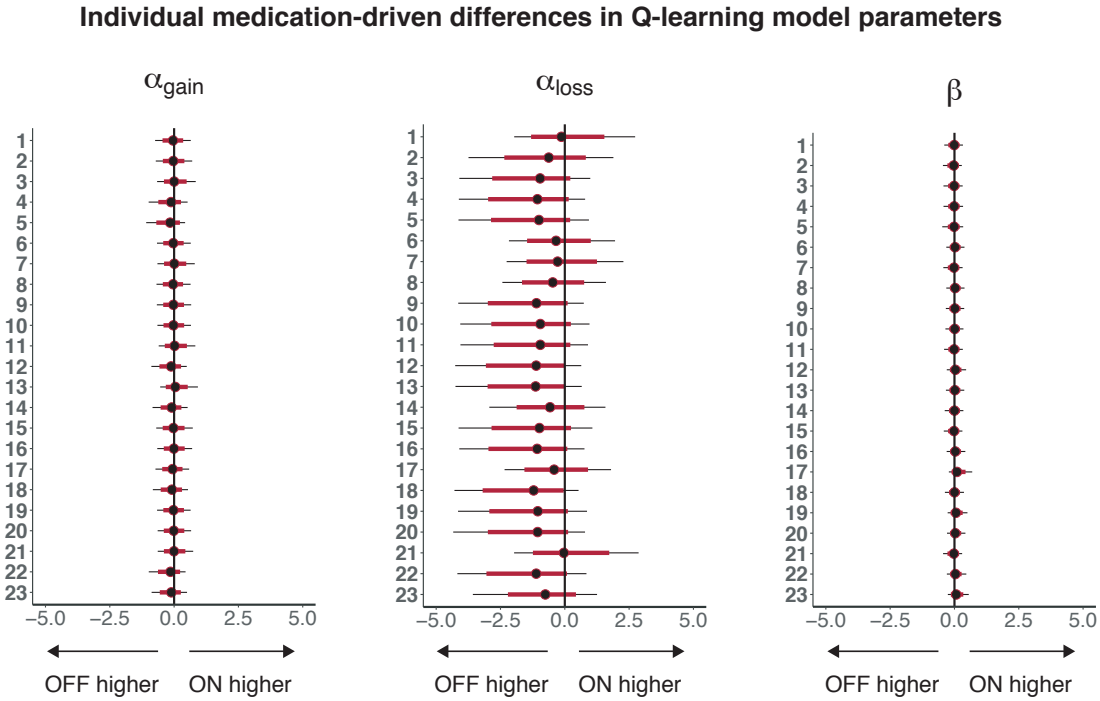

Q-learning model evaluation
A
Posterior
α_{gain}
α_{loss}
Parameter estimate
β
Parameter estimate
B
HC
PD ON
PD OFF
P(A|AB)
P(C|CD)
P(E|EF)
data
simulation
Trials



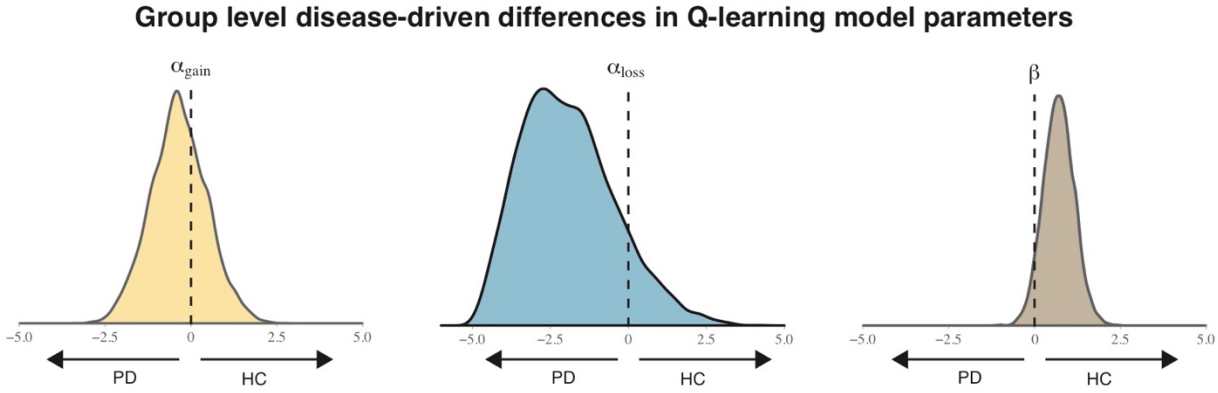
Supplementary Figure 2 | Bayesian hierarchical computational model assessment and simulation. A) Group-level parameter estimate distributions. Similar to other studies using separate learning rates for positive and negative events, we found higher learning rates for positive compared to negative feedback, termed "optimism bias" (Lefebvre *et al.*, 2017; Jahfari *et al.*, 2018; Van Slooten *et al.*, 2018). **B)** Simulation of the fitted model. To check whether the model sufficiently captured actual choice behavior of participants, we simulated the probability of choosing the best option using the posterior distributions of the fitted free parameters of each participant. Plots of the modeled against empirical data across each group and stimulus pair show that the model is a good representation of overall learning. **C**) Final Q-values across each stimulus and group from the fitted model, showing that the model did a good job in capturing declining Q-value according to decreasing reward contingency.



Supplementary Figure 3 | Behavioral performance based on experimental session order in patients (not dopamine medication manipulation). See accompanying Supplementary Information: *Session order effects*. PD patients performed the task twice (ON or OFF medication, counterbalanced across first/second session). HC performed the task once; HC results here are the same as in manuscript Fig. 1 and Fig. 5A. **A)** Learning phase accuracy per group in choosing the better option of each stimulus pair. **B)** Transfer phase accuracy per group in correctly choosing the best A stimulus when it is presented and correctly avoiding the worst B stimulus when it is presented.

