

Computational modelling and neural correlates of reinforcement learning following three-week escitalopram: A double-blind, placebo-controlled semi-randomised study.

Christelle Langley^{1,2*}, Graham K. Murray^{1,3}, Sophia Armand^{4,5}, Franziska Knolle^{1,6}, Rudolf N. Cardinal^{1,3}, Annette Johansen⁴, Peter S. Jensen⁴, Jianfeng Feng^{7,8}, Dea S. Stenbæk^{4,5}, Gitte M. Knudsen^{4,9}, Patrick M. Fisher^{4,10}, Barbara J. Sahakian^{1,2}

Affiliations:

1. Department of Psychiatry, University of Cambridge, Cambridge, UK
2. Behavioural and Clinical Neuroscience Institute, University of Cambridge, Cambridge, UK
3. Cambridgeshire and Peterborough NHS Trust, Cambridge, UK
4. Neurobiology Research Unit, Copenhagen University Hospital Rigshospitalet, Copenhagen, Denmark
5. Department of Psychology, University of Copenhagen, Copenhagen, Denmark
6. Department of Diagnostic and Interventional Neuroradiology, Technical University of Munich, Munich, Germany.
7. Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University, Shanghai, China.
8. Department of Computer Science, University of Warwick, Coventry, United Kingdom.
9. Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark
10. Department of Drug Design and Pharmacology, University of Copenhagen, Denmark

*Corresponding Author:

Christelle Langley, Herchel Smith Building for Brain and Mind Sciences, Forvie Site, Robinson Way, Cambridge, CB20SZ, cl798@medschl.cam.ac.uk

Wordcount: 4683

Keywords: Serotonin, Selective Serotonin Reuptake Inhibitor (SSRI), Cognition, Reinforcement Learning, Psychopharmacology

Abstract

Reinforcement learning is a fundamental aspect of adaptive behaviour, since it involves the acquisition and updating of associations between actions and their outcomes based on the rewarding or punishing consequences. Acute experimental manipulations of serotonin have provided compelling evidence for its role in reinforcement learning. However, it remains unknown how more chronic manipulation of serotonin, which holds greater clinical relevance, affects reinforcement learning and the underlying neural mechanisms. Consequently, we aimed to investigate the effect of a three-week administration of the SSRI, escitalopram, on a reinforcement learning paradigm during functional magnetic resonance imaging. The study used a double-blind, placebo-controlled design with 64 healthy volunteers. Participants were semi-randomised, ensuring matched groups for age, sex and intelligence quotient (IQ), to receive either 20mg of escitalopram (n=32) or placebo (n=32) for at least 21 days. We analysed group differences in reinforcement learning using both analysis of covariance as well as innovative hierarchical Bayesian modelling of the reinforcement learning task. Escitalopram reduced learning from punishment during punishment trials. A key novel finding was that there was decreased activation of the intraparietal sulcus in the escitalopram group when compared to the placebo group during reward trials. The involvement of the intraparietal sulcus suggests that escitalopram affects the encoding of value outcome, which may lead to reduced reinforcement sensitivity, and thereby impacting adaptive learning from feedback. Understanding these mechanisms may help to optimize SSRI treatment to mitigate clinical symptoms and improve quality of life for neuropsychiatric patients, by elucidating serotonin's effects on affect, cognition, and behaviour.

Introduction

Serotonin or 5-hydroxytryptamine (5-HT) is a monoamine neurotransmitter known for its multifaceted role in regulating mood, cognition and behaviour (1, 2). Reinforcement learning is a key component of adaptive behaviour, which is essential in everyday life, therefore, understanding the role of serotonin in this process is of particular importance. Abnormal response to reinforcing feedback has been noted in neuropsychiatric conditions and in particular in major depressive disorder (MDD) (3-5). Specifically, individuals with MDD often show exaggerated responses to negative feedback, whereas they fail to respond to rewards appropriately (3, 5, 6). This suggests that there is an impairment in responding to rewarding and punishing feedback. Similarly, individuals with MDD seem to have a negative bias to emotional stimuli, where they respond faster to negative compared to positive emotional stimuli (7, 8). Studies have shown that serotonin ameliorates these negative biases in depressed patients, and also enhance the positive bias in healthy volunteers (9). Considering that drugs targeting the serotonin system such as selective serotonin reuptake inhibitors (SSRIs) are the first-line pharmacological treatments for MDD (10), it is of importance to determine whether serotonin plays a role in learning from reward and punishment. Understanding how serotonin modulates reinforcement learning processes has significant implications for elucidating the mechanisms underlying various neuropsychiatric disorders characterized by dysregulated reinforcement learning.

Reinforcement learning is a fundamental aspect of adaptive behaviour. It involves the acquisition and updating of associations between actions and their outcomes based on the rewarding or punishing consequences (11). The involvement of serotonin in reward learning is evident through the widespread distribution of serotonergic receptors in brain regions implicated in reward processing, including the mesolimbic dopamine system, prefrontal cortex, and amygdala (12). Moreover, serotonergic projections innervate key nodes of the reward circuitry, exerting intricate regulatory effects on synaptic transmission and neural activity.

Experimental manipulations of serotonin have provided compelling evidence for the involvement of serotonin in reinforcement learning, yet findings vary depending on the specific manipulation employed and the duration of administration. Studies utilizing tryptophan depletion, the precursor to serotonin, have consistently demonstrated impaired performance on reinforcement and reversal learning tasks (13-16). Similarly, investigations into SSRI administration, both acute and chronic, have revealed significant effects on reinforcement learning processes (17-21).

