

A Selective Role for Dopamine in Learning to Maximize Reward But Not to Minimize Effort: Evidence from Patients with Parkinson's Disease

Vasilisa Skvortsova,^{1,2,3}  Bertrand Degos,⁴ Marie-Laure Welter,^{2,3,4} Marie Vidailhet,⁴ and Mathias Pessiglione^{1,2,3}

¹Motivation, Brain and Behavior Laboratory, Brain and Spine Institute, Paris, 75013, France, ²INSERM U1127, Centre National de la Recherche Scientifique, Unité Mixte de Recherche 7225, Paris, 75013, France, ³Université Pierre et Marie Curie, Paris 6, 75013, Paris, France, and ⁴Neurology Department, Centre Inter-Régional de Coordination de la Maladie de Parkinson, Hôpital de la Pitié-Salpêtrière, Assistance Publique Hôpitaux de Paris, 75013, Paris, France

Instrumental learning is a fundamental process through which agents optimize their choices, taking into account various dimensions of available options such as the possible reward or punishment outcomes and the costs associated with potential actions. Although the implication of dopamine in learning from choice outcomes is well established, less is known about its role in learning the action costs such as effort. Here, we tested the ability of patients with Parkinson's disease (PD) to maximize monetary rewards and minimize physical efforts in a probabilistic instrumental learning task. The implication of dopamine was assessed by comparing performance ON and OFF prodopaminergic medication. In a first sample of PD patients ($n = 15$), we observed that reward learning, but not effort learning, was selectively impaired in the absence of treatment, with a significant interaction between learning condition (reward vs effort) and medication status (OFF vs ON). These results were replicated in a second, independent sample of PD patients ($n = 20$) using a simplified version of the task. According to Bayesian model selection, the best account for medication effects in both studies was a specific amplification of reward magnitude in a Q-learning algorithm. These results suggest that learning to avoid physical effort is independent from dopaminergic circuits and strengthen the general idea that dopaminergic signaling amplifies the effects of reward expectation or obtainment on instrumental behavior.

Key words: dopamine; effort learning; modeling; Parkinson's disease; reinforcement learning; reward learning

Significance Statement

Theoretically, maximizing reward and minimizing effort could involve the same computations and therefore rely on the same brain circuits. Here, we tested whether dopamine, a key component of reward-related circuitry, is also implicated in effort learning. We found that patients suffering from dopamine depletion due to Parkinson's disease were selectively impaired in reward learning, but not effort learning. Moreover, anti-parkinsonian medication restored the ability to maximize reward, but had no effect on effort minimization. This dissociation suggests that the brain has evolved separate, domain-specific systems for instrumental learning. These results help to disambiguate the motivational role of prodopaminergic medications: they amplify the impact of reward without affecting the integration of effort cost.

Introduction

Instrumental learning is a fundamental process through which animals get to obtain rewards and avoid punishments. Simple

formal models of instrumental learning assume that action selection is based on the comparison of hidden values, which are updated following a delta rule when choice outcomes are delivered (Sutton and Barto, 1998). A delta rule means that value update is proportional to prediction error (actual value minus expected value of the outcome). Numerous studies have investigated the neural underpinnings of such instrumental learning processes in mammals. One key finding is that midbrain dopamine neurons seem to signal reward prediction error (RPE)

Received June 29, 2016; revised Jan. 24, 2017; accepted Jan. 26, 2017.

Author contributions: M.P. designed research; V.S. performed research; V.S. analyzed data; V.S., B.D., M.-L.W., M.V., and M.P. wrote the paper.

This work was supported by the European Research Council (Starting Grant BioMotiv). V.S. received funding from the Ecole de Neurosciences de Paris (ENP) and the Fondation Schlumberger. We thank Stefano Palminteri for insightful suggestions and Sophie Aix for the help with data collection.

The authors declare no competing financial interests.

Correspondence should be addressed to Mathias Pessiglione, Institut du Cerveau et de la Moelle épinière, Hôpital de la Pitié-Salpêtrière, 47 Boulevard de l'Hôpital, 75013 Paris, France. E-mail: mathias.pessiglione@gmail.com.

DOI:10.1523/JNEUROSCI.2081-16.2017

Copyright © 2017 the authors 0270-6474/17/376087-11\$15.00/0

across species (Hollerman and Schultz, 1998; Zaghoul et al., 2009; Eshel et al., 2015). Moreover, direct activation of dopamine neurons using optogenetics or microstimulation is sufficient to reinforce the behavior (Tsai et al., 2009; Steinberg et al., 2013; Arsenault et al., 2014). Consistently, systemic administration of drugs boosting or mimicking dopamine release (levodopa or dopamine receptor agonists) improves learning from positive but not negative outcome (Pessiglione et al., 2006; Rutledge et al., 2009; Schmidt et al., 2014). Prodopaminergic medication was even shown to have detrimental effects on punishment learning, in stark contrast to the beneficial effects on reward learning (Frank et al., 2004; Bódi et al., 2009; Palminteri et al., 2009).

Although differential implication of dopamine in reward versus punishment learning is well established, less is known about how dopamine could affect other dimensions of choice options, notably physical effort. In principle, the same computations (delta rule) could be used to learn how much effort is associated with each choice option. Whether dopamine participates in this effort learning process has never been assessed to our knowledge. In decision-making tasks that involve choosing between high reward/high effort and low reward/low effort options, dopamine depletion shifts preferences toward making less effort, whereas dopamine enhancers produce the opposite effect in both rodents (Denk et al., 2005; Salamone et al., 2007; Floresco et al., 2008) and humans (Wardle et al., 2011; Treadway et al., 2012; Chong et al., 2015). These observations suggest that dopamine might be involved in promoting willingness to produce effort and, by extension, that dopamine could play a role in effort learning. However, a modulation of reward weight in cost/benefit trade-offs would be sufficient to explain the effects on willingness to work (Le Bouc et al., 2016). This parsimonious explanation would be consistent with voltammetric or electrophysiological recordings in mid-brain dopaminergic nuclei showing higher sensitivity to reward than effort (Gan et al., 2010; Pasquereau and Turner, 2013; Varazzani et al., 2015). In instrumental learning paradigms, fMRI studies have reported correlates of reward prediction and prediction errors in dopaminergic projection targets such as ventral striatum and ventromedial prefrontal cortex (O'Doherty et al., 2003; Pessiglione et al., 2006; Rutledge et al., 2010; Chowdhury et al., 2013), whereas correlates of punishment prediction and prediction errors have been observed in the anterior insula and dorsal anterior cingulate (Seymour et al., 2004; Nitschke et al., 2006; Palminteri et al., 2012; Harrison et al., 2016). Because these punishment-related regions overlap with the network activated in relation to effort (Crosson et al., 2009; Kurniawan et al., 2010; Prévost et al., 2010; Skvortsova et al., 2014; Scholl et al., 2015), it could be inferred that, if anything, dopamine should have a detrimental effect on effort learning, as was reported for punishment learning.

To clarify the role of dopamine in effort learning, we compared the performance of PD patients with and without dopaminergic medication ("ON" and "OFF" states). Patients performed an instrumental learning task that involves both maximizing reward and minimizing effort. The task was used in a previous fMRI study (Skvortsova et al., 2014), which showed that although reward and effort learning can be accounted for by similar computational mechanisms relying on partially distinct neural circuits. Here, we found in a first study using the exact same task that dopaminergic medication improves reward but not effort learning. This result was replicated in a second study for which the task was simplified and thus was better adapted to patients. Finally, we used computational modeling and Bayesian model selection

Table 1. Demographic and clinical data

	Study 1	Study 2	
	PD	PD	Controls
Laterality (R/L)	14/1	18/2	17/1
Sex (F/M)	1/14	7/13	7/11
Age (y)	59.75 (6.48)	57.7 (10.18)	59.5 (7.8)
Education (y)	5.8 (1.9)	4.84 (2.03)	5.4 (2.4)
Disease duration (y)	9.8 (2.51)	10.35 (2.72)	—
MMS score	27.73 (2.71)	28.25 (1.94)	29.2 (0.71)
UPDRS III score OFF	30.27 (5.84)	32.85 (9.46)	—
UPDRS III score ON	5.73 (4.83)	7.25 (3.68)	—
MDRS	138 (8.50)	140.3 (4.38)	138 (4.2)
WCST score	16.73 (5.82)	17.7 (2.74)	—
Starkstein apathy scale score (ON state)	6 (2.04)	7.8 (3.72)	7.94 (3.0)
Levodopa equivalent dosage (mg)	1121.8 (337.32)	1159.2 (331.82)	—

MMS, Mini-Mental Status; MDRS, Mattis Dementia Rating Scale; WCST, Wisconsin Card Sorting Test. Values in parentheses are SD.

(BMS) procedures to better characterize the differential implication of dopamine in reward versus effort learning.

Materials and Methods

Subjects. All patients (Study 1: $n = 15$, 1 female, age 59.75 ± 6.48 years; Study 2: $n = 20$, 7 female, age 57.7 ± 10.18 years) gave their informed consent to participate in the study. Patients were screened by a certified neurologist, psychiatrist, and neuropsychologist for the absence of dementia, depression, and neurological disorders other than PD. Clinical and demographic data are summarized in Table 1. All patients were candidates for deep-brain stimulation as an alternative treatment for their Parkinson's disease (PD). Presurgery screening included an overnight withdrawal from dopaminergic medication followed by levodopa tolerance test in the next morning. All patients were tested twice during the morning (~8:00 A.M.) after the overnight withdrawal ("OFF" state) and after intake of their usual dose of dopaminergic medication ("ON" state). The last dose was therefore taken at least 10 h before testing (before 10:00 P.M. on the preceding day) in the OFF state and ~1 h before testing (~7:00 A.M. on the same day) in the ON state. Session order (ON–OFF or OFF–ON) was counterbalanced between patients. Most of the patients (10/15 in Study 1 and 16/20 in Study 2) received a daily combined medication that consisted of both levodopa and various dopamine agonists. The rest of the patients were treated with levodopa only.