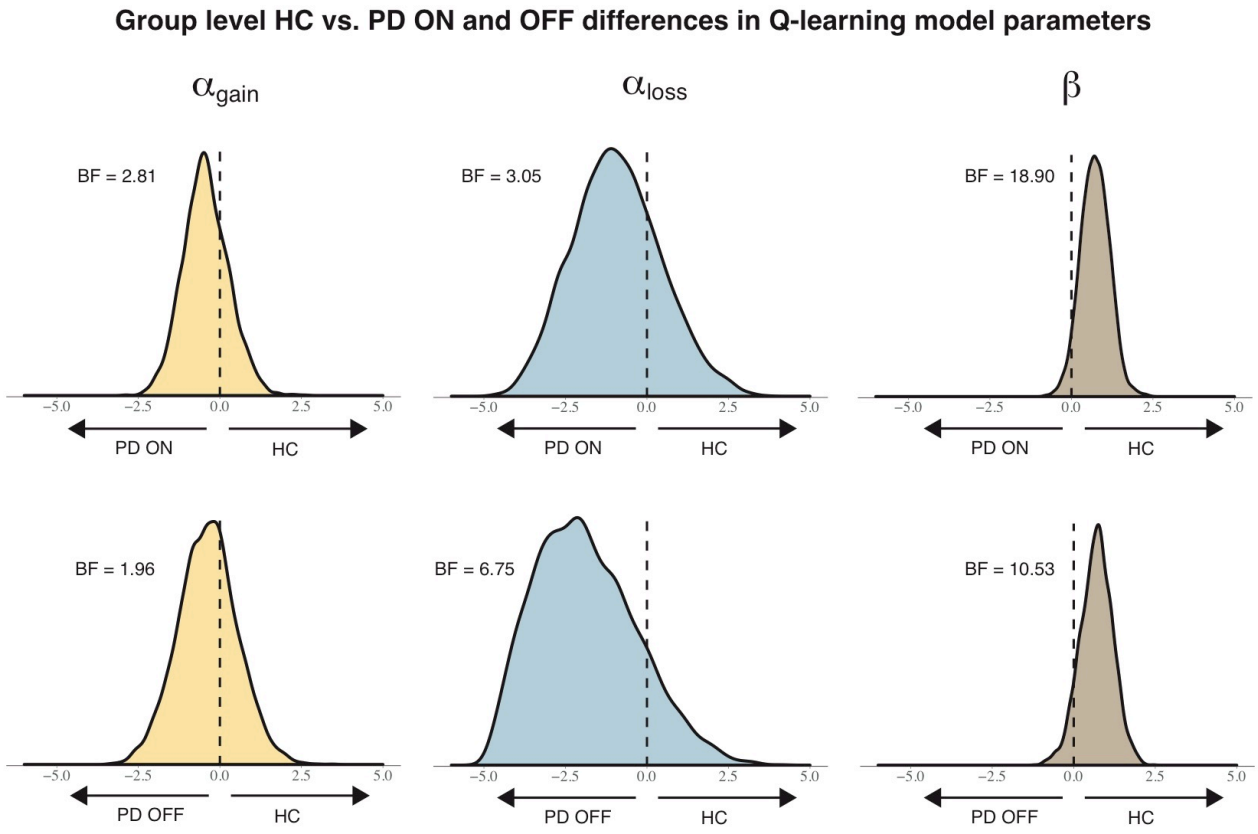


Supplementary Figure 4 | Individual medication differences in Parkinson's patients in computational model parameters. Medians are displayed, with red and black horizontal bars denoting 85% and 95% highest density intervals (HDI), respectively. Note that wherever individual medication difference parameters are used, we take the mode rather than median of the distributions, as this point has the highest posterior density of all parameter values.

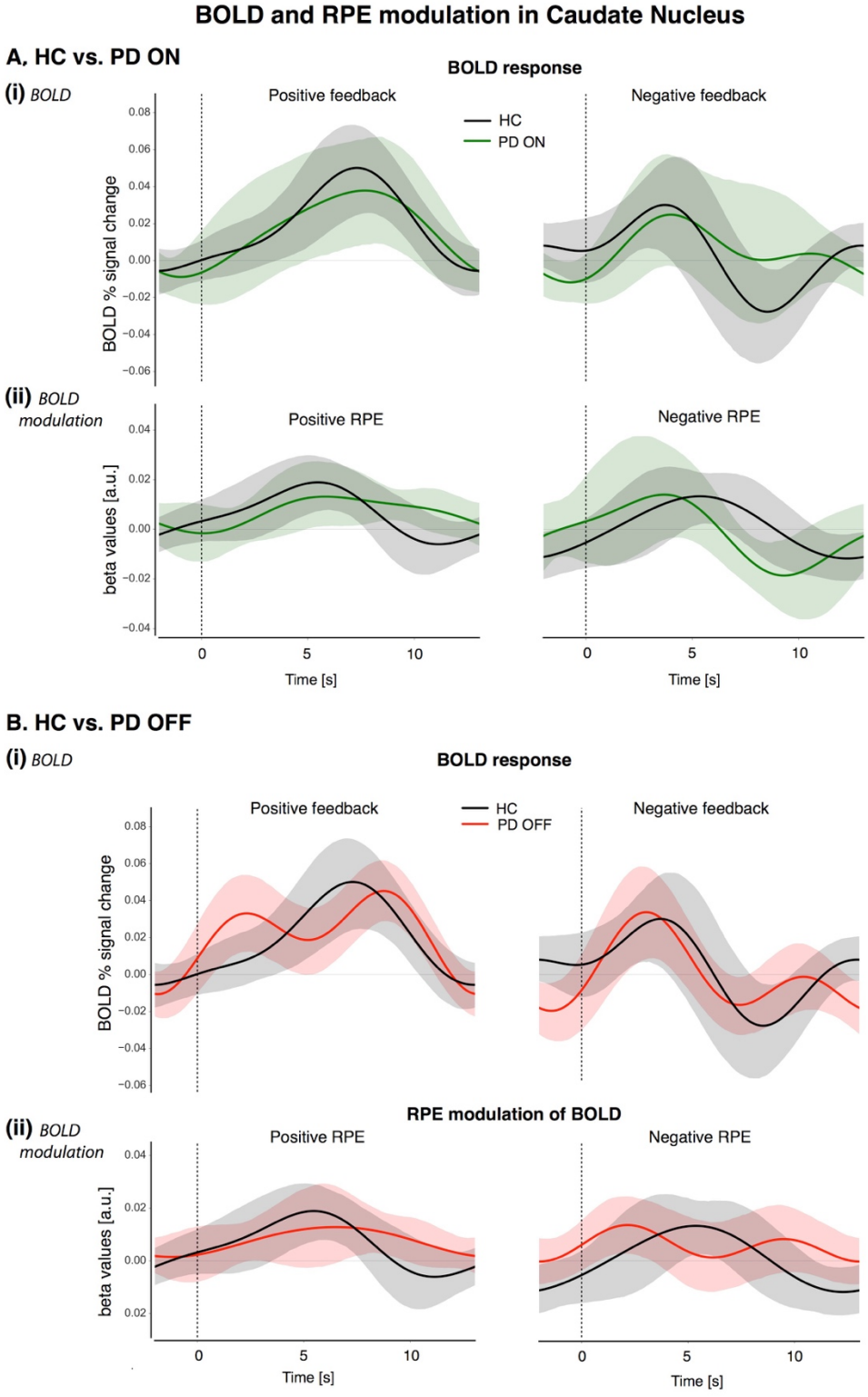


Parameter	Mean	SEM	95% HDI	BF_{10}
β	0.707	0.014	[-0.185, 1.565]	16.59
α_{gain}	-0.354	0.021	[-1.944, 1.298]	2.05
α_{loss}	-1.928	0.038	[-4.500, 1.065]	7.89