Remarkably, serotonin may exert differential effects on learning from rewards and punishments. For instance, studies have indicated that tryptophan depletion selectively affects punishment prediction, sparing reward prediction (13). On the other hand, after boosting serotonin, Michely et al. (2022) demonstrated that sub-chronic (7 day) SSRI administration enhanced learning from punishment while reducing learning from reward. However, contrasting findings from animal studies suggest an opposing pattern, where increasing serotonin in rats decreased sensitivity to negative feedback but heightened reward sensitivity (22). Furthermore, some investigations propose that there may be no inherent asymmetry in learning following serotonin manipulation in healthy humans (20, 21). For instance, Scholl et al. (2017) found that two weeks of SSRI administration improved learning from both rewards and punishments. Recent research by Langley et al. (2023) has shed further light on the effects of three-week SSRI administration on reinforcement learning

using a reversal learning task. Contrary to previous studies, they reported that reinforcement learning rates remained unaffected by a 3-week SSRI regimen. However, they did observe a reduction in reinforcement sensitivity following three-week SSRI administration. Overall, these findings underscore the intricate role of serotonin in reinforcement learning, highlighting the nuanced effects of serotonin manipulation on reward and punishment processing. The variability in results across studies underscores the complexity of serotonin's involvement in cognitive processes and underscores the need for further investigation to elucidate its precise mechanisms.

One potential reason for the mixed results may stem from the different experimental manipulations used. Acute tryptophan depletion is thought to temporarily reduce brain serotonin levels, whereas SSRI administration is generally assumed to increase extracellular serotonin. However, acute SSRI as a single dose can actually decrease serotonin in some brain regions (1). This complexity arises partly from the potential variations in pre- and post-synaptic actions of these medications. Luo et al. (2024) re-examined a number of studies that manipulated serotonin in both rats and humans and while the results show a clear role of serotonin in learning and behavioural flexibility, it does show the complex interactions of the different doses and administration duration (23). Furthermore, evidence suggests that the neuroplasticity effects of SSRIs might only manifest following more prolonged administration periods, typically spanning 14 to 21 days (24, 25). Indeed, Johansen et al. (2023), using positron emission tomography (PET), have recently shown that neuroplasticity was related to the duration spent on escitalopram (26). Consequently, chronic SSRI administration may yield more reliable outcomes. Notably, chronic SSRI administration serves as an experimental paradigm closely resembling the treatment regimen for MDD, where clinical improvement is usually only seen after several weeks of treatment.

Reinforcement learning involves a complex interplay among various brain regions. A number of brain regions involved in reinforcement learning, include the ventromedial prefrontal cortex (vmPFC), the anterior cingulate cortex (ACC), and subcortical structures such as the striatum and nucleus accumbens (21, 27-32). For example, Eldar et al. (2016) demonstrated that the function and structure of the striatum were predictive of individual harm avoidance behaviours, emphasizing its role in reinforcement learning processes. Similarly, Niv et al. (2012) highlighted the involvement of the nucleus accumbens in risk sensitivity, which plays a crucial role in learning.

Furthermore, additional brain regions such as the orbital frontal cortex and parietal cortex are also implicated in the processing of reward value (33, 34). Importantly, Guo et al. (2013) revealed that uncertain rewards elicited more widespread activation in the brain, indicating the involvement of multiple regions when the reward was probabilistic or uncertain. Serotonin has also been shown to modulate neural mechanisms during reinforcement learning tasks. For example, Scholl et al. (2017) found increased reward and effort learning signals in the vmPFC and ACC, respectively, following 2 weeks of SSRI administration. Notably, the rewards and punishments used in their task differed, with rewards being monetary and punishments involving effort, once again highlighting the complexity of serotonin's influence on reinforcement learning. These findings emphasise the distributed nature of reinforcement learning processes, with multiple brain regions contributing to different aspects of reward and punishment processing, decision-making, and learning. Understanding the interactions among these regions and neurotransmitter systems is crucial

for elucidating the neural mechanisms underlying reinforcement learning and related behaviours.

In the present study, we used a double-blind placebo-controlled design to examine the effects of the SSRI escitalopram administered on average for 26 days, on reinforcement learning. In addition, we examine the neural substrates, using task-based functional magnetic resonance imaging (fMRI), that are associated with the effect of escitalopram on reinforcement learning. Escitalopram was chosen for the present study as it shows very high selectivity for the serotonin transporter and is thus the best choice for testing pharmacologic actions of SSRIs (35-37). Moreover, escitalopram is one of the best-tolerated SSRIs (35-37). Given the previous literature on serotonin potentially differentially affecting reward and punishment, our reinforcement learning paradigm included both reward and punishment trials. We hypothesised that SSRI treatment would affect reinforcement-related behaviour. In addition, the SSRI treatment would affect the brain regions previously identified as being involved in reinforcement learning. This is the first study to examine the effects of three-week SSRI using a reinforcement learning paradigm.

Methods

Participants

This pre-registered study used a double-blind placebo-controlled design with 64 healthy volunteers (**Table 1**) of whom 32 received 20 mg daily of escitalopram and 32 received placebo for at least 21 days (escitalopram, mean(s.d.)= 26.06(2.78) days; placebo, mean(s.d.)= 26.06(3.34) days; $t(60.03)=.00$; $p=1.00$; $d=.00$). Participants were semi-randomised (by a staff member not involved with the participants or the data analysis) into the two groups, which were matched for age, sex and intelligence quotient (IQ) (Reynolds Intellectual Screening Test, RIST). Participants aged between 18 and 45 were recruited from an established database of healthy volunteers at the Neurobiology Research Unit at the Copenhagen University Hospital Rigshospitalet. The study was pre-registered on clinicaltrials.gov (NCT04239339). Participants underwent a medical screening prior to enrolment in the study to ensure they were eligible for inclusion. The study was conducted between May 2020 and October 2021. Exclusion criteria are detailed in the **Supplementary Material**.

