ORIGINAL INVESTIGATION

5-HT 2A and 5-HT 2C receptor antagonism differentially modulate reinforcement learning and cognitive flexibility: behavioural and computational evidence

Mona El- Sayed Hervig[1,2,3] · Katharina Zühlsdorff[1,2,4] · Sarah F. Olesen[5] · Benjamin Phillips[1] · Tadej Božič[5] · Jeffrey W. Dalley[1,2,6] · Rudolf N. Cardinal[2,6,7] · Johan Alsiö[1,2,8] · Trevor W. Robbins[1,2]

Abstract

Rationale Cognitive flexibility, the ability to adapt behaviour in response to a changing environment, is disrupted in several neuropsychiatric disorders, including obsessive–compulsive disorder and major depressive disorder. Evidence suggests that flexibility, which can be operationalised using reversal learning tasks, is modulated by serotonergic transmission. However, how exactly flexible behaviour and associated reinforcement learning (RL) processes are modulated by 5-HT action on specific receptors is unknown.

Objectives We investigated the effects of 5-HT$_{2A}$ receptor (5-HT$_{2A}$R) and 5-HT$_{2C}$ receptor (5-HT$_{2C}$R) antagonism on flexibility and underlying RL mechanisms.

Methods Thirty-six male Lister hooded rats were trained on a touchscreen visual discrimination and reversal task. We evaluated the effects of systemic treatments with the 5-HT$_{2A}$R and 5-HT$_{2C}$R antagonists M100907 and SB-242084, respectively, on reversal learning and performance on probe trials where correct and incorrect stimuli were presented with a third, probabilistically rewarded, stimulus. Computational models were fitted to task choice data to extract RL parameters, including a novel model designed specifically for this task.

Results 5-HT$_{2A}$R antagonism impaired reversal learning only after an initial perseverative phase, during a period of random choice and then new learning. 5-HT$_{2C}$R antagonism, on the other hand, impaired learning from positive feedback. RL models further differentiated these effects. 5-HT$_{2A}$R antagonism decreased punishment learning rate (i.e. negative feedback) at high and low doses. The low dose also decreased reinforcement sensitivity (beta) and increased stimulus and side stickiness (i.e., the tendency to repeat a choice regardless of outcome). 5-HT$_{2C}$R antagonism also decreased beta, but reduced side stickiness.

Conclusions These data indicate that 5-HT$_{2A}$ and 5-HT$_{2C}$Rs both modulate different aspects of flexibility, with 5-HT$_{2A}$Rs modulating learning from negative feedback as measured using RL parameters and 5-HT$_{2C}$Rs for learning from positive feedback assessed through conventional measures.

Keywords Cognitive flexibility · Reinforcement learning · 5-HT receptors · Stickiness · Reward learning · Punishment learning

Introduction

The monoamine neurotransmitter serotonin (5-hydroxy-tryptamine; 5-HT) system is implicated in several neuropsychiatric disorders, including major depressive disorder (MDD), obsessive–compulsive disorder (OCD)

and schizophrenia, disorders in which cognitive flexibility and reinforcement learning (RL) are altered (Chamberlain et al. 2006; Clevenger et al. 2018; Zhu et al. 2021). Drugs that target the 5-HT system are often the first-line pharmacological treatment for these disorders, such as selective serotonin reuptake inhibitors (SSRIs) for MDD and OCD (APA 2010; Fineberg et al. 2020). Emerging therapies such as the 5-HT agonist psilocybin and other psychedelics are thought to hold promising treatment potential to ameliorate symptoms such as cognitive inflexibility and anhedonia

(Andersen et al. 2021; Carhart-Harris and Friston 2019; Doss et al. 2021; Stroud et al. 2018). Thus, understanding the role of serotonergic modulation mediated by specific 5-HT receptors is critical for developing future therapies for disorders characterized by inflexible behaviour and diminished RL.

5-HT contributes to various cognitive processes across species, including RL (Den Ouden et al. 2013; Iigaya et al. 2018) and cognitive flexibility (Alsiö et al. 2021; Barlow et al. 2015; Clarke et al. 2004). Cognitive flexibility is defined as the ability to adapt behaviour in response to changes in the environment. Inflexible behaviour can manifest itself as compulsive behaviour, e.g. excessively perseverative actions that are independent of outcome–value associations (Berlin and Hollander 2014; Jentsch and Taylor 2001; Koob and Volkow 2016). Moreover, the ability to adjust behaviour to changes in the environment is closely linked to underlying RL processes, which integrate positive and negative feedback from the environment to maximise rewards and minimise punishment (Sutton and Barto 1998).

