Supplemental Methods

Computational modelling

As described in the main methods, five models of increasing complexity were fitted to quantify putative mechanisms that may drive motivational bias.

For all models, action weights (w) are estimated for each response option (a) for all trials (t) per cue (s). Choice probabilities are computed using a softmax function based on these action weights:

$$p(a_t|s_t) = \left[\frac{\exp\left(w(a_t, s_t)\right)}{\sum_{a'} \exp\left(w(a', s_t)\right)}\right] \quad (1)$$

In the simplest, Rescorla-Wagner, model (M1) the action weights are fully determined by the learned action values (Q-values). These action values are learned through a standard delta-rule learning with two free parameters: a learning rate (ε) which scales the update term, and feedback sensitivity (ρ) scaling the outcome value (comparable to the softmax temperature):

$$Q_t(a_t, s_t) = Q_{t-1}(a_t, s_t) + \varepsilon(\rho r_t - Q_{t-1}(a_t, s_t)) \quad (2)$$

Outcomes are reflected by r, which incorporates negative, neutral and positive outcomes: $r \in (-1,0,1)$. As cue valence was instructed (through green and red cue edges), initial Q-values (Q_0) are set to $\rho * 0.5$ for Win cues and $\rho * -0.5$ for Avoid cues.

In M2 we add a go bias parameter (b) to allow for a differential 'base rate' of Go responding across individuals, independent of valence.

$$w(a_t, s_t) = \begin{matrix} Q(a_t, s_t) + b & if\ a = Go \\ Q(a_t, s_t) & else \end{matrix} \quad (3)$$

In models M3-5, we implement different mechanisms through which motivational valence could affect choice. In M3 a motivational bias parameter π is added that modulates action weights according to cue valence (V). For positive values of π, action weights for Go actions are increased the weight of Go responses for Win and decreased for Avoid cues:

$$w(a_t, s_t) = \begin{matrix} Q(a_t, s_t) + \pi V_{(S)} + b & if\ a = Go \\ Q(a_t, s_t) & else \end{matrix} \quad (4)$$

$$V_{(S)} = \ 0.5 \quad if\ s = win\ cue$$
$$V_{(S)} = -0.5 \quad if\ s = avoid\ cue$$

Note that V is fixed because cue valence is instructed.

In model M4 and M5 we extend M3 to explore whether there is additional evidence for differential learning based as a function of outcome or cue valence, to test for previously observed effects of dopaminergic medication on reward versus punishment learning (Frank *et al.*, 2004; Cools *et al.*, 2006) . In M4, the learning rate depends on sign of the prediction error; any outcome that is better than expected results in a positive learning rate ε_{win} (i.e. a neutral outcome for Avoid cues, or a win for Win cues), while impact of outcomes that are worse than expected, will be governed by for ε_{loss}.

$$\varepsilon = \begin{cases} \varepsilon_{win}: & if\ r = 1 \ \ \& \ \ s = win\ cue \ \ \text{OR} \ \ if\ r = 0 \ \ \& \ s = avoid\ cue \\ \varepsilon_{loss}: & if\ r = -1 \ \ \& \ s = avoid\ cue \ \ \text{OR} \ \ if\ r = 0 \ \ \& \ \ s = win\ cue \end{cases} \quad (5)$$

In contrast, in model M5, the two learning rates are based on *cue valence*, so that patients may learn differently from outcomes on Win trials relative to Avoid trials:

$$\varepsilon = \begin{cases} \varepsilon_{win}: if \ \ s = win\ cue \\ \varepsilon_{avoid}: if\ s = avoid\ cue \end{cases} \quad (6)$$

To estimate model parameters and model fit, we used an MCMC sampling- based method for hierarchical Bayesian estimation of group-level and participant-level parameters. Here, group-level parameters (X) serve as priors for the individual-level parameters(x), such that $x \sim \mathcal{N}(X, \sigma)$. The hyperpriors for s are specified by a half-Cauchy (Gelman, 2006) with a scale of 2. The hyperpriors for X are centered around 0 (with the exception of (X_ρ) and weakly informative: $X_\rho \sim \mathrm{N}\mathcal{N}(2,3)$, $X_\varepsilon \sim \mathrm{N}\mathcal{N}(0,2)$, $X_{b,\pi} \sim \mathrm{N}\mathcal{N}(0,3)$. Parameters b, π are unconstrained, ρ was constrained to be positive through and exponential transform, learning rates ε were constrained to [0 1] through an inverse logit transform.