To give a reference point for patients' performance, we also tested 18 age-matched control subjects (7 females, age 59.5 ± 7.8 years) in Study 2. Control participants were recruited from the community and were not taking any dopaminergic medication at the time of the study. They were screened for the absence of acute depression and dementia using questionnaires (Table 1). Control participants were tested only once on the second version of the task. For both patients and controls, monetary rewards received during the tasks were purely virtual. Note that virtual money was shown to induce qualitatively similar effects as real rewards in reinforcement learning, intertemporal choice, or incentive motivation paradigms (Frank et al., 2004; Bickel et al., 2009; Chong et al., 2015).

Behavioral data. Physical effort was implemented using homemade power grips with two mold wood cylinders compressing an air tube, which was connected to a transducer. Isometric compression was translated into a differential voltage signal proportionally to the exerted force and read by custom-made MATLAB functions. Task presentation was coded using the Cogent 2000 toolbox for MATLAB (Wellcome Trust for Neuroimaging, London). Feedback on the produced force was translated visually to subjects via a mercury level moving up and down within a thermometer drawn on the computer screen. Before the task, patients were required to squeeze the handgrip as hard as possible during a 5 s period, twice with their left and right hands. Maximal force for each hand (F_{max}) was computed as the average over the data points above the median across the two contractions. These maximal force measures were acquired separately for ON and OFF states and were used to normalize the forces produced during the task.

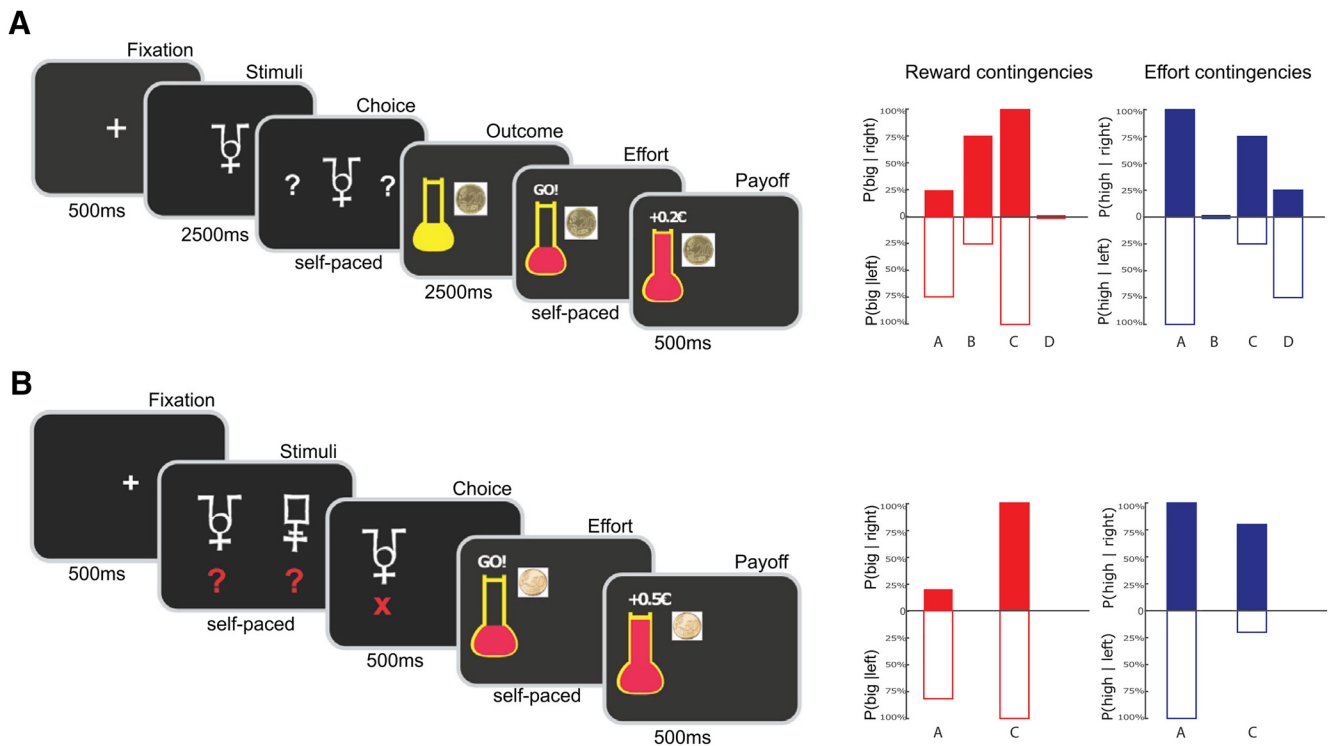


Figure 1. Behavioral tasks used in Study 1 (A) and Study 2 (B). Successive screenshots from left to right illustrate the timing of stimuli and responses for one example trial. When interrogation dots appeared on screen, subjects had to choose between right and left options. Each option was associated with both monetary reward and physical effort. For the chosen option only (left in the example), reward level was indicated by an image of corresponding coin and effort level by the horizontal bar on the thermometer. At the GO! signal, subjects had to squeeze the chosen handgrip (left in the example) until the red fluid level reached the horizontal bar. At that moment, subjects were notified that the reward was added to their cumulative payoff. Changes between the two studies relate to the symbolic cues (one per condition in Study 1 vs one per option in Study 2), the timing (subjects had to wait for a fixed delay before responding in Study 1, whereas they could choose and squeeze as soon as cues and outcomes were displayed in Study 2), and the reward levels (10¢ and 20¢ in Study 1 vs 10¢ and 50¢ in Study 2). Effort levels (20% and 80% of maximal force) were unchanged. Bar graphs on the right illustrate the contingencies between symbolic cues and both reward (red) and effort (blue) outcomes (left and right graphs). Bars indicate the probability of getting high/effort outcomes (or one minus the probability of getting low reward/effort outcomes) for left and right options (empty and filled bars, below and above the x-axis). In Study 1, there were four different contingency sets cued by four different symbols (A–D), whereas in Study 2, there were only two contingency sets cued by two different pairs of symbolic cues (A, C). In each contingency set, one dimension (reward or effort) was fixed such that learning was only possible for the other dimension. Reward learning was assessed by red sets (A, B in Study 1; A in Study 2) and effort learning by blue sets (C, D in Study 1; C in Study 2). The illustration only applies to one task session. Contingencies were fully counterbalanced across the four sessions.

Task version 1. The first version of the task (version 1) was identical to the task previously used in a published fMRI study (Skvortsova et al., 2014) except for suppression of temporal jitters, which served to optimize the fit of hemodynamic response. It was a probabilistic instrumental learning task with binary choices (left or right) and the following four possible outcomes: two reward levels (20¢ or 10¢) times two effort levels (80% and 20% of F_{max}). Every choice option was paired with both a monetary reward and a physical effort. Patients were encouraged to accumulate as much money as possible and to avoid making unnecessary effort. Every trial started with a fixation cross, followed by an abstract symbol (a letter from the Agathodaimon font) displayed at the center of the screen (Fig. 1A). When two interrogation dots appeared on the screen, patients had to make a choice between left and right options by slightly squeezing either the left or the right power grip. Once a choice was made, patients were informed about the outcome (i.e., the reward and effort levels materialized, respectively, as a coin image and a visual target). Next, the command “GO!” appeared on the screen and the bulb of the thermometer turned blue, triggering effort exertion. Patients were required to squeeze the grip until the mercury level reached the target (horizontal bar). At this moment, the current reward was displayed with Arabic digits and added to the cumulative total payoff. The two behavioral responses (choice and force exertion) were self-paced. Therefore, participants needed to produce the required force to proceed further, which they managed to achieve on every trial.

Patients were given no explicit information about the stationary probabilistic contingencies that were associated with left and right options, which they had to learn by trial and error. The contingencies were varied across contextual cues such that reward and effort learning were separa-

ble as follows: for each cue, the left and right options differed either in the associated reward or in the associated effort (Fig. 1A). For reward learning cues, the two options had distinct probabilities of delivering the big reward (75% vs 25% and 25% vs 75%), whereas probabilities of having to produce the high effort were identical (either 100% or 0%). Symmetrical contingencies were used for effort learning cues, with distinct probabilities of high effort (75% vs 25% and 25% vs 75%) and unique probability of big reward (100% or 0%). The four contextual cues were associated with the following contingency sets: reward learning with high effort, reward learning with low effort, effort learning with big reward, and effort learning with small reward. The best option was on the right for one reward learning and one effort learning cues and on the left for the two other cues. The associations between response side and contingency set were counterbalanced across sessions and subjects. Each of the three sessions contained 24 presentations of each cue randomly distributed over the 96 trials and lasted ~15 min. The four symbols used to cue the contingency sets changed for each new session and thus had to be learned from scratch.