Ethics

The study was approved by the ethics committee for the Capital Region of Copenhagen, Denmark (H-18038352) and written informed consent was obtained from all participants. All methods were performed in accordance with the relevant guidelines and regulations.

Experimental Procedure

After obtaining written informed consent, participants underwent screening for somatic illness, which included a medical examination, blood screening for somatic disease, an electrocardiogram (ECG), and assessment for the presence of psychiatric conditions using the Danish translation version 6.0.0 of the Mini-International Neuropsychiatric Interview (Sheehan et al., 1998). Eligible participants were then semi-randomly assigned to receive either an effective clinical dose of escitalopram (20 mg daily in capsules of 10 mg) or placebo in identical capsules provided by the Capital Region Pharmacy, for a duration of

three to five weeks. Both the participants and the investigators involved in data acquisition and analysis were blinded to the intervention type until completion of data analysis.

Blinded medical personnel provided participants with both oral and written instructions on taking escitalopram, including potential side effects. Participants were directed to take 10 mg daily for the initial three days, followed by an increase to 20 mg daily from the fourth day until the last day of examination, coinciding with the neuropsychological testing visit and fMRI scanning session. Before the cognitive visit, participants completed several self-report questionnaires to assess their psychological state. The results of the cognitive and neuropsychiatric analyses are detailed in a separate article (20).

To ensure treatment compliance, the capsule container was examined during visits, and participants provided blood samples both at the halfway point and during the cognitive visit, usually in the morning (see **Figure S1**). Additionally, participants maintained a daily medication logbook, which was reviewed during the follow-up assessment. Participants were directed to ingest the drug capsule after providing the blood sample to ensure the measurement of steady-state serum escitalopram levels. A medical professional oversaw participant management, maintaining regular contact throughout the study period.

At the end of the study, participants were asked to state whether they believed they had received escitalopram or placebo. Among those in the escitalopram group, 53% accurately guessed that they received escitalopram, while 16.12% of participants in the placebo group thought they received escitalopram. A comparison between the two groups revealed a significant disparity in the ability to correctly discern group membership ($\chi^2(1, N = 63) = 9.48, p = .01$ [two-tailed]). Thus, the accuracy of guessing the correct allocation in the escitalopram group was at chance level.

Reinforcement Learning Paradigm

During the fMRI scan, subjects carried out a probabilistic learning paradigm that required making choices to maximize wins and minimize losses (**Figure 1**), adapted from previous similar tasks (28-30, 38). In each trial, one of three possible probabilistic pairs of abstract pictures was randomly presented: a rewarding, punishing, or neutral pair (32 trials of each valence). For each trial, the subject used a button press to indicate a choice of picture (the picture on the right or left). When viewing the probabilistic rewarding pair, selection of one of the pictures led to a financial win with a 70% probability and of a no-change outcome with 30% probability, whereas the selection of the other picture led to gain with only 30% probability. The probabilistic punishing pair led to a financial loss on 70% and 30% trials depending on stimulus choice, and the neutral pair led to no change. The participants completed 3 blocks with new stimuli in each block. All blocks were analysed together to increase the number of trials and thereby, the power in the analyses, particularly the imaging analysis.

Behavioural analysis

All statistical analyses were conducted in R, version 4.1.1 (R Foundation for Statistical Computing).

The group comparisons for the accuracy and reaction time were conducted using analysis of covariance with both within group (feedback: reward vs punishment) and between group (administration: placebo vs escitalopram) factors controlling for age, sex and IQ. The analysis was conducted using the `aov` function in R. Multiple comparisons correction was

conducted using the Benjamini-Hochberg false discovery rate (FDR) with $q=.05$. The p -values reported are uncorrected.

To analyse the reinforcement learning paradigm, we used sophisticated computational modelling approach, by fitting families of hierarchical Bayesian reinforcement learning models to trial-by-trial task data (20, 39, 40). Model comparison was conducted between four models using a bridge sampling estimate of the marginal likelihood using the *bridgesampling* (41) function in R. Model 1 included a reward learning rate, punishment learning rate and reinforcement sensitivity; Model 2 included a reward learning rate, punishment learning rate, reinforcement sensitivity and stimulus stickiness, Model 3 included a combined learning rate, reinforcement sensitivity and stimulus stickiness; and Model 4 used an experience weighted approach (42) which includes learning rate, inverse temperature and experience weight. In Model 4, learning from reinforcement is modulated by an “experience weight” for a stimulus; the experience weight for a stimulus is updated every time it is chosen, and its change over time is governed by a decay factor. In this model, the softmax inverse temperature was also a parameter able to vary.

We analysed the differences in parameter values between groups by first calculating group mean differences (MDs) per parameter. The 90% and 95% highest density intervals (HDIs) of the posterior distribution per MD were then calculated and inspected to check whether they included zero (evidence for no difference between groups). Full details of model formulation, model fitting, and parameter recovery are provided in the **Supplementary Material**.

Code Availability

The code supporting the modelling analyses are available at <https://github.com/RudolfCardinal/rlib>.

Data Availability

Applications to use data from the CIMBI database can be made to the Neurobiology Research Unit, Rigshospitalet, Copenhagen University Hospital.

Image Acquisition

At the end of the intervention period, participants completed an MRI scan on a 3T Siemens Magnetom Prisma scanner (Erlangen, Germany) using a 32-channel head coil. We acquired a high-resolution, whole-brain, T1-weighted MPRAGE structural scan (inversion time = 972 ms, repetition time = 2000 ms, echo time = 2.58 ms, flip angle = 8°, in-plane matrix = 256x256 mm², in-plane resolution = 0.9x0.9 mm², 224 slices, slice thickness = 0.9 mm). We also acquired blood oxygen level-dependent (BOLD) fMRI scans during the reinforcement learning paradigm using a T2*-weighted gradient echo-planar imaging (EPI) sequence (TR = 2000 ms, TE = 30 ms, flip angle = 70°, in-plane matrix = 76x76 mm², in-plane resolution=3x3 mm², 35 slices (thickness = 3.0 mm, gap between slices= 0.6 mm). We acquired a gradient field map to minimise spatial distortions in the EPI BOLD fMRI acquisition.