Model estimation procedure was identical to (Swart *et al.*, 2017), using Stan software in R (RStan) (Stan-Development-Team, 2016). Stan provides full Bayesian inference with Markov chain Monte Carlo (MCMC) sampling methods (Metropolis *et al.*, 1953). The number of Markov chains was set at 4, with 200 burn-in iterations and 1000 post burn-in iterations per chains (4000 total). Model convergence was considered when the potential scale reduction factor R^ < 1.1 for all parameters (Gelman and Rubin, 1992). Model comparison was evaluated using the Watanabe-Akaike Information Criteria (WAIC) (Watanabe, 2010). WAIC is an estimate of the likelihood of the data given the model parameters, penalized for the effective number of parameters to adjust for overfitting. Lower (i.e. more negative) WAIC values indicate better model fit. As WAIC is reported on a deviance scale (Gelman *et al.*, 2014), a difference in WAIC value of 2–6 is considered positive evidence, 6–10 strong evidence, and >10 very strong evidence (Kass and Raftery, 1995).

Nuisance and Confound analyses

To quantify relevant clinical or demographic differences between groups (controls vs. patients, and tremor vs. non-tremor patients) we used a series of two-tailed T-tests for our continuous variables – cf. table 1 – main article). When a difference between groups was detected, we followed this up with extra control analyses dedicated to this particular variable, which will be reported here. Specifically, we observed a difference in the delay between medication intake and behavioural testing. To assess whether this difference, putatively leading to a difference in effective dopamine levels, could explain patient group differences, we reanalysed the behavioural data using a smaller subset of tremor patients that was matched with respect to the delay. The short-delay tremor subgroup consisted of 15 patients, with a mean delay of 139 mins (std 17, range : 101-155). The delay in this matched subgroup did not significantly differ from the non-tremor patients (2-sample t-test t(33) = 0.84 , p = .4). Note that for this analysis, we collapsed across levodopa responsive and non-responsive Tremor patients, so attain large enough subgroups. This is warranted by the absence of any significant differences between these groups. We further compare behaviour between the matched and non-matched subsets, to see whether these significantly differ. Finally, to assess the robustness of our findings, our main ANOVA was extended with nuisance variables drug-delay, LEDD and age (covariates), and gender (factor).

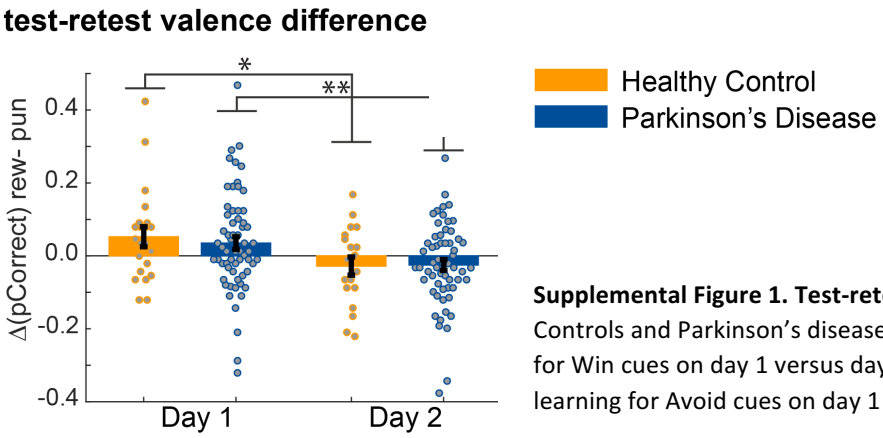
Supplemental Results

Supplemental table 1. Data availability per group and testing session.

