

Similar Improvement of Reward and Punishment Learning by Serotonin Reuptake Inhibitors in Obsessive-Compulsive Disorder

Stefano Palminteri, Anne-Hélène Clair, Luc Mallet, and Mathias Pessiglione

Background: The role of dopamine in reinforcement learning has been extensively studied, but the role of other major neuromodulators, particularly serotonin, remains poorly understood. An influential theory has suggested that dopamine and serotonin represent opponent systems respectively driving reward and punishment learning.

Methods: To test this theory, we compared two groups of patients with obsessive-compulsive disorder, one unmedicated ($n = 12$) and one treated with serotonin reuptake inhibitors (SRI; $n = 13$). To avoid confounding basic reinforcement learning with strategic conscious reasoning, we used a subliminal conditioning task that involves subjects learning to associate masked cues with gambling outcomes to maximize their payoff. The same task was used in a previous study to demonstrate opposite effects of dopaminergic medication on reward and punishment learning.

Results: Unmedicated obsessive-compulsive disorder patients exhibited an instrumental learning deficit that was fully alleviated under SRI treatment. Contrary to dopaminergic medication, SRIs similarly modulated reward and punishment learning.

Conclusions: Thus, departing from the opponency model, our results support a beneficial role of serotonin in instrumental learning that is independent of outcome valence.

Key Words: Instrumental learning, obsessive-compulsive disorder, punishment, reward, serotonin, subliminal perception

The phenomenon of reinforcement learning is pervasive in everyday life and affected in numerous psychiatric diseases and treatments. This phenomenon has been described in great details by behaviorist pioneers (1–3) and the underlying mechanisms later formalized as a set of computational operations in the machine learning literature (4,5). Reinforcement learning theories aim at explaining how outcomes shape vegetative or instrumental responses to the environment. In instrumental conditioning, the agent typically learns to link stimuli, actions, and outcomes. Positive outcomes (rewards) facilitate the selection of preceding actions, whereas negative outcomes (punishments) will impede it, when the same stimuli appear again. A standard computational solution to model reinforcement processes consists in updating the value of chosen actions proportional to prediction errors, which are defined as the difference between obtained and expected outcome values. A large body of evidence in human and nonhuman primates suggests that during instrumental learning, reward prediction errors are signaled by midbrain dopamine neurons (6–8). Many pharmacologic studies in humans have consistently shown that reward learning is improved by dopamine enhancers such as levodopa and impaired by dopamine blockers such as neuroleptics (9–14). Hence, the role of dopamine in driving re-

ward-seeking behavior seems both well established and well formalized.

In contrast, the neural underpinnings of punishment avoidance appear less clear. Some authors have suggested that dips in the activity of dopaminergic neurons could encode negative prediction errors, in particular, when rewards are expected but not obtained (6,9). It has been argued, however, that the low firing rate of dopaminergic neurons would not allow dips to cover a large range of punishment values (7). A physiologic solution may be the intervention of an opponent system (15) that would respond to punishments with increased firing rates. Some authors have provided evidence that separate dopaminergic neurons could respond to punishments, but their importance remains controversial (16–21). Other authors have proposed serotonin as an opponent neuromodulator because it has been implicated in punishment, disengagement, inhibition, and avoidance (22–25). Moreover, a direct antagonism has been established between the dopaminergic and serotonergic systems at the neural level, consistent with the idea of a functional opponency (26,27). In the field of reinforcement learning, the role of serotonin as an opponent to dopamine has been formalized in an influential model (28) assuming that the phasic activity of serotonergic neurons encodes punishment prediction errors, which in principle could drive avoidance learning. However, direct evidence of dopamine-serotonin opponency for reward-punishment learning is still lacking. Indeed, pharmacologic studies testing the effects of serotonin depletion on reward-based learning and decision making, using primarily acute tryptophan depletion, yielded disparate results (29–33). To our knowledge, the effects of serotonin reuptake inhibitors (SRIs) on instrumental learning have been poorly explored in humans. Furthermore, there has thus far been no attempt to contrast dopaminergic and serotonergic drugs in the same instrumental learning task.

To test the implication of serotonin in reward and punishment learning, we exploited a previously validated subliminal conditioning task (34). The task involves subjects choosing to gamble or not, following cues that they do not perceive consciously. The risky response (gambling) can be either a Go (button press) or a No-Go,

From the Motivation, Brain and Behavior team (SP, MP); and Institut du Cerveau et de la Moelle épinière (SP, A-HC, LM, MP), Institut National de la Santé et de la Recherche Médicale unité 975, Centre National de la Recherche Scientifique unité 7225, Université Pierre et Marie Curie (Paris 6), Groupe Hospitalier Pitié-Salpêtrière, Paris, France.

Authors SP and A-HC contributed equally to this work.

Address correspondence to Mathias Pessiglione, Ph.D., Institut du Cerveau et de la Moelle épinière, Groupe Hospitalier Pitié-Salpêtrière, 47 Boulevard de l'Hôpital, 75013 Paris, France; E-mail: mathias.pessiglione@gmail.com.

