Mind matters: Placebo enhances reward learning in Parkinson's disease

Liane Schmidt[1], **Erin Kendall Braun**[1], **Tor D. Wager**[2,*], and **Daphna Shohamy**[1,3,*]

[1]Psychology Department, Columbia University, New York, New York

[2]Psychology Department, University of Colorado at Boulder, Colorado

[3]Kavli Center for Brain Science, Columbia University, New York, New York

Abstract

Expectations have a powerful influence on how we experience the world. Neurobiological and computational models of learning suggest that dopamine is crucial for shaping expectations of reward and that expectations alone may influence dopamine levels. However, because expectations and reinforcers are typically manipulated together, the role of expectations per se has remained unclear. Here, we separated these two factors using a placebo dopaminergic manipulation in Parkinson's patients. We combined a reward learning task with fMRI to test how expectations of dopamine release modulate learning-related activity in the brain. We found that the mere expectation of dopamine release enhances reward learning and modulates learning-related signals in the striatum and the ventromedial prefrontal cortex. These effects were selective to learning from reward: neither medication nor placebo had an effect on learning to avoid monetary loss. These findings suggest a neurobiological mechanism by which expectations shape learning and affect.

Expectations are profound incentives for behavior. They affect perception, decision making, and action. Perhaps the most striking example of the power of expectations is the "placebo effect," wherein the mere expectation of medical treatment leads to physiological effects that mimic the benefits of the treatment itself. These effects pose a crucial challenge for clinical trials, as they can obscure the benefits of an effective intervention. However, if harnessed, they offer the promise of enhancing treatment by combining psychological and pharmacological factors. Thus, understanding the neurobiological basis of how placebo treatment affects behavior is of great importance.

In Parkinson's disease (PD) patients, preliminary data from clinical trials[1–3] and a few landmark experimental studies[4–6] suggest that placebo effects may be substantial and

involve modulation of dopamine release in the striatum, presumably via nerve terminals whose origins are located in the midbrain. PD is characterized by changes in motor, cognitive, and emotional systems. Dopaminergic treatment can improve motor symptoms, but it also has broader consequences for behavior, including restored (and sometimes excessive) reward-driven motivation and learning[7]. Although there is evidence for placebo effects on motor function[1,4], placebo effects on the cognitive and affective aspects of PD have not been examined.

Here we sought to investigate the effects of placebo on reward learning in PD and its neurobiological mechanisms. Building on prior data suggesting that placebo treatment may cause endogenous recruitment of brain mechanisms that underpin reward learning[5,8–10], we combined placebo and pharmacological manipulations with a well-characterized instrumental learning paradigm[7,11] and used functional magnetic resonance imaging (fMRI) to link learning behavior to learning- and value-related brain activation.

The ability to learn from rewarding outcomes is known to depend on midbrain dopamine neurons and their striatal and prefrontal targets. In particular, computational models of reinforcement learning have highlighted two key variables related to such learning: *expected value* and *prediction error*[12]. The expected value of an option is assumed to guide choices on each trial. After an outcome is received, this value is updated based on the prediction error, which quantifies how much the outcome deviated from what was expected based on past experience[12–15]. FMRI studies in healthy participants have shown that expected value correlates with blood oxygen level-dependent (BOLD) activity in the ventromedial prefrontal cortex (vmPFC)[16–18], whereas reward prediction errors correlate with activity in the striatum[19–22]. Moreover, the loss of midbrain dopamine neurons due to PD is associated with changes in these reward-related processes[23–25].

Guided by the well-established role of the vmPFC and the ventral striatum in feedback learning, we focused our hypotheses on these regions and asked: (1) Behaviorally, does placebo mimic the effects of dopaminergic medication on learning in PD patients? (2) Does placebo affect value representation in the vmPFC and prediction error representation in the ventral striatum? (3) Does placebo differentially affect learning from reward versus punishment? We predicted that placebo, like dopaminergic medication, will facilitate reward learning in PD; that this effect on behavior will be reflected in modulation of value coding in the vmPFC and reward prediction errors in the ventral striatum; and that the effects of placebo, like actual dopaminergic drugs, would be specific to learning to predict rewards but not learning to predict losses.

A within-patient design allowed us to selectively test the effects of dopaminergic medication (comparing on vs. off drug) and whether or not placebo had a similar effect (comparing placebo vs. off drug) (see Supplementary Fig. 1 for a depiction of conditions and experimental design). Each patient was tested under three conditions: off drug (no treatment), placebo (sham treatment), and on drug (standard dopaminergic medication). In the off drug condition, participants were withdrawn overnight from dopaminergic medication. In the on drug condition, participants received their standard dopaminergic medication crushed into orange juice prior to scanning. In the placebo condition, participants

watched their verum medication being crushed and dissolved in a glass of orange juice. Unbeknownst to them, they were administered a glass of orange juice containing crushed placebo (starch) pills.

While being scanned with fMRI, PD patients performed an instrumental learning task, (Fig. 1) previously shown to be influenced by dopaminergic medication[7,11]. Participants learned by trial-and-error to maximize monetary payoff. For each of a series of trials, participants had to choose between two shapes to obtain a monetary feedback: In the GAIN condition, the correct choice was usually reinforced with a reward of $10, while the incorrect choice was usually not reinforced ($0). This condition tested learning from monetary reward. In the LOSS condition, the correct choice was usually not reinforced ($0), and the incorrect choice was usually punished with a loss of $10. This condition tested learning from monetary punishment and served as a comparison condition against which to test the specificity of effects to reward learning. For each of the two pairs of shapes (i.e. gains, losses) feedback probabilities were either 0.75 or 0.25.

RESULTS

Placebo mimics effects of medication enhancing reward learning

To test the effects of dopaminergic medication and placebo on learning, we estimated learning curves for each patient. These curves reflect the increase in the proportion of correct choices across time as learning progressed. Multilevel linear regression tested the effects of drug and placebo (vs. off drug). PD patients showed significant learning across all conditions. Importantly, there were fundamental differences in observed learning curves between conditions that were also captured by fitted learning curves derived from a standard reinforcement learning model (Fig. 2).

As predicted, reward learning was modulated by both pharmacological (drug) and psychological (placebo) manipulations. Reward learning was enhanced when patients were on drug (block*On>Off drug: $t_{14}=1.8$, p<.05). Critically, reward learning was also enhanced by placebo (block*Placebo>Off drug: $t_{14}=2.3$, p<.05), with no difference between placebo and on drug conditions (block*On>Placebo: $t_{14}=-0.2$, p=.39). An exploratory analysis of individual differences in drug and placebo effects revealed that these learning effects paralleled drug and placebo effects on motor symptoms (Supplementary Fig. 2).

Finally, the effect of placebo on learning was selective to the gain condition, as indicated by a significant three-way interaction (block*Placebo>Off*Gains>Losses: $t_{11}=1.9$, p<.05). In the loss condition, the slope of learning over time was not significantly enhanced by drug or placebo (block*On>Off drug: $t_{14}=0$, p=0.49; block*Placebo>Off drug: $t_{14}=-0.53$, p=.3).

We next analyzed the behavioral data by applying a standard reinforcement learning model[12,14]. Quantification of model fits across the different conditions indicated that the model provided a good fit to the data in all conditions, with no differences between them (see **Online Methods**, Supplementary Table 1). Further, the model-derived learning curves across trials were consistent with the observed learning curves (Fig. 2) confirming that the reinforcement learning model captured the qualitative pattern of performance improvements

detailed above. Additionally, a comparison of model parameters revealed parallel effects of placebo and drug on the model-derived learning rate parameter (Supplementary Table 1). Specifically, in the gain condition the average learning rate under placebo was significantly lower than the learning rate in the off drug condition (t_{14}= 2.6, p<0.05, two-tailed, paired t-test). The average learning rate on drug was numerically but not significantly lower than in the off drug condition (t_{14}=1.5, p=0.16, two-tailed). There were no significant differences in learning rate between placebo and on drug (t_{14}=1.24, p=0.23, two-tailed). In the loss condition, there were no differences in learning rates between treatment conditions (Off vs. Placebo: t_{14}= −0.73, p=0.47; Off vs. On: t_{14}=−1.79, p=0.10; On vs. Placebo: t_{14}=0.97, p=0.34). The finding of reduced learning rates in the reinforcement learning model suggests greater integration of reward information across trials, which is beneficial for this kind of probabilistic learning task.

Placebo enhances value representation in the vmPFC

Having established that placebo selectively enhances reward learning in PD, we next investigated the neural substrates of this effect. To determine whether placebo enhances BOLD correlates of learning, trial-by-trial measures of expected values and prediction errors were derived from the reinforcement learning model[12,14] and regressed against the BOLD signal at the time of choice and feedback, respectively.

Multilevel linear regression tested the effects of drug and placebo (vs. off drug) on parameter estimates from regions of interest in the vmPFC and the ventral striatum. Expected value was associated with vmPFC increases across all conditions (Fig. 3). Crucially, we found that vmPFC activation related to the parameter of expected value was enhanced both by drug (On>Off: t_{11}=2.8, p<0.05) and by placebo (Placebo>Off: t_{11}=2.0, p<0.05), with no difference between drug and placebo (On>Placebo: t_{11}=0.52, p=0.31). The enhancing effect of placebo on value representation in vmPFC was selective to the reward learning condition: vmPFC correlates of expected value during loss learning were not affected by either drug or placebo (On>Off: t_{11}=0.1, p=0.46; Placebo>Off: t_{11}=1.3, p=0.11). Similar results were obtained for correct vs. incorrect responses at time of choice, outside of the reinforcement learning framework, with stronger vmPFC responses during drug and placebo relative to off drug (Supplementary Fig. 3a).