Task version 2. Because of the complexity of this first version of the task, many patients failed to learn the contingencies in Study 1 (Fig. 2A). Therefore, in Study 2, we ran a simplified version of the task on a different group of PD patients and on a group of age-matched controls. Each session now included only two contingency sets (instead of four), one for reward and one for effort learning (Fig. 1B). Furthermore, each learning type was reduced to 20 trials, resulting in a total of 40 trials per session instead of 96. To accentuate the difference between the two options, probabilities were changed to 80% versus 20% (instead of 75% vs 25%) for both reward and effort contingencies. To help in building associa-

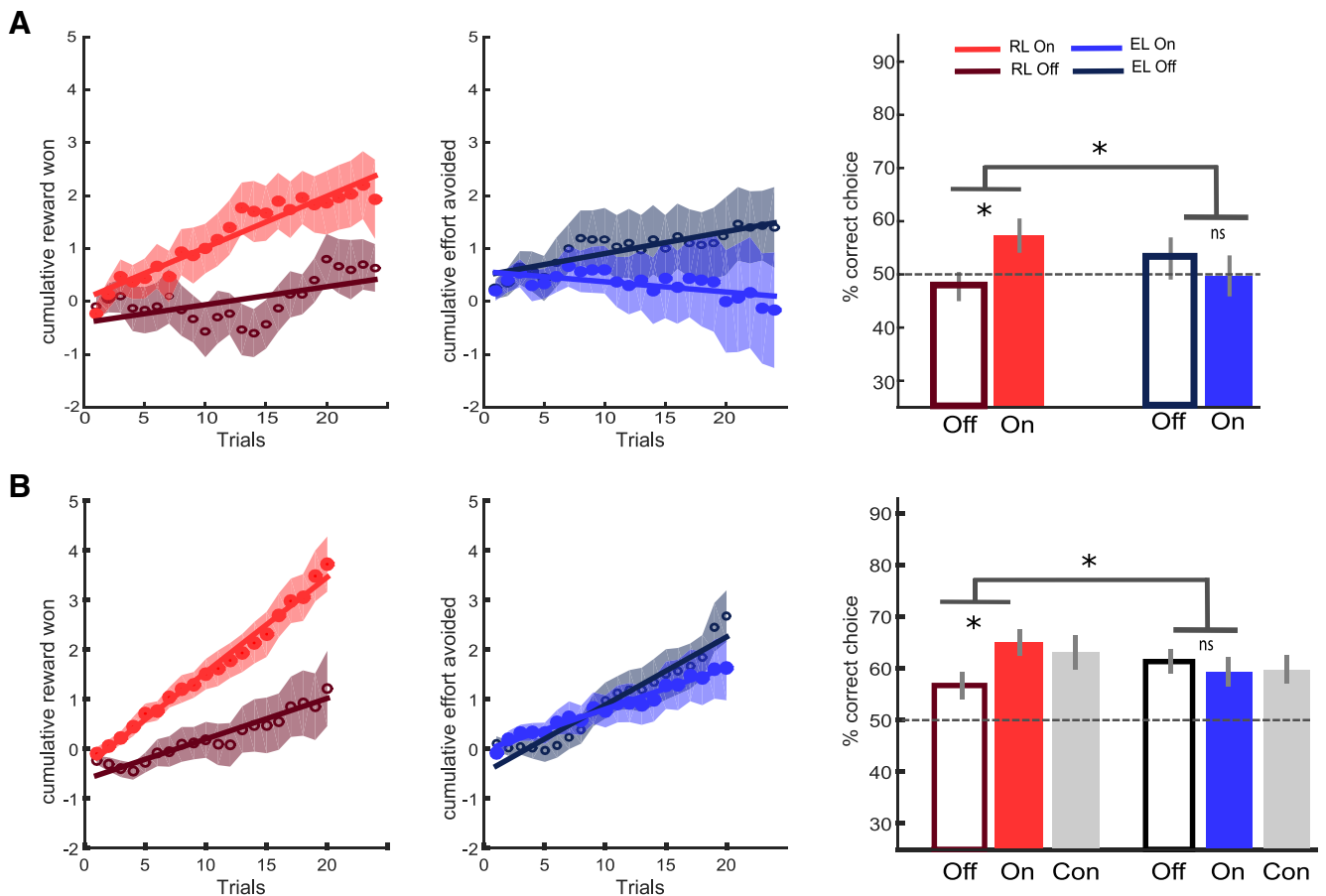


Figure 2. Behavioral results of Study 1 (**A**) and Study 2 (**B**). Left, Learning curves show cumulative scores, that is, money won (big minus small reward) in the reward context (empty and filled red circles for OFF and ON medication states, respectively) and effort avoided (low minus high effort outcomes) in the effort context (empty and filled blue circles for OFF and ON states, respectively). Shaded areas represent trial-by-trial intersubject SEM. Lines indicate linear regression fit. Right, Bar graphs show mean correct response rates (same color coding as for the learning curves, with grey bars for the control group in Study 2). Dotted lines correspond to chance-level performance. Error bars indicate \pm intersubject SEM. Stars indicate significant main effects of treatment and interaction with learning condition ($p < 0.05$). CON, Control; EL, effort learning; RL, reward learning.

tions between options and outcomes, the two options were now represented visually on the screen with two abstract cues. In addition, to make the choice more salient, a red cross was presented for 500 ms below the selected cue while the unchosen alternative was removed from the screen. As in the original design, the probabilities associated with left and right options of a same cue pair differed only on one dimension (reward or effort). To enhance the difference between reward outcomes, the big reward was increased to 50¢ (instead of 20¢) while keeping the small reward unchanged (10¢). We also kept the effort levels unchanged (20% vs 80% of F_{max}) because they were already well discriminated. Once the effort was fulfilled, the monetary gain and the full thermometer were presented for 500 ms before the onset of the next trial.

Statistical analysis. We analyzed two dependent behavioral variables, correct choice rate and squeezing time, with the same series of ANOVAs followed by *post hoc t* tests using Stata Statistical Software: Release 13 (StataCorp) in both studies. The correct option was the one with highest probability of big reward in the reward learning condition and that with the lowest probability of high effort in the effort learning condition. To answer our primary question of interest, we ran two-way repeated-measures ANOVA with learning type (reward vs effort) and medication state (ON vs OFF) as within-subject factors. To assess the influence of potential confounds, we performed three-way ANOVA with side (left vs right hand) or session order as additional factors. We also compared performance between levels in the irrelevant dimension (high vs low effort for reward learning and big vs small reward for effort learning). Significant results were then qualified by *post hoc* paired *t* tests, one-tailed when comparing with chance level and two-tailed when comparing between experimental conditions. Finally, we searched for correlations between

behavioral effects in the task and demographic or clinical variables [age, disease duration, levodopa equivalent dosage and Unified Parkinson's Disease Rating Scale Part III (UPDRS-III) score] by testing Pearson's rho coefficient.

We took a similar approach to analyze squeezing time, defined as the time needed to reach the target from the onset of GO signal. As a manipulation check, we verified that effort production did not depend on medication status, meaning that patients were equally able to adjust their squeezing performance to imposed force targets in the ON and OFF states. Two-way repeated-measures ANOVA on normalized force with medication status (ON vs OFF) and effort level (high vs low) showed a main effect of effort level (Study 1: $F_{(1,14)} = 35.73$, $p < 0.001$; Study 2: $F_{(1,19)} = 162$, $p < 0.0001$), but no significant effect of medication or interaction between medication and effort level (both $p > 0.1$).

Computational modeling. The same Q-learning model was used to fit choices made by PD patients in the two studies. This model was shown previously to best account for learning performance of healthy subjects in the same task (Skvortsova et al., 2014). On every trial expectations attached to the chosen option were updated according to a delta-rule: $Q_R(t+1) = Q_R(t) + \alpha_R * PE_R(t)$ and $Q_E(t+1) = Q_E(t) + \alpha_E * PE_E(t)$, where $Q_R(t)$ and $Q_E(t)$ are the expected reward and effort at trial t , α_R and α_E the reward and effort learning rates, $PE_R(t)$ and $PE_E(t)$ the reward and effort prediction errors calculated as $PE_R(t) = R(t) - Q_R(t)$ and $PE_E(t) = E(t) - Q_E(t)$, and $R(t)$ and $E(t)$ are the reward and effort outcomes obtained at trial t . To avoid specifying an arbitrary common currency, reward levels were coded as 1 for big reward and 0 for small reward and effort levels were coded as -1 for high effort and 0 for low effort. This ensured that reward and effort were treated equally by the model. In

Study 1, there were four contextual sets (two for reward learning and two for effort learning) times two response sides (left and right), resulting in eight option values, whereas in Study 2, there were only two contextual sets times two response sides resulting in four option values. All reward Q -values were initiated at 0.5 and all effort Q -values at -0.5 , which are the true means of all possible reward and effort outcomes. We used a linear formulation for the net value that was shown previously to better model subjects' performance in the same task compared with a hyperbolic discount function (Skvortsova et al., 2014). The net value for each choice option was computed as follows: $Q(t) = Q_R(t) + \gamma * Q_E(t)$, where γ is a positive linear discount factor. The two net values corresponding to the two options on a given trial were compared using a softmax decision rule, which estimates the probability of each choice as a sigmoid function of the difference between the net values of left and right options: $P_{\text{left}}(t) = 1/(1 + \exp((Q_{\text{right}}(t) - Q_{\text{left}}(t))/\beta))$, with β being a positive temperature parameter that captures choice stochasticity.

Overall, the null model (no medication effect) included four free parameters: reward and effort learning rates (α_R and α_E), discount factor (γ), and choice temperature (β). In principle, each of these parameters could potentially be affected by dopaminergic treatment: change in α_R or α_E would alter the updating process for one dimension specifically, change in γ would shift the relative weight of reward and effort expectations at the decision stage, and change in β would make choices closer to random (or exploratory) behavior regardless of the dimension. We also hypothesized that dopamine could affect the subjective valuation of reward or effort outcomes. To test these hypotheses, we included two multiplicative modulations of the objective reward and effort (R and E) outcomes by two additional non-negative parameters (kR and kE): $r = kR * R$ and $E = kE * E$. This makes a total of six parameters that could or not be affected by dopaminergic modulation.