Received Jun 7, 2011; revised and accepted Dec 23, 2011.

depending on the subject. One learning session employs three cues that are deterministically associated with a positive (+1 euro), negative (−1 euro), or neutral outcome (0 euros). Outcomes are explicitly shown for subjects to learn the predictive value of the subliminal cues and optimize their choices. This instrumental conditioning task has several advantages: 1) because cues are subliminal, learning is not confounded with conscious reasoning abilities; 2) positive and negative monetary payoffs offer equivalent counterparts for rewards and punishments; and 3) reinforcement learning effects are orthogonal to the propensity to gamble (choice impulsivity) and to make Go responses (motor impulsivity). We previously used this task to examine the effects of neuroleptic (NL) medications in patients with Gilles de la Tourette (GTS) syndrome (14). Results showed that NL-treated patients (GTS On) were better at punishment learning, whereas unmedicated patients (GTS Off) were better at reward learning, in keeping with the putative role of dopamine in instrumental learning (9). Now examining the impact of serotonin, we administered the same task to two cohorts of patients with obsessive-compulsive disorders (OCD), one unmedicated (OCD Off) and one treated with an SRI (OCD On). GTS and OCD share common pathophysiological features (basal ganglia dysfunction), are both characterized by production of maladaptive behaviors (tics and compulsions), and can manifest in the same patients (35–37). Thus, we used similar procedures in two closely related neuropsychiatric diseases, GTS being treated with dopamine blockers (NL) and OCD with serotonin enhancers (SRI). Reward and punishment learning have been previously contrasted in OCD patients but without controlling for medication status, which precludes any conclusion about SRI effects (38). Here we examined whether the role of serotonin in instrumental learning is symmetrical to that of dopamine by comparing performance of On versus Off OCD patients. To facilitate the comparison with dopamine, we illustrate in the figures the NL effects on GTS patients that were previously reported (14).

Methods and Materials

Subjects

This study was approved by the Ethics Committee for Biomedical Research of the Pitié-Salpêtrière Hospital, where the study was conducted. Sixty-one subjects (25 patients and 36 controls) were included in the study. All subjects gave written informed consent before participation. They were not paid for their voluntary participation and were told that the money won in the task would be virtual. Previous studies have shown that using real money is not mandatory to obtain robust motivational or conditioning effects in both patients and controls (9,39). In our case, using real money would be unethical because it would mean paying patients depending on their handicap or treatment. In total, 25 patients with OCD were included in the study. We also tested 36 healthy control subjects, who were screened for any history of neurologic or psychiatric conditions. Healthy subjects were selected to match OCD patients in age [statistical comparison: $t(58) = -.7, p > .5$, two-tailed t test]. See Table 1 for a summary of demographic data.

Table 1. Demographic Data

Demographic Features	Controls (<i>n</i> = 36)	OCD Off (<i>n</i> = 12)	OCD On (<i>n</i> = 13)
Age (Years)	35.1 ± 3.2	37.5 ± 4.8	39.6 ± 3.5
Sex (Female/Male)	18/18	7/5	7/6
Education (Years)	15.5 ± .4	13.7 ± .8	13.6 ± .9

OCD, obsessive-compulsive disorder; Off, off serotonin reuptake inhibitor; On, on serotonin reuptake inhibitor.

Table 2. Clinical Data

Clinical Features	OCD Off	OCD On	<i>p</i> Value	<i>t</i> Value
Disease Duration (Years)	14.7 ± 3.7	22.5 ± 2.9	.1	1.7
Time Spent with Current Medication (Months)	—	8.7 ± 2.1	—	—
Time Elapsed Since Last Medication (Months) ^a	20.0 ± 5.9	—	—	—
Y-BOCS	22.5 ± 1.3	19.8 ± 1.0	.1	1.2
Y-BOCS (Obsessions)	11.0 ± .9	9.5 ± .5	.1	1.5
Y-BOCS (Compulsions)	11.5 ± .5	10.3 ± .7	.2	1.4
HADS (Depression)	5.6 ± 1.2 ^b	7.0 ± .9	.4	1.0
HADS (Anxiety)	10.4 ± 1.6 ^b	9.2 ± 1.2	.6	.6

HADS, Hospital Anxiety and Depression Scale; OCD, obsessive-compulsive disorder; Off, off serotonin reuptake inhibitor; On, on serotonin reuptake inhibitor; Y-BOCS, Yale-Brown Obsessive-Compulsive Scale.

^aSix patients had never been treated with an SRI.

^bData missing for two patients.

OCD patients were included to obtain two equivalent groups that differed only in their medication status (see Table 2 for clinical details). All patients fulfilled the DSM-IV criteria of OCD as assessed with the Mini International Neuropsychiatric Interview and were free of Axis I comorbidities (40). In the first group (OCD On, *n* = 13), patients were treated with SRIs in monotherapy, whereas the other group (OCD Off, *n* = 12) received no medication. Among OCD-On patients, 11 were on selective serotonin reuptake inhibitors (SSRI: sertraline, fluoxetine, paroxetine, escitalopram), and 2 were treated with both serotonin and norepinephrine reuptake inhibitors (SNRI: venlafaxine and milnacipran). We kept these two SNRI patients in the analysis because their data did not differ from the SSRI patients. Among OCD-Off patients, six had never received any SRI medication, whereas the other six had been medicated in the past, with an average wash out of 20.0 ± 5.9 months (range 1–24 months) at the moment of the experiment. They had interrupted their treatment because SRIs had poor beneficial effects (*n* = 5) or caused adverse side effects (*n* = 1). The OCD-On patients had been on medication for 8.7 ± 2.1 months (range 1–24 months) when included in the experiment. Only two patients (one in each group) were participating in cognitive and behavioral therapy.

The two OCD groups did not differ in age [$t(23) = .4, p > .5$, two-tailed t test], sex ($\chi^2 = .1; p > .5, \chi^2$ test), disease duration [$t(23) = 1.7, p > .1$, two-tailed t test] and education [$t(23) = .0, p > .9$, two-tailed t test]. Scores on the Yale-Brown Obsessive-Compulsive Scale (Y-BOCS) (41) were similar in the two groups [$t(23) = 1.2, p > .1$, two-tailed t test], as were obsession [$t(23) = 1.5, p > .1$, two-tailed t test] and compulsion subscores [$t(23) = 1.4, p > .1$, two-tailed t test] taken separately. The OCD patients were also assessed with the Hospital Anxiety and Depression Scale (HADS) (42); some presented depressive or anxiety symptoms, but on average, the HADS scores were not different between On and Off patients [anxiety: $t(23) = .6, p > .5$, depression $t(23) = 1.0, p > .1$; two-tailed t test].

Behavioral Tasks

The behavioral tasks are the same as those used in our previous study (14).