Placebo attenuates striatal prediction error responses

Next, we tested how drug and placebo affected striatal prediction errors at time of feedback. Reward prediction errors were associated with increased activation in the ventral striatum across all conditions (Fig. 4). Prior studies in healthy participants suggest that dopamine increases the correlation between striatal BOLD and reward prediction error[11,23–25]. Interestingly, when we looked at the gain trials we found the opposite: a robust striatal prediction error response in the off drug condition which was weakened by dopaminergic medication (Off>On: t_{11}=2.06, p<0.05) as well as by placebo (Off>Placebo: t_{11}=2.13, p<0.05) (Fig. 4b). As in the effect on expected value, effects on prediction error were found only for reward learning and not for loss learning (Off>Placebo*Gains>Losses: t_8=2.58, p<0.05; Off>On: t_{11}=0.19, p=0.42; Off>Placebo: t_{11}=−1.38, p=0.09). Here too a similar pattern of results was obtained for an analysis of actual observed correct vs. incorrect

feedback at time of outcome, outside of the reinforcement learning framework (Supplementary Fig. 3b): Reward-related responses were reduced in both on drug and placebo responses relative to off drug.

An analysis breaking down the prediction error into its algebraic components[26–28] (expected value and reward, at time of outcome) found that placebo effects on learning signals in the brain were underpinned by a significant effect on the reward component, but not the value component (**Online Methods** and Supplementary Fig. 4). Striatal responses to reward were significantly weaker under placebo than off drug (Placebo<Off: $t_{14}=2.14$, p<0.05, two-tailed, paired t-test) and numerically but not significantly weaker responses in on drug compared to off drug (On<Off: $t_{14}=1.63$, p=0.12, two-tailed, paired t-test), with no difference between placebo and on drug (Placebo>On: $t_{14}=-0.21$, p=0.83, two-tailed, paired t-test). By contrast, there were no differences across treatment conditions related to choice value (Placebo<Off: $t_{14}=1.14$, p=0.27; On<Off: $t_{14}=0.45$, p=0.65, and Placebo > On: $t_{14}=-0.82$, p=0.42, two-tailed, paired-t-test).

In summary, both drug and placebo were associated with better behavioral performance, enhanced value signals in the vmPFC, and weaker reward responses in the striatum. Notably, the selectivity of placebo and drug effects to the reward condition, in contrast to the null-effect on the loss condition, suggests that the placebo effects are not due to global effects of treatment on any broad psychological or physiological variables (see **Online Methods** for a detailed description of control analyses).

DISCUSSION

Our results indicate a robust effect of placebo on brain and behavior. These findings suggests that mere expectations about dopamine release can be harnessed to enhance treatment by facilitating reward learning as well as motor symptoms, with potentially broad implications for appetitive motivation.

Expectations have long been known to have a substantial influence on how we experience the world[29,30]. Neurobiological and computational models of learning suggest that expected reward should lead to neurobiological changes in midbrain dopamine systems and their prefrontal and striatal targets. However, since a) the vast majority of studies manipulate expectations and reward at the same time, and b) motivational priming in the striatum can occur without conscious expectation[31,32], the role of expectations in and of themselves has remained unclear. Here we show that the mere expectation of dopaminergic treatment enhances reward learning, as we manipulated beliefs about treatment independent of drug-associated cues (both drug and placebo were delivered in orange juice, not in pills). Thus, it was the belief in treatment that influenced reward processing and learning, paralleling effects of instructions on affective processes in other domains[33].

Expectations of dopaminergic effects without actual drug mimicked the effects of drug-induced dopaminergic modulation on both model-independent learning curves and model-based learning rates. These findings suggest that part of the benefit of open-label drug treatment delivered in standard clinical settings may be caused by the expectations of the

patient. Moreover, our finding that responses to drug may be partly related to effects of expectation suggests that one of the basic assumptions of standard randomized controlled trials is violated. Attempts to screen out placebo responders or to subtract away placebo responses to reveal drug effects may be misguided, as the drug response will be eliminated along with the placebo[34]. Much prior work on these principles has been done in pain and depression[35], and one study has shown that placebo- and levodopa-induced dopamine effects in PD patients can covary[6]. Our results extend these findings to appetitive processes and suggest that concerns about the non-independence of placebo and drug mechanisms extends to PD as well.

Importantly, the observed effects of placebo were selective to learning from gains, consistent with previous findings demonstrating asymmetry in the effects of dopamine modulation on appetitive vs. aversive learning[11]. However, while prior studies have reported that in some cases dopaminergic medication can impair loss learning[36–38], here we find a lack of effect on learning to avoid losses, both for medication and for placebo. This may be related to the specific design of our study, in which positive and negative prediction errors are present in both the reward and loss conditions and in which the lack of a loss outcome could be interpreted as a reward. Future studies constraining participants' interpretation of gains vs. losses and comparing performance of value-learning and actor-critic models may be better suited to addressing this issue[39,40].

Our findings demonstrated decreased ventral striatal BOLD responses to reward and prediction error under both dopaminergic medication and placebo treatment. The direction of this effect is somewhat surprising, given that dopaminergic medication (and, putatively, placebo) is known to increase striatal dopamine. One possible explanation relates to the mechanism by which dopamine modulates corticostriatal synaptic activity. It has been shown that dopamine can increase the signal to noise ratio in corticostriatal synapses via a combination of inhibition and excitation[41–43]. Given the resolution of the BOLD response, such modulation could lead to a net decrease in reward related BOLD activity, while still reinforcing the behaviorally relevant synapses. Notably, reduced striatal responses related to rewards and prediction errors in PD have also been reported previously, though these effects may depend on the patient characteristics[7]. Though the precise relationships between dopamine release, circuit dynamics, and BOLD activity are largely unknown, these findings suggest that in PD patients, increased dopamine in the striatum may be accompanied by decreased prediction error-related BOLD signals.

Finally, placebo effects are widely thought to be created both by expectations and by learning, as conditioning has been shown to elicit some of the most robust placebo responses[44,45]. Here, we demonstrate that in addition to the effects of expectancy and learning on placebo effects in brain and behavior, there are also significant effects of expectancy induced by a placebo on learning and learning-related signals in the brain. Together with the assumption that placebo and drug effects in PD might be mediated by the same dopaminergic mechanism in the brain, this finding creates the potential for feedback cycles in which expectations and experienced reward are mutually reinforcing. Thus, the learned association between a drug and its outcome could be reinforced by expectations induced by verbal suggestions and other elements of the treatment context (e.g. the physical

location of treatment and social interactions with care providers). Such synergy between expectations, pharmacology and behavior could have a positive cumulative effect on patients' motivation, response to treatment, and quality of life.

ONLINE METHODS

Participants

We recruited twenty-one patients with mild-to-moderate Parkinson's disease (PD; 13/8 male/female) through neurologists at Columbia University's Center for Parkinson's Disease and Other Movement Disorders and advertisements on the websites of the Michael J. Fox Foundation and the Parkinson's Disease Foundation. Columbia University Medical Center Institutional Review Board approved the study procedures, and we obtained informed written consent from all participants. During recruitment, patients were told that the study aim was to test how PD medications, such as carbidopa/levodopa, affects learning, memory and decision making and the way the brain functions. Thus until a final debriefing session at the end of the experiment, participants remained blind to the placebo. Patients were remunerated for their time ($20 per hour), as well as 5% of their winnings averaged across the six learning task runs.

Before participation, all of the patients were screened using the following inclusion criteria: Diagnosis of mild-to-moderate PD (Hoehn and Yahr stages 1 to 3), age between 50 and 85 years, treatment of PD by carbidopa/levodopa for at least 3 months, responsive to PD medication, medication schedule with at least 2 and up to 4 daily doses (e.g. morning, lunch, afternoon, bedtime), no report of psychiatric history or neurological disease other than PD, normal to corrected-to-normal vision, and right-handedness. It has been shown that even at the initial diagnosis, PD patients have already lost a significant amount of striatal dopamine[46]. Thus for this initial study, we deliberately recruited patients at a stage with substantial dopamine loss and enough experience with the medication to elicit expectations about its effects, while having mild enough symptoms to perform a cognitive task within an MRI scanner. A formal power analysis was not used to select the sample size a priori, as the effect sizes could not be estimated prior to the study. The sample sizes was chosen to be similar to those generally employed in the field.

All patients were treated with levodopa/carbidopa. Additionally patients were prescribed the following PD medications: Selegiline (5 patients), Rasagiline (5 patients), Amantadine (4 patients), Entacapone (1 patient), and Pramiprexole (1 patient). Because these medications were part of the PD treatment, they were crushed and dissolved in orange juice together with the levodopa/carbidopa medication. Some patients also were prescribed antidepressant and anxiolytic treatment that included Nortriptyline (2 patients), Amitriptyline (1 patient), Citalopram (1 patient), Clonazepam (2 patients), and Lorazepam (2 patients). On the morning of the experiment, patients did not take any antidepressant or anxiolytic medications.

We excluded six patients from analysis due to: a) extreme anxiety related to scanning (n=4, 2 of these 4 patients agreed to participate behaviorally, the other 2 patients withdrew completely), b) abnormally fast overall reaction times (3 SD below the group average

response times), suggesting a failure to follow task instructions (n=1, excluded from both behavioral and imaging data), and c) poor fMRI image quality (n=1, included in the behavioral data analysis). In sum, we included 18 patients in the behavioral data analysis and 15 patients in the fMRI data analysis.

Experimental design

To minimize variance due to individual differences, we employed a within-subject design in which each patient was scanned three times: off drug, on placebo and on drug (see Supplementary Fig. 1a). Patients refrained from taking their morning dose of PD medication, following conventional procedures for comparing PD patients on vs. off medication[47–50]. Patients were withdrawn from their medication for at least 16 hours. The order of the off drug and placebo sessions was counterbalanced across patients; however, the on drug condition was always last, since the half-life of levodopa/carbidopa combination requires at least 7 hours to be metabolized and excreted. Placebo and levodopa/carbidopa treatment were each administered 30 minutes prior to scanning.