Image Pre-processing

All pre-processing was conducted in SPM 12 (<https://www.fil.ion.ucl.ac.uk/spm/software/spm12/>). Functional images were slice-timing corrected, realigned and unwrapped to correct for head movements and EPI distortions; co-registered and segmented to normalise images into standard space based on the MNI template, for group level analysis; and smoothed with an 8mm full-width half-maximum

(FWHM) Gaussian kernel, to account for residual inter-subject differences. The default SPM12 steps were used. We used the FSL motion outliers' function to determine the framewise displacement of each image. We determined that participants with mean framewise displacement (FD) > .20 mm would be excluded (43). Movement was small in the cohort and no participants were excluded.

Neuroimaging Analysis

Statistical imaging analysis was also performed in SPM 12. The data from each participant was analysed using general linear models (GLM) and analyses were performed according to an event-related design. The explanatory variables (EVs) that we used were the onset times of the cue stimuli presentation and the feedback stimuli presentation for both reward and punishment trials. They were modelled as 0s duration events. The 6 realignment parameters were included in the design matrix to correct for signal changes due to head movement. An additional set of harmonic regressors was used to account for any temporal low-pass frequency variance within the data that is typical to fMRI signal with a cut-off of 1/128 Hz. All regressors were convolved with the canonical haemodynamic response function.

In the first-level analysis, the GLMs were used to generate contrast images for our four effects of interest, which included cue presentation and feedback presentation for both reward and punishment conditions. For each contrast, we tested whether the parameter estimates (activation levels) at each voxel were significantly greater than zero.

For the group level analysis to compare between the placebo and escitalopram groups a two-sample t-test was conducted for the reward and punishment trial separately. The analysis was conducted as a whole brain analysis. Voxel-wise results were thresholded at $p < .05$, corrected for family-wise error (FWE) at the peak level, to control for multiple comparisons across the entire brain volume as per random field theory in SPM.

Results

Demographics

The analysis confirmed that the two groups were well matched and there were no significant differences in age, sex or IQ (see Table 1).

Behavioural Results

There was no main effect of group between the escitalopram and placebo groups for accuracy ($F = .29$, $p = .59$, $\eta^2 < .01$) or reaction times ($F = 1.71$, $p = .19$, $\eta^2 < .01$). There was a significant main effect of feedback (reward and punishment) for accuracy ($F = 5.83$, $p = .016$, $\eta^2 = .04$) and reaction time ($F = 70.67$, $p < .001$, $\eta^2 = .36$). The results showed that accuracy was better in the reward trials than the punishment trials, and that reaction time was significantly faster in reward trials compared to punishment trials. There was no significant interaction effect between group*task for accuracy ($F = .93$, $p = .34$, $\eta^2 < .01$) or reaction time ($F = .52$, $p = .47$, $\eta^2 < .01$). Performance is displayed in **Figure 2**.

Computational Modelling Results

Hierarchical Bayesian modelling showed that Model 2 was the best model. This revealed that for reward trials, there were no group differences between the escitalopram and the placebo group at the credible difference level of 95% for any of the four model parameters: reward learning (mean difference (MD) = -.08 [95% HDI -.27 to .10], punishment learning (MD = .02 [95% HDI -.06 to .10]), stimulus stickiness (MD = .16 [95% HDI -.11 to .43]) or

reinforcement sensitivity (MD= .36 [95% HDI -1.00 to 1.71]). The results are represented in **Figure 3** and the model comparison is presented in **Table S1**.

For the punishment trials, the escitalopram group had lower punishment learning rate than the placebo group at the credible difference level of 90% (MD= -.15 [90% HDI -.31 to -.01]). There were no differences between the groups for any of the other three model parameters: reward learning (MD= -.08 [95% HDI -.19 to .04]), stimulus stickiness (MD= -.08 [95% HDI -.24 to .06]) or reinforcement sensitivity (MD= .02 [95% HDI -.83 to .92]). The results are represented in **Figure 3** and the model comparison is presented in **Table S2**.

Neuroimaging Results

The neuroimaging results showed no group differences for cue presentation for either the reward or punishment trials. For feedback, the escitalopram group had reduced activation in the intraparietal sulcus compared to the placebo group during the reward trials. There were no differences for feedback on punishment trials. The results are shown in **Table 2** and **Figure 4**.

Discussion

In this double-blind, placebo-controlled study, a relatively large group of healthy volunteers received either escitalopram or placebo for an average of 26 days. This is the first study to examine the effects of three-week SSRI using a reinforcement learning paradigm. The key findings were that the escitalopram group had a lower punishment learning rate when compared to the placebo group during punishment trials. In addition, the escitalopram group showed reduced activation in the intraparietal sulcus compared to the placebo group during reward trials.

Behavioural Results

Using the standard analysis of covariance accuracy and reaction time, results showed no main effect of group, suggesting no differences between the escitalopram and placebo groups for either reward or punishment trials. There was a significant main effect of feedback demonstrating that accuracy was decreased, and reaction time was increased in punishment trials compared to reward trials. There was no interaction between feedback and group, suggesting that the placebo and escitalopram group both performed worse in the punishment compared to reward trials. These results together suggest that participants find it more difficult to learn from punishment than from reward, but that escitalopram does not have an effect on accuracy or reaction time.