The instrumental conditioning task involved choosing between pressing or not pressing a key, in response to masked cues (Figure 1). After showing the fixation cross and the masked cue, the response interval was indicated on the computer screen by a question mark. The interval was fixed to 3 sec, and the response was taken at the end: “Go” if the key was being pressed, “No-Go” if the key was released. The response was written on the screen as soon as the delay had elapsed. Subjects were told that one response was

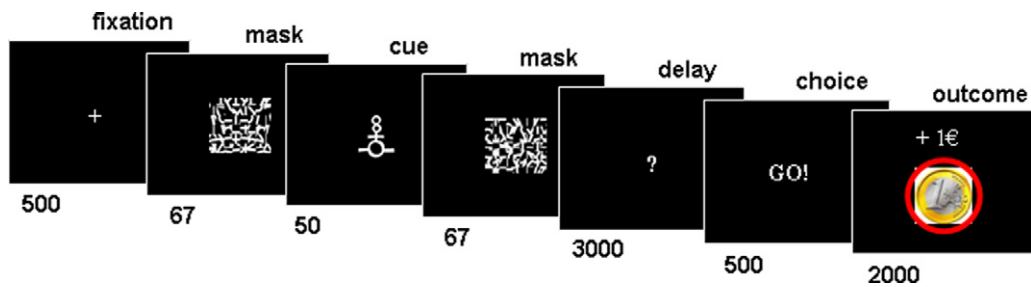


Figure 1. Subliminal learning task. Successive screen shots displayed during a given trial are shown from left to right, with durations in milliseconds. After seeing a masked contextual cue flashed on a computer screen, subjects choose to press or not press a response key and then observe the outcome. In this example, “Go” appears on the screen because the subject has pressed the key following the cue associated with reward (winning 1 euro).

safe (you do not win or lose anything), whereas the other was risky (you can win 1 euro, lose 1 euro, or win or lose nothing). The risky response was assigned to “Go” for half the subjects and to “No-Go” for the other half, such that motor aspects were counterbalanced between reward and punishment conditions in each group. Subjects were also told that the outcome of the risky response would depend on the cue that was displayed between the mask images. In fact, three cues were used: one was rewarding (+1 euro), one was punishing (−1 euro), and one was neutral (0 euro). Because subjects were not informed about the associations, they could only learn them by observing the outcome, which was displayed at the end of the trial. This was either a circled coin image (meaning +1 euro), a barred coin image (meaning −1 euro), or a gray square (meaning 0 euro).

The perceptual discrimination task was used as a control for perceptual learning and awareness at the end of conditioning sessions. In this task, subjects were flashed two masked cues, 3 sec apart, displayed on the center of a computer screen, each following a fixation cross. Subjects had to report whether they perceived any difference between the two visual stimulations. Importantly, subjects had no opportunity to see the cues unmasked, so they could not receive any prior information about what these cues looked like. See Supplement 1 for additional information about behavioral procedures.

Statistical Analysis

From the conditioning task, we extracted the percentage of Go and risky responses, which can be taken as indirect measures of motor and choice impulsivity, respectively. We also extracted the number of correct choices, which is equivalent to the monetary payoff. To display reinforcement learning progression, we plotted the cumulative money won across trials. Individual payoff was then split into euros won for the reward condition and euros not lost for the punishment condition. To correct for motor and choice impulsivity bias, we subtracted the risky choices made in the neutral condition, which captures both the propensity to gamble and the propensity to make a “Go” versus a “No-Go” response, irrespective of reinforcements. We also calculated a “reward bias” by subtracting the correct responses (money not lost) in the punishment condition to the correct responses (money won) in the reward condition.

From the visual discrimination task we calculated a sensitivity index (d') as the difference between normalized rates of hits (correct “different” response) and false alarms (incorrect “different” responses). As for reinforcement learning, we illustrated perceptual learning by plotting the cumulative percentage of correct responses across trials.

All data (demographic, clinical, or experimental) are reported as mean \pm between-subject SEM. To assess instrumental condition-

ing, we used one-tailed paired t tests comparing individual performance with chance level (which corresponds to a zero payoff). Similarly, to assess visual discrimination, we compared individual d' with chance level (which is also zero) using one-tailed paired t tests. We assessed medication effects by comparing dependent variables between On and Off groups with unpaired two-tailed t tests. To assess disease effects relative to control subjects, we also performed between-group comparisons using unpaired two-tailed t tests. To assess significance of linear correlations between reinforcement learning (payoff) and visual discrimination (d') performance, we estimated Pearson’s coefficients. For all statistical tests the threshold for significance was set at $p < .05$. Finally, to explicitly control for possible confounds, we conducted two general linear model analyses aiming at explaining payoff (the main dependent measure) with medication status (On coded 1 and Off coded 0) plus covariates of no interest. To discard experimental confounds, we included as covariates the other dependent measures (d' , percentage of “Go” and “risky” responses). To discard clinical confounds, we included as covariates disease duration, Y-BOCS scores, and HADS scores.

Results

When testing direct comparisons between patients and control subjects, we found no significant disease effect in any of our experimental measure. We do not report these negative results in this section but illustrate the results of healthy control subjects in the figures to provide reference points for the different measures.

All dependent measures observed in the various groups have been summarized in Table 3. We first analyzed the effects of treatments on choice impulsivity (percentage of risky responses) and motor impulsivity (percentage of Go responses) measures (Figure 2, top). The percentage of risky choices was not significantly affected by medication status in OCD patients [$t(23) = -.78, p > .1$]. The same negative results were observed concerning the percentage of Go responses [$t(23) = -.53, p > .5$].