Placebo administration

On the morning of the experiment and prior to scanning, patients were interviewed regarding their daily medication schedule. In particular, patients were asked: (1) how many daily doses they take, (2) at what time of the day, and (3) if and when they feel an improvement from the drug. Patients were told that because of the initial time to 'kick in' between 30 to 45 minutes, their typical daily dose of medication would be administered 30 minutes prior to scanning and thus, together with the time it will take to place them in the scanner (15 minutes), they could expect to be at their best at start of the learning task.

All patients were off their PD medication during the placebo scan session, but they believed that they had taken their medication. Thirty minutes before the placebo scan session started, the patients observed the experimenter crushing and dissolving their daily morning PD medication into a glass of orange juice. Importantly, to avoid carry over into subsequent sessions, we told patients that the reason for crushing the medication (i.e. destroying its coating) and dissolving it into orange juice was to make it 'faster acting' and 'shorter lasting'. Unbeknownst to the patients, we replaced the glass of orange juice containing their medication with a glass of orange juice containing placebo (starch) pills. Thus, participants were led to expect that they were taking their typical medication, and it was implied that the medication would help them function optimally on all tasks, without providing specific expectancies about performance. Participants were not informed that their medication, or dopamine in general, might specifically improve reward–based learning; however, in presenting patients with the set of tasks we did, it is likely that participants expected that performance on all tasks (gain–based learning, loss–based learning, motor) might depend on dopamine and be affected by their daily medication in some positive way.

All patients enrolled in this study were on medication schedules of two to four doses per day (e.g. morning, lunch, afternoon, bedtime). If the placebo session was the first scan session, patients were told they were taking their morning dose; if the placebo session was the second scan session, patients were told they were taking a late morning/lunch dose; and for

the third on drug scan, patients were told that they were taking a late lunch/early afternoon dose.

The rationale for this placebo manipulation was that patients may benefit from conditioned responses that have built up over the duration of their Parkinson's therapy. Our goal was to address the question: Does the expectation of relief from PD symptoms associated to taking a medication affect reward learning and learning–related signals in the brain? By making the patients believe they were taking their daily medication, we intended to induce patient-specific expectations and associations, which are built up over months and years of experience with the positive and negative contingencies of a daily medication and cannot be elicited by an unfamiliar study drug.

Drug administration

The same procedure was repeated 30 minutes before the third scan session (on drug session). Again, the patients observed the experimenter crushing and dissolving their second dose of daily PD medication into orange juice; however, this time glasses were not switched, and the patients were administered their medication.

Timing of sessions

Since we administered both the placebo and drug 30 minutes before scanning, each scan session last for 60 minutes and there was a 60 minutes break, the placebo session and the following session were separated by 2 hours and 30 minutes. This timing follows the design used by a PET study of placebo effects on dopamine release[51].

Although we cannot completely rule out the possibility of a persistence of the placebo effect into subsequent scan sessions (i.e. off or on drug), we believe it is unlikely to confound our findings. The critical comparison between the placebo and off drug conditions was counterbalanced across patients. Any persistence of placebo effects into subsequent off drug sessions would decrease differential effects (i.e. placebo vs. off drug) on behavior and brain activity. Moreover, recent work by Benedetti et al. on placebo in PD suggests that placebo effects on PD last on the order of 30 minutes[52,53], though much remains to be learned about the time course of placebo effects across samples and placebo inductions.

Debriefing on placebo manipulation

At the end of the experiment, we informed patients that one of the two drug doses administered during testing was a placebo, and they were asked to guess which dose was the placebo. Five patients guessed incorrectly (i.e. they thought the placebo was the real drug), five patients guessed correctly, and seven patients reported that they could not guess.

Effects of treatment on clinical symptoms

After each scan session, emotional symptoms were assessed using the Starkstein Apathy Scale[54] and the State–Trait Anxiety Inventory (STAI)[55]. Working memory was assessed with the Digit Span, and motor symptoms were assessed with the Unified Parkinson's Disease Rating Scale III (UPDRS III)[56]. Additionally, expectancy ratings were collected before each scan session: patients indicated how much they expected their PD symptoms to

improve following treatment using a visual analogous scale ranging from 0 (no improvement) to 100 (maximum improvement).

Paired one–sample t–tests between treatment conditions were not significant for apathy, anxiety or digit span measures. However, expectancy scores indicated that patients expected significantly more improvement after having been administered a placebo or their drug, relative to the off drug condition (Placebo>Off: $t_{17}=2.9$, $p<0.05$; On>Off: $t_{17}=3.7$, $p<0.05$, two–tailed, paired t–test) (Supplementary Table 2).

We videotaped patients while the UPDRS III was assessed (no rigidity measures were collected), and an experimenter blind to treatment rated motor symptoms. We observed a significant improvement due to being on drug in comparison to being either on placebo (On>Placebo: $t_{17}=-5.1$, $p<0.001$, two–tailed) or off drug (On>Off: $t_{17}=-4.1$, $p<0.001$, two–tailed). Since different individuals typically show medication-related improvements on a unique set of patient-specific motor behaviors, and patients were in the mild-to-moderate stages of the disease that may restrict drug effects to specific symptoms, we also calculated UPDRS III scores for off and placebo across the specific subset of items that showed a drug effect (items where On – Off < 0) for each patient. We found that for the individualized sub-items that showed an improvement with medication we also found an improvement with placebo (Placebo<Off: 7.86 ± 1.21 vs. 10.2 ± 1.59: $t_{17}=4.2$, $p<0.001$, two-tailed, paired t-test).

Behavioral task

In each of the three fMRI scan sessions, patients performed two runs of an instrumental learning task. The aim of the task was to maximize monetary payoff. Each run consisted of 64 trials: 32 gain trials and 32 loss trials. Gain and loss trials were randomly intermixed within a run. On each trial, two shapes were displayed on a computer screen, and patients chose one of them to receive a monetary feedback. Choices were followed by one of three possible outcomes. Subjects could win $10, get nothing ($0), or lose $10. In the GAIN condition choosing the optimal (i.e. correct) shape was reinforced with a reward of $10 on 75% of trials (25% of trials were non-reinforced), while choosing the non-optimal (i.e. incorrect) shape was not reinforced ($0) in 75% of trials (25% of the trials were rewarded $10). In the LOSS condition, the optimal choice (i.e. correct) was not reinforced ($0) on 75% of trials (but was punished on 25% of trials –$10), and the non-optimal (i.e. incorrect) choice was punished with a loss of $10 on 75% of trials (but 25% of trials were not punished $0). Thus, in gain trials we tested the ability to learn from financial rewards (reward learning), and in loss trials we tested the ability to avoid financial punishment (loss learning).

We jittered the duration of inter-trial intervals (2 to 6 s) and of the choice-feedback interval (0.5 to 3 s) by drawing from an exponential distribution. We paid the patients according to their performance (5% of the payoff averaged across the 6 learning task runs), in addition to compensation for their participation.

Individual differences in drug and placebo effects

The effects of drug and placebo on the slope of reward learning trended towards a correlation across subjects, (zero-order correlation of placebo and drug effects: $r=0.59$, $p<$.

01; partial correlation controlling for Off drug: r=0.34, p=.08, one–tailed) (Supplementary Fig. 2). An exploratory analysis of the UPDRS III scores found parallel effects of placebo and drug for clinically relevant motor symptoms of PD (zero-order correlation placebo and drug effects: r=0.68, p<0.001; partial correlation controlling for Off drug: r=0.59, p<0.01).

Computational model

We used a standard temporal difference learning algorithm to calculate the expected value of choices (i.e. choice value) and prediction error based on individual trial-by-trial choices and feedback[57]. The expected value (i.e. choice value), termed a Q-value, corresponded to the expected reward obtained by choosing a particular cue. Q-values were set to zero at the beginning of each run. After each trial, t>0, the value of the chosen cue was updated according to the rule, $Q_{chosen\_cue}(t+1) = Q_{chosen\_cue}(t) + \alpha*\delta(t)$. Central to this algorithm is prediction error, $\delta(t)$, which is defined as $\delta(t) = R(t) - Q_{chosen\_cue}(t)$ or the difference between expected feedback $Q_{chosen\_cue}(t)$ and actual feedback $R(t)$. The reward magnitude was set +1, 0 and −1 for outcomes +$10, $0 and −$10, respectively.

Given the Q-values, the probability of selecting each action was calculated by implementing the softmax rule:

$$P_{chosen\_cue}(t)=\exp(Q_{chosen\_cue}(t)/\beta)/(\exp(Q_{chosen\_cue}(t)/\beta)+\exp(Q_{unchosen\_cue}(t)/\beta)).$$

The two free parameters, α (learning rate) and β (temperature), were fit individually for each subject to maximize the probability of actual choices under the model. To improve the model's subsequent fit to fMRI data, we used the average estimates for α and β calculated across patients in each treatment condition.

Note that we tested potential confounds due to differences in learning rate and temperature between gain and loss trials. In a control analysis, one learning rate (α=0.2) and one temperature (β=0.4) fitted across all subjects, treatment and trial conditions was used to calculate trial-by-trial expected values and prediction errors. The control fMRI results were consistent with main findings and are reported below (see Effects of scaling expected value and prediction error).