All possible combinations of modulations including the null model resulted in $2^6 = 64$ models that were inverted using a variational Bayes analysis (VBA) approach under the Laplace approximation (Daunizeau et al., 2014). Model inversions were implemented using the MATLAB VBA toolbox (available at <http://mbb-team.github.io/VBA-toolbox/>). This iterative algorithm provides a free-energy approximation for the model evidence, which represents a natural trade-off between model accuracy (goodness of fit) and complexity (degrees of freedom) (Friston et al., 2007; Penny, 2012; Rigoux et al., 2014). In addition, the algorithm provides an estimate of the posterior density over the free parameters starting with Gaussian priors. Log model evidence was then taken to group-level random-effect analysis and compared by families for each parameter of interest using a BMS procedure (Penny et al., 2010). This procedure results in an exceedance probability (XP) that measures how likely it is that a given family of models is implemented more frequently than the others in the population (Stephan et al., 2009; Rigoux et al., 2014). Therefore, for each of the six parameters of interest, we compared two halves of the model space: the family of models in which the considered parameter is affected by dopaminergic medication and the family in which it is not. The exceedance probability of the latter family can be interpreted as the p -value of the null hypothesis: if lower than 0.05, then the medication effect is significant. This is equivalent to exceedance probability of the former family (with medication effect) being >0.95 .

Some of the computational parameters are obviously not independent. To assess whether medication effects on specific parameters could be identified nonetheless, we performed a parameter recovery simulation focusing on Study 2. For each patient, we took the best-fitting parameters of the computational model in the OFF state. To simulate a potential medication effect on each parameter of interest (reward and effort sensitivity, reward and effort learning rates, discount factor, and choice temperature), we changed its value by 40% in the direction of the observed medication effect. We chose 40% because it was the dopaminergic medication effect size that was observed for the parameter kR in Study 2. We next fitted these simulated data using the full model, in which modulations of all parameters by dopaminergic medication were allowed simultaneously. We performed these simulations and fits 100 times for each patient and each parameter of interest. Then, we compared the parameter values between the OFF state (fitted on actual data) and the ON state (fitted on simulated data) using a two-tailed paired t test for

each target parameter. A significant difference ($p < 0.05$) between the initial value of the target parameter in the OFF state and the recovered value of the parameter in the ON state indicates a good sensitivity of our fitting procedure for capturing the simulated specific modulation. In contrast, any difference between the initial OFF state and the recovered ON state values for the nontarget parameters indicates a lack of specificity. The target modulation was significantly recovered in 70–90% of the simulations and the proportion of false alarms varied between 0% and 20% of simulations depending on the considered parameter. Therefore, medication effects on each of the six parameters could be recovered independently and with reasonable sensitivity and specificity.

Results

Learning performance

To illustrate learning dynamics, we computed trial-by-trial cumulative scores for earned money and avoided effort. On every trial, the cumulative score was increased by one when the outcome was big reward or low effort and decreased by one when it was small reward or high effort. To show whether the balance was progressing over trials, we regressed these cumulative scores on trial number separately for reward and effort learning contexts (Fig. 2, left).

The main working hypothesis was that dopamine is involved differentially in reward and effort learning. Consistently, we found in both studies that dopaminergic medication affected reward but not effort learning (Fig. 2, right): two-way repeated-measures ANOVA on the correct choice rate revealed a significant interaction (Study 1: $F_{(1,14)} = 4.97, p = 0.043$; Study 2: $F_{(1,19)} = 8.16, p = 0.01$) between medication status (ON vs OFF) and learning type (reward vs effort). Specifically, two-tailed paired t tests showed that patients learned better from rewards while on compared with off medication in both Study 1 ($R_{\text{ON}}: 57.29 \pm 2.84\%, R_{\text{OFF}}: 47.71 \pm 2.36\%, t_{(14)} = 2.82, p = 0.014$) and Study 2 ($R_{\text{OFF}}: 56.4 \pm 2.20\%, R_{\text{ON}}: 65.06 \pm 2.15\%, t_{(19)} = 3.55, p = 0.002$). In contrast, there was no medication effect on effort learning in either Study 1 ($E_{\text{OFF}}: 52.99 \pm 3.58\%, E_{\text{ON}}: 49.72 \pm 3.47\%, t_{(14)} = 0.64, p = 0.53$) or Study 2 ($E_{\text{OFF}}: 59.82 \pm 1.92\%, E_{\text{ON}}: 59.14 \pm 2.47\%, t_{(19)} = 0.28, p = 0.78$).

We also conducted similar two-way repeated-measures ANOVA on the regression slopes of cumulative curves. Consistent with the analysis on average correct choice rate, the interaction between learning type and medication status was borderline in Study 1 ($F_{(14)} = 3.45, p = 0.084$) but significant in Study 2 ($F_{(19)} = 13.41, p = 0.002$). Two-tailed paired t tests showed a marginally significant medication effect for the reward slope in Study 1 ($t_{(14)} = 1.95, p = 0.072$), which was significant in Study 2 ($t_{(19)} = 3.47, p = 0.0025$). In contrast, the effort slope was not significantly improved by medication in either Study 1 ($t_{(14)} = -1.07, p = 0.3$) or Study 2 ($t_{(19)} = -1.183, p = 0.15$).

In our view, the replication provided by the second study, with simplification of the design, was necessary to conclude in favor of the dissociation. This is because, in Study 1, patients might have been overwhelmed by the complexity of the task, precluding any effect on effort learning. Indeed, correct choice rate in Study 1 only differed from chance level (50%) for reward learning in the "ON" state ($t_{(14)} = 2.84, p = 0.007$). In Study 2, performance was above chance level in all conditions (all p -values < 0.0125 , corresponding to Bonferroni-corrected threshold for four tests). In addition, Study 2 allowed comparison with healthy controls to assess disease effect. As expected, OFF patients were impaired relative to controls in reward learning ($t_{(36)} = 2.10, p = 0.042$) but not effort learning ($t_{(36)} = 0.61, p = 0.54$). There was no difference between ON patients and controls (reward learning: $t_{(36)} = 0.40, p = 0.69$; effort learning: $t_{(36)} = 0.1, p = 0.92$),

suggesting that dopaminergic medication simply normalized performance in the reward learning condition.

We ran additional analyses to assess potential confounds. Importantly, performance pooled across medication states did not differ between reward and effort learning contexts ($t_{(14)} = 0.63$, $p = 0.54$ in Study 1 and $t_{(19)} = 0.74$, $p = 0.47$ in Study 2). This suggests that medication did not affect learning from feedback in general, but learning from reward outcome specifically. In addition, the two conditions appeared to be matched in difficulty because there was no significant difference between reward and effort learning performance in healthy subjects either in this study ($t_{(17)} = 1.45$, $p = 0.17$) or in the previous study (Skvortsova et al., 2014).

We next examined the possible interactions with session order, side of correct response, and irrelevant dimension (effort level in reward learning and reward level in effort learning). Mixed-effect three-way ANOVA with learning context (reward vs effort) and medication status (ON vs OFF) as within-subject factors and session order (ON-OFF vs OFF-ON) as a between-group factor showed no main effect of session order (Study 1: $F_{(1,13)} = 2.24$, $p = 0.08$; Study 2: $F_{(1,18)} = 0.21$, $p = 0.65$) or any two-way or three-way interactions (Study 1: all p -values >0.15 ; Study 2: all p -values $= 0.2$). In any case, session order could not have been a confound because it was orthogonal to the other factors. Similarly, a three-way interaction with hand (left vs right) as an additional within-subject factor showed no side effect (Study 1: $F_{(1,14)} = 2.01$, $p = 0.18$; Study 2: $F_{(1,19)} = 1.9$, $p = 0.18$) or any interaction (all p -values >0.37). This suggests that medication affected learning processes at a central cognitive stage and were not influenced by potentially asymmetrical motor disturbances. To test the effect of the irrelevant dimension, we computed the relative performance score (difference between high and low effort for reward learning and difference between big and small reward for effort learning). This relative performance was not different from zero and was not affected by medication (Study 1, all p -values >0.13 , Study 2 all p -values >0.41). Similarly, we did not find any significant relative performance difference in healthy aged controls: (reward learning $t_{(17)} = 1.04$, $p = 0.31$; effort learning $t_{(17)} = 1.21$, $p = 0.24$). This is consistent with our previous findings in young healthy adults performing the same task (Skvortsova et al., 2014) and suggests that reward and effort learning performance was not affected by the irrelevant condition.

Finally, we searched for correlations between medication effects and relevant clinical variables. Medication effects on reward and effort learning were uncorrelated across patients in both studies (Study 1: $R^2 = 0.15$, $p = 0.6$; Study 2: $R^2 = 0.26$, $p = 0.27$; studies pooled together: $R^2 = 0.13$, $p = 0.47$). This strengthens the notion that the two learning processes are underpinned by different brain circuits, one involving dopamine and the other not. Even when pooling the two groups to increase statistical power, we could not detect any significant correlation between reward learning improvement (ON minus OFF state) and age ($R^2 = -0.07$, $p = 0.9$), daily levodopa equivalent dose (LED) ($R^2 = 0.23$, $p = 0.19$), disease duration ($R^2 = 0.14$, $p = 0.41$), or medication effect on UPDRS III score ($R^2 = 0.05$, $p = 0.76$).

Squeezing performance

ANOVA performed on squeezing time (the time that it took patients to reach the target) with reward level, effort level, or medication status as within-subject factors only revealed a significant main effect of effort level in both studies (Study 1: $F_{(1,14)} = 12.04$, $p = 0.004$; Study 2: and $F_{(1,19)} = 40.48$, $p < 0.0001$). This means

that patients were reliably faster to reach 20% F_{\max} compared with 80% F_{\max} (difference: 636 ± 154.11 ms in Study 1 and 505 ± 85 ms in Study 2). This delay was probably hard to notice and, in any case, was not different between medication states, which might not be surprising given that effort level was recalibrated to F_{\max} in each state. The absence of interaction between reward level and medication status in squeezing time suggests that dopamine might not be involved in any reward-related process, but more specifically in instrumental behavior.