We next examined reinforcement learning effects in the sublim-

Table 3. Experimental Data

Behavioral Measures	Controls ($n = 36$)	OCD Off ($n = 12$)	OCD On ($n = 13$)
Monetary Payoff (Euro)	1.7 \pm .7	−.3 \pm .7	1.9 \pm .8
Visual Discrimination (d')	.02 \pm .09	−.04 \pm .07	.25 \pm .08
Go Responses (%)	46.6 \pm 3.2	50.1 \pm 5.7	46.0 \pm 5.3
Risky Choices (%)	64.6 \pm 2.2	58.6 \pm 5.1	63.4 \pm 5.3
Reward Bias (Euro)	.6 \pm .8	1.2 \pm .9	.7 \pm 1.1

OCD, obsessive-compulsive disorder; Off, off serotonin reuptake inhibitor; On, on serotonin reuptake inhibitor.

inal conditioning task (Figure 2, bottom). On patients showed positive monetary payoff that was significantly different from zero [$t(12) = 2.45, p < .05$; one-tailed t test], whereas Off patients showed no significant conditioning effect [$t(11) = -.42, p > .5$; one-tailed t test]. We also verified that, within the OCD-Off patients, the behavioral performance did not differ between those who had never been treated ($n = 6$) and those who had interrupted the treatment ($n = 6$). We found no significant difference [$t(10) = 1.4, p = .2$, two-tailed t test], which argues against potential long-term effect of SRI withdrawal, although statistical power may be too weak when comparing these small subgroups. Direct comparison between OCD-Off and OCD-On groups was significant [$t(23) = 2.10, p < .05$, two-tailed t test], indicating that SRI improved reinforcement learning deficits in these patients. The comparison remained significant after excluding the two subjects treated with SNRI from the On group [average payoff: 1.97 euros, $t(21) = 2.32, p < .05$]. We also compared patients with a short history of medication (less than 7 months, $n = 7$) to patients who never took SRI ($n = 6$): the difference was still significant [$t(11) = 2.46, p < .01$].

The reward bias (Figure 2 bottom), that is, the differential performance between reward and punishment learning, was not significantly different from zero in both groups [OCD Off: $t(11) = 1.30, p > .1$; OCD On: $t(12) = .72, p > .1$; one-tailed t test]. Moreover, the reward bias was not affected by SRI in OCD patients [$t(23) = .30, p > .5$; two-tailed t test].

The discrimination sensitivity index (d') was found positive in both OCD groups and significantly different from zero in On patients [$t(12) = 3.09, p < .01$; one-tailed t test]. We found a significant effect of SRI treatment on d' in OCD patients [$t(23) = -2.58, p < .05$; two-tailed t test]. This raises the possibility that reinforcement

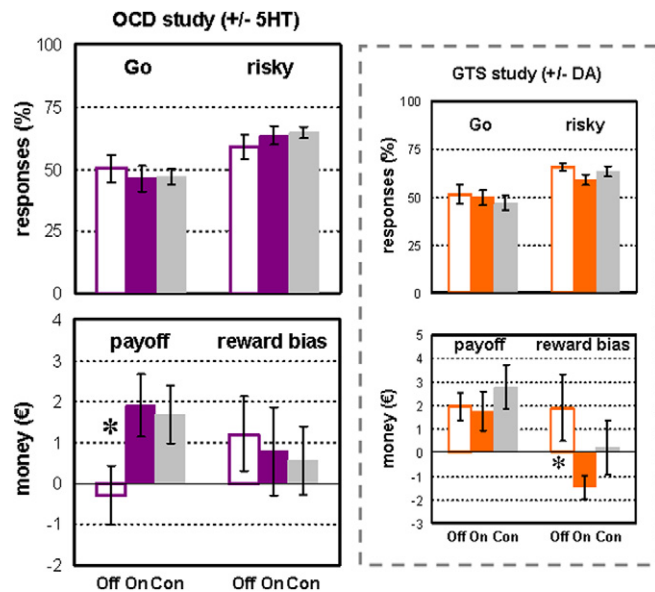


Figure 2. Behavioral performance in the subliminal conditioning task. (Top) Impulsivity measures. Percentage of Go and risky responses reflect motor and choice impulsivity, respectively. (Bottom) Reinforcement learning effects. Monetary payoff is directly proportional to percentage of correct responses. Reward bias corresponds to the difference between reward and punishment learning. Data displayed within the dotted gray square come from a previously published study (14) (Nref) that used the same behavioral task to assess GTS patients On and Off neuroleptics. Colored empty and full bars represent Off and On patients, respectively. Gray bars represent healthy controls (Con). Error bars indicate SEM. * $p < .05$, two-tailed t test. 5HT, serotonin; Con, healthy controls; DA, dopamine; GTS, Gilles de la Tourette syndrome; OCD, obsessive compulsive disorder.

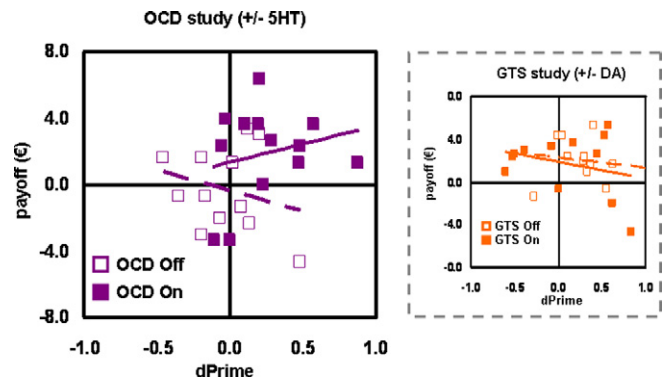


Figure 3. Correlation between reinforcement learning and visual discrimination performance. Data from the present study testing serotonin (5HT) reuptake inhibitors in OCD patients are contrasted to a previous study testing neuroleptics in GTS patients (14). In both studies, monetary payoff was plotted against discrimination sensitivity (d'). Empty and full squares represent Off and On patients, respectively. Dotted and full lines represent linear regressions for Off and On patients, respectively. DA, dopamine; GTS, Gilles de la Tourette syndrome; OCD, obsessive compulsive disorder.

learning improvement induced by SRI was due to enhanced visual discrimination.