Comparison of RL model fits between treatments

To rule out potential confounds due to qualitative differences in model fit between off drug, placebo, and on drug, the following standard measures of model fit were calculated for each subject in each treatment condition: (1) the Bayesian Information Criterion defined by BIC=2*LLE+2*log(n), (2) the Akaike Information Criterion defined by AIC = 2 * LLE + 2 * 2, (3) pseudo R^2 values defined by pseudoR^2=1−LLE$_{model}$/LLE$_{chance}$. Moreover, we tested how well the RL model fits the observed data compared to chance by performing a likelihood ratio test defined by D=2*LLE$_{model}$−2*LLE$_{chance}$. LLE$_{model}$ corresponds to the maximum logarithmic likelihood of the observed choices under the model. LLE$_{chance}$ corresponds to the logarithmic likelihood of choices at chance (LLE$_{chance}$=n*log(0.5), with n=number of trials). The average of each measure across patients in each condition and treatment are reported in Supplementary Table 1.

Analysis of variance compared pseudo-R2 values across treatments (off drug, placebo, on drug,) and across learning conditions (gains, losses) and revealed no main effect of treatment ($F_{(2,107)}=0.79$, $p=0.45$), no main effect of gains vs. losses ($F_{(1,107)}=2.38$, $p=0.12$), and no interaction of treatment by gains vs. losses ($F_{(2,107)}=0.16$, $p=0.85$). Post hoc, one-tailed t-tests further tested differences in pseudo-R2 values across treatments. These too revealed no significant differences between treatments (pseudo-R2 values: 'On >Off' gains: $t_{17}=1.16$, $p=0.12$; losses: $t_{17}=0.50$, $p=0.30$; 'Placebo > Off' gains: $t_{17}=0.26$, $p=0.39$; losses: $t_{17}=-0.34$, $p=0.63$; 'On>Placebo' gains: $t_{17}=0.83$, $p=0.20$; losses: $t_{17}=1.17$, $p=0.13$).

Treatment effects on reinforcement learning parameters

To investigate individual differences between treatment conditions, we individually estimated the free parameters, learning rate, and inverse softmax temperature, which are also referred to as random effect estimations (i.e. one set of parameters for each subject for each treatment condition). The average learning rates and temperature across subjects in each treatment condition are summarized in Supplementary Table 1. Analysis of variance tested for main effects of treatment, gains vs. losses, and the interaction of treatment by gains vs. losses in learning rates and softmax inverse temperature, separately.

For the learning rates, we found no main effect of treatment ($F_{(2,89)}=0.94$, $p=0.39$), a significant main effect of trial type ($F_{(1,89)}=3.65$, $p<0.05$), and a trend interaction treatment by gains vs. losses ($F_{(2,89)}=2.4$, $p=0.09$). When comparing the treatments during learning from gains, patients on placebo had significantly smaller learning rates compared to off drug (Off>Placebo: $t_{14}=2.6$, $p<0.05$) and smaller learning rates on drug compared to off drug (Off>On: $t_{14}=1.5$, $p=0.16$), with no differences between placebo and on drug (On>Placebo: $t_{14}=1.24$, $p=0.11$). The numerical differences in learning rate across treatment conditions, although less robust statistically, were consistent with behavioral performances and suggest that placebo led to more incremental learning. We did not find differences between treatment conditions for learning rates in the loss condition (Off>Placebo: $t_{14}=-0.73$, $p=0.47$; Off>On: $t_{14}=-1.79$, $p=0.18$; On>Placebo: $t_{14}=0.97$, $p=0.34$) (Supplementary Table 1).

Behavioral data analysis

Data collection and analysis were not performed blind to the conditions of the experiment, but all analyses were conducted using automated methods and a priori statistical tests to avoid the possibility of experimenter bias. All of the motor scores (UPDRS III) were coded by a rater blind to the experimental conditions.

To assess learning, the 32 gain and the 32 loss trials were each binned into 4 blocks of 8 trials (1st, 2nd, 3rd, and 4th block). Within each block, we calculated the percentage of optimal choices, resulting in a learning curve composed of these 4 optimal choice scores for each patient. These curves reflected the increase in the proportion of correct choices across time (i.e. 1st, 2nd, 3rd, and 4th block) as learning progressed. Multilevel general linear models (GLM) using iterative generalized least squares (IGLS)[58] tested for drug and placebo effects on reward and loss learning.

GLM1 tested for a drug effect and included the following 3 regressors:

Block, On > Off drug, Block*On > Off drug.

Block (dummy coded: −3, −1, 1, −3) tested the increase of learning across time (i.e. the slope of learning over the 4 blocks of 8 trials). On>Off drug tested for a main effect of drug compared to Off drug (coded: −1 for Off drug, 1 for On drug). Block*On>Off drug tested if the slope of learning On drug was enhanced compared to Off drug.

GLM2 tested for a placebo effect, analogous to GLM1:

Block, Placebo > Off drug, Block*Placebo > Off drug.

GLM3 tested for the selectivity of effects on reward learning compared to punishment learning. It included the following 6 regressors:

Block, Placebo > Off, Gains > Losses, Block*(Gains > Losses), (Placebo > Off drug)*Gains > Losses, Block*(Placebo > Off drug)*Gains > Losses.

We fit individual regression coefficients obtained for each regressor into a second-level group analysis to provide population inferences against zero. We entered the order of placebo administration as a second level covariate (see Order effect section, below, for further control analyses of potential confounds due to treatment order). We report the significance of regression coefficients at a threshold of $p<0.05$, one-tailed.

We applied an analogous analysis to parameter estimates extracted from regions of interest in the brain (see Analysis of parameter estimates) and reaction times for choices (see Reaction time).

To test if drug and placebo effects on reward learning covaried on the individual level, we obtained individual slopes of reward learning by multiplying the 4 trial blocks (dummy coded: −3, −1, 1, 3) by the percentage of optimal choices in each block. The differences in learning slopes, on – off drug and placebo – off drug, were then correlated across all patients by using Pearsons's correlation coefficient and partial correlations controlling for off drug (Supplementary Fig. 2). An analogous correlation analysis was done in order to test for individual placebo (i.e. placebo – off drug) and drug effects (i.e. on – off drug) on motor symptoms measured by the UPDRS III (Supplementary Fig. 2).

All statistical tests were conducted with the Matlab Statistical Toolbox (Matlab 2010b; The MathWorks) and a Matlab implementation of iterative generalized least squares (IGLS). No statistical methods were used to predetermine sample sizes, but our sample sizes are similar to those generally employed in the field. Data distribution was assumed to be normal and was examined visually to check for normality. Statistical tests for normality were not conducted, because they are generally inconclusive for sample sizes in the range employed here.

Reaction times

Reaction times (RT) measured the time patients took to choose between the two shape choices on each trial. Average RTs for each patient were calculated for gain and loss trials

separately. To test if there were significant differences in RTs across trials, we applied the same behavioral analysis used to assess learning to the RT data. RTs for gain and loss trials were averaged across blocks of 8 trials separately. Two GLMs, analogous to the GLMs that analyzed optimal choices, tested for placebo and drug effects, respectively (see behavioral analysis). RTs decreased with time in the gain condition ($t_{14}=-6.1$, $p<0.05$), but not in the loss condition ($t_{14}=-0.64$, $p=0.5$) (Supplementary Fig. 5). We found no significant effect of either drug or placebo on RTs.

Learning in the subset of fMRI patients

Behavioral results in the 15 fMRI patients were consistent with the findings observed across the whole patient group. Learning from gains was enhanced by dopaminergic medication (Block*On>Off: $t_{11}=1.8$, $p<0.05$) and by placebo (Block*Placebo>Off: $t_{11}=2.3$, $p<0.05$) (Supplementary Figs. 6 and 7). We did not observe treatment effects for learning from losses in the subset of fMRI patients either.

Order effects

Although the order of placebo and off drug sessions was counterbalanced across patients, the on drug session was always the last session of the experiment. To control for order effects, we regressed the order (dummy coded: 1, 2, 3) of the scan sessions against optimal choices and parameter estimates from the ventromedial prefrontal cortex (vmPFC) and ventral striatum, separately. For both the gain and the loss condition, no significant effect of order was observed (on optimal choices in gain trials: $t_{16}=0.89$, $p>0.05$; in loss trials: $t_{16}=1.6$, $p>0.05$; on parameter estimates from the vmPFC for choice value in gain trials: $t_{13}=0.65$, $p>0.05$; in loss trials: $t_{13}=-0.10$, $p>0.05$; on parameter estimates from the ventral striatum for prediction error in gain trials: $t_{13}=-1.16$, $p>0.05$; in loss trials: $t_{13}=-0.2$, $p>0.05$). To further test if the order of the placebo administration interacted with placebo effects, order (dummy coded: -1 for placebo first, 1 for placebo second) was a second level covariate in the GLMs testing for placebo effects on optimal choices and parameter estimates from vmPFC and ventral striatum.

No second level interactions between order and block ($t_{14}=0.6$, $p>0.05$), Placebo>Off drug ($t_{14}=0.09$, $p>0.05$), or Block*Placebo>Off drug ($t_{14}=-0.16$, $p>0.05$) were found for learning from gains. Similarly, no second–level interactions between order and block ($t_{14}=-1.46$, $p>0.05$), Placebo>Off drug ($t_{14}=1.1$, $P>0.05$) or Block*Placebo>Off drug ($t_{14}=2.3$, $p>0.05$) were observed for learning from losses.