For the more sophisticated computational modelling, our results showed that the escitalopram group had a lower punishment learning rate during punishment trials compared to placebo controls. There were no group differences for learning during reward trials. In addition, there were no group differences for reinforcement sensitivity or stimulus stickiness between groups for reward or punishment trials. Our results suggest that escitalopram has a greater effect on learning from punishment than learning from reward. Previous studies have shown a similar asymmetry where serotonin affects learning from punishment to a greater extent (13, 19, 22, 44). The results from the present study suggest that three-week escitalopram impairs punishment learning, but reward learning remains stable. It is worth noting that these studies differ not only in the exact reinforcement learning task used, but also the serotonin manipulation used. The present study and that of Michely et al. (2022) for example differ in the dosage and length of SSRI administration. In our study, we used at

least 3 weeks of SSRI treatment, whereas Michely et al. (2022) used only one week. Given the findings that the neuroplasticity effects can take 14-35 days (24-26), this may be one possible reason for the difference in findings. For this reason, in studies of chronic SSRI effects, the duration should be at least 21 days. In addition, given the importance of learning from reinforcing feedback in everyday life and the large number of individuals on SSRI treatment it is crucial to understand the effects of chronic SSRI to optimise treatment and quality of life for patients with neuropsychiatric conditions. Moreover, understanding the neural mechanisms through which serotonin exerts its effects on affect, cognition and behaviour is of importance to reduce the severity of clinical symptoms.

Neuroimaging Results

Our fMRI results showed a key novel finding namely that there was decreased activation in the intraparietal sulcus in the escitalopram group compared to the placebo group during performance on reward trials. The same region showed decreased activation during punishment trials, but this did not survive the peak-level FWE correction. There were no increased regions of activation in the escitalopram group when compared to the placebo group. Several studies have implicated the intraparietal sulcus in reward processing. Specifically it has been shown to play a role in both the decision-making during uncertainty (34, 45-47), as well as being involved in a 'value-driven attention network' where attentional resources are allocated to valuable choices (48-51).

For example, it has been demonstrated that neurons in the parietal cortex of non-human primates encode the probability of rewards, particularly through eye-movements, suggesting its involvement in reward-based decision-making processes (45). Neuroimaging studies in humans have also implicated the intraparietal sulcus in reward processing. One fMRI study showed increased activity in the parietal cortex during probabilistic reward anticipation tasks, thereby suggesting a role in both the anticipation of rewarding outcomes, but also during uncertainty (46), as well as when the reward is uncertain (34). A further fMRI study has shown that the IPS is associated with state prediction error which is a measure the degree of surprise in encountering a new state, based on the subjects' current understanding of the probabilities that link one state to another following specific actions (47). Therefore, our finding of decreased IPS activation in the escitalopram group potentially reflect an aberrant response to the probabilistic and uncertain elements of the task. This pattern could imply that escitalopram affects neural mechanisms involved in processing state prediction errors or adapting to probabilistic task demands, thereby impacting decision-making under uncertainty.

In terms of the involvement of the intraparietal sulcus in a 'value-driven attention network', it has been shown in non-human primates that the intraparietal sulcus responds to rewards during visual attention tasks, indicating its role in integrating reward-related information with attentional processes (48). Similarly, it has been shown to respond preferentially to stimuli that predict available reward (51). In humans, the intraparietal sulcus has been shown to bias attentional resources to stimuli associated with reward (49). Our finding of reduced learning from punishment in the escitalopram group may be due to the fact that escitalopram disrupts the encoding of the value of reinforcing feedback. These results make sense when considering that SSRIs are used to treat the negative bias and responses to abnormal feedback in MDD. It may be that SSRIs blunt reinforcement encoding, and in fact blunted affect has often been reported by patients (52, 53) and previously demonstrated in research (20). This may lead to reduced sensitivity to reinforcers, which thereby impacts adaptive

learning from feedback. This is in line with previous findings of disrupted reinforcement sensitivity and learning following serotonin manipulations (17, 18, 20, 23). However, this may be an advantage in MDD as it reduces the abnormal response to negative feedback.

Taken together, the involvement of the intraparietal sulcus in reward processing and the impact of serotonin on reinforcement learning provide a plausible framework for understanding the observed differences in intraparietal sulcus activation between the placebo and escitalopram groups. The interplay between serotonin, reward processing, and the intraparietal sulcus highlights the complex neurobiological mechanisms underlying the modulation of reward-related behaviours.

We did not see differences between the groups in more traditional brain regions associated with reward, for example the nucleus accumbens and the anterior cingulate cortex. However, many of these regions are involved in neuroimaging studies examining the role of dopamine not the role of serotonin. Studies have also demonstrated that when outcomes are uncertain there is a much more widespread pattern of activation in the brain (34). The outcomes in the reinforcement paradigm used in the present study were indeed uncertain. One previous study of the effects of citalopram on reinforcement learning showed enhanced reward and effort learning signals in a widespread network of brain regions, including ventromedial prefrontal and anterior cingulate cortex (21). In addition, we did not see group differences in neuroimaging prediction error signals (see Supplementary Material). However, as we have noted in the mixed behavioural findings, the specific manipulations and durations of serotonin seem to affect the results. Further research which examines the chronic effects of serotonin is needed in both healthy volunteers and patients with MDD. This is particularly important considering that some national guidelines, for example the National Institute for Health and Care Excellence (NICE) in the UK and the American Psychological Association (APA) in the USA, suggest that if there is no treatment response to increase the dosage or even change the drug, after 4-6 weeks (54, 55). Given that SSRIs are administered chronically to patients with neuropsychiatric disorders, the present findings hold more clinical relevance than acute studies.