To rule out the possibility that differences in reinforcement learning performance were driven by differences in visual discrimination sensitivity, we calculated the correlation between d' and payoffs (Figure 3). Pearson's coefficients were small and not significant in both groups (OCD Off: $R = -.25, p > .1$; OCD On: $R = .22, p > .1$). To verify further that reinforcement learning could not be reduced to perceptual learning, we plotted the cumulative money won (averaged across sessions and conditions) in the subliminal conditioning task and the cumulative percentage of correct responses in the perceptual discrimination task (Figure 4, top). Reinforcement learning effects progressively appeared and peaked at the end, whereas above-chance perceptual discrimination transiently appeared at the beginning and rapidly vanished. This temporal dissociation suggests that above-chance visual discrimination and efficient reinforcement learning were two distinct phenomena. Finally, general linear model analyses showed that medication status remained a significant predictor of the payoff [$t(19) = 2.1, p < .05$], even when we controlled for other clinical variables, such as Y-BOCS and HADS scores and for other experimental variables such as the d' , which were not significant predictors (all t values < 1.0 and p values $> .3$).

Another possible negative interpretation of our results is that OCD-Off patients were simply not engaged in the experiment, because they presented null payoff and d' . This seems unlikely because their rate of risky responses, which denotes subjects trying to win money, was not different from Off patients (see results presented earlier in this section) and significantly above 50% [$t(11) = 1.89, p < .05$; one-tailed t test].

Discussion

Here we show that 1) positive modulation of serotonergic transmission with SRI restored subliminal instrumental learning performance in OCD patients, 2) SRI-induced improvement was not due to affecting the propensity to initiate movement (motor impulsivity) or to engage in gambling (choice impulsivity), and 3) SRI-induced improvement was not specific to outcome valence, contrary to previously observed effects of manipulating dopamine transmis-

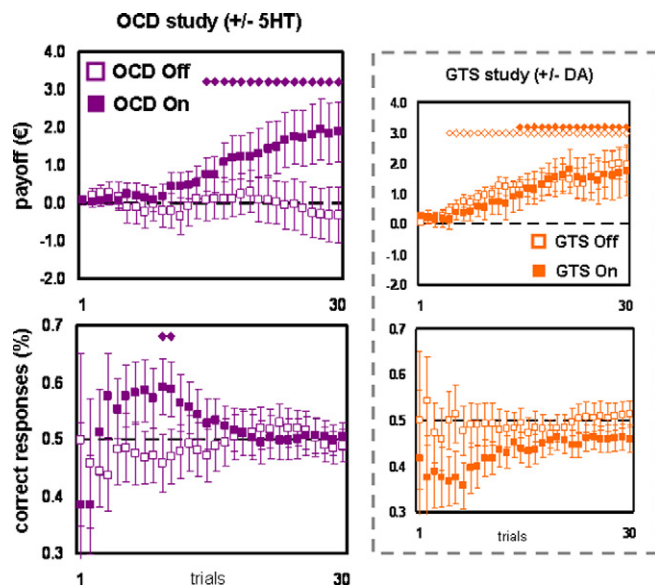


Figure 4. Learning curves. Cumulative learning curves are shown for both the subliminal conditioning (top) and visual discrimination (bottom) tasks. Data displayed within the dotted gray square come from a previously published study (14) that used the same behavioral task to assess GTS patients On and Off neuroleptics. Colored empty and full squares represent Off and On patients, respectively. Empty and full diamonds indicate trials in which performance was significantly superior to chance level for Off and On patients, respectively ($p < .05$, one-tailed paired t test). 5HT, serotonin; DA, dopamine; GTS, Gilles de la Tourette syndrome; OCD, obsessive compulsive disorder.

sion in GTS patients. We examine these three points in the following paragraphs.

Our finding of an impaired instrumental learning performance is consistent with the dysfunction of frontal cortex–basal ganglia loops that has been documented in OCD (43–46). Few studies have examined instrumental learning in these patients, and most have not properly controlled for medication status and outcome valence. One previous study used functional magnetic resonance imaging (fMRI) to examine reversal learning in unmedicated OCD patients (47). Results revealed impaired behavioral performance and reduced outcome-related activation in the striatum and prefrontal cortex, with no specific deficit for punishment. Reduced cortical activations during probabilistic reversal learning were confirmed in unaffected relatives of OCD patients (48). Another study used electroencephalography in unmedicated OCD patients and found altered error-related negativity correlating with impaired performance in probabilistic reversal learning (49). Finally, a recent study showed blunted reward-related signals in OCD patients, although during a behavioral task that did not involve any learning (50).

What we contribute here is the presence of a learning deficit independent of conscious strategic reasoning because learning performance was independent of cue visibility in our data. Indeed, we found no correlation between monetary payoffs and discrimination sensitivity. Moreover, reinforcement effects gradually built up across trials, whereas discrimination performance converged to chance level. It is of interest that even subliminal learning is deficient in OCD, as the above-cited studies used reversal learning tasks, which might involve conscious representations of changes in cue–outcome contingencies. In accordance with our results, another study in OCD patients reported an instrumental learning deficit even in the absence of reversals (51). In a recent fMRI study (52), we also found an instrumental learning deficit in GTS patients

with OCD comorbidity (but not in other GTS patients), which correlated with blunted activity in the ventral prefrontal cortex. Thus, we conclude that instrumental learning deficit may represent a specific endophenotype for OCD. However, the connection between reinforcement learning theory and obsessive-compulsive symptoms remains to be articulated. We only speculate here that repetitive behaviors might come from a failure to incorporate prediction errors when updating action values.