Consistent with the behavioral findings, order of placebo administration did not interact with the effect of Block (gains: $t_{11}=0.2$, $p>0.05$; losses: $t_{11}=0.76$, $p>0.05$), Placebo>Off drug (gains: $t_{11}=-0.2$, $p>0.05$; losses: $t_{11}=-1.3$, $p=0.05$), On>Off (gains: $t_{11}=1.1$, $p>0.05$; losses: $t_{11}=0.96$, $p>0.05$), Block*Placebo>Off drug (gains: $t_{11}=0.4$, $p>0.05$; losses: $t_{11}=-0.4$, $p>0.05$), or Block*On>Off drug (gains: $t_{11}=0.7$, $p>0.05$; losses: $t_{11}=-1.1$, $p>0.05$), choice value activation in the vmPFC or prediction error activation in the ventral striatum (Block (gains: $t_{11}=0.7$, $p>0.05$; losses: $t_{11}=0.9$, $p>0.05$), Placebo>Off drug (gains: $t_{11}=-0.7$, $p>0.05$; losses: $t_{11}=0.2$, $p>0.05$), On>Off (gains: $t_{11}=0.06$, $p>0.05$; losses: $t_{11}=0.5$, $p>0.05$),

Block*Placebo>Off drug (gains: $t_{11}=-0.7$, $p>0.05$; losses: $t_{11}=0.5$, $p>0.05$), or Block*On>Off drug (gains: $t_{11}=-0.3$, $p>0.05$; losses: $t_{11}=-0.3$, $p>0.05$)).

Pleasantness and liking ratings of cues

Following each run of the learning task, patients were asked to rate their preference and liking of the shapes that were presented during learning. Patients rated preferences on a 4-point Lickert scale and liking on a visual analog scale that ranged from most unpleasant to most pleasant. Ratings were subjected to a three-way analysis of variance to test for effects of treatment (off, placebo, on drug), shape (optimal, non-optimal), and task condition (gains, losses). Analysis of variance revealed that patients preferred ($F(1,211)=92.04$, $p<0.001$) and liked ($F(1,211)=60.2$, $p<0.001$) the optimal shape more than the non–optimal shape. This was selective to the gain condition (preferences: $F(1,211)=30.7$; $p<0.001$, liking: $F(1,211)=14.97$, $p<0.001$) as indicated by a significant interaction by task condition (preferences: $F(1,211)=2.85$, $p=0.09$; liking: $F(1,211)=6.04$, $p<0.05$). No main effect of treatment (preferences: $F(2,211)=0.66$, $p=0.51$), liking: ($F(2,211)=0.76$, $p=0.47$) or any other interaction were significant.

Image acquisition

We acquired T2*-weighted echo spiral in/out images with blood oxygenation level dependent (BOLD) contrast using a 1.5 tesla scanner (GE Medical Systems). Each volume comprised of 26 interleaved, axial slices of 4 mm thickness and 3.5 mm in-plane resolution, using the following parameters: TR=2000 ms, TE=35 ms, flip angle=84°, FOV=192 mm, 64×64 matrix. For each subject a total number of 600 volumes were obtained, 300 in each learning task run, corresponding to 10 minutes of scanning for each run. Additionally, a single high resolution T1-weighted structural image (spoiled gradient echo sequence (SPGR)) was acquired for each patient at the beginning of the on drug scan session, prior to performing the on drug learning task runs of the experiment.

FMRI analysis

All functional images were screened on quality using in house software. Specifically, spike and high movement time points were identified and corrected by interpolating adjacent time points. In addition to this, images were screened on fast movement (1 mm/TR) and corrected using the Art Repair toolbox[59].

Analysis of fMRI data was conducted using the statistical parametric mapping software SPM8 (Wellcome Department of Imaging Neuroscience, London, UK). We normalized the structural scans to a standard T1 template and created a normalized structural average image upon which the t-maps were superimposed for anatomical localization. Preprocessing of the functional scans consisted of spatial realignment to the first volume, spatial normalization to a standard T2* template with a resampled voxel size of 2×2 mm, and spatial smoothing using a Gaussian kernel with a full width at half maximum of 6 mm. We analyzed the functional images in an event–related manner, within a general linear model (GLM), and included subject specific realignment parameters into each GLM as covariates to control for motion artifacts.

First, we tested if expected value activated the vmPFC and if prediction error activated the ventral striatum irrespective of trial condition (i.e. gains, losses) or time point (1st, 2nd, 3rd, or 4th block). Two onset regressors (i.e. onset choice, onset feedback) were respectively modulated by expected value and prediction error, obtained for each trial by a standard temporal difference learning algorithm that assigns a value to each cue based on the subjects' previous choices and reward received. Delta functions of onset regressors and parametric modulators were convolved with the canonical hemodynamic response function and regressed against each subject's fMRI data. Linear contrasts were fit into a second-level random effects analysis, which used one-sample t-tests to identify brain regions underpinning expected value and prediction error in each treatment condition. We applied small volume correction to determine the effects of expected value and prediction error in each treatment condition with a lenient initial threshold of $p < 0.01$, uncorrected.

Having identified vmPFC responses to expected value and ventral striatum responses to prediction error, we investigated the specificity of these activations. Building on the behavioral results that suggested that placebo and drug selectively enhance the slope of learning from gains, we used a second GLM to model fMRI data in each subject similarly to the behavioral analysis. Specifically, trials at time of choice and feedback were binned into four regressors per trial condition (i.e. gain trials, loss trials). Each onset regressor reflected a given time during the task (i.e. the 1st, 2nd, 3rd, and 4th block of 8 trials) and was modulated by expected value (at choice onset) and prediction error (at feedback onset), respectively. Delta functions of the 16 onset regressors and the 16 parametric modulators were convolved with the canonical hemodynamic response function and regressed against each subject's fMRI data. Average parameter estimates for the parametric modulators were extracted from regions of interest (ROIs) located in the vmPFC (for expected value) and the ventral striatum (for prediction error). Thus, we obtained, per patient for both expected value and prediction error, 4 parameter estimates for the gain condition and 4 parameters estimates for the loss condition.

Analysis of parameter estimates

A multilevel linear regression model using iterative generalized least squares (IGLS)[58] tested for drug and placebo effects. Analogous to the analysis of optimal choices, the parameter estimates from each ROI were fit by GLM1 (testing for drug effects), GLM2 (testing for placebo effects), and GLM3 (testing for interaction with trial type: gains, losses) from the behavioral analysis. Individual betas were entered into second–level group analyses to provide population inferences against zero. We entered the order of placebo administration (dummy coded: −1 for placebo first, 1 for placebo second) as a second–level covariate. Similarly to our behavioral findings, order did not confound drug and placebo effects. We used a statistical significance threshold of $p < 0.05$, one-tailed, to evaluate regression coefficients.

Note, we observed significant main effects of drug (On>Off drug) and placebo (Placebo>Off drug) on both vmPFC and the ventral striatum parameter estimates. To better display the main effects, parameter estimates were averaged across patients and blocks (i.e. the 1st, 2nd, 3rd, and 4th block of 8 trials) in Figs. 5b and 6b.

Regions of interest

Regions of interest in the vmPFC and the ventral striatum were defined a priori from independent fMRI studies. The vmPFC ROI was defined by the Montreal Neurological Institute (MNI) space coordinates [x=−1, y=27, z=−18] reported for value by Hare et al. 2008[60]. The ventral striatum ROI was defined by the MNI coordinates [x=−10, y=12, z=−8] reported for striatal prediction errors by Pessiglione et al. 2006[61]. Average parameter estimates (i.e. betas) were extracted from a 10 mm diameter sphere centered on the vmPFC and ventral striatal coordinates.

Direct comparison of treatment conditions

Direct comparisons of treatment conditions did not reveal any differences in whole brain activation for expected value or prediction error at an uncorrected threshold of $p<0.001$. See Supplementary Fig. 8 for whole brain activations at $p<0.05$ uncorrected.

Effects of scaling choice value and prediction error

One possible interpretation of our findings is that the observed brain activations are due to individual variance in learning. For example, choice values or prediction errors may reach a different maximum and vary on different scales between treatment conditions and subjects. Thus, to rule out the possibility these differences in parametric modulators contributed to the observed drug and placebo effects, we z-scored expected values and prediction error regressors. The temporal difference learning algorithm used to obtain trial-by-trial expected values and prediction errors used one optimal learning rate (α=0.2) and one optimal temperature (β=0.4), which were fit across all subjects, treatment conditions, and trials. This was done in order to avoid additional confounds due to differences in learning model between subjects, treatment, and trial conditions (i.e. gains, losses).

Consistent with our main findings, this supplemental GLM analysis, which controlled for scaling and learning model differences, revealed that the placebo and on drug conditions enhanced choice value representation in the vmPFC (Placebo>Off drug: t_{11}=1.4, p=0.08; On>Off drug: t_{11}=1.7, p<0.05) compared to off drug, with no difference between placebo and on drug (On>Placebo: t_{11}=0.75, p>0.05). Prediction error activation of the ventral striatum was decreased (Placebo<Off drug: t_{11}=2.1, p<0.05; On<Off drug: t_{11}=1.8, p<0.05) when patients were treated by a placebo or dopaminergic medication compared to off drug with no difference between placebo and on drug (On<Placebo: t_{11}=0.23, p>0.05). The observed placebo and drug effects were selective to learning from gains. No significant differences in treatment conditions were observed for learning from losses. Therefore, we concluded that placebo and drug effects on vmPFC and ventral striatum BOLD responses were not confounded by individual differences in magnitude and scale of learning.

Learning related brain responses outside the RL framework

To test for learning related responses in vmPFC and ventral striatum outside of a reinforcement learning model framework, we used raw observed responses and replaced the expected value regressor at time of choice with actual observed choices (dummy coded: 1 for optimal choice, −1 for non-optimal choice) and replaced the prediction error regressor at time of outcome by actual observed feedback (dummy coded: 1 for correct, −1 for

incorrect). Delta functions of the onset regressors and parametric modulators were convolved with the canonical hemodynamic responses function and regressed against each individual's fMRI data. We then used paired t-tests to examine the effects of placebo and drug on parameter estimates in independently defined regions of interest in the vmPFC and the ventral striatum.