Conclusion

In this double-blind placebo-controlled design of escitalopram administered on average for 26 days to healthy individuals, we showed that there was reduced learning from punishment in the escitalopram group. Using fMRI, our key and novel finding was that there was reduced activation in the intraparietal sulcus in the escitalopram group during reward trials compared to the placebo group. Given the role of the intraparietal sulcus in reinforcement learning, specifically in uncertainty and outcome value it may be that reinforcement learning is altered by serotonin through its effects on encoding of value outcomes. This may lead to reduced sensitivity to reinforcers, which thereby impacts adaptive learning from feedback. In addition, it seems that escitalopram specifically disrupts the processing of probabilistic feedback. These novel findings provide strong evidence for a key role of serotonin and the intraparietal sulcus in reinforcement learning. Given the importance of learning from reinforcing feedback in everyday life and the large number of patients on SSRI treatment it is crucial to understand the effects of chronic SSRIs to optimise treatment and quality of life for patients with neuropsychiatric conditions. Moreover, understanding the neural mechanisms through which serotonin exerts its effects on affect, cognition and behaviour is of importance for clinical treatment of MDD.

Funding and Disclosures

This study was funded by a Lundbeck Foundation Grant to Professor Barbara J Sahakian of the University of Cambridge in collaboration with Professor Gitte Moos Knudsen of the Copenhagen University Hospital Rigshospitalet.

Acknowledgements

We would like to thank all the participants of the study. We also thank the students Tina Segerberg, Anna Søndergaard, Elisabeth B. Pedersen, Nanna Svart, Oliver Overgaard-Hansen, and Ida Likaj for their assistance in the study. We would like to thank the Kirsten & Freddy Johansen (KFJ) Foundation for funding the MR scanner. All research at the Department of Psychiatry in the University of Cambridge is supported by the NIHR Cambridge Biomedical Research Centre (NIHR203312) and the NIHR Applied Research Collaboration East of England.

Competing Interests

CL, GKM, SA, FK, AJ, PSJ, JF, DSS, PF have no conflicts of interest.

GMK served as a speaker for Angelini, Abbvie, Cybin, and H. Lundbeck, as an advisor for Sanos, Onsero, Pangea Botanica, Gilgamesh, Pure Technologies, and her lab is a research site for Reunion and Delix Therapeutics.

RNC receives royalties from Routledge, Cambridge University Press, and Cambridge Enterprise, and has consulted for Campden Instruments.

BJS consults for Cambridge Cognition.

Author Contributions

CL was responsible for study design, study set-up, study oversight, data analyses and writing of the manuscript. GKM was involved in the fMRI paradigm design and data analyses and editing of the manuscript. SA was involved in data collection, study oversight and editing of the manuscript. FK was involved in the fMRI paradigm design and data analyses. RNC was involved in oversight of the computational modelling and editing of the manuscript. AJ was responsible for medical oversight, blood sample collection and editing of the manuscript. PSJ was involved in study oversight and editing of the manuscript. JF was involved in data analyses and editing of the manuscript. DSS was involved in study design, study oversight and editing of the manuscript. GMK was involved in study design, study oversight, allocation to treatment group, statistical advice and editing of the manuscript. PF was involved in fMRI data collection, data analyses and editing of the manuscript. BJS was involved in study conception, study design, study oversight, statistical advice and writing of the manuscript.

References

1. Cools R, Roberts AC, Robbins TW. Serotonergic regulation of emotional and behavioural control processes. *Trends in cognitive sciences*. 2008;12(1):31-40.
2. Roiser JP, Elliott R, Sahakian BJ. Cognitive mechanisms of treatment in depression. *Neuropsychopharmacology*. 2012;37(1):117-36.
3. Elliott R, Sahakian B, Herrod J, Robbins T, Paykel E. Abnormal response to negative feedback in unipolar depression: evidence for a diagnosis specific impairment. *Journal of Neurology, Neurosurgery & Psychiatry*. 1997;63(1):74-82.
4. Beats B, Sahakian BJ, Levy R. Cognitive performance in tests sensitive to frontal lobe dysfunction in the elderly depressed. *Psychological medicine*. 1996;26(3):591-603.