BMS

To investigate which of the Q-learning model parameters was primarily affected by dopaminergic treatment, we performed a factorial BMS. For each parameter of interest, we compared families with and without modulation by dopaminergic medication (meaning with same or different parameters in the ON and OFF states). In total, we performed six comparisons (Fig. 3, top) to test the effects of dopaminergic medication on: reward sensitivity kR , effort sensitivity kE , reward learning rate αR , effort learning rate αE , discount factor γ , and choice temperature β . In both studies, the highest exceedance probability for treatment effect was obtained for parameter kR (XP = 0.88 for study1 and XP = 0.95 for Study 2). Therefore, selective modulation kR is the best model under the parsimonious hypothesis that medication only affected one parameter. We also compared the “ kR -only” model with the five models in which one other parameter is modulated (kR/kE , $kR/\alpha R$, $kR/\alpha E$, kR/γ , and kR/β). The kR only was the most plausible in both Study 1 (XP = 0.96) and Study 2 (XP = 0.81).

In the winning model family, kR was set to 1 in the OFF state and allowed a different (fitted) value in the ON state. In both studies, the posterior mean of kR was significantly different from 1 at the group level (Study 1: 1.25 ± 0.08 , $t_{(14)} = 2.58$, $p = 0.02$; Study 2: 1.41 ± 0.06 , $t_{(19)} = 3.94$, $p = 0.001$, two-tailed paired t test). This means that reward outcomes were systematically amplified under dopaminergic modulation (see insets in Fig. 3, top). We verified that no other parameter showed a similar systematic shift from the “OFF” to the “ON” states (all 5 p -values >0.26). We have also tested two-way interactions between kR and medication effects on the 5 other parameters estimated in the full model (in which all parameters are allowed to vary with medication status). The 5 interactions were significant in both Study 1 (all $p < 0.01$) and Study 2 (all $p < 0.006$), showing that medication effect on kR was higher than any other possible effect, even after Bonferroni correction for multiple comparisons. This is consistent with the fact that the kR -only model was far more plausible than the full model (XP = 1).

To assess the fit of the winning model to the data, we computed Pearson’s correlations between observed and predicted choice rates separately for reward and effort learning contexts and for ON and OFF medication states (Fig. 3, bottom). In Study 1, correlation was highly significant when subjects were performing above chance level (reward learning, ON state): $R^2 = 0.86$, $p < 0.0001$. Correlation failed to survive Bonferroni-corrected threshold in the other conditions ($p = 0.05/4 = 0.0125$). In Study 2, all correlations were significant ($p < 0.006$) with at least $R^2 = 0.6$ of the variance explained.

Finally, we performed simulations of the winning kR -only model using the posterior estimates of parameters fitted at the individual level and the outcome sequences generated for actual participants. We performed 15 simulations using the parameters from Study 1 and 20 simulations using the parameters from Study 2. We then plotted the cumulative learning curves and analyzed the percentage of correct choices averaged across trials

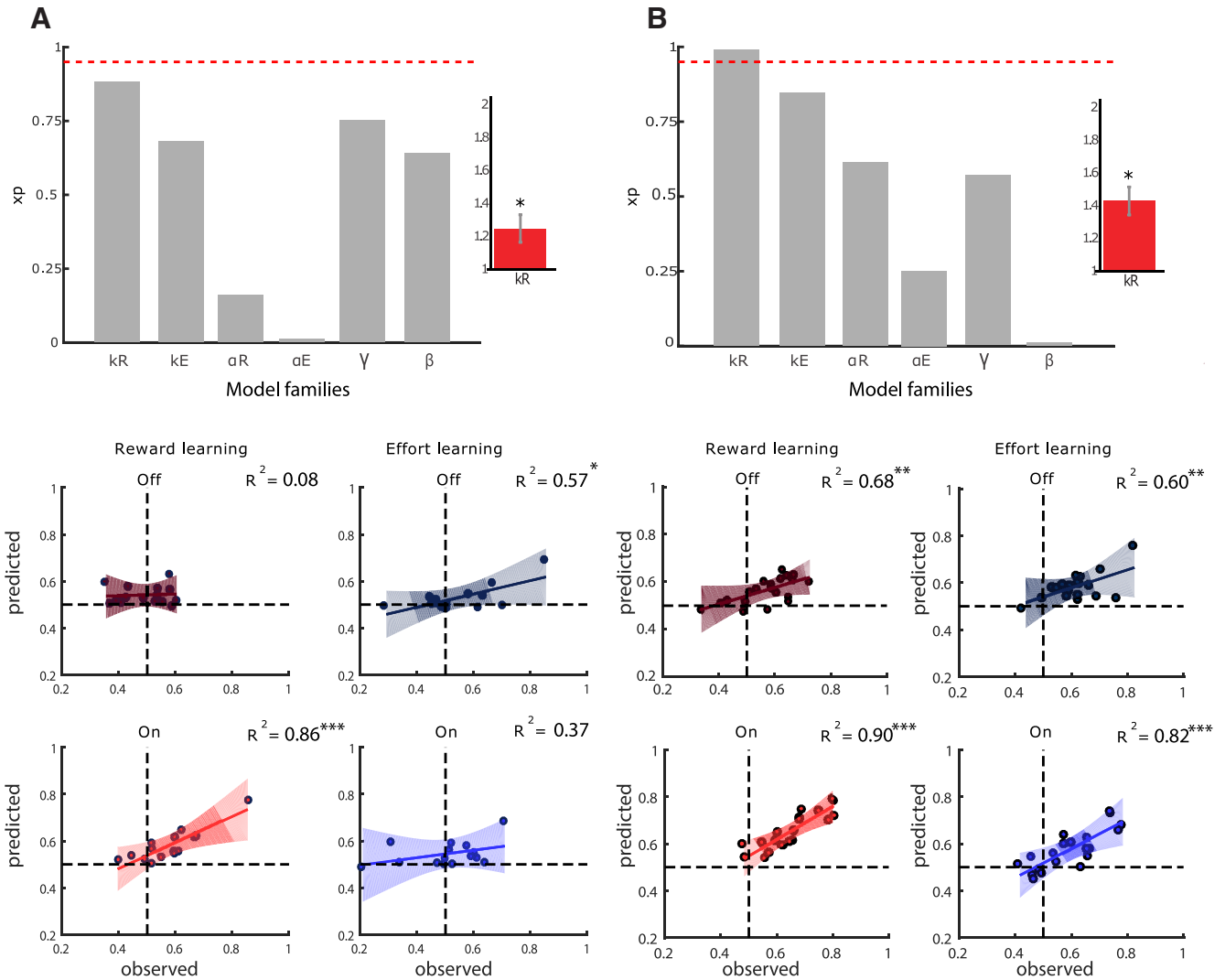


Figure 3. Model comparison based on data from Study 1 (A) and Study 2 (B). Top, Bayesian comparisons of model families. For each parameter, two families corresponding to two halves of the model space were compared: all models including versus all models excluding the possibility of medication effect on the considered parameter. Bars show XP obtained for the presence of medication effect on six model parameters: kR , kE , αR , αE , γ , and β . Note that the XP for the null hypothesis (absence of medication effect) is simply one minus that shown on the graph. Red dotted line corresponds to significance threshold (exceedance probability of 0.95). Insert, Mean posterior estimate for reward sensitivity parameter $kR \pm$ intersubject SEM. Star indicates significant difference from one ($p < 0.05$). Bottom, Scatter plots of interpatient correlations between observed correct choices and correct choices predicted from the kR -only model for the two learning conditions and medication states. Each dot represents one subject. Shaded areas indicate 95% confidence intervals on linear regression estimates.

to compare the simulated data with observed data (Fig. 4 vs Fig. 2). Two-way repeated-measures ANOVA on the correct choice rate revealed a borderline or significant interaction (Study 1: $F_{(1,14)} = 3.52$, $p = 0.082$; Study 2: $F_{(1,19)} = 5.86$, $p = 0.026$) between medication status (ON vs OFF) and learning type (reward vs effort). Specifically, two-tailed paired t tests showed that patients learned better from rewards while on compared with the off medication in both Study 1 (R_{ON} : $58.69 \pm 1.59\%$, R_{OFF} : $54.94 \pm 2.29\%$, $t_{(14)} = 2.01$, $p = 0.064$) and Study 2 (R_{OFF} : $60.19 \pm 2.16\%$, R_{ON} : $64.56 \pm 2.44\%$, $t_{(19)} = 2.37$, $p = 0.029$). In contrast, there was no medication effect on effort learning in either Study 1 (E_{OFF} : $55.35 \pm 1.42\%$, E_{ON} : $56.01 \pm 1.45\%$, $t_{(14)} = 0.86$, $p = 0.41$) or Study 2 (E_{OFF} : $59.31 \pm 1.52\%$, E_{ON} : $57.19 \pm 2.01\%$, $t_{(19)} = 1.01$, $p = 0.33$). Therefore, the model was able to recapitulate the pattern of results observed in patients' choices.

Discussion

We compared learning to maximize rewards and learning to minimize efforts in two groups of PD patients with and without do-

paminergic medication. *A priori*, these two learning processes could follow the same algorithmic principles and therefore be implemented in the same neural circuits. However, dopaminergic treatment improved reward-based but not effort-based learning in both groups of patients. To our knowledge, this is the first demonstration that dopamine is implicated selectively in reward learning by opposition to effort learning.