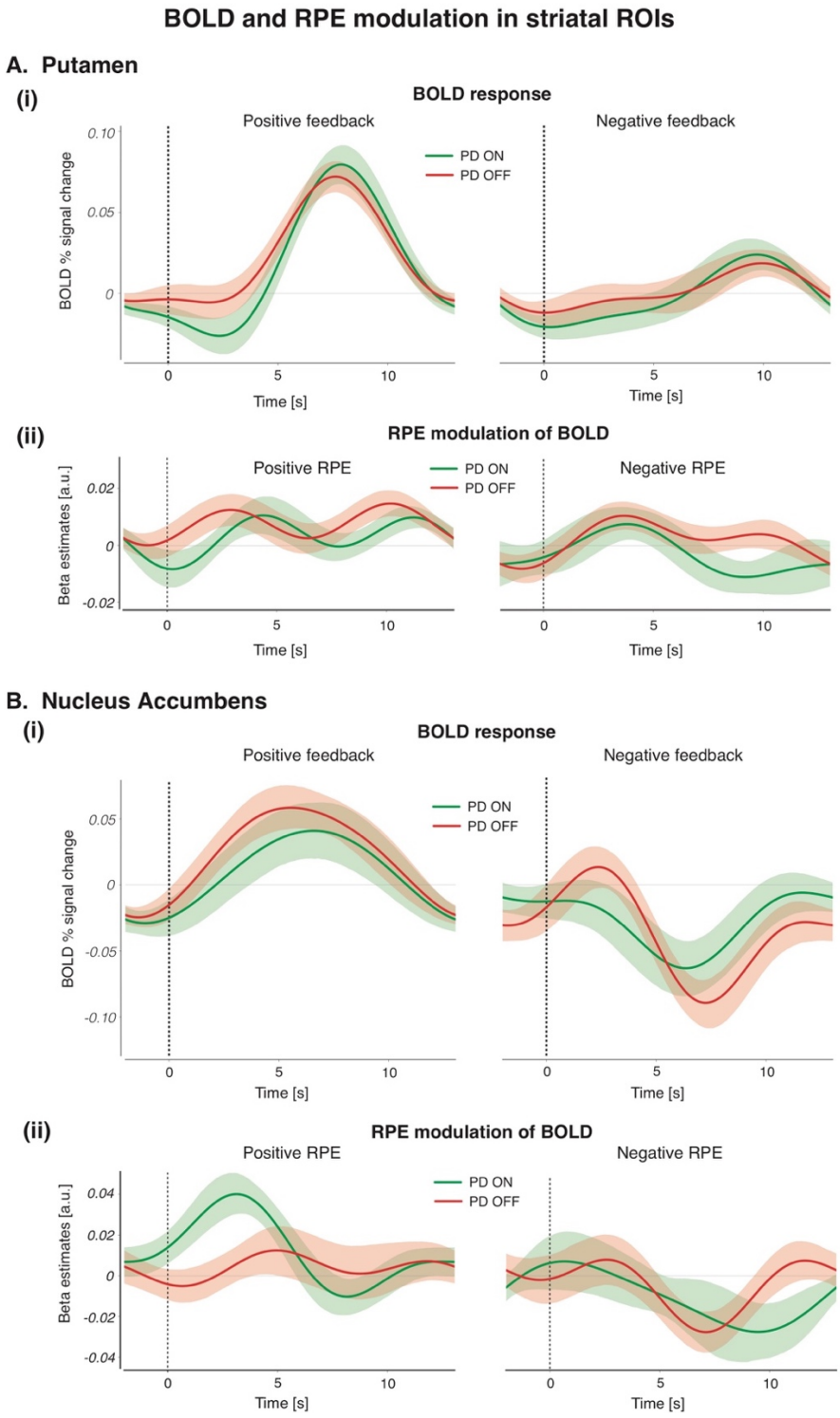
Supplementary Figure 5 | Disease-related differences in computational model parameters. Parameter distributions are shown for HC v PD (N=46). We found moderate evidence for greater α_{loss} negative learning in PD compared to HC (BF = 7.89), and strong evidence for greater exploitation in HC compared to PD patients (BF = 16.89).



Supplementary Figure 6 | Disease-related differences in computational model parameters, separately per PD medication status. To assess whether both PD ON and OFF contributed in a similar or differential way to the disease-related difference distributions shown in Supplementary Figure 4, two additional smaller models were run using a subset of data entered into the main model: HC vs. PD ON and HC vs. PD OFF. We ran 3000 samples each (discarding the first 1500 of each chain for burn-in). Checks for model convergence via visual inspection and $\hat{r}$ statistics was carried out in the same way as in the main model. The effect of disease on group-level parameters are displayed above for HC vs. PD ON (top row) and HC vs. PD OFF (bottom row). There is strong evidence for greater exploitation in HC compared to both PD ON and OFF, as indicated by a rightward shift in the β parameter distribution and BFs > 10.

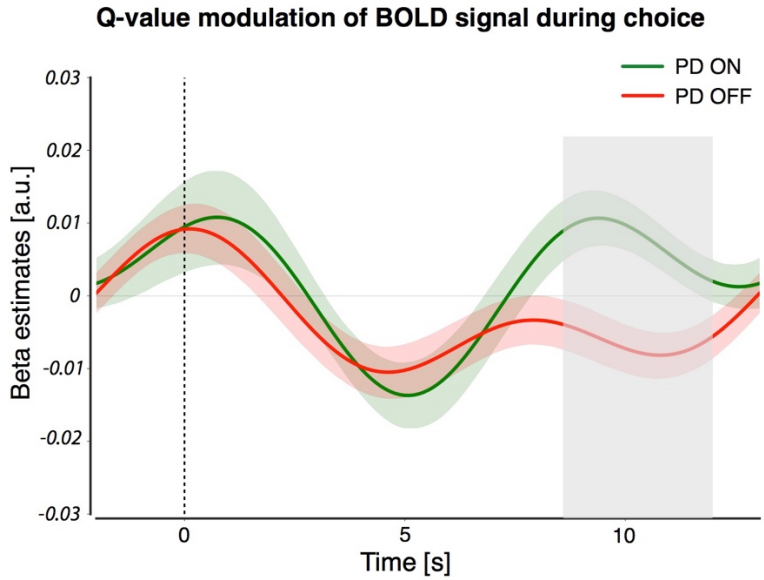


Supplementary Figure 7 | HC vs. PD ON or OFF comparisons of caudate nucleus feedback-related BOLD response and modulation of the BOLD signal by positive and negative RPEs. A) BOLD percent signal change (i) and RPE modulation of BOLD (ii) for HC and PD ON. **(B)** BOLD percent signal change (i) and RPE modulation of BOLD (ii) for HC and PD OFF. The shaded regions represent 95% confidence intervals, so visual inspection alone is enough to detect significant differences between groups, i.e. non-overlapping bands. The PD ON/OFF plots are the same as those shown in Fig. 4 of the manuscript (caudate nucleus ROI). Similar group comparison analyses of the putamen and nucleus accumbens ROIs (not shown) also show no group differences.