Group	total N	Data available		
		both sessions	session 1 only	session 2 only
healthy control	22	19	3	0
tremor nonresponsive	20	19	0	1
tremor responsive	23	21	2	0
non tremor	20	18	2	0

Test-retest effect

The degree to which participants were able to learn from reward versus punishment feedback changed over testing days (valence x day: [$F(1,76)=14.4$, $\eta^2=0.16$, $p<.001$], see Supplemental Figure 1), such that participants learnt better for Win cues on day 1 (Day 1: Win vs. Avoid, [$F(1,76)=8.1$, $\eta^2=0.10$, $p=.006$] and better for Avoid cues on day 2 (Day 2: Win vs. Avoid, [$F(1,76)=4.3$, $\eta^2=0.05$, $p=0.042$]. This interaction was significant for both patients (Valence x Day; [$F(1,57)=9.2$, $\eta^2=0.14$, $p=0.004$] and controls (Valence x Day; $F(1,18)=5.1$, $\eta^2=0.22$, $p=.036$]. Because of this significant difference between performance on day 1 and 2 in terms of valence effects, we added 'testing order' as a factor in the within-participant analysis of patient data to assess whether potential test-retest effects interacted with effects of medication. There was a Valence x Medication x TestOrder interaction [$F(1,54)= 8.9$, $\eta^2=0.14$, $p=.004$], in addition to Valence x Action x a Medication x TestOrder interaction [$F(1,54)= 5.3$, $\eta^2=0.09$, $p=.025$]. Here, it should be noted that a Medication x TestOrder interaction is mathematically identical to a main effect of testing day. Given the presence of such an effect of testing day in the healthy controls, it is not possible to meaningfully interpret the Medication x Testorder interaction. Therefore, we restricted our analysis to data from day 1, considering Medication as a between-participants factor instead of a within-participants factor.

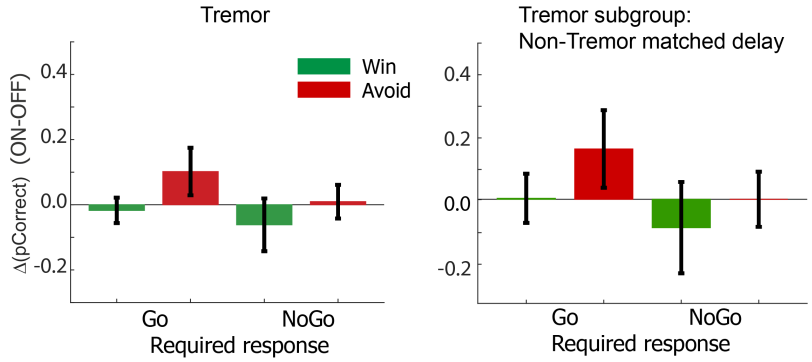


Supplemental Figure 1. Test-retest differences. Both Healthy Controls and Parkinson's disease patients show better learning for Win cues on day 1 versus day 2, while showing worse learning for Avoid cues on day 1 relative to day 2.

Robustness analyses

Drug-delay matched subgroup analysis

As reported in the main, manuscript, there was a difference in average "task-delay" (i.e. the delay between medicine administration and the onset of the behavioural task) of approximately 30 minutes. This is a potential confound for group comparison, as longer time between drug intake and task onset could affect levodopa levels during the task. To assess this potential confound (in addition to adding task-delay as a covariate of no interest, as reported above), we test whether the main results are altered when restricting analysis to a subgroup of tremor patients with delay that matches the delay of the non-tremor group. As shown in supplemental figure 1, the results do not change. Using this matched subgroup, the impact of medication on valence was still strongly modulated by patient-group [Tremor Group x Medication x Valence: $F(1,31)= 12.0$, $\eta^2=0.28$, $p=.002$], where the tremor sub-group still showed a significant modulation by levodopa of performance on Win versus Avoid cues [Medication x Valence $F(1,13)= 8.05$, $\eta^2=0.38$, $p=0.014$). Furthermore, there is still no main effect of levodopa medication on the Win versus the Avoid trials [$F(1,31)= 0.05$, $\eta^2=0.001$, $p=.8$].



Supplemental Figure 1. Differential performance associated with levodopa administration (positive = higher performance On levodopa). There is no difference in effects between the overall tremor patient group and the subgroup of patients who were matched in terms of delay between drug intake and behavioural testing.