In healthy controls, as well as in PD patients, when averaging across medication states, performance was similar in reward and effort learning. Therefore, the dissociation does not seem to be driven by one dimension being more difficult to discriminate than the other. This dissociation suggests that subjects treated reward and effort learning as different dimensions, which was not trivial in our context. Indeed, patients could have recoded each outcome in terms of correct versus incorrect response. Therefore, the dissociation indicates that our task was efficient in triggering at least partially separate processes, in keeping with our previous fMRI study showing partial dissociation between brain networks

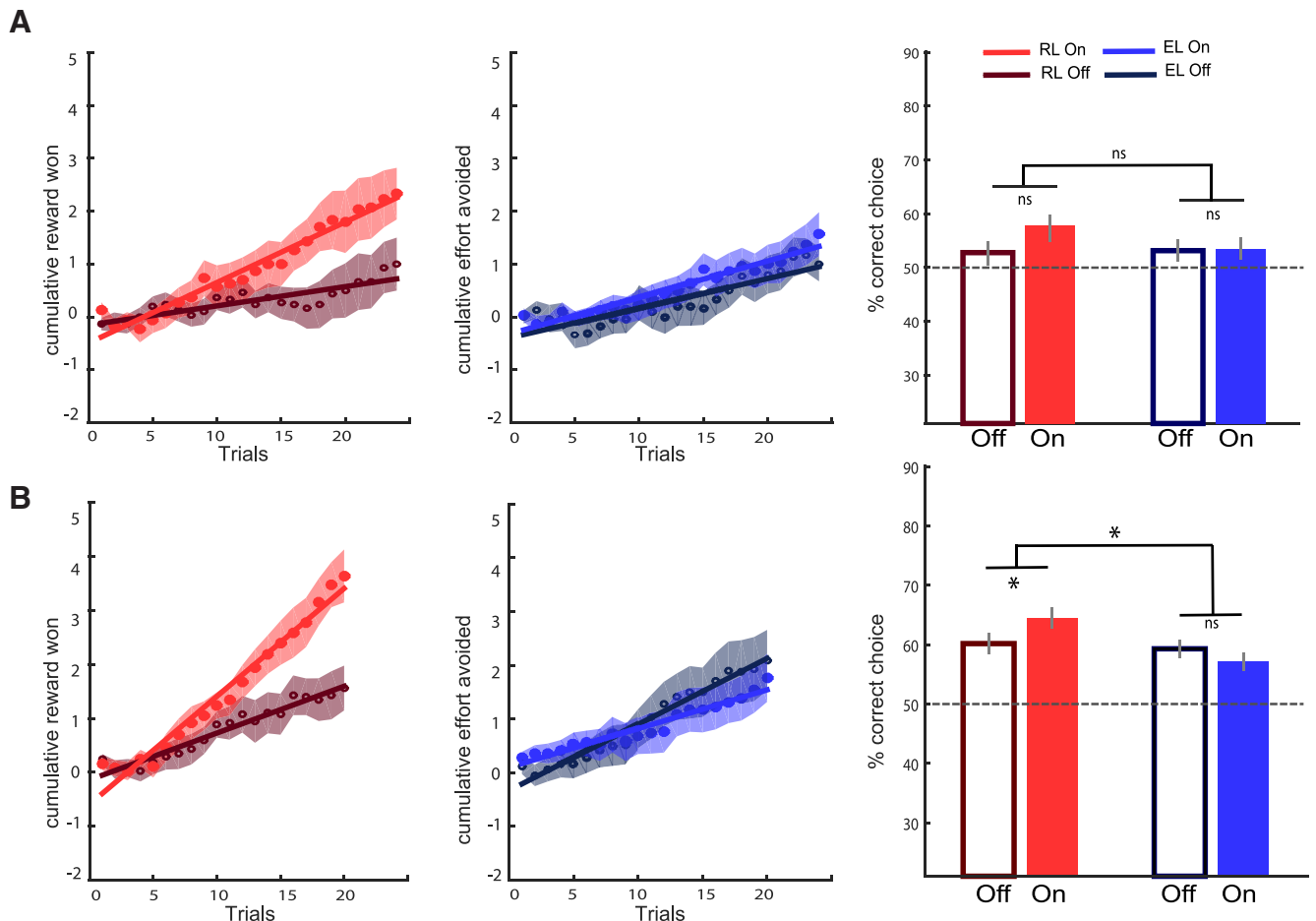


Figure 4. Model simulations for Study 1 (**A**) and Study 2 (**B**). Left, Simulated learning curves show cumulative scores, that is, money won (big minus small reward) in the reward context (empty and filled red circles for OFF and ON medication states, respectively) and effort avoided (low minus high effort outcomes) in the effort context (empty and filled blue circles for OFF and ON states, respectively). Shaded areas represent trial-by-trial intersubject SEM. Lines indicate linear regression fit. Right, Bar graphs show mean correct response rates (same color coding as for the learning curves, with gray bars for the control group in Study 2). Dotted lines correspond to chance-level performance. Error bars indicate \pm intersubject SEM. Stars indicate significant main effects of treatment and interaction with learning condition ($p < 0.05$). EL, Effort learning; RL, reward learning.

involved in reward and effort learning (Skvortsova et al., 2014). The valence-specific effect of medication also speaks against the hypothesis that dopamine might signal motivational salience, which would apply equally to appetitive and aversive stimuli (Kapur et al., 2005; Bromberg-Martin et al., 2010).

The dissociation might also help disambiguate previous findings in decision-making tasks that manipulating dopamine can shift effort/reward trade-off in both rodents (Denk et al., 2005; Salamone et al., 2007; Floresco et al., 2008; Bardgett et al., 2009; Mai et al., 2012) and humans (Wardle et al., 2011; Chong et al., 2015). In cost/benefit decision-making tasks, reward and effort levels are confounded: higher efforts are always paired with higher rewards. This creates an ambiguity in the interpretation of drug effects because both increased sensitivity to reward and decreased sensitivity to effort could result in the same shift of preference. Our results clearly suggest an interpretation in terms of reward sensitivity. We characterized the effects of dopaminergic medication on effort/reward decision making (without learning) in a recent study using computational modeling (Le Bouc et al., 2016). Bayesian model comparison showed that the bias toward big reward, high effort options under dopamine enhancers is best captured by increasing sensitivity to reward prospect and not by decreasing sensitivity to effort cost. This analysis aligns well with the present result, which was also best captured by enhancing

reward sensitivity. Combining these two findings, one may conclude that dopamine amplifies the effects of reward on behavioral output.

Bayesian model comparison in the present study showed that only reward sensitivity is affected by dopaminergic medication, not effort sensitivity, learning rates, discount factors, or choice temperature. This means that the reward value of the best cue was amplified such that both the slope and plateau of learning curves were enhanced. This effect is subtly different from that postulated in the OpAL model (Collins and Frank, 2014), which also intended to integrate the incentive and learning functions of dopamine. There are two levels where reward and effort would be processed differently in the OpAL model, first during update with two different learning rates and second during choice with two different temperature parameters in the softmax function. The reward learning rate only controls the performance slope (not the plateau), in contrast to our reward sensitivity parameter. The reward learning plateau could be controlled by a specific temperature parameter, but this mechanism would still be slightly different from what was observed here. This is because the modulation by dopaminergic treatment was multiplicative and reward outcome coded 0 or 1. Therefore, it only affected the chosen cue value when the outcome was a big reward. Note that we included both a relative weight (discount factor) and a generic

temperature parameter in our model, which is equivalent to having separate temperature parameters for reward and effort learning. However, parameter recovery simulation and BMS showed that specific medication effect on reward sensitivity could be identified. This effect is closer to a selective modulation of positive prediction error, which generally follows big reward outcomes, as was implemented in some previous model-based analyses (Cools et al., 2006; Frank et al., 2007b; Rutledge et al., 2009). Modulation of reward sensitivity suggests that dopaminergic medication affected the update process, although this remains hard to disentangle from the choice process using only behavioral data. However, it is consistent with a previous fMRI study showing that enhancement of reward sensitivity by levodopa treatment was underpinned by amplified representation of RPE in the striatum (Pessiglione et al., 2006). This striatal prediction error signal was generated at the time when choice outcome was delivered; that is, when the update process, not the choice process, should occur.

The dissociation of dopamine implication in reward versus effort learning is reminiscent of that repeatedly observed between reward and punishment learning and explained by the opposite effect of dopamine release on the D1 (GO) and D2 (NO-GO) pathways of the basal ganglia (Collins and Frank, 2014). Although pharmacological studies in humans converge on the conclusion that dopamine helps with reward learning, there was some discrepancy regarding punishment learning: a detrimental effect was observed in some cases (Frank et al., 2004; Bódi et al., 2009; Palminteri et al., 2009), but just an absence of effect in others (Pessiglione et al., 2006; Rutledge et al., 2009; Schmidt et al., 2014). Here, the trend was toward a detrimental effect of dopaminergic medication on effort learning, but this did not reach significance. We cannot rule out that a more sensitive task, or higher number of patients, could have yielded significant effects. However, our results show that medication effects on effort learning are weaker and opposite (detrimental and not beneficial) relative to effects on reward learning. The parallel with punishment learning raises the question of how these two dimensions (effort and punishment) differ from each other. Theoretically, effort belongs to the action space, whereas punishment belongs to the outcome space. However, producing an effort such as squeezing the handgrip might trigger, not only effort sensation, but also joint or muscular pain, which might be closer to the notion of punishment although still different from the financial losses or negative (“incorrect”) feedback used in previous studies. That effort and punishment belong to the same domain is also supported by the observation that brain regions involved in effort processing, such as the dorsal anterior cingulate cortex and the anterior insula (Rudebeck et al., 2006; Croxson et al., 2009; Hosokawa et al., 2013; Kurniawan et al., 2013; Skvortsova et al., 2014), have also been implicated in punishment processing (Büchel et al., 1998; Seymour et al., 2005; Nitschke et al., 2006; Samanez-Larkin et al., 2008; Palminteri et al., 2015).