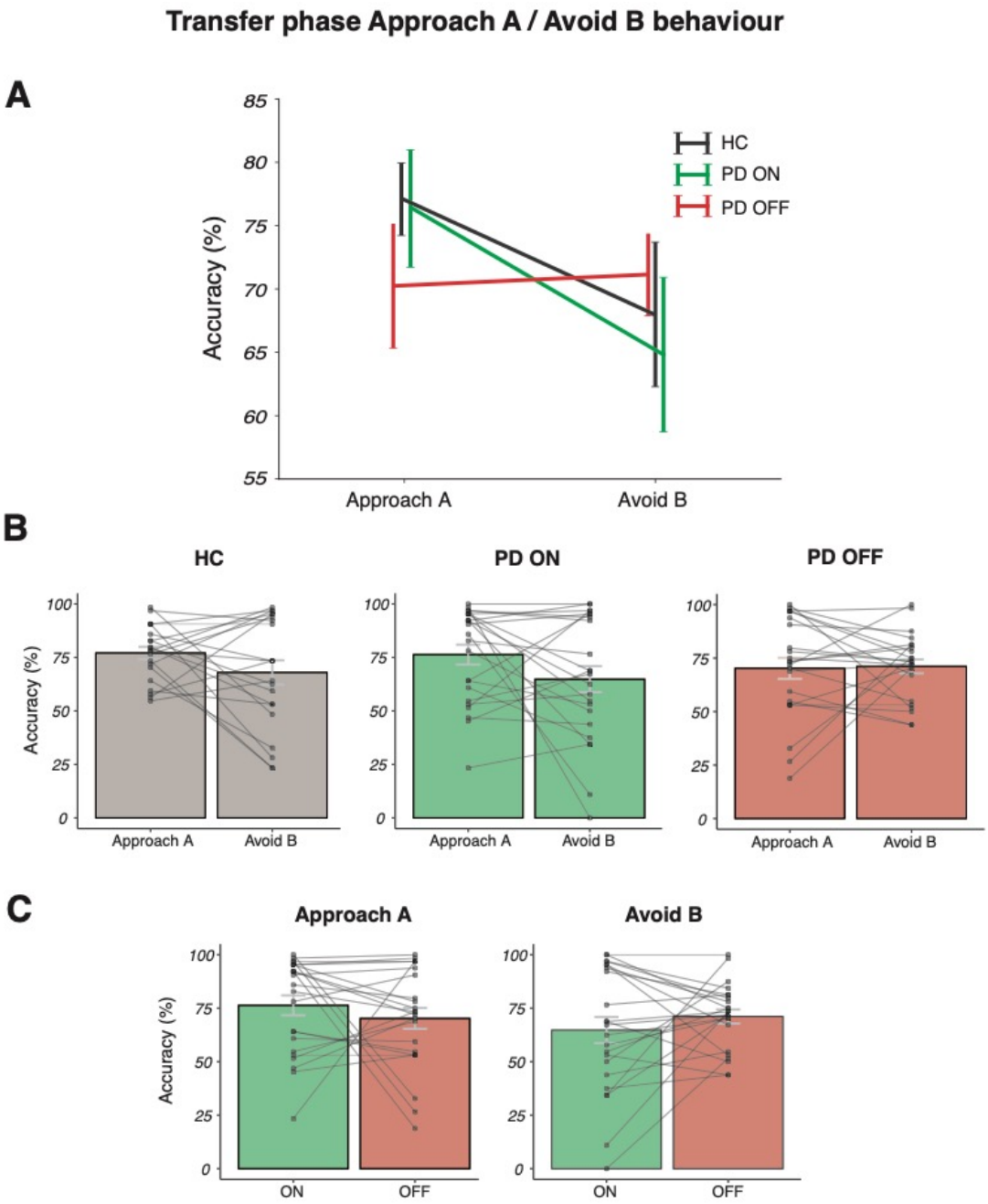


Supplementary Figure 8 | BOLD percent signal change and RPE modulation of BOLD for positive and outcomes in striatal ROIs of PD patients (related to Fig. 4). (A) BOLD percent signal change (i) and RPE modulation of BOLD (ii) in the putamen. **(B)** BOLD percent signal change (i) and RPE modulation of BOLD (ii) in the nucleus accumbens (NAc). Although no medication difference in RPE was found for ventral striatal NAc activity, it was included here for informational purposes since it has been implicated in several previous studies on the effects of dopamine on learning (Breiter *et al.*, 2001; McClure *et al.*, 2003; O'Doherty *et al.*, 2003; Cools *et al.*, 2007). In NAc, there appears to be a quantitative PD ON > OFF group level difference in positive RPE, however this difference is not statistically significant when within-subject differences were cluster-corrected across multiple timepoints. Positive and negative RPEs in each ROI were z-scored around that