To our knowledge, SRI effects on instrumental learning have not been examined in OCD patients. Our finding that SRIs improved learning performance concurs with the repeated observation that prefrontal serotonin depletion alters reversal learning in marmosets (53,54) but seems to contradict a previous report that SRI increased perseverations following reversals in humans (55). Beyond differences in paradigms (subliminal instrumental vs. probabilistic reversal task) and subjects (OCD patients vs. healthy volunteers), several factors might explain this discrepancy, such as precise medications and doses as well as underlying genotypes, which are known to influence pharmacologic effects (56). We should stress that we were testing here a very basic, likely subconscious, reinforcement process that does not necessitate higher-order, model-based reasoning. However, the contradictory results also raise the possibility that in our case, a beneficial effect on other functions that are sensitive to SRI might have indirectly improved learning performance. In particular, SRIs are known to influence impulsivity, which may relate to the suggested role of serotonin in response inhibition (25,57–59). However, we controlled for two forms of impulsivity, the propensity to make Go responses and the propensity to gamble, which were orthogonal to reinforcement effects in our design. We found no significant effect of SRI on both impulsivity measures, in accordance with the fact that serotonergic manipulations do not affect inhibition performance in tasks in which Go and No-Go responses are not directly associated with rewards and punishments (32,55,60). Thus, the observed SRI effects on reinforcement learning are unlikely to come from a bias applied on impulsive responses.

SRI effects were independent of outcome valence in our data: the reward bias, defined as the difference between reward and punishment learning, was similar in On- and Off-OCD patients. This result accords well with a recent fMRI finding that SRIs affect brain sensitivity to both rewards and punishments (61). Most experiments investigating outcome valence effects used acute tryptophan depletion and yielded inconsistent results, with some studies showing more influence on punishment avoidance, some more influence on reward seeking, and others similar effects with rewards and punishments (30–33,62–64). Here, we challenge the theory that serotonin and dopamine have opponent roles (28) by testing SRI effects on a behavioral paradigm that proved sensitive to dopaminergic manipulation. Indeed, we previously showed that enhancing dopamine transmission with levodopa impairs punishment learning, whereas blocking dopamine transmission with NLS impairs reward learning (14). These results fulfilled the predictions of models assuming that phasic dopamine releases and dips encode positive and negative prediction errors, respectively (6,9). We may now contrast SRI and NL effects in OCD and GTS, which are two neuropsychiatric diseases that share pathophysiologic features (dysfunction of frontal cortex–basal ganglia circuits) and that are frequently associated in the same patients (35–37). Although NLS did reverse the reward bias without affecting overall learning performance, SRIs improved learning performance without affecting the reward bias. This dissociation suggests that serotonin does not play a symmetrical role for punishments to that of dopamine for rewards.

Because of certain limitations, however, we do not claim that our findings alone falsify the dopamine–serotonin opponency theory. For instance, valence effects might be dose-dependent, and we only tested the doses that were prescribed to OCD patients for clinical purposes. Also, SRI effects were assessed in patients who had been medicated for a long time, raising the possibility that other systems compensated or blurred serotonergic effects. Furthermore, the direction of SRI effects cannot be taken as certain: some authors have argued that they could reduce (not enhance) serotonin release via the stimulation of autoreceptors (65,66). In any case, the changes induced by SRI were tonic, and the opponency model assumes that tonic serotonin release does not encode punishments but the average reward rate (28). However, even with this sophistication, the minimal prediction of the opponency model is still an asymmetry in SRI effects between reward and punishment learning, which we did not find in our data. Our conclusions therefore accord with a body of literature suggesting other roles for serotonin, as, for instance, in the temporal discounting of rewards (67,68).

Finally, one could argue that our task was not sensitive enough to observe differences introduced with outcome valence. For instance, a flooring effect might have concealed a learning asymmetry in Off-OCD patients, which could have been compensated for by an asymmetrical improvement with SRIs. This double assumption is, however, less parsimonious than the straight interpretation that SRIs had a similar impact on reward and punishment learning. It remains possible that a valence-specific effect of SRIs might have been found with truly painful punishments or with visible cues. Note, however, that with the subliminal conditioning task, performance remains far away from a ceiling effect, leaving room for improvement. Indeed, we did observe an improvement of learning performance with SRIs, and we previously found a reversion of reward bias with different dopaminergic medications (14), which argue against sensitivity issues. Moreover, double dissociations between outcome valence and medication status, which we demonstrated in GTS patients, have been found by others researchers using various instrumental learning tasks in several pathologic conditions treated with either dopamine blockers or enhancers (9–13). Thus, we conclude that the specificity of serotonin regarding outcome valence is at least not clear-cut compared with that of dopamine.

We thank Véronique Desbaumes, Maël Lebreton, and Helen Bates for their help testing patients and Alina Strasser for checking the English. We also thank Yulia Worbe and Andreas Hartmann for providing clinical data. A-HC received a Ph.D. fellowship from the Ministère de la Recherche. SP received a Ph.D. fellowship of the Neuropôle de Recherche Francilien (NERF).

The authors report no biomedical financial interests or potential conflicts of interest.

Supplementary material cited in this article is available online.