As was suggested for punishment, it could be postulated that avoidance of effort cost is learned through the impact of dopaminergic dips on the D2 pathway in the basal ganglia (Collins and Frank, 2014). However, this hypothesis would predict a detrimental effect of prodopaminergic medication on effort learning, which was not significant in our data. This may call for the existence of other opponent systems that would operate aversive learning. In addition to the cortical regions already mentioned (anterior cingulate and insula), other neuromodulators might play a role in effort learning. Notably, serotonin was suggested to regulate the weight of both punishment (Daw et al., 2002; Niv et

al., 2007; Boureau and Dayan, 2011; Cohen et al., 2015) and delay (Doya, 2002; Miyazaki et al., 2014; Fonseca et al., 2015). Therefore, serotonin seems to be involved in processing different kinds of costs, which could be generalized to effort (Meyniel et al., 2016). Alternatively, effort processing might be controlled by noradrenaline because locus ceruleus activity was related to both the mental and physical effort required by upcoming action and reflected in pupil dilation (Alnæs et al., 2014; Joshi et al., 2016; Varazzani et al., 2015).

We acknowledge that our demonstration has limitations, in particular regarding the exact mechanisms underlying the effects of dopaminergic medication at the cellular and molecular level. Superficially, our results agree well with a wealth of physiological recordings linking dopamine to reward prediction or RPE (Schultz et al., 1997; Waelti et al., 2001; Satoh et al., 2003; Bayer and Glimcher, 2005; Tobler et al., 2005) and showed that dopaminergic signals are more sensitive to reward than effort (Gan et al., 2010; Pasquereau and Turner, 2013; Varazzani et al., 2015). However, how systemic medications affect dopaminergic signals remains unclear; in particular, whether they only enhance tonic dopaminergic activity or also boost phasic responses is still a matter of debate. The issue is complicated by the fact that we pooled patients who were taking medications with different modes of action, such as metabolic precursors (levodopa) and receptor agonists (notably pramipexole and bromocriptine). In addition, the polymorphism of dopamine-related genes is known to induce some variability in medication effects (Frank et al., 2007a; Klein et al., 2007; den Ouden et al., 2013), which was not assessed here because we did not genotype the patients.

Finally, we could not establish any link between medication effect (improved reward learning) and clinical variables, which might question the clinical interest of our finding. The absence of correlation with anti-parkinsonian effect (i.e., change in UPDRS score) suggests that the experimental result is not due to the alleviation of motor symptoms. Unfortunately, we did not measure the anti-apathetic effect (i.e., change in Starkstein score), which correlated with enhanced reward sensitivity in our previous study (Le Bouc et al., 2016). The amplification of reward sensitivity by dopaminergic medication, which we replicated here, might not only enhance incentive motivation (when potential reward is presented before action initiation), but also positive reinforcement (when actual reward is presented after action completion). A better sensitivity to positive feedback might have pervasive effects, not only on mood, but also on behavioral adaptation to motor, cognitive, or social challenges. These putative consequences still need to be evaluated properly.

References

- Alnæs D, Sneve MH, Espeseth T, Endestad T, van de Pavert SH, Laeng B (2014) Pupil size signals mental effort deployed during multiple object tracking and predicts brain activity in the dorsal attention network and the locus coeruleus. *J Vis* 14: pii: 1. [CrossRef Medline](#)
- Arsenault JT, Rima S, Stemmann H, Vanduffel W (2014) Role of the primate ventral tegmental area in reinforcement and motivation. *Curr Biol* 24:1347–1353. [CrossRef Medline](#)
- Bardgett ME, Depenbrock M, Downs N, Points M, Green L (2009) Dopamine modulates effort-based decision making in rats. *Behav Neurosci* 123:242–251. [CrossRef Medline](#)
- Bayer HM, Glimcher PW (2005) Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47:129–141. [CrossRef Medline](#)
- Bickel WK, Pitcock JA, Yi R, Angtuaco EJ (2009) Congruence of BOLD response across intertemporal choice conditions: fictive and real money gains and losses. *J Neurosci* 29:8839–8846. [CrossRef Medline](#)
- Bódi N, Kéri S, Nagy H, Moustafa A, Myers CE, Daw N, Dibó G, Takáts A, Bereczki D, Gluck MA (2009) Reward-learning and the novelty-seeking

- personality: a between- and within-subjects study of the effects of dopamine agonists on young Parkinson's patients. *Brain* 132:2385–2395. [CrossRef Medline](#)
- Boureau YL, Dayan P (2011) Opponency revisited: competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology* 36:74–97. [CrossRef Medline](#)
- Bromberg-Martin ES, Matsumoto M, Hikosaka O (2010) Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron* 68:815–834. [CrossRef Medline](#)
- Büchel C, Morris J, Dolan RJ, Friston KJ (1998) Brain systems mediating aversive conditioning: an event-related fMRI study. *Neuron* 20:947–957. [CrossRef Medline](#)
- Chong TT, Bonnelle V, Manohar S, Veromann KR, Muhammed K, Tofaris GK, Hu M, Husain M (2015) Dopamine enhances willingness to exert effort for reward in Parkinson's disease. *Cortex* 69:40–46. [CrossRef Medline](#)
- Chowdhury R, Guitart-Masip M, Lambert C, Dayan P, Huys Q, Düzel E, Dolan RJ (2013) Dopamine restores reward prediction errors in old age. *Nat Neurosci* 16:648–653. [CrossRef Medline](#)
- Cohen JY, Amoroso MW, Uchida N (2015) Serotonergic neurons signal reward and punishment on multiple timescales. *Elife* 4. [CrossRef Medline](#)
- Collins AG, Frank MJ (2014) Opponent actor learning (OpAL): Modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. *Psychol Rev* 121:337–366. [CrossRef Medline](#)
- Cools R, Altamirano L, D'Esposito M (2006) Reversal learning in Parkinson's disease depends on medication status and outcome valence. *Neuropsychologia* 44:1663–1673. [CrossRef Medline](#)
- Croxson PL, Walton ME, O'Reilly JX, Behrens TE, Rushworth MF (2009) Effort-based cost-benefit valuation and the human brain. *J Neurosci* 29:4531–4541. [CrossRef Medline](#)
- Daunizeau J, Adam V, Rigoux L (2014) VBA: a probabilistic treatment of nonlinear models for neurobiological and behavioural data. *PLoS Comput Biol* 10:e1003441. [CrossRef Medline](#)
- Daw ND, Kakade S, Dayan P (2002) Opponent interactions between serotonin and dopamine. *Neural Netw* 15:603–616. [CrossRef Medline](#)
- den Ouden HE, Daw ND, Fernandez G, Elshout JA, Rijpkema M, Hoogman M, Franke B, Cools R (2013) Dissociable effects of dopamine and serotonin on reversal learning. *Neuron* 80:1090–1100. [CrossRef Medline](#)
- Denk F, Walton ME, Jennings Ka, Sharp T, Rushworth MF, Bannerman DM (2005) Differential involvement of serotonin and dopamine systems in cost-benefit decisions about delay or effort. *Psychopharmacology (Berl)* 179:587–596. [CrossRef Medline](#)
- Doya K (2002) Metalearning and neuromodulation. *Neural Netw* 15:495–506. [CrossRef Medline](#)
- Eshel N, Bukwich M, Rao V, Hemmelder V, Tian J, Uchida N (2015) Arithmetic and local circuitry underlying dopamine prediction errors. *Nature* 525:243–246. [CrossRef Medline](#)
- Floresco SB, Tse MT, Ghods-Sharifi S (2008) Dopaminergic and glutamatergic regulation of effort- and delay-based decision making. *Neuropsychopharmacology* 33:1966–1979. [CrossRef Medline](#)
- Fonseca MS, Murakami M, Mainen ZF (2015) Activation of dorsal raphe serotonergic neurons promotes waiting but is not reinforcing. *Curr Biol* 25:306–315. [CrossRef Medline](#)
- Frank MJ, Seeberger LC, O'Reilly RC (2004) By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* 306:1940–1943. [CrossRef Medline](#)
- Frank MJ, Moustafa AA, Haughey HM, Curran T, Hutchison KE (2007a) Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proc Natl Acad Sci U S A* 104:16311–16316. [CrossRef Medline](#)
- Frank MJ, Samanta J, Moustafa AA, Sherman SJ (2007b) Hold your horses: impulsivity, deep brain stimulation, and medication in parkinsonism. *Science* 318:1309–1312. [CrossRef Medline](#)
- Friston K, Mattout J, Trujillo-Barreto N, Ashburner J, Penny W (2007) Variational free energy and the Laplace approximation. *Neuroimage* 34:220–234. [CrossRef Medline](#)
- Gan JO, Walton ME, Phillips PE (2010) Dissociable cost and benefit encoding of future rewards by mesolimbic dopamine. *Nat Neurosci* 13:25–27. [CrossRef Medline](#)
- Harrison NA, Voon V, Cercignani M, Cooper EA, Pessiglione M, Critchley HD (2016) A neurocomputational account of how inflammation enhances sensitivity to punishments versus rewards. *Biol Psychiatry* 80:73–81. [CrossRef Medline](#)
- Hollerman JR, Schultz W (1998) Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat Neurosci* 1:304–309. [CrossRef Medline](#)
- Hosokawa T, Kennerley SW, Sloan J, Wallis JD (2013) Single-neuron mechanisms underlying cost-benefit analysis in frontal cortex. *J Neurosci* 33:17385–17397. [CrossRef Medline](#)
- Joshi S, Li Y, Kalwani RM, Gold JI (2016) Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. *Neuron* 89:221–234. [CrossRef Medline](#)
- Kapur S, Mizrahi R, Li M (2005) From dopamine to salience to psychosis—linking biology, pharmacology and phenomenology of psychosis. *Schizophr Res* 79:59–68. [CrossRef Medline](#)
- Klein TA, Neumann J, Reuter M, Hennig J, von Cramon DY, Ullsperger M (2007) Genetically determined differences in learning from errors. *Science* 318:1642–1645. [CrossRef Medline](#)
- Kurniawan IT, Seymour B, Talmi D, Yoshida W, Chater N, Dolan RJ (2010) Choosing to make an effort: the role of striatum in signaling physical effort of a chosen action. *J Neurophysiol* 104:313–321. [CrossRef Medline](#)
- Kurniawan IT, Guitart-Masip M, Dayan P, Dolan RJ (2013) Effort and valuation in the brain: the effects of anticipation and execution. *J Neurosci* 33:6160–6169. [CrossRef Medline](#)
- Le Bouc R, Rigoux L, Schmidt L, Degos B, Welter ML, Vidailhet M, Daunizeau J, Pessiglione M (2016) Computational dissection of dopamine motor and motivational functions in humans. *J Neurosci* 36:6623–6633. [CrossRef Medline](#)
- Mai B, Sommer S, Hauber W (2012) Motivational states influence effort-based decision making in rats: the role of dopamine in the nucleus accumbens. *Cogn Affect Behav Neurosci* 12:74–84. [CrossRef Medline](#)
- Meyniel F, Goodwin GM, Deakin JF, Klinge C, MacFadyen C, Milligan H, Mullings E, Pessiglione M, Gaillard R (2016) A specific role for serotonin in overcoming effort cost. *eLife* 5:e17282. [CrossRef Medline](#)
- Miyazaki KW, Miyazaki K, Tanaka KF, Yamanaka A, Takahashi A, Tabuchi S, Doya K (2014) Optogenetic activation of dorsal raphe serotonin neurons enhances patience for future rewards. *Curr Biol* 24:2033–2040. [CrossRef Medline](#)
- Nitschke JB, Dixon GE, Sarinopoulos I, Short SJ, Cohen JD, Smith EE, Kosslyn SM, Rose RM, Davidson RJ (2006) Altering expectancy dampens neural response to aversive taste in primary taste cortex. *Nat Neurosci* 9:435–442. [CrossRef Medline](#)
- Niv Y, Daw ND, Joel D, Dayan P (2007) Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl)* 191:507–520. [CrossRef Medline](#)
- O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ (2003) Temporal difference models and reward-related learning in the human brain. *Neuron* 38:329–337. [CrossRef Medline](#)
- Palminteri S, Lebreton M, Worbe Y, Grabli D, Hartmann A, Pessiglione M (2009) Pharmacological modulation of subliminal learning in Parkinson's and Tourette's syndromes. *Proc Natl Acad Sci U S A* 106:19179–19184. [CrossRef Medline](#)
- Palminteri S, Clair AH, Mallet L, Pessiglione M (2012) Similar improvement of reward and punishment learning by serotonin reuptake inhibitors in obsessive-compulsive disorder. *Biol Psychiatry* 72:244–250. [CrossRef Medline](#)
- Palminteri S, Khamassi M, Joffily M, Coricelli G (2015) Contextual modulation of value signals in reward and punishment learning. *Nat Commun* 6:8096. [CrossRef Medline](#)
- Pasquereau B, Turner RS (2013) Limited encoding of effort by dopamine neurons in a cost-benefit trade-off task. *J Neurosci* 33:8288–8300. [CrossRef Medline](#)
- Penny WD (2012) Comparing dynamic causal models using AIC, BIC and free energy. *Neuroimage* 59:319–330. [CrossRef Medline](#)
- Penny WD, Stephan KE, Daunizeau J, Rosa MJ, Friston KJ, Schofield TM, Leff AP (2010) Comparing families of dynamic causal models. *PLoS Comput Biol* 6:e1000709. [CrossRef Medline](#)
- Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD (2006) Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* 442:1042–1045. [CrossRef Medline](#)
- Prévost C, Pessiglione M, Météreau E, Cléry-Melin ML, Dreher JC (2010) Separate valuation subsystems for delay and effort decision costs. *J Neurosci* 30:14080–14090. [CrossRef Medline](#)