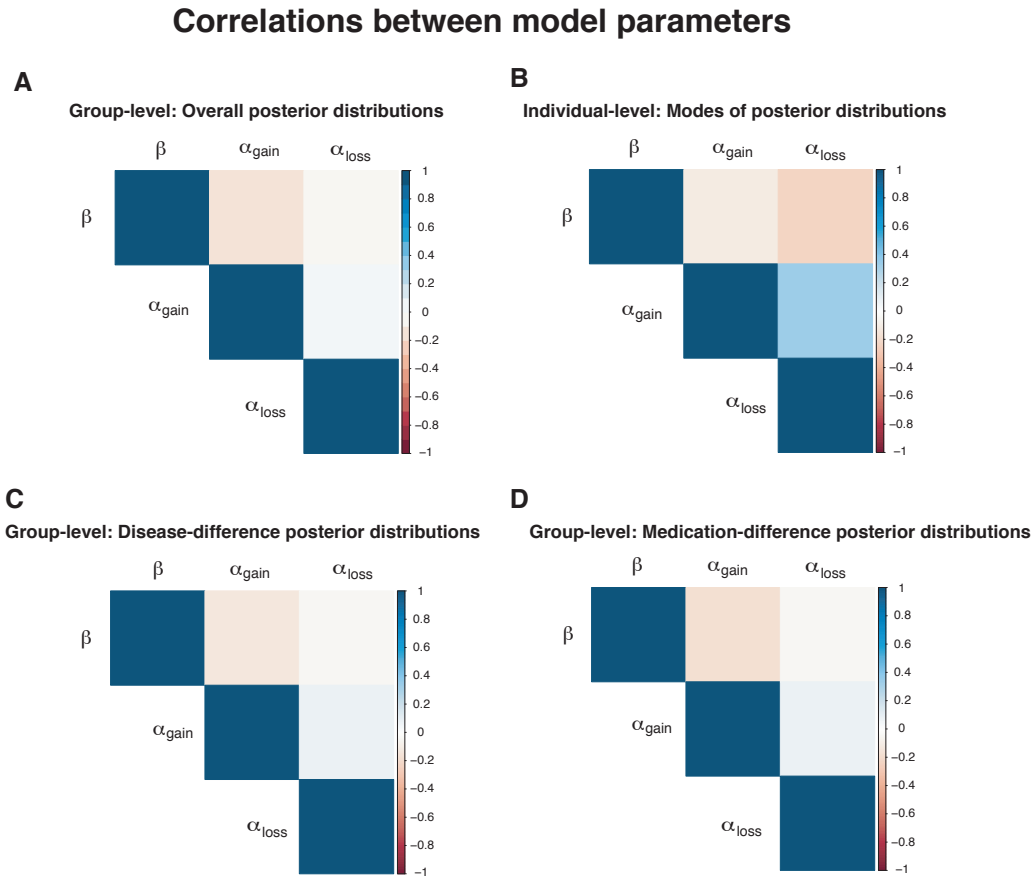
ROI's BOLD response for positive and negative events, respectively. Colored bands represent 68% confidence intervals (±1 SEM).



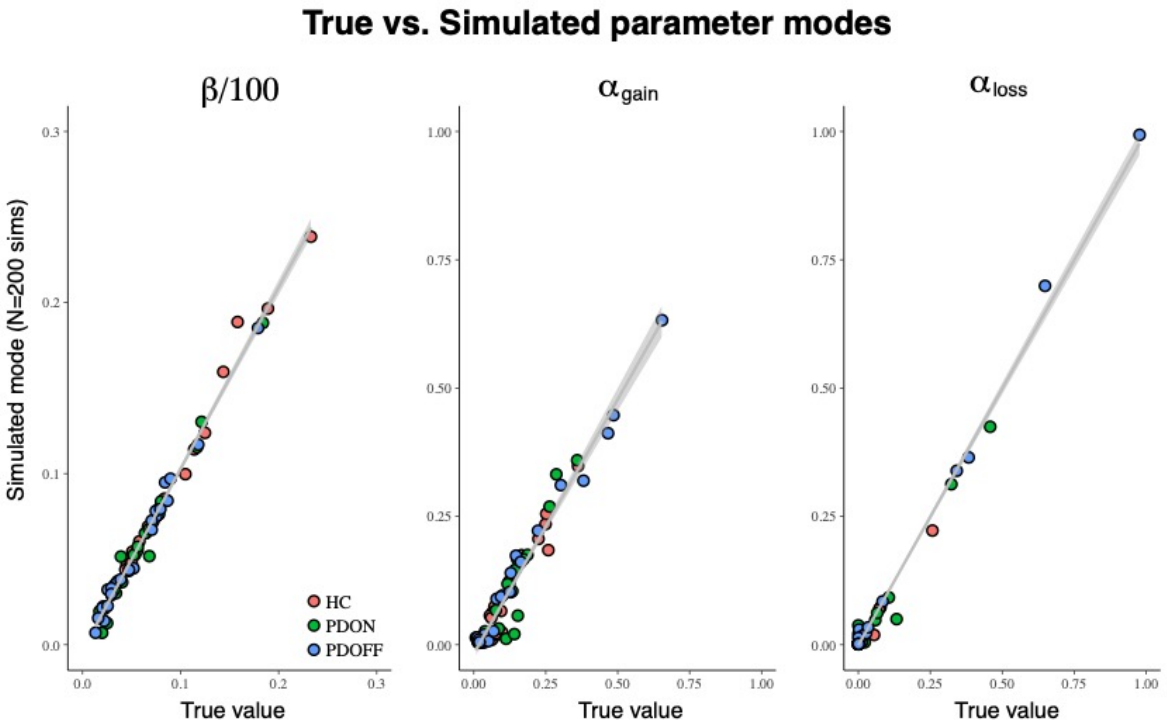
Supplementary Figure 9 | Modulation of BOLD signal in the putamen by Q-value of the chosen stimulus, time-locked to the onset of the stimulus options. Q-values were z-scored around stimulus onset events. A significant difference between PD ON and OFF is represented by the grey shaded region, which passed cluster-correction for multiple comparisons across timepoints ($p<.05$). Colored bands represent 68% confidence intervals (±1 SEM). See Supplementary section "*Deconvolution analysis of choice interval and Q-values"* for analysis description.



Supplementary Figure 10 | Accuracy in transfer phase approach and avoid behaviour for each group and subject. (A) Fig. 4 from manuscript with HC group included. **(B)** Averages per condition and within-subject plots, separately for HC, PD ON and PD OFF. **(B)** Within-subject plots across PD medication session, separately per Approach/Avoid condition.



Supplementary Figure 11 | Correlation between fitted model parameters. Correlations between full posterior distributions (group-level) or the modes of the posterior distributions (individual-level) of the three α_{gain}, α_{loss}, β model parameters, at each level of the hierarchical model. **(A)** Correlations at the overall group level, i.e. using the full posterior distributions of each parameter displayed in Supplementary Fig. 2A. **(B)** Individual- (session-) level correlations, using the modes of individual parameter distributions. **(C, D)** Parameter correlations between full posteriors of disease- and medication-difference distributions, respectively. These evaluations show that the parameters used in the model are only weakly related (Pearson $r < .4$) and indicates that these parameters can capture different aspects of the observed behaviour during learning.