The results of this analysis are strikingly similar to what we observed with the model–derived regressors for expected value and prediction error. Choice–related activation in the vmPFC was enhanced both by drug and by placebo compared to off drug (On>Off drug: $t_{14}=2.9$, $p<0.05$; Placebo>Off: $t_{14}=2.38$, $p<0.05$, two–tailed, paired t–test), with no differences between drug and placebo (On>Placebo: $t_{14}=1.05$, $p=0.3$, two-tailed) (Supplementary Fig. 3). Moreover, these effects were selective to learning from gains, as indicated by a trend interaction (Gains>Losses*Placebo>Off drug: $t_{14}=1.96$, $p=0.06$, two–tailed). The interaction was significant for (Gains>Losses*On>Off: $t_{14}=2.9$, $p<0.05$, two–tailed). No significant differences were found for learning from losses (Supplementary Fig. 3a).

Activation to correct vs. incorrect feedback in the ventral striatum was also similar to the results from RL-model derived analyses: we found greater activation for correct vs. incorrect feedback during learning from gains when patients were off drug compared to on drug (Off>On: $t_{14}=1.97$, $p=0.06$, two–tailed) or placebo (Off>Placebo: $t_{14}=2.27$, $p<0.05$, two-tailed). No differences were observed between on drug and placebo (On>Placebo: $t_{14}=0.63$, $p=0.5$, two-tailed). Again, this finding was selective to learning from gains: the interactions (Gains>Losses*Off>Placebo: $t_{14}=2.9$, $p<0.05$, two–tailed) and (Gains>Losses*Off>On: $t_{14}=1.7$, $p=0.10$, two-tailed) were significant. No effects were found for learning from losses (Supplementary Fig. 3b).

Striatal responses to components of the prediction error

To test how treatment affects different aspects of the reinforcement learning signal, we broke down the prediction error signal to its separate algebraic components – expectation and reward (following[62–64]) – and examined how each of these signals were affected by treatment. Specifically, we built a GLM comprising two onset regressors: time of choice and time of outcome. Additionally, at time of outcome, two parametric modulators were included that corresponded to the algebraic components of the prediction error: expected value of the choice (derived from the RL model) and reward magnitude (dummy coded: 1 for a gain of $10, 0 for $0, and −1 for a loss of $10). Delta functions of onset regressors and parametric modulators were convolved with the canonical hemodynamic response function and regressed against each subject's fMRI data. Individual parameter estimates for expected value and reward were extracted from the independently defined ventral striatum ROI (MNI coordinates=[−10, 12, −8] from Pessiglione et al. 2006[61]). Two–tailed paired t–tests were used to test for treatment effects on both expected value and reward related activation. While there were no significant treatment effects for expected value related ventral striatum activation (On>Off: $t_{14}=1.14$; Placebo>Off: $t_{14}=0.45$ and On>Placebo: $t_{14}=0.82$, $p>0.05$), reward magnitude elicited significantly stronger responses off drug compared to placebo (Off>Placebo: $t_{14}=2.14$, $p<0.05$) and numerically, but not significantly, stronger responses

compared to on drug (Off>On: t_{14}=1.63, p=0.12), with no differences between placebo and on drug (On>Placebo: t_{14}=0.21, p>0.05) (Supplementary Fig. 4). We concluded from these finding that placebo and drug alter the reward component of the prediction error at time of outcome.

Control analyses

If placebo and drug affected BOLD activity in general, rather than specifically affecting learning–related activity, then effects of treatment would be expected (1) at task–irrelevant events (e.g. fixation), and/or (2) in other brain regions.

1. Main effect of treatment condition at time of fixation—To test the possibility of a global effect of treatment on BOLD activity, we examined brain activation at time of fixation. The GLM included one regressor at onset fixation. Boxcar functions for the whole fixation duration were convolved with the hemodynamic response function in each trial. Individual contrast images were entered into a second–level random effects analysis. A one-way ANOVA revealed no main effect of treatment conditions (F(2,42)=8.2, p>0.001, uncorrected).

2. Treatment effects in control ROIs—To test the possibility of treatment effects on BOLD activity in other brain regions, we defined control regions of interest (ROI) in the primary motor cortex and the primary visual cortex. ROIs were defined anatomically using the AAL package implemented in the WFU-Pickatlas software of SPM8[65]. The motor cortex ROI comprised the left precentral gyrus and the visual cortex ROI comprised the calcarine sulcus. We applied small volume correction to determine the effects of value and prediction error in each treatment condition with a lenient initial threshold of p<0.01 uncorrected. We found no effect of drug or placebo on BOLD activity in these control regions for either value or prediction error related activity.

Task instructions

The game consists of 2 sessions, each comprising 64 trials and lasting for 10 minutes. The first session will be preceded by a training session outside the scanner comprising 32 trials.

The word 'ready' will be displayed 6 seconds before the game starts.

In each trial you have to choose between the two cues displayed on the screen, on the left and right of a central cross.

- to choose the right cue, press the right mouse button
- to choose the left cue, press the left mouse button

As soon as you press the button, a white arrow displayed at the bottom of the selected cue will indicate your choice.

As an outcome of your choice you may

- get nothing ($0)
- win a ten dollar bill (+$10)

- lose a ten dollar bill (–$10)

To have a chance to win, you must make a choice and press one of the two buttons. If you do nothing, the white arrow will remain at the central cross.

The two cues displayed on a same screen are not equivalent in terms of outcome: with one you are more likely to get nothing than with the other. Each cue has got its own meaning, regardless of where (left or right of the central cross) or when it is displayed.

The aim of the game is to win as much money as possible.

Good luck!

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Goetz CG, et al. Placebo influences on dyskinesia in Parkinson's disease. Mov Disord. 2008; 23:700–707. [PubMed: 18175337]

2. Goetz CG, et al. Placebo response in Parkinson's disease: Comparisons among 11 trials covering medical and surgical interventions. Movement Disorders. 2008; 23:690–699. [PubMed: 18228568]

3. McRae C, et al. Effects of perceived treatment on quality of life and medical outcomes in a double-blind placebo surgery trial. Arch Gen Psychiatry. 2004; 61:412–420. [PubMed: 15066900]

4. Benedetti F, et al. Placebo-responsive Parkinson patients show decreased activity in single neurons of subthalamic nucleus. Nature Neuroscience. 2004; 7:587–588. [PubMed: 15146189]

5. de la Fuente-Fernandez R, et al. Expectation and dopamine release: mechanism of the placebo effect in Parkinson's disease. Science. 2001; 293:1164–1166. [PubMed: 11498597]

6. Lidstone SC, et al. Effects of expectation on placebo-induced dopamine release in Parkinson disease. Arch Gen Psychiatry. 2010; 67:857–865. [PubMed: 20679593]

7. Voon V, et al. Mechanisms underlying dopamine-mediated reward bias in compulsive behaviors. Neuron. 2010; 65:135–142. [PubMed: 20152119]

8. de la Fuente-Fernandez R, et al. Dopamine release in human ventral striatum and expectation of reward. Behav Brain Res. 2002; 136:359–363. [PubMed: 12429397]

9. Schweinhardt P, Seminowicz DA, Jaeger E, Duncan GH, Bushnell MC. The anatomy of the mesolimbic reward system: a link between personality and the placebo analgesic response. J Neurosci. 2009; 29:4882–4887. [PubMed: 19369556]

10. Zubieta JK, Stohler CS. Neurobiological mechanisms of placebo responses. Ann N Y Acad Sci. 2009; 1156:198–210. [PubMed: 19338509]

11. Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature. 2006; 442:1042–1045. [PubMed: 16929307]

12. Sutton, RS.; Barto, AG. Reinforcement learning: an introduction. A Bradford Book; 1998.

13. Barto AG. Reinforcement learning and dynamic programming. Analysis, Design and Evaluation of Man-Machine Systems 1995. 1995; 1 and 2:407–412.

14. Daw, ND. Decision making, affect, learning: attention and performance xxiii. Delgado, MR.; Phelps, EA.; Robbins, TW., editors. Vol. 1. Oxford University Press; 2011.

15. Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ. Cortical substrates for exploratory decisions in humans. Nature. 2006; 441:876–879. [PubMed: 16778890]

16. Hare TA, O'Doherty J, Camerer CF, Schultz W, Rangel A. Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. J Neurosci. 2008; 28:5623–5630. [PubMed: 18509023]

17. Levy DJ, Glimcher PW. The root of all value: a neural common currency for choice. Curr Opin Neurobiol. 2012; 22

18. Plassmann H, O'Doherty JP, Rangel A. Appetitive and aversive goal values are encoded in the medial orbitofrontal cortex at the time of decision making. J Neurosci. 2010; 30:10799–10808. [PubMed: 20702709]

19. McClure SM, Berns GS, Montague PR. Temporal prediction errors in a passive learning task activate human striatum. Neuron. 2003; 38:339–346. [PubMed: 12718866]

20. O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ. Temporal difference models and reward-related learning in the human brain. Neuron. 2003; 38:329–337. [PubMed: 12718865]

21. Daw ND, Doya K. The computational neurobiology of learning and reward. Curr Opin Neurobiol. 2006; 16:199–204. [PubMed: 16563737]

22. Rushworth MF, Summerfield C. General mechanisms for decision making. Current Opinion in Neurobiologie. 2009; 19(75):83.