- Skinner BF (1991): *The Behavior of Organisms*. Acton, MA: Copley Publishing Group.
- Thorndike EL (2010): *Animal intelligence; experimental studies*. Charleston, SC: Nabu Press.
- Pavlov IP (2003): *Conditioned Reflexes*. Mineola, NJ: Dover.
- Rescorla RA (1988): Behavioral studies of Pavlovian conditioning. *Annu Rev Neurosci* 11:329–352.
- Sutton RS, Barto AG (1998): *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.
- Schultz W, Dayan P, Montague PR (1997): A neural substrate of prediction and reward. *Science* 275:1593–1599.
- Bayer HM, Glimcher PW (2005): Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47:129–141.
- Zaghloul KA, Blanco JA, Weidemann CT, McGill K, Jaggi JL, Baltuch GH, *et al.* (2009): Human substantia nigra neurons encode unexpected financial rewards. *Science* 323:1496–1499.
- Frank MJ, Seeberger LC, O'Reilly RC (2004): By carrot or by stick: Cognitive reinforcement learning in parkinsonism. *Science* 306:1940–1943.
- 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.
- Cools R, Frank MJ, Gibbs SE, Miyakawa A, Jagust W, D'Esposito M (2009): Striatal dopamine predicts outcome-specific reversal learning and its sensitivity to dopaminergic drug administration. *J Neurosci* 29:1538–1543.
- Bódi N, Kéri S, Nagy H, Moustafa A, Myers CE, Daw N, *et al.* (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.
- 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.
- 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* 106:19179–19184.
- Grossberg S (1988): *Neural Networks and Natural Intelligence*. Cambridge, MA: MIT Press.
- Guarraci FA, Kapp BS (1999): An electrophysiological characterization of ventral tegmental area dopaminergic neurons during differential pavlovian fear conditioning in the awake rabbit. *Behav Brain Res* 99:169–179.
- Coizet V, Dommert E, Redgrave P, Overton P (2006): Nociceptive responses of midbrain dopaminergic neurones are modulated by the superior colliculus in the rat. *Neuroscience* 139:1479–1493.
- Matsumoto M, Hikosaka O (2009): Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* 459:837–841.
- Mirenovic J, Schultz W (1996): Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature* 379:449–451.
- Ungless MA, Magill PJ, Bolam JP (2004): Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. *Science* 303:2040–2042.
- Bromberg-Martin ES, Matsumoto M, Hikosaka O (2010): Distinct tonic and phasic anticipatory activity in lateral habenula and dopamine neurons. *Neuron* 67:144–155.
- Soubrie P (1986): Reconciling the role of central serotonin neurons in human and animal behavior. *Behav Brain Res* 9:319–335.
- Deakin JFW, Graeff FG (1991): 5-HT and mechanisms of defence. *J Psychopharmacol* 5:305–315.
- Abrams JK (2004): Anatomic and functional topography of the dorsal raphe nucleus. *Ann N Y Acad Sci* 1018:46–57.
- Cools R, Roberts AC, Robbins TW (2008): Serotonergic regulation of emotional and behavioural control processes. *Trends Cogn Sci* 12:31–40.
- Lorrain DS, Riolo JV, Matuszewich L, Hull EM (1999): Lateral hypothalamic serotonin inhibits nucleus accumbens dopamine: Implications for sexual satiety. *J Neurosci* 19:7648–7652.
- Jones S, Kauer JA (1999): Amphetamine depresses excitatory synaptic transmission via serotonin receptors in the ventral tegmental area. *J Neurosci* 19:9780–9787.
- Daw ND, Kakade S, Dayan P (2002): Opponent interactions between serotonin and dopamine. *Neural Netw* 15:603–616.
- Evers EAT, Cools R, Clark L, van der Veen FM, Jolles J, Sahakian BJ, *et al.* (2005): Serotonergic modulation of prefrontal cortex during negative feedback in probabilistic reversal learning. *Neuropsychopharmacology* 30:1138–1147.
- Finger EC, Marsh AA, Buzas B, Kamel N, Rhodes R, Vythilingham M, *et al.* (2006): The impact of tryptophan depletion and 5-HTTLPR genotype on passive avoidance and response reversal instrumental learning tasks. *Neuropsychopharmacology* 32:206–215.
- Cools R, Robinson OJ, Sahakian B (2007): Acute tryptophan depletion in healthy volunteers enhances punishment prediction but does not affect reward prediction. *Neuropsychopharmacology* 33:2291–2299.
- Crockett MJ, Clark L, Robbins TW (2009): Reconciling the role of serotonin in behavioral inhibition and aversion: Acute tryptophan depletion