Supplementary Figure 12 | Parameter recovery of fitted model parameters for each participant. The mode of each participant's posterior parameter distributions was taken and used to simulate 200 new datasets per participant. The actual parameter value (β, $\alpha_{gain,}$ and α_{loss}) used for 200 simulated sets on the x-axis ("True value") is plotted against the parameter modes of the 200 simulation fits on the y-axis ("Simulated mode") for each participant (denoted by a dot). Posterior modes obtained from the hierarchical Bayesian model taken for each participant (HC=red, PD ON=green, PD OFF=blue) represent the true value with which data was simulated per participant and used for recovery. As can be seen, the true and simulated modes were highly correlated across all three parameters (all $p<.001$), showing that this model is well able to recover the actual (original) parameters used for data simulations.

Supplementary Tables

Characteristic	Healthy Controls (HC) N=23	Parkinson patients (PD) N=23	HC vs. PD (difference)
Age (years)	60.35 (8.72)	63.30 (8.24)	$t(44) = 1.182$, p=.244
Gender	8 females	6 females	$\chi^2(1) = .411$, p =.522
Education level (Verhage)	6.09 (0.85)	5.26 (1.14)	$t(44) = 2.793$, p=.008**
MoCA score	27.91 (1.88)	26.96 (1.92)	$t(44) = 1.708$, p=.095
BDI	4.09 (3.03)	6.07 (4.31)	$t(44) = 1.802$, p=.078
BAI	23.17 (2.27)	31.48 (6.15)	$t(44) = 6.077$, p<.001***
Digit Span Backwards	7.00 (2.09)	6.35 (2.39)	$t(44) = .986$, p=.329
NLV reading score	90.96 (8.47)	89.57 (7.36)	$t(43) = .588$, p=.560
Verbal fluency	43.91 (11.26)	37.91 (11.01)	$t(43) = 1.808$, p=.078
Disease duration (years)	-	4.00 (3.18)	-
UPDRS III on	-	17.87 (7.77)	-
UPDRS III off	-	20.65 (8.22)	-
Hoehn & Yahr	-	2.11 (0.58)	-
LEDD (mg)	-	790 (629)	-

Supplementary Table 1 | Demographic and clinical characteristics of participants. Executive functioning was assessed using the following tests: the Montreal Cognitive Assessment (MoCA), the Dutch version (NLV) of the National Adult Reading Test (NART) as a measure of pre-morbid IQ, the Stroop color-word task to assess effects of interference, verbal (category) fluency tests, and the rule-shift cards test of the Behavioural Assessment of the Dysexecutive Syndrome (BADS), to assess mental flexibility (Wilson *et al.*, 1997). The Complex Figure of Rey (CFR) was used as a measure of visuospatial memory. Verbal memory was assessed using the Dutch version of the Rey Auditory Verbal Learning Test (AVLT), testing both short- and long-term verbal memory (Saan and Deelman, 1986). Digit span forwards and backwards in short form (WAIS) was used to assess working memory. Participants also completed several self-report questionnaires: Beck Depression Inventory (BDI), Beck Anxiety Inventory (BAI), and Monetary Choice Questionnaire (MCQ). PD patients additionally completed the Wearing-Off Questionnaire (WOQ-Q10; related to the wearing off of DA medication), and the Questionnaire for Impulsive-Compulsive Disorders in Parkinson's Disease-Rating Scale (QUIP-RS). The motor part of the Unified Parkinson's Disease Rating Scale (UPDRS III) was carried out before each fMRI session. An overview of several test scores is provided in the table below. These assessments were not examined in the current study but are discussed in greater detail elsewhere (Engels *et al.*, 2018a, b). Quantities are presented as the mean across the sample, with brackets denoting 1 standard deviation.

Patient (Nr.)	Age (years)	Disease duration (years)	LEDD (mg)	Medication information			Time to scan since medication (hours)	
				Levodopa	DA-agonist	Other	OFF	ON
1	55	3.5	564	Yes			15.0	1.0
2	73	2.0	752	Yes			17.0	1.0
3	67	10.0	564	Yes			20.5	1.0
4	72	3.0	375	Yes			16.5	1.5
5	68	1.0	828	Yes			16.0	2.5
6	56	2.0	375	Yes			26.5	2.0
7	68	2.0	378	Yes			19.5	5.5
8	65	8.0	850	Yes		MAO-B inhibitor (Rasagiline) COMT inhibitor (Entacapone)	12.5	1.5
9	62	4.0	2780	Yes		COMT inhibitor (Entacapone)	14.5	3.0
10	64	5.5	982	Yes	Pramipexol		14.5	2.0
11	68	2.0	125	Yes			15.5	2.5
12	69	1.0	500	Yes			15.0	8.0
13	73	5.0	375	Yes			13.5	3.5
14	70	3.0	1548	Yes		COMT inhibitor (Entacapone)	15.5	1.5
15	71	6.0	1038	Yes	Pramipexol		8.5	5.5
16	47	6.0	1428	Yes	Ropinirol		14.0	1.0
17	48	0.5	1000	Yes			15.0	1.5
18	56	6.0	935	Yes	Pramipexol	MAO-B inhibitor (Rasagiline)	13.0	1.5
19	66	1.0	90	Yes			16.5	2.0
20	53	5.0	615	Yes	Ropinirol		19.5	1.5
21	72	13.0	1150	Yes	Pramipexol	MAO-B inhibitor (Rasagiline)	16.0	2.0
22	57	1.0	106	No	Pramipexol		18.5	4.5
23	51	6.0	1645	Yes	Ropinirol	Amantadine	14.0	2.5
24	61	1.5	108	No	Pramipexol		27.0	12.0

Supplementary Table 2 | Medication information for PD patients. Patient information regarding levodopa, dopamine agonists and any other dopaminergic medication. Levodopa Equivalent Daily Dosage (LEDD) was calculated according to Tomlinson *et al.*, 2010**.**

Model	Parameters	LOOIC (estimate ± SE)
1	α, β	12432 ± 778
2	α_{gain}, α_{loss}, β	12167 ± 789
3	α_{gain}, α_{loss}, β, π	26325 ± 1476