Nuisance variables

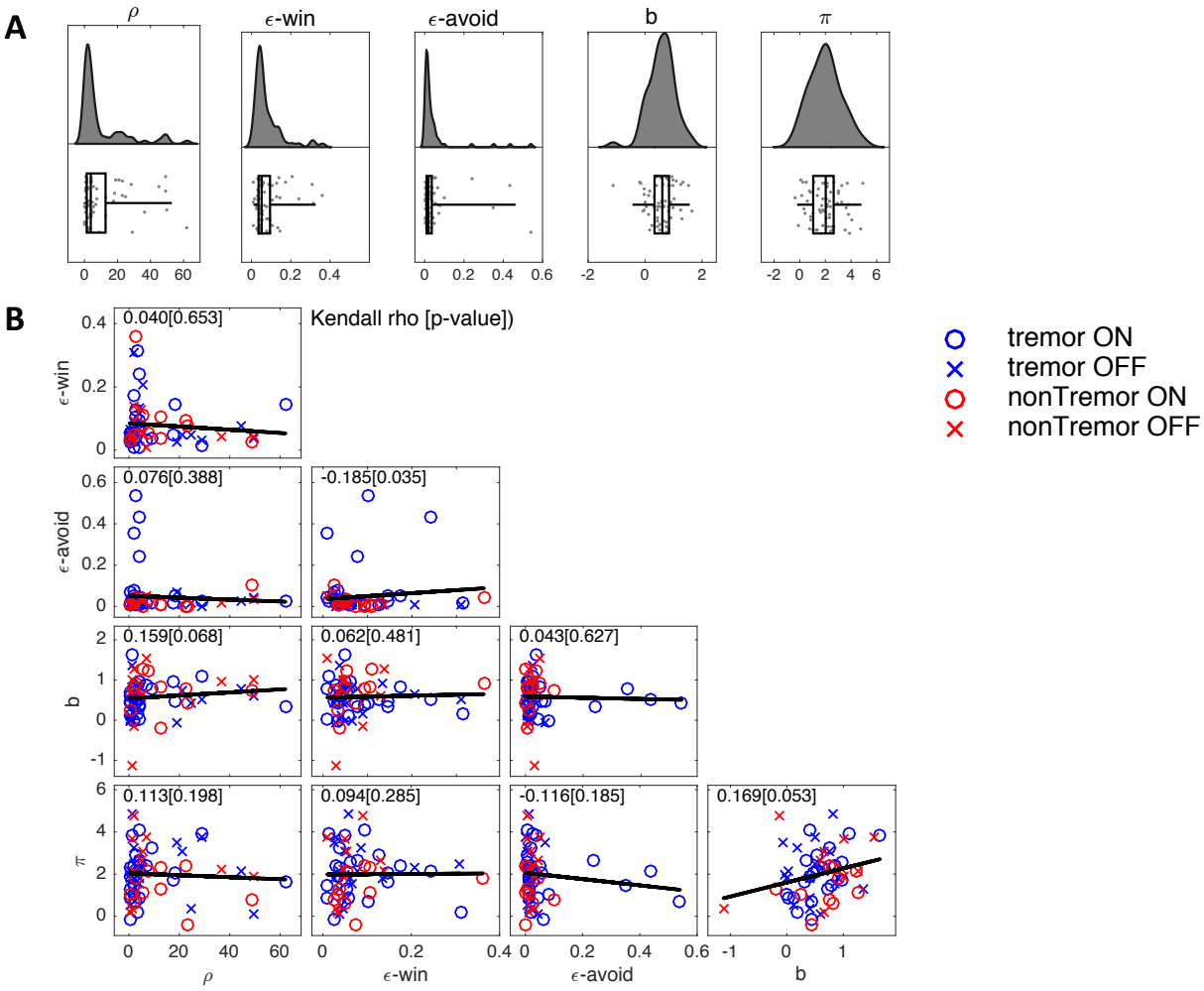
The main ANOVA with factors Action, Valence, Medication, Patient Group (3 levels) was repeated using the nuisance covariates age, drug-delay and LEDD, and nuisance factor gender. Most importantly, inclusion of these nuisance variables did not change our main findings: The impact of medication on valence was still strongly modulated by patient-group [Group x Medication x Valence: $F(1,47)= 6.2$, $\eta^2=0.21$, $p=.004$, while globally, levodopa medication did not significantly alter performance on Win versus the Avoid trials [Medication x Valence: $F(1,47)= 0.2$, $\eta^2=0.004$, $p=.7$]. Regarding the nuisance variables themselves, accuracy decreased with increasing age [$F(1,47)=12.0$, $\eta^2=0.20$, $p<.001$]. There were no significant main effects of gender [$F(1,47)=1.2$, $\eta^2=0.03$, $p=.28$], drug-delay [$F(1,47)=0.06$, $\eta^2=0.001$, $p=.8$], or LEDD ([$F(1,47)=0.3$, $\eta^2=0.006$, $p=.6$]). Given that LEDD may reflect baseline hypodopaminergic state, it could potentially have a direct effect on dopamine-sensitive behaviour. We therefore also assessed whether LEDD predicted performance independent of medication, testing the hypothesis that variability in LEDD would have a similar effect to medication administration. LEDD did not interact with Valence ([$F(1,47)=1.4$, $\eta^2=0.029$, $p=.25$]).