- abolishes punishment-induced inhibition in humans. *J Neurosci* 29:11993–11999.
33. Tanaka SC, Shishida K, Schweighofer N, Okamoto Y, Yamawaki S, Doya K (2009): Serotonin affects association of aversive outcomes to past actions. *J Neurosci* 29:15669–15674.
 34. Pessiglione M, Petrovic P, Daunizeau J, Palminteri S, Dolan RJ, Frith CD (2008): Subliminal instrumental conditioning demonstrated in the human brain. *Neuron* 59:561–567.
 35. Bloch MH, Leckman JF (2009): Clinical course of Tourette syndrome. *J Psychosom Res* 67:497–501.
 36. Grados MA, Mathews CA (2009): Clinical phenomenology and phenotype variability in Tourette syndrome. *J Psychosom Res* 67:491–496.
 37. Worbe Y, Mallet L, Golmard J, Béhar C, Durif F, Jalenques I, *et al.* (2010): Repetitive behaviours in patients with Gilles de la Tourette syndrome: Tics, compulsions, or both? *PLoS ONE* 5:e12959.
 38. Endrass T, Kloft L, Kaufmann C, Kathmann N (2011): Approach and avoidance learning in obsessive-compulsive disorder. *Depress Anxiety* 28:166–172.
 39. Schmidt L, d'Arc BF, Lafargue G, Galanaud D, Czernecki V, Grabli D, *et al.* (2008): Disconnecting force from money: Effects of basal ganglia damage on incentive motivation. *Brain* 131:1303–1310.
 40. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, *et al.* (1998): The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry* 59(20 suppl):22–33.
 41. Goodman WK, Price LH, Rasmussen SA, Mazure C, Fleischmann RL, Hill CL, *et al.* (1989): The Yale-Brown Obsessive Compulsive Scale. I. Development, use, and reliability. *Arch Gen Psychiatry* 46:1006–1011.
 42. Zigmond AS, Snaith RP (1983): The Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand* 67:361–370.
 43. Aouizerate B, Guehl D, Cuny E, Rougier A, Bioulac B, Tignol J, *et al.* (2004): Pathophysiology of obsessive-compulsive disorder: A necessary link between phenomenology, neuropsychology, imagery and physiology. *Prog Neurobiol* 72:195–221.
 44. Chamberlain S, Blackwell A, Fineberg N, Robbins T, Sahakian B (2005): The neuropsychology of obsessive compulsive disorder: The importance of failures in cognitive and behavioural inhibition as candidate endophenotypic markers. *Neurosci Biobehav Rev* 29:399–419.
 45. Huey ED, Zahn R, Krueger F, Moll J, Kapogiannis D, Wassermann EM, *et al.* (2008): A psychological and neuroanatomical model of obsessive-compulsive disorder. *J Neuropsychiatry Clin Neurosci* 20:390–408.
 46. Rotge J, Langbour N, Guehl D, Bioulac B, Jaafari N, Allard M, *et al.* (2009): Gray matter alterations in obsessive-compulsive disorder: An anatomic likelihood estimation meta-analysis. *Neuropsychopharmacology* 35:686–691.
 47. Remijne PL, Nielen MMA, van Balkom AJLM, Cath DC, van Oppen P, Uylings HBM, *et al.* (2006): Reduced orbitofrontal-striatal activity on a reversal learning task in obsessive-compulsive disorder. *Arch Gen Psychiatry* 63:1225–1236.
 48. Chamberlain SR, Menzies L, Hampshire A, Suckling J, Fineberg NA, del Campo N, *et al.* (2008): Orbitofrontal dysfunction in patients with obsessive-compulsive disorder and their unaffected relatives. *Science* 321:421–422.
 49. Cavanagh JF, Gründler TO, Frank MJ, Allen JJ (2010): Altered cingulate sub-region activation accounts for task-related dissociation in ERN amplitude as a function of obsessive-compulsive symptoms. *Neuropsychologia* 48:2098–2109.
 50. Figeo M, Vink M, de Geus F, Vulink N, Veltman DJ, Westenberg H, *et al.* (2011): Dysfunctional reward circuitry in obsessive-compulsive disorder. *Biol Psychiatry* 69:867–874.
 51. Nielen MM, den Boer JA, Smid HGOM (2009): Patients with obsessive-compulsive disorder are impaired in associative learning based on external feedback. *Psychol Med* 39:1519–1526.
 52. Worbe Y, Palminteri S, Hartmann A, Vidailhet M, Lehericy S, Pessiglione M (2011): Reinforcement learning and Gilles de la Tourette syndrome: Dissociation of clinical phenotypes and pharmacological treatments. *Arch Gen Psychiatry* 68:1257–1266.
 53. Clarke HF, Dalley JW, Crofts HS, Robbins TW, Roberts AC (2004): Cognitive inflexibility after prefrontal serotonin depletion. *Science* 304:878–880.
 54. Clarke HF, Walker SC, Crofts HS, Dalley JW, Robbins TW, Roberts AC (2005): Prefrontal serotonin depletion affects reversal learning but not attentional set shifting. *J Neurosci* 25:532–538.
 55. Chamberlain SR, Müller U, Blackwell AD, Clark L, Robbins TW, Sahakian BJ (2006): Neurochemical modulation of response inhibition and probabilistic learning in humans. *Science* 311:861–863.
 56. Ullsperger M (2010): Genetic association studies of performance monitoring and learning from feedback: The role of dopamine and serotonin. *Neurosci Biobehav Rev* 34:649–659.
 57. Eagle DM, Bari A, Robbins TW (2008): The neuropsychopharmacology of action inhibition: Cross-species translation of the stop-signal and go/no-go tasks. *Psychopharmacology* 199(3):439–456.
 58. Pattij T, Vanderschuren LJ (2008): The neuropharmacology of impulsive behaviour. *Trends Pharmacol Sci* 29:192–199.
 59. Rogers RD (2011): The roles of dopamine and serotonin in decision making: Evidence from pharmacological experiments in humans. *Neuropsychopharmacology* 36:114–132.
 60. Cools R, Blackwell A, Clark L, Menzies L, Cox S, Robbins TW (2005): Tryptophan depletion disrupts the motivational guidance of goal-directed behavior as a function of trait impulsivity. *Neuropsychopharmacology* 30:1362–1373.
 61. McCabe C, Mishor Z, Cowen PJ, Harmer CJ (2010): Diminished neural processing of aversive and rewarding stimuli during selective serotonin reuptake inhibitor treatment. *Biol Psychiatry* 67:439–445.
 62. Rogers RD, Tunbridge EM, Bhagwagar Z, Drevets WC, Sahakian BJ, Carter CS (2002): Tryptophan depletion alters the decision-making of healthy volunteers through altered processing of reward cues. *Neuropsychopharmacology* 28:153–162.
 63. Talbot PS, Watson DR, Barrett SL, Cooper SJ (2005): Rapid tryptophan depletion improves decision-making cognition in healthy humans without affecting reversal learning or set shifting. *Neuropsychopharmacology* 31:1519–1525.
 64. Blair KS, Finger E, Marsh AA, Morton J, Mondillo K, Buzas B, *et al.* (2007): The role of 5-HTTLPR in choosing the lesser of two evils, the better of two goods: examining the impact of 5-HTTLPR genotype and tryptophan depletion in object choice. *Psychopharmacology* 196:29–38.
 65. Artigas F, Romero L, de Montigny C, Blier P (1996): Acceleration of the effect of selected antidepressant drugs in major depression by 5-HT_{1A} antagonists. *Trends Neurosci* 19:378–383.
 66. Hjorth S, Bengtsson HJ, Kullberg A, Carlzon D, Peilto H, Auerbach SB (2000): Serotonin autoreceptor function and antidepressant drug action. *J Psychopharmacol* 14:177–185.
 67. Schweighofer N, Bertin M, Shishida K, Okamoto Y, Tanaka SC, Yamawaki S, Doya K (2008): Low-serotonin levels increase delayed reward discounting in humans. *J Neurosci* 28:4528–4532.
 68. Miyazaki K, Miyazaki KW, Doya K (2011): Activation of dorsal raphe serotonin neurons underlies waiting for delayed rewards. *J Neurosci* 31:469–479.