Supplementary Table 3 | Model comparison. In order to compare the validity of the reported model with three free parameters α_{gain}, α_{loss}, and β, we evaluated two additional hierarchical Bayesian models and used the leave-one-out cross-validation information criterion (LOOIC) procedure for model comparisons. LOOIC is highly recommended for model comparisons of hierarchical Bayesian structures that use MCMC sampling (Vehtari *et al.,* 2017). All models were set up in the same fully hierarchical way as the model shown in the manuscript (Model 2), either in a reduced form (Model 1) to include only two free base parameters (α and β), i.e. with a learning rate that is updated on every trial regardless of a positive or negative outcome, or in extended form (Model 3) to include one additional perseverance ("stickiness") parameter (π), to account for any bias in choosing the same stimulus of a pair regardless of the reward outcome (Kable and Glimcher, 2007; Schönberg *et al.*, 2007). π was included in the softmax equation and was bounded as [-5, 5] in accordance with previous research using this parameter (Wunderlich *et al.*, 2012). A lower LOOIC score indicates a better-fitting model. LOOIC estimates pointwise out-of-sample prediction accuracy from a fitted Bayesian model using the log-likelihood evaluated at the posterior simulations of the parameter values and may be extracted using the "loo" package in R (Gelman et al. 2013, Vehtari et al. 2017, Yao et al., 2017, Ahn et al. 2017). The log-likelihoods were calculated per subject in the Stan model using the log categorical probability mass function. This was updated on a trial-by-trial basis to reflect whether a trial was correct given the probability of choosing that option. These log-likelihoods were then extracted from the fitted model on a per subject basis and used to calculate the LOOIC. Model 2 with three free parameters α_{gain}, α_{loss}, and β (as presented in the manuscript) was found to be the best fitting model.

Explanatory variable	BIC (behavior)	BIC (brain)
$k\alpha_{loss}$ only	231.38	33.36
$k\alpha_{gain} + k\alpha_{loss}$	234.34	36.17
$k\alpha_{gain}$ only	234.48	37.92

Supplementary Table 4 | Summary of BIC values for the role of medication-related shifts in learning rate parameters in subsequent medication-related changes in transfer phase behavior and BOLD activity. BIC values relating to transfer phase behavior (second column) describe the explanatory power of within-patient medication-related shifts in learning rate parameters ($k\alpha_{gain}$, $k\alpha_{loss}$) in the transfer phase medication-related interaction in approach/avoidance behavioral accuracy. BIC values relating to brain activity (third column) describe the explanatory power of the same within-patient medication-related shifts in learning rate parameters in the transfer phase medication-related interaction in caudate nucleus BOLD activity during approach versus avoid trials. Overall, in both brain and behavior, the medication-related shift in only the negative learning rate, α_{loss}, parameter best explained subsequent medication-related changes in approach/avoidance trials.

Parameter	Mean	SEM	95% HDI	BF_{10}
β	0.015	0.003	[-0.213, 0.268]	1.24
α_{gain}	-0.048	0.005	[-0.505, 0.407]	1.39
α_{loss}	-0.960	0.021	[-2.566, 0.338]	11.40

Supplementary Table 5 | Summary of learning phase PD medication differences in posterior distributions of the Bayesian model group parameters. Bayes factors (BF) for medication differences in learning parameter distributions represent the BF according to direction of the visible shift in the posterior difference, i.e., the α_{loss} parameter in Fig. 2B is shifted to the left (higher OFF medication), so the BF represents the probability of OFF > ON being greater than zero. HDI = highest density interval.

# voxels	log10(p)	MAX X (mm)	MAX Y (mm)	MAX Z (mm)	COPE-MAX X (mm)	COPE-MAX Y (mm)	COPE-MAX Z (mm)	Areas included
529	25.6	12.3	-82.8	47.4	35.2	-72.9	54	lateral occipital cortex, precuneous (R)
233	13.7	5.71	-89.3	-15.3	2.43	-79.5	-8.7	V1, occipital fusiform gyrus, occ. pole (R)
101	6.75	-33.7	-43.4	-45	-30.4	-43.4	-45	cerebellum (L)
79	5.35	-17.3	-13.9	11.1	-7.41	9.08	4.5	caudate nucleus, putamen, glob. pallidus, thalamus (L)
75	5.08	-43.5	-30.3	44.1	-27.1	-10.6	67.2	postcent. gyrus, supramarg. gyrus, sup. parietal lobe (L)
64	4.31	32	-69.7	-45	-4.13	-63.1	-48.3	cerebellum (R)
44	2.78	51.6	35.3	30.9	38.5	45.2	37.5	middle front. gyrus, frontal pole (R)
42	2.62	61.5	-43.4	-8.7	61.5	-46.7	-8.7	middle temp. gyrus, inf. temp. gyrus (R)
42	2.62	28.7	-7.32	50.7	32	-4.04	50.7	precent. gyrus, middle front. gyrus, sup. front. gyr (R)
40	2.45	2.43	25.5	54	5.71	18.9	47.4	sup. front. gyr. (R)
39	2.37	38.5	-4.04	44.1	38.5	2.52	37.5	precent. gyrus, middle front. gyrus. sup. front. gyr (R)

Supplementary Table 6 | Medication difference in whole brain RPE signal. Clusters of group-level PD OFF > ON medication difference ($p < .01$, $z = 2.3$, cluster-corrected).

Supplementary References

Abraham A, Pedregosa F, Eickenberg M, Gervais P, Muller A, Kossaifi J, et al. Machine Learning for Neuroimaging with Scikit-Learn. Front. Neuroinform. 2014; 8: 1–10.

Ahn W-Y, Haines N, Zhang L. Revealing neuro-computational mechanisms of reinforcement learning and decision-making with the hBayesDM package. Comput. Psychiatry 2017; 1: 24–57.

Avants BB, Epstein CL, Grossman M, Gee JC. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. Med. Image Anal. 2008; 12: 26–41.

Behzadi Y, Restom K, Liau J, Liu TT. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. Neuroimage 2007; 37: 90–101.

Breiter HC, Aharon I, Kahneman D, Dal A, Shizgal P. Functional Imaging of Neural Resposes to Monetary Gains and Losses. Neuron 2001; 30: 619-.

Brown J, Bullock D, Grossberg S. How the basal ganglia use parallel excitatory and inhibitory learning pathways to selectively respond to unexpected rewarding cues. J. Neurosci. 1999; 19: 10502–10511.

Cools R, Lewis SJG, Clark L, Barker RA, Robbins TW. L-DOPA disrupts activity in the nucleus accumbens during reversal learning in Parkinson's disease. Neuropsychopharmacology 2007; 32: 180–189.

Cox SML, Frank MJ, Larcher K, Fellows LK, Clark CA, Leyton M, et al. Striatal D1 and D2 signaling differentially predict learning from positive and negative outcomes. Neuroimage 2015; 109: 95–101.

Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis: I. Segmentation and surface reconstruction. Neuroimage 1999; 9: 179–194.

Daw ND, Kakade S, Dayan P. Opponent interactions between serotonin and dopamine. Neural Networks 2002; 15: 603–616.

Engels G, McCoy B, Vlaar A, Theeuwes J, Weinstein H, Scherder E. Clinical pain and functional network topology in Parkinson's disease : a resting-state fMRI study. J. Neural Transm. 2018a; 0: 0.