23. Chowdhury R, et al. Dopamine restores reward prediction errors in old age. Nat Neurosci. 2013; 16:648–653. [PubMed: 23525044]

24. Schonberg 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. [PubMed: 18032658]

25. Schonberg T, et al. Selective impairment of prediction error signaling in human dorsolateral but not ventral striatum in Parkinson's disease patients: evidence from a model-based fMRI study. Neuroimage. 2010; 49:772–781. [PubMed: 19682583]

26. Behrens TE, Hunt LT, Woolrich MW, Rushworth MF. Associative learning of social value. Nature. 2008; 456:245–249. [PubMed: 19005555]

27. Li J, Delgado MR, Phelps EA. How instructed knowledge modulates the neural systems of reward learning. Proc Natl Acad Sci U S A. 2011; 108:55–60. [PubMed: 21173266]

28. Niv Y, Edlund JA, Dayan P, O'Doherty JP. Neural prediction errors reveal a risk-sensitive reinforcement-learning process in the human brain. J Neurosci. 2012; 32:551–562. [PubMed: 22238090]

29. Kirsch I. Response expectancy as a determinant of experience and behavior. American Psychologist. 1985

30. Wager TD, et al. Placebo-induced changes in FMRI in the anticipation and experience of pain. Science. 2004; 303:1162–1167. [PubMed: 14976306]

31. Pessiglione M, et al. Subliminal instrumental conditioning demonstrated in the human brain. Neuron. 2008; 59:561–567. [PubMed: 18760693]

32. Schmidt L, Palminteri S, Lafargue G, Pessiglione M. Splitting motivation: unilateral effects of subliminal incentives. Psychol Sci. 2010; 21:977–983. [PubMed: 20511391]

33. Benedetti F. How the doctor's words affect the patient's brain. Eval Health Prof. 2002; 25:369–386. [PubMed: 12449081]

34. Fournier JC, et al. Antidepressant drug effects and depression severity: a patient-level meta-analysis. JAMA. 2010; 303:47–53. [PubMed: 20051569]

35. Kirsch I, Sapirstein G. Listening to Prozac but hearing placebo: A meta-analysis of antidepressant medication. Prevention & Treatment. 1998; 1

36. Bodi N, et al. Reward-learning and the novelty-seeking personality: a between- and within-subjects study of the effects of dopamine agonists on young Parkinson's patients. Brain. 2009; 132:2385–2395. [PubMed: 19416950]

37. Frank MJ, Seeberger LC, O'Reilly RC. By carrot or by stick: cognitive reinforcement learning in parkinsonism. Science. 2004; 306:1940–1943. [PubMed: 15528409]

38. Palminteri S, et al. Pharmacological modulation of subliminal learning in Parkinson's and Tourette's syndromes. Proc Natl Acad Sci U S A. 2009; 106:19179–19184. [PubMed: 19850878]

39. Maia TV. Reinforcement learning, conditioning, and the brain: Successes and challenges. Cogn Affect Behav Neurosci. 2009; 9:343–364. [PubMed: 19897789]

40. Gold JM, et al. Negative symptoms and the failure to represent the expected reward value of actions: behavioral and computational modeling evidence. Arch Gen Psychiatry. 2012; 69:129–138. [PubMed: 22310503]

41. Bamford NS, et al. Heterosynaptic dopamine neurotransmission selects sets of corticostriatal terminals. Neuron. 2004; 42:653–663. [PubMed: 15157425]

42. Wang W, et al. Regulation of prefrontal excitatory neurotransmission by dopamine in the nucleus accumbens core. J Physiol. 2012; 590:3743–3769. [PubMed: 22586226]

43. Bamford NS, et al. Dopamine modulates release from corticostriatal terminals. J Neurosci. 2004; 24

44. Colloca L, et al. Learning potentiates neurophysiological and behavioral placebo analgesic responses. Pain. 2008; 139:306–314. [PubMed: 18538928]

45. Voudouris NJ, Peck CL, Coleman G. The role of conditioning and verbal expectancy in the placebo response. Pain. 1990; 43:121–128. [PubMed: 2277714]

46. Kordower JH, et al. Disease duration and the integrity of the nigrostriatal system in disease. Brain. 2013; 136:2419–2431. [PubMed: 23884810]

47. Cools R, Barker RA, Sahakian BJ, Robbins TW. Enhanced or impaired cognitive function in Parkinson's disease as a function of dopaminergic medication and task demands. Cereb Cortex. 2001; 11:1136–1143. [PubMed: 11709484]

48. Frank MJ, Seeberger LC, O'Reilly RC. By carrot or by stick: cognitive reinforcement learning in parkinsonism. Science. 2004; 306:1940–1943. [PubMed: 15528409]

49. Shohamy D, Myers CE, Geghman KD, Sage J, Gluck MA. L–dopa impairs learning, but spares generalization, in Parkinson's disease. Neuropsychologia. 2006; 44:774–784. [PubMed: 16150469]

50. Shohamy D, Myers CE, Grossman S, Sage J, Gluck MA. The role of dopamine in cognitive sequence learning: evidence from Parkinson's disease. Behav Brain Res. 2005; 156:191–199. [PubMed: 15582105]

51. de la Fuente–Fernandez R, et al. Expectation and dopamine release: mechanism of the placebo effect in Parkinson's disease. Science. 2001; 293:1164–1166. [PubMed: 11498597]

52. Benedetti F, et al. Placebo–responsive Parkinson patients show decreased activity in single neurons of subthalamic nucleus. Nat Neurosci. 2004; 7:587–588. [PubMed: 15146189]

53. Benedetti F, et al. Electrophysiological properties of thalamic, subthalamic and nigral neurons during the anti–parkinsonian placebo response. J Physiol. 2009; 587:3869–3883. [PubMed: 19546163]

54. Starkstein SE, et al. Reliability, validity, and clinical correlates of apathy in Parkinson's disease. J Neuropsychiatry Clin Neurosci. 1992; 4:134–139. [PubMed: 1627973]

55. Spielberger, CD. State–Trait Anxiety Inventory: Bibliography. 2. Consulting Psychologists Press; 1989.

56. Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease. The Unified Parkinson's Disease Rating Scale (UPDRS): Status and Recommendations. Movement Disorders. 2003; 18:738–750. [PubMed: 12815652]

57. Sutton, RS.; Barto, AG. Reinforcement learning: an introduction. A Bradford Book; 1998.

58. Lindquist MA, Spicer J, Asllani I, Wager TD. Estimating and testing variance components in a multi–level GLM. Neuroimage. 2012; 59:490–501. [PubMed: 21835242]

59. Mazaika, P.; Hoeft, F.; Glover, GH.; Reiss, AL. Human Brain Mapping. 2009.

60. Hare TA, O'Doherty J, Camerer CF, Schultz W, Rangel A. Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. J Neurosci. 2008; 28:5623–5630. [PubMed: 18509023]

61. Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD. Dopamine–dependent prediction errors underpin reward–seeking behaviour in humans. Nature. 2006; 442:1042–1045. [PubMed: 16929307]

62. Behrens TE, Hunt LT, Woolrich MW, Rushworth MF. Associative learning of social value. Nature. 2008; 456:245–249. [PubMed: 19005555]

63. Niv Y, Edlund JA, Dayan P, O'Doherty JP. Neural prediction errors reveal a risk–sensitive reinforcement–learning process in the human brain. J Neurosci. 2012; 32:551–562. [PubMed: 22238090]

64. Li J, Delgado MR, Phelps EA. How instructed knowledge modulates the neural systems of reward learning. Proc Natl Acad Sci U S A. 2011; 108:55–60. [PubMed: 21173266]

65. Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH. An automated method for neuroanatomic and cytoarchitectonic atlas–based interrogation of fMRI data sets. Neuroimage. 2003; 19
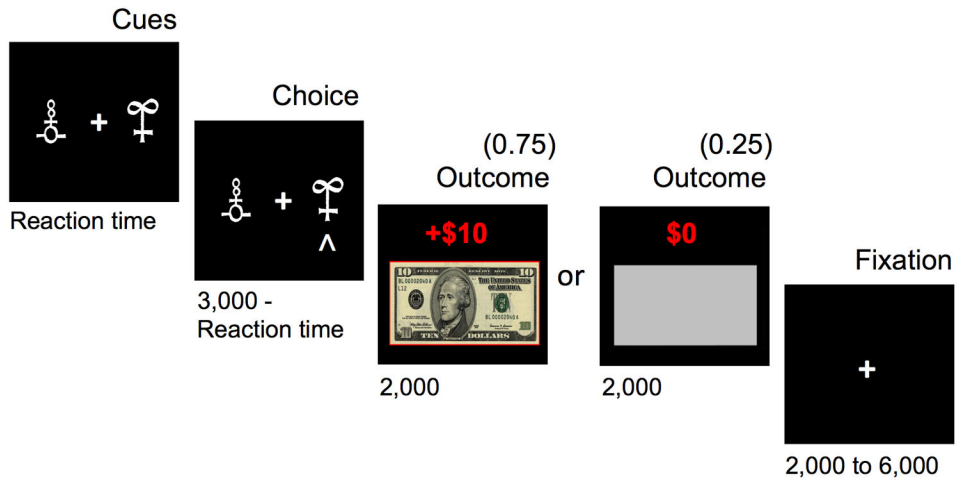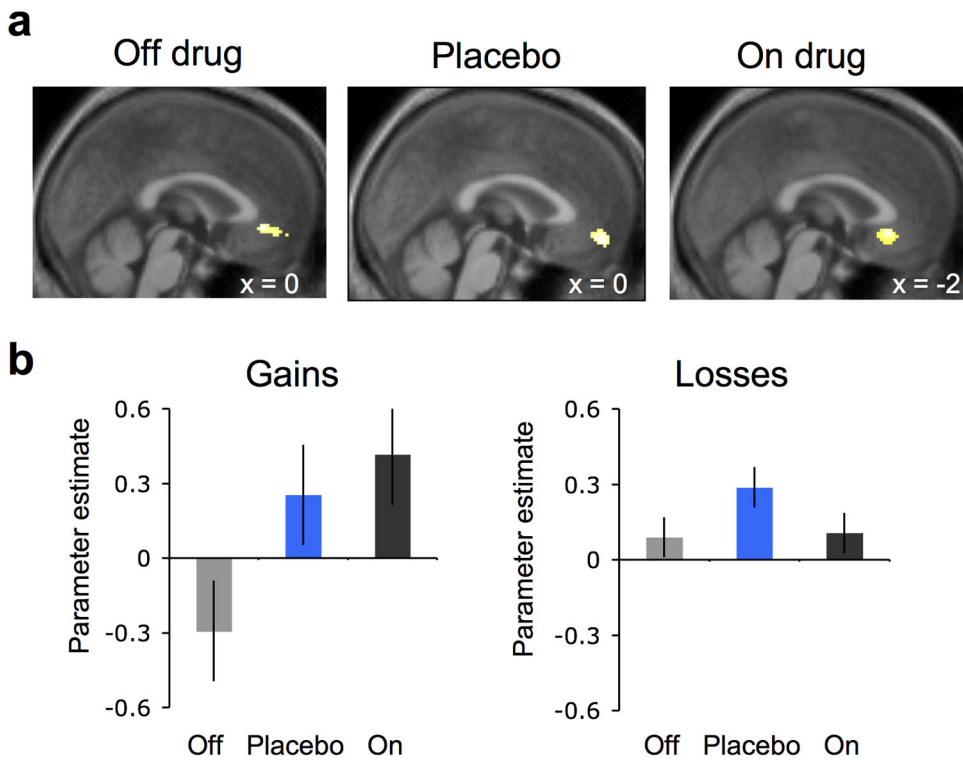
Figure 1.
Instrumental learning task. Successive screenshots are displayed for one trial with durations in milliseconds. On each trial, participants chose between two shapes and subsequently observed an outcome. Participants were instructed to maximize their winnings by learning the optimal shape. In this example (a gain trial), the choice is optimal, because 75% of the time it would lead to a monetary gain ($10) and only 25% of the time would lead to no reward ($0). During an alternative loss trial (not displayed), the optimal choice would lead to a monetary loss (−$10) only 25% of the time and would lead to no reward ($0) 75% of the time. Gain and loss trials, distinguished by unique pairs of symbols, were randomly intermixed within a learning session.
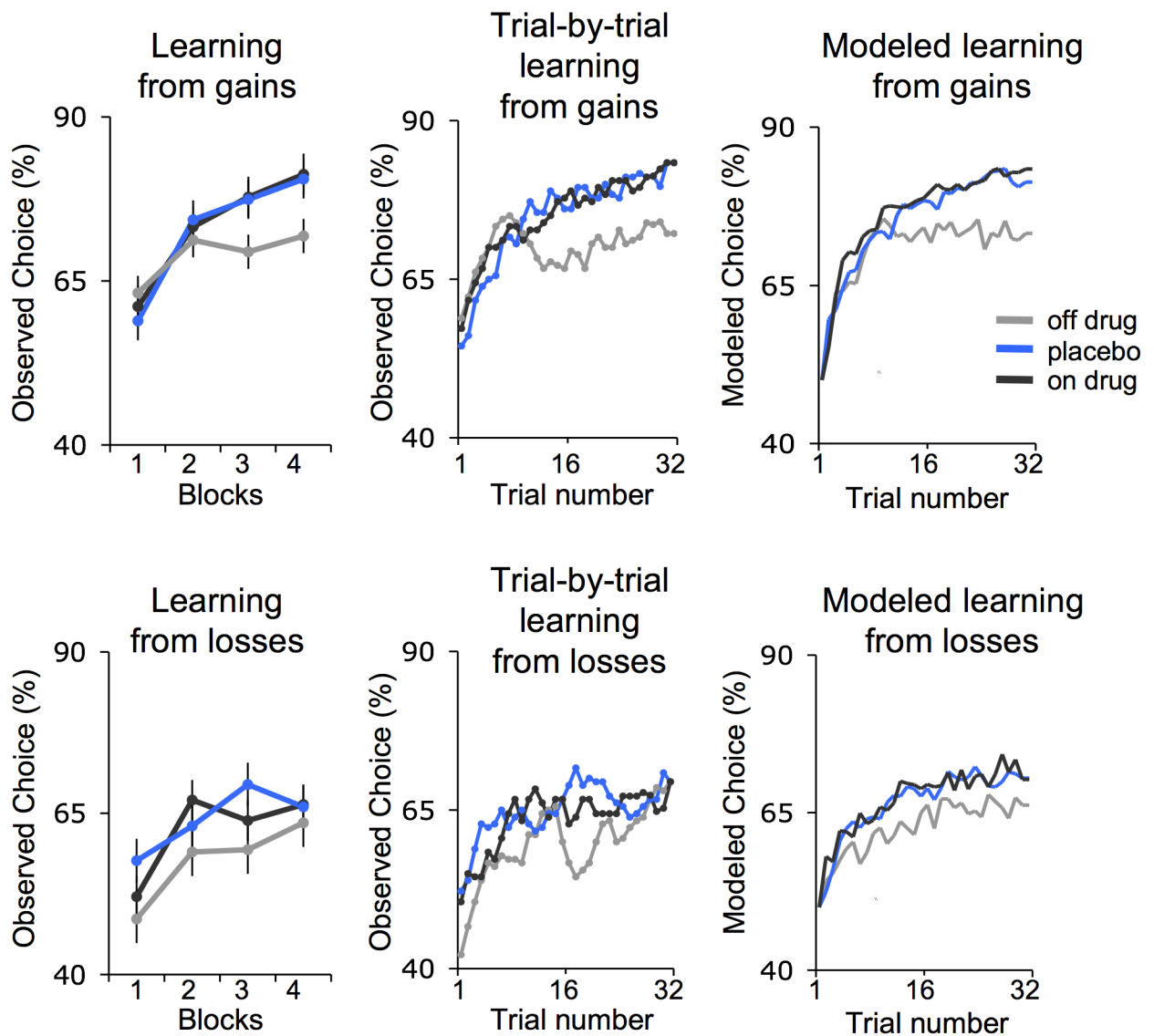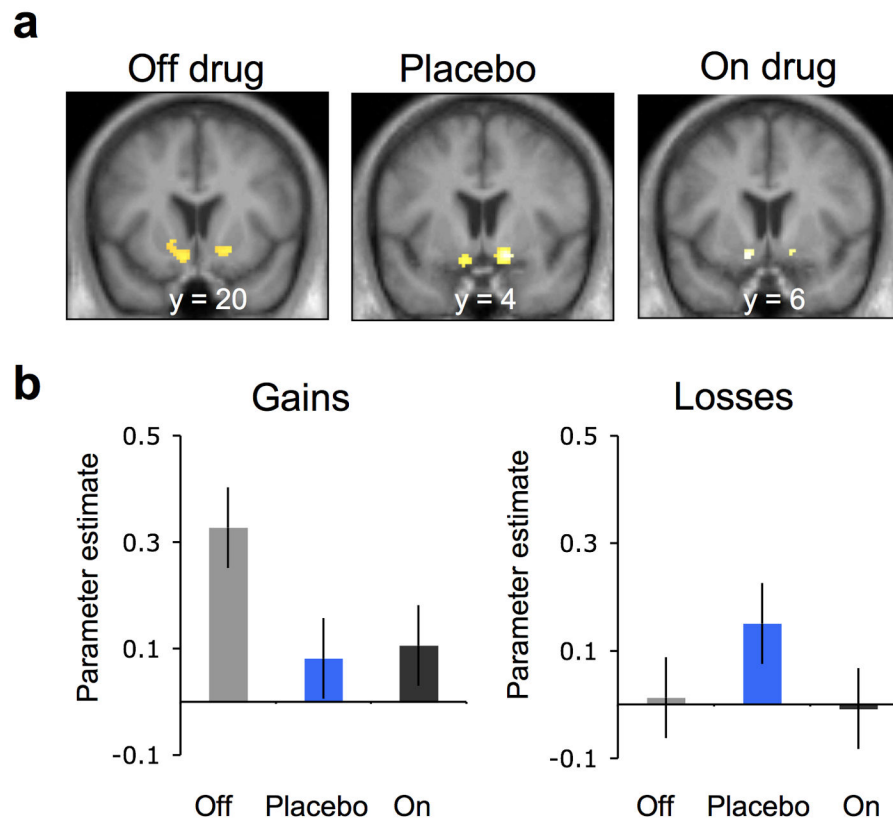
Figure 2.
Behavioral results (n = 18). Percentage of optimal choices averaged across blocks (1st, 2nd, 3rd, and 4th) of 8 trials (left panel), smoothed (middle panel), and fitted by a standard reinforcement learning model (right panel) in the gain (top) and loss condition (bottom). The learning curves depict how often subjects chose the 75% rewarding cue during the gain condition (block: $t_{14} = 5.4$, $p < 0.001$, multilevel linear regression), and the 75% nothing cue during the loss condition (block: $t_{14} = 3.5$, $p < 0.005$, multilevel linear regression). Error bars represent within subject standard errors.

Figure 3.
Ventromedial prefrontal cortex (vmPFC) value activity (n = 15). **(a)** BOLD activity in the vmPFC correlated with expected value ($p^{FWE} < 0.05$, small volume corrected), collapsed across gain and loss conditions during learning. Statistical parametric maps (SPMs) are shown for each treatment condition at $p < 0.005$, uncorrected and masked for the vmPFC. **(b)** Parameter estimates (betas) for expected value from the vmPFC region of interest in the gain (left) and loss condition (right) for off drug (gray), placebo (blue), and on drug (black) treatments. Error bars represent within subject standard errors.

Figure 4.
Striatal prediction error activity (n = 15). **(a)** BOLD activity in the ventral striatum correlated with prediction error during the off drug and the placebo sessions ($p^{FWE} < 0.05$, small volume corrected), collapsed across gain and loss conditions during learning. SPMs are shown for each treatment condition at p < 0.005, uncorrected and masked for the ventral striatum. **(b)** Parameter estimates (betas) for prediction error from the ventral striatum region of interest in the gain (left) and loss condition (right) for off drug (gray), placebo (blue), and on drug (black) conditions. Error bars represent within subject standard errors.