5. Murphy F, Michael A, Robbins T, Sahakian B. Neuropsychological impairment in patients with major depressive disorder: the effects of feedback on task performance. *Psychological medicine*. 2003;33(3):455-67.
6. Elliott R, Sahakian BJ, McKay A, Herrod J, Robbins TW, Paykel E. Neuropsychological impairments in unipolar depression: the influence of perceived failure on subsequent performance. *Psychological medicine*. 1996;26(5):975-89.
7. Dam V, Stenbæk DS, Köhler-Forsberg K, Ip C, Ozenne B, Sahakian B, et al. Hot and cold cognitive disturbances in antidepressant-free patients with major depressive disorder: a NeuroPharm study. *Psychological Medicine*. 2021;51(14):2347-56.
8. Erickson K, Drevets WC, Clark L, Cannon DM, Bain EE, Zarate Jr CA, et al. Mood-congruent bias in affective go/no-go performance of unmedicated patients with major depressive disorder. *American Journal of Psychiatry*. 2005;162(11):2171-3.
9. Harmer CJ, Shelley NC, Cowen PJ, Goodwin GM. Increased positive versus negative affective perception and memory in healthy volunteers following selective serotonin and norepinephrine reuptake inhibition. *American Journal of Psychiatry*. 2004;161(7):1256-63.
10. Clevenger SS, Malhotra D, Dang J, Vanle B, IsHak WW. The role of selective serotonin reuptake inhibitors in preventing relapse of major depressive disorder. *Therapeutic advances in psychopharmacology*. 2018;8(1):49-58.
11. Dayan P, Balleine BW. Reward, motivation, and reinforcement learning. *Neuron*. 2002;36(2):285-98.
12. Kranz G, Kasper S, Lanzenberger R. Reward and the serotonergic system. *Neuroscience*. 2010;166(4):1023-35.
13. Cools R, Robinson OJ, Sahakian B. Acute tryptophan depletion in healthy volunteers enhances punishment prediction but does not affect reward prediction. *Neuropsychopharmacology*. 2008;33(9):2291-9.
14. Kanen JW, Apergis-Schoute AM, Yellowlees R, Arntz FE, van der Flier FE, Price A, et al. Serotonin depletion impairs both Pavlovian and instrumental reversal learning in healthy humans. *Molecular Psychiatry*. 2021;26(12):7200-10.
15. Murphy F, Smith K, Cowen P, Robbins T, Sahakian B. The effects of tryptophan depletion on cognitive and affective processing in healthy volunteers. *Psychopharmacology*. 2002;163:42-53.
16. Seymour B, Daw ND, Roiser JP, Dayan P, Dolan R. Serotonin selectively modulates reward value in human decision-making. *Journal of Neuroscience*. 2012;32(17):5833-42.
17. Skandali N, Rowe JB, Voon V, Deakin JB, Cardinal RN, Cormack F, et al. Dissociable effects of acute SSRI (escitalopram) on executive, learning and emotional functions in healthy humans. *Neuropsychopharmacology*. 2018;43(13):2645-51.
18. Chamberlain SR, Muller U, Blackwell AD, Clark L, Robbins TW, Sahakian BJ. Neurochemical modulation of response inhibition and probabilistic learning in humans. *Science*. 2006;311(5762):861-3.
19. Michely J, Eldar E, Erdman A, Martin IM, Dolan RJ. Serotonin modulates asymmetric learning from reward and punishment in healthy human volunteers. *Communications Biology*. 2022;5(1):812.
20. Langley C, Armand S, Luo Q, Savulich G, Segerberg T, Søndergaard A, et al. Chronic escitalopram in healthy volunteers has specific effects on reinforcement sensitivity: a double-blind, placebo-controlled semi-randomised study. *Neuropsychopharmacology*. 2023;48(4):664-70.
21. Scholl J, Kolling N, Nelissen N, Browning M, Rushworth MF, Harmer CJ. Beyond negative valence: 2-week administration of a serotonergic antidepressant enhances both reward and effort learning signals. *PLoS biology*. 2017;15(2):e2000756.
22. Bari A, Theobald DE, Caprioli D, Mar AC, Aidoo-Micah A, Dalley JW, Robbins TW. Serotonin modulates sensitivity to reward and negative feedback in a probabilistic reversal learning task in rats. *Neuropsychopharmacology*. 2010;35(6):1290-301.
23. Luo Q, Kanen JW, Bari A, Skandali N, Langley C, Knudsen GM, et al. Comparable roles for serotonin in rats and humans for computations underlying flexible decision-making. *Neuropsychopharmacology*. 2024;49(3):600-8.
24. Nibuya M, Morinobu S, Duman RS. Regulation of BDNF and trkB mRNA in rat brain by chronic electroconvulsive seizure and antidepressant drug treatments. *Journal of Neuroscience*. 1995;15(11):7539-47.
25. De Foubert G, Carney S, Robinson C, Destexhe E, Tomlinson R, Hicks C, et al. Fluoxetine-induced change in rat brain expression of brain-derived neurotrophic factor varies depending on length of treatment. *Neuroscience*. 2004;128(3):597-604.