Engels G, Vlaar A, McCoy B, Scherder E, Douw L. Dynamic functional connectivity and symptoms of Parkinson's disease: a resting-state fMRI study. Front. Aging Neurosci. 2018b; 10

Fonov VS, Evans AC, McKinstry RC, Almli CR, Collins DL. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. Neuroimage 2009; 47

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

Gelman A, Rubin DB. Inference from Iterative Simulation Using Multiple Sequences. Stat. Sci. 1992; 7: 457–472.

Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB (2013). Bayesian data analysis; CRC press.

Gorgolewski K, Burns CD, Madison C, Clark D, Halchenko YO, Waskom ML, et al. Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python [Internet]. Front. Neuroinform. 2011; 5Available from: http://journal.frontiersin.org/article/10.3389/fninf.2011.00013/abstract

Gorgolewski K, Esteban O, Burns C, Zeigler E, Pinsard B, Madison C, et al. Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in Python. 0.13.0 [Internet]. Zenodo 2017Available from: 10.5281/zenodo.581704

Gramfort A, Luessi M, Larson E, Engemann DA, Strohmeier D, Brodbeck C, et al. MEG and EEG data analysis

with MNE-Python. Front. Neurosci. 2013; 7: 1–13.

Gramfort A, Luessi M, Larson E, Engemann DA, Strohmeier D, Brodbeck C, et al. MNE software for processing MEG and EEG data. Neuroimage 2014; 86: 446–460.

Greve DN, Fischl B. Accurate and robust brain image alignment using boundary-based registration. Neuroimage 2009; 48: 63–72.

Hazy TE, Frank MJ, O'Reilly RC. Neural mechanisms of acquired phasic dopamine responses in learning. Neurosci. Biobehav. Rev. 2010; 34: 701–720.

Hoffman MD, Gelman A. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. J. Mach. Learn. Res. 2016; 15: 1351–1381.

de Hollander G, Knapen T. nideconv [Internet]. 2017Available from: https://response-fytter.readthedocs.io/en/latest/index.html

Huntenburg JM, Gorgolewski KJ, Anwander A, Margulies DS. Evaluating nonlinear coregistration of BOLD EPI and T1 images [Internet]. Proc. Organ. Hum. Brain Mapp. 2012Available from: http://hdl.handle.net/11858/00-001M-0000-002B-1CB5-A

Jahfari S, Ridderinkhof KR, Collins AGE, Knapen T, Waldorp LJ, Frank MJ. Cross-task contributions of fronto-basal ganglia circuitry in response inhibition and conflict-induced slowing. Cereb. Cortex 2018: 1–15.

Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 2002; 17: 825–841.

Kable JW, Glimcher PW. The neural correlates of subjective value during intertemporal choice. Nat. Neurosci. 2007; 10: 1625–1633.

Klein A, Ghosh SS, Bao FS, Giard J, Häme Y, Stavsky E, et al. Mindboggling morphometry of human brains. 2017.

Lefebvre G, Lebreton M, Meyniel F, Bourgeois-Gironde S, Palminteri S. Behavioural and neural characterization of optimistic reinforcement learning. Nat. Hum. Behav. 2017; 1: 1–9.

McClure SM, Berns GS, Montague PR. Temporal prediction errors in a passive learning task activate human striatum. Neuron 2003; 38: 339–346.

Mumford J a, Turner BO, Ashby FG, Poldrack R a. Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. Neuroimage 2012; 59: 2636–43.

O'Doherty J, Dayan P, Friston K, Critchley H, Dolan RJ. Temporal Difference Models and Reward-Related Learning in the Human Brain. 2003; 28: 329–337.

O'Doherty JP, Cockburn J, Pauli WM. Learning, Reward, and Decision Making. Annu. Rev. Psychol. 2017; 68: 73–100.

den Ouden HE, Daw ND, Fernandez G, Elshout JA, Rijpkema M, Hoogman M, et al. Dissociable effects of dopamine and serotonin on reversal learning. Neuron 2013; 80: 1090–100.

Pauli WM, Nili AN, Tyszka JM. A high-resolution probabilistic in vivo atlas of human subcortical brain nuclei. Sci. Data 2018: 1–13.

Power JD, Mitra A, Laumann TO, Snyder AZ, Schlaggar BL, Petersen SE. Methods to detect, characterize, and remove motion artifact in resting state fMRI. Neuroimage 2014; 84: 320–341.

Saan RJ, Deelman BG. De Nieuwe 15-Woordentest (A en B). Een Handleiding. 1986.

Schönberg T, Daw ND, Joel D, O'Doherty JP. Reinforcement learning signals in the human striatum distinguish

learners from nonlearners during reward-based decision making. J. Neurosci. 2007; 27: 12860–12867.

Schultz W, Dayan P, Montague PR. A Neural Substrate of Prediction and Reward. Science (80-.). 1997; 275: 1593–1600.

Van Slooten JC, Jahfari S, Knapen T, Theeuwes J. How pupil responses track value-based decision-making during and after reinforcement learning. PLoS Comput. Biol. 2018; 14: 1–24.

Stan Development Team. Stan Modeling Language User ' s Guide and Reference Manual (v. 2.6.2). 2015

Tomlinson CL, Stowe R, Patel S, Rick C, Gray R, Clarke CE. Systematic review of levodopa dose equivalency reporting in Parkinson's disease. Mov. Disord. 2010; 25: 2649–2653.

Treiber JM, White NS, Steed TC, Bartsch H, Holland D, Farid N, et al. Characterization and correction of geometric distortions in 814 Diffusion Weighted Images. PLoS One 2016; 11: 1–9.

Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, et al. N4ITK: Improved N3 bias correction. IEEE Trans. Med. Imaging 2010; 29: 1310–1320.

Vehtari A, Gelman A, Gabry J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Statistics and Computing 2017; 27: 1413–1432.

Wang S, Peterson DJ, Gatenby JC, Li W, Grabowski TJ, Madhyastha TM. Evaluation of Field Map and Nonlinear Registration Methods for Correction of Susceptibility Artifacts in Diffusion MRI. Front. Neuroinform. 2017; 11: 1–9.

Wunderlich K, Smittenaar P, Dolan RJ. Dopamine Enhances Model-Based over Model-Free Choice Behavior. Neuron 2012; 75: 418–424.

Yao Y, Vehtari A, Simpson D, Gelman A. Using stacking to average Bayesian predictive distributions. Bayesian Analysis 2017; doi: 10.1214/17-BA1091.

Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans. Med. Imaging 2001; 20: 45–57.