- Rigoux L, Stephan KE, Friston KJ, Daunizeau J (2014) Bayesian model selection for group studies—revisited. *Neuroimage* 84:971–985. [CrossRef Medline](#)
- Rudebeck PH, Walton ME, Smyth AN, Bannerman DM, Rushworth MF (2006) Separate neural pathways process different decision costs. *Nat Neurosci* 9:1161–1168. [CrossRef Medline](#)
- Rutledge RB, Lazzaro SC, Lau B, Myers CE, Gluck MA, Glimcher PW (2009) Dopaminergic drugs modulate learning rates and perseveration in parkinson's patients in a dynamic foraging task. *J Neurosci* 29:15104–15114. [CrossRef Medline](#)
- Rutledge RB, Dean M, Caplin A, Glimcher PW (2010) Testing the reward prediction error hypothesis with an axiomatic model. *J Neurosci* 30:13525–13536. [CrossRef Medline](#)
- Salamone JD, Correa M, Farrar A, Mingote SM (2007) Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology (Berl)* 191:461–482. [CrossRef Medline](#)
- Samanez-Larkin GR, Hollon NG, Carstensen LL, Knutson B (2008) Individual differences in insular sensitivity during loss anticipation predict avoidance learning. *Psychol Sci* 19:320–323. [CrossRef Medline](#)
- Satoh T, Nakai S, Sato T, Kimura M (2003) Correlated coding of motivation and outcome of decision by dopamine neurons. *J Neurosci* 23:9913–9923. [Medline](#)
- Schmidt L, Braun EK, Wager TD, Shohamy D (2014) Mind matters: placebo enhances reward learning in Parkinson's disease. *Nat Neurosci* 17:1793–1797. [CrossRef Medline](#)
- Scholl J, Kolling N, Nelissen N, Wittmann MK, Harmer CJ, Rushworth MF (2015) The good, the bad, and the irrelevant: neural mechanisms of learning real and hypothetical rewards and effort. *J Neurosci* 35:11233–11251. [CrossRef Medline](#)
- Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* 275:1593–1599. [CrossRef Medline](#)
- Seymour B, Doherty JP, Dayan P, Koltzenburg M, Jones AK, Dolan RJ, Friston KJ, Frackowiak RS (2004) Temporal difference models describe higher-order learning in humans. *Nature* 429:664–667. [CrossRef Medline](#)
- Seymour B, O'Doherty JP, Koltzenburg M, Wiech K, Frackowiak R, Friston K, Dolan R (2005) Opponent appetitive-aversive neural processes underlie predictive learning of pain relief. *Nat Neurosci* 8:1234–1240. [CrossRef Medline](#)
- Skvortsova V, Palminteri S, Pessiglione M (2014) Learning to minimize efforts versus maximizing rewards: computational principles and neural correlates. *J Neurosci* 34:15621–15630. [CrossRef Medline](#)
- Steinberg EE, Keiflin R, Boivin JR, Witten IB, Deisseroth K, Janak PH (2013) A causal link between prediction errors, dopamine neurons and learning. *Nat Neurosci* 16:966–973. [CrossRef Medline](#)
- Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ (2009) Bayesian model selection for group studies. *Neuroimage* 46:1004–1017. [CrossRef Medline](#)
- Sutton RS, Barto AG (1998) Reinforcement learning: an introduction. Cambridge, MA: MIT. Available from: <https://pdfs.semanticscholar.org/aa32/c33e7c832e76040edc85e8922423b1a1db77.pdf>. Accessed June 7, 2017.
- Tobler PN, Fiorillo CD, Schultz W (2005) Adaptive coding of reward value by dopamine neurons. *Science* 307:1642–1645. [CrossRef Medline](#)
- Treadway MT, Buckholz JW, Cowan RL, Woodward ND, Li R, Ansari MS, Baldwin RM, Schwartzman AN, Kessler RM, Zald DH (2012) Dopaminergic mechanisms of individual differences in human effort-based decision-making. *J Neurosci* 32:6170–6176. [CrossRef Medline](#)
- Tsai HC, Zhang F, Adamantidis A, Stuber GD, Bonci A, de Lecea L, Deisseroth K (2009) Phasic firing in dopaminergic neurons. *Science* 324:1080–1084. [CrossRef Medline](#)
- Varazzani C, San-Galli A, Gilardeau S, Bouret S (2015) Noradrenaline and dopamine neurons in the reward/effort trade-off: a direct electrophysiological comparison in behaving monkeys. *J Neurosci* 35:7866–7877. [CrossRef Medline](#)
- Waelti P, Dickinson A, Schultz W (2001) Dopamine responses comply with basic assumptions of formal learning theory. *Nature* 412:43–48. [CrossRef Medline](#)
- Wardle MC, Treadway MT, Mayo LM, Zald DH, de Wit H (2011) Amping up effort: effects of d-amphetamine on human effort-based decision-making. *J Neurosci* 31:16597–16602. [CrossRef Medline](#)
- Zaghloul KA, Blanco JA, Weidemann CT, McGill K, Jaggi JL, Baltuch GH, Kahana MJ (2009) Human substantia nigra neurons encode unexpected financial rewards. *Science* 323:1496–1499. [CrossRef Medline](#)