26. Johansen A, Armand S, Plavén-Sigraý P, Nasser A, Ozenne B, Petersen IN, et al. Effects of escitalopram on synaptic density in the healthy human brain: a randomized controlled trial. *Molecular Psychiatry*. 2023:1-8.
27. Eldar E, Hauser TU, Dayan P, Dolan RJ. Striatal structure and function predict individual biases in learning to avoid pain. *Proceedings of the National Academy of Sciences*. 2016;113(17):4812-7.
28. Murray GK, Knolle F, Ersche KD, Craig KJ, Abbott S, Shabbir SS, et al. Dopaminergic drug treatment remediates exaggerated cingulate prediction error responses in obsessive-compulsive disorder. *Psychopharmacology*. 2019;236:2325-36.
29. Bernacer J, Corlett PR, Ramachandra P, McFarlane B, Turner DC, Clark L, et al. Methamphetamine-induced disruption of frontostriatal reward learning signals: relation to psychotic symptoms. *American Journal of Psychiatry*. 2013;170(11):1326-34.
30. Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*. 2006;442(7106):1042-5.
31. O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *science*. 2004;304(5669):452-4.
32. Niv Y, Edlund JA, Dayan P, O'Doherty JP. Neural prediction errors reveal a risk-sensitive reinforcement-learning process in the human brain. *Journal of Neuroscience*. 2012;32(2):551-62.
33. Elliott R, Dolan RJ, Frith CD. Dissociable functions in the medial and lateral orbitofrontal cortex: evidence from human neuroimaging studies. *Cerebral cortex*. 2000;10(3):308-17.
34. Guo Z, Chen J, Liu S, Li Y, Sun B, Gao Z. Brain areas activated by uncertain reward-based decision-making in healthy volunteers. *Neural regeneration research*. 2013;8(35):3344-52.
35. Stahl SM. *Stahl's essential psychopharmacology: neuroscientific basis and practical applications*: Cambridge university press; 2021.
36. Garnock-Jones KP, McCormack PL. Escitalopram: a review of its use in the management of major depressive disorder in adults. *CNS drugs*. 2010;24:769-96.
37. Cipriani A, Furukawa TA, Salanti G, Chaimani A, Atkinson LZ, Ogawa Y, et al. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. *The Lancet*. 2018;391(10128):1357-66.
38. Ermakova AO, Knolle F, Justicia A, Bullmore ET, Jones PB, Robbins TW, et al. Abnormal reward prediction-error signalling in antipsychotic naive individuals with first-episode psychosis or clinical risk for psychosis. *Neuropsychopharmacology*. 2018;43(8):1691-9.
39. Marzuki AA, Tomić I, Ip SHY, Gottwald J, Kanen JW, Kaser M, et al. Association of environmental uncertainty with altered decision-making and learning mechanisms in youths with obsessive-compulsive disorder. *JAMA network open*. 2021;4(11):e2136195-e.
40. Kanen JW, Ersche KD, Fineberg NA, Robbins TW, Cardinal RN. Computational modelling reveals contrasting effects on reinforcement learning and cognitive flexibility in stimulant use disorder and obsessive-compulsive disorder: remediating effects of dopaminergic D2/3 receptor agents. *Psychopharmacology*. 2019;236:2337-58.
41. Gronau QF, Sarafoglou A, Matzke D, Ly A, Boehm U, Marsman M, et al. A tutorial on bridge sampling. *Journal of mathematical psychology*. 2017;81:80-97.
42. Den Ouden HE, Daw ND, Fernandez G, Elshout JA, Rijpkema M, Hoogman M, et al. Dissociable effects of dopamine and serotonin on reversal learning. *Neuron*. 2013;80(4):1090-100.
43. Power JD, Mitra A, Laumann TO, Snyder AZ, Schlaggar BL, Petersen SE. Methods to detect, characterize, and remove motion artifact in resting state fMRI. *Neuroimage*. 2014;84:320-41.
44. Malamud J, Lewis G, Moutoussis M, Duffy L, Bone J, Srinivasan R, et al. The selective serotonin reuptake inhibitor sertraline alters learning from aversive reinforcements in patients with depression: evidence from a randomized controlled trial. *Psychological Medicine*. 2024:1-13.
45. Sugrue LP, Corrado GS, Newsome WT. Matching behavior and the representation of value in the parietal cortex. *science*. 2004;304(5678):1782-7.
46. Tom SM, Fox CR, Trepel C, Poldrack RA. The neural basis of loss aversion in decision-making under risk. *Science*. 2007;315(5811):515-8.
47. Gläscher J, Daw N, Dayan P, O'Doherty JP. States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*. 2010;66(4):585-95.
48. Platt ML, Glimcher PW. Neural correlates of decision variables in parietal cortex. *Nature*. 1999;400(6741):233-8.
49. Anderson BA, Laurent PA, Yantis S. Value-driven attentional priority signals in human basal ganglia and visual cortex. *Brain Research*. 2014;1587:88-96.

50. Anderson BA. Reward processing in the value-driven attention network: reward signals tracking cue identity and location. *Social cognitive and affective neuroscience*. 2017;12(3):461-7.
51. Peck CJ, Jangraw DC, Suzuki M, Efem R, Gottlieb J. Reward modulates attention independently of action value in posterior parietal cortex. *Journal of Neuroscience*. 2009;29(36):11182-91.
52. Barnhart WJ, Makela EH, Latocha MJ. SSRI-induced apathy syndrome: a clinical review. *Journal of Psychiatric Practice*. 2004;10(3):196-9.
53. Marazziti D, Mucci F, Tripodi B, Carbone MG, Muscarella A, Falaschi V, Baroni S. Emotional blunting, cognitive impairment, bone fractures, and bleeding as possible side effects of long-term use of SSRIs. *Clinical neuropsychiatry*. 2019;16(2):75.
54. American Psychological Association. APA Clinical Practice Guideline for the Treatment of Depression Across Three Age Cohorts. In: Association AP, editor. 2019.
55. National Institute for Health and Care Excellence. Depression in adults: treatment and management. In: National Institute for Health and Care Excellence, editor. 2022.

Figure Legends

Figure 1. Schematic of the Reward Learning Paradigm. A) shows the three different trial types and feedback probabilities. B) demonstrates the timeline of the experimental task. Participants are first presented with a fixation cue for 0.5s, followed by the presentation cues and they have a maximum of 2s to respond, their choice is then displayed for 3s until the feedback is presented for 1s.

Figure 2. Accuracy and reaction time performance during reward and punishment trials. A) depicts the accuracy performance, B) depicts the reaction time performance.

Figure 3. Computational modelling results for the reward and punishment trials. Error bars indicate no credible differences in posterior distributions between placebo and escitalopram groups for which the 90/95% highest density intervals (HDI) excluded 0. A) depicts the results for the reward trials, B) depicts the results for the punishment trials. Error bars in orange indicate credible differences in posterior distributions between placebo and escitalopram groups for which the 90% highest density interval (HDI) excluded 0.

Figure 4. Decreased BOLD activation in the escitalopram group vs the placebo group during reward trials. MNI Coordinates (28, -66, 32). $p < .05$ peak-level FWE corrected.

Tables

Table 1. Demographics.

	Placebo (n=32)	Escitalopram (n=32)	t or χ^2	p	Cohen's d or phi
Age	25.38 (5.77)	24.25 (5.56)	t=.79	.43	d=.20
Sex	21 Females (65.63%)	21 Females (65.63%)	$\chi^2=.00$	1.00	$\phi=.00$
IQ	111.50 (9.51)	112.22 (9.30)	t=-.31	.76	d=.08

Table 2. Reduced BOLD activation in the escitalopram group compared to the placebo group for feedback on reward trials.

Brain Region	BA	x	y	z	z-value	Cohen's d
Intraparietal Sulcus	19	28	-66	32	5.03	.63

Note: BA: Brodmann's Area