Day 2 analyses

For completeness, we also performed the primary ANOVA on data of day 2 only (3 patient levels). The main, medication-independent results on the task replicate, with strong evidence for good learning of the task (RequiredAction: $F(1,53)= 83.1$, $\eta^2=0.61$, $p< .001$), and presence of a motivational bias (RequiredAction x Valence $F(1,53)= 133.0$, $\eta^2=0.71$, $p<.001$). Furthermore, as on day 1, data failed to replicate previous reports that levodopa medication improved performance on the Win versus the Avoid trials [Valence x Medication $F(1,53)= 0.04$, $\eta^2=0.001$, $p=.8$]. However, in contrast to day 1, there was no modulation by patient group of the interaction of medication and valence (Patient Group x Medication x Valence: $F(1,53)= 1.5$, $\eta^2=0.05$, $p=.23$). Absence of this effect on day 2 may be due to the observed test-retest differences in Valence-dependent learning described in the main manuscript, i.e. a shift from better learning from Win cues on day 1 to better learning from Avoid cues on day 2. Importantly, this shift in baseline performance was also observed in healthy controls, and may mask or affect any effects of medication.

Computational modelling – parameter distributions

We assessed whether parameter estimates were normally distributed using a Kolmogorov-Smirnov test on the (z-scored) parameter estimates. The learning rate and outcome sensitivity parameter distributions deviated significantly from normal (ε_{win}: $p = 4e^{-5}$; ε_{avoid}: $p = .005$; ρ: $p = 7e^{-7}$), while the motivational and go bias parameters did not significantly deviate (π: p=.9; gobias p=.6; see supplemental figure 2A). To assess the independence of the parameters, we report the covariance between each parameter pair. Given the non-parametric distributions of the parameters (for 9/10 parameter pairs), we report non-parametric Kendall correlations. Only the learning rates ε_{win} and ε_{avoid} were significantly correlated (p = 0.035). Given that these learning rates are estimated on independent data (Win vs Avoid trials), this correlation does not stem from a problem with identifiability of the parameters, but rather reflects a true underlying property of the individual; some people learn faster than others. However, note that this p-value does not survive correction for multiple comparisons for 10 correlations.



Supplemental figure 2. A. parameter distributions for the winning model M5. density plots and individual data point for all parameters. ε_{win}, ε_{avoid} and ρ distributions deviated significantly from normal. **B. Parameter covariance for the winning model M5.** The only significant correlation is between ε_{win} and ε_{avoid}. Correlations are computed across all patients, but markers illustrate different patients group/conditions.

Supplemental References

Gelman A, Hwang J, Vehtari A. Understanding predictive information criteria for Bayesian models. Stat Comput 2014; 24: 997–1016.

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

Kass RE, Raftery AE. Bayes Factors. J Am Stat Assoc 1995; 90: 773–795.

Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of State Calculations by Fast Computing Machines. J Chem Phys 1953; 21: 1087–1092.

Swart JC, Froböse MI, Cook JL, Geurts DE, Frank MJ, Cools R, et al. Catecholaminergic challenge uncovers distinct Pavlovian and instrumental mechanisms of motivated (in)action. eLife 2017; 6: e22169.

Watanabe S. Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory. J Mach Learn Res 2010; 11: 3571–3594.