Flexible responding can be assessed using reversal learning paradigms across species (Uddin 2021). During reversal learning tasks, initially learned stimulus contingencies change and the subject needs to update behaviour accordingly. Substantial evidence suggests that 5-HT is involved in the modulation of reversal learning, as shown through 5-HT depletion in the orbitofrontal cortex (OFC) in monkeys (Clarke et al. 2004, 2005; Rygula et al. 2015) and rats (Alsiö et al. 2021; Izquierdo et al. 2012). In humans, acute tryptophan depletion (reducing 5-HT levels due to a reduction in its amino-acid precursor tryptophan) increases outcome-independent choice perseveration (Seymour et al. 2012) and impairs reversal learning (Kanen et al. 2021). 5-HT also modulates RL processes underlying flexible behaviour, possibly through distinct mechanisms (Bari et al. 2010; Seymour et al. 2012). In healthy human participants, short-term administration of the SSRI citalopram results in increased punishment learning and reduced reward learning (Michely et al. 2022). In patients with MDD, SSRIs impairs learning from negative feedback, while having negligible effects on learning from positive feedback (Herzallah et al. 2013). In rats, acute low-dose citalopram improves negative feedback sensitivity, while acute high-dose citalopram impairs negative feedback sensitivity, similarly to observations in human studies (Bari et al. 2010).

While it is evident that 5-HT is a key modulator of behavioural flexibility, it targets a broad range of receptor subtypes with diverse actions, exerting both excitatory and inhibitory transmission depending on receptor subtype and localisation (Alvarez et al. 2021). Thus, it is vital to understand the modulatory role of 5-HT through different receptors on cognition and RL. In particular, the excitatory 5-HT$_{2A}$Rs, which are primarily localized on excitatory pyramidal neurons, and inhibitory 5-HT$_{2C}$Rs, found primarily on inhibitory parvalbumin neurons, seem to be involved in reversal learning – possibly with dissociable roles (Aghajanian and Marek 1999; Amargós-Bosch et al. 2004; Liu et al. 2007; Santana et al. 2004). Systemic 5-HT$_{2A}$R blockade impairs spatial reversal learning performance, whereas systemic blockade of 5-HT$_{2C}$Rs improves performance (Boulougouris et al. 2008). Moreover, high levels of perseveration in rats have been found to be associated with decreased levels of 5-HT$_{2A}$R in the OFC (Barlow et al. 2015), consistent with decreased levels of 5-HT$_{2A}$R density in the OFC and PFC predicting clinical severity in OCD patients (Perani et al. 2008). Recent findings also suggest that psilocybin improves cognitive flexibility through a mechanism dependent on 5-HT$_{2A}$Rs, but not 5-HT$_{2C}$Rs (Torrado Pacheco et al. 2023). Less is known about the effects of 5-HT$_{2A}$R and 5-HT$_{2C}$R stimulation and blockade on component processes of reversal learning, including sensitivity to feedback and subsequent action selection.

To investigate the specific roles of 5-HT receptors in flexibility and RL, we employed the valence-probe visual discrimination (VPVD) task (Alsiö et al. 2019) and combined this task with RL modelling to gain a deeper insight into the latent processes underlying behaviour. We recently employed RL computational modelling to assess effects of 5-HT depletion and SSRI treatment in a different, probabilistic reversal task (Luo et al. 2023). We thus aimed in this study to extend this analysis to specific 5-HT receptor agents. Such models are fitted to trial-by-trial data and allow for extraction of parameters such as value-dependent (i.e., dependent on wins/losses on the previous trial) positive and negative learning rates, the 'reinforcement sensitivity' parameter, as well as the value-independent side and stimulus stickiness parameters, which reflect repeated responses to the same side or stimulus, respectively, regardless of the outcome on the previous trial (Daw 2009). Stickiness differs from perseveration as it provides a measure of the overall tendency to repeat a choice based on all previous trials, whereas perseveration is usually measured as the number of responses to the previously correct stimulus after a reversal. These parameters reflect different aspects of flexibility and RL, separating value-dependent from value-independent components. We examined whether these parameters contribute to choice behaviour on the VPVD task and if they were affected by 5-HT$_{2A}$R or 5-HT$_{2C}$R blockade. We hypothesized that 5-HT$_{2A}$R blockade would increase stickiness parameters, and that 5-HT$_{2C}$R blockade would lead to higher learning rates, as previous studies (summarized

above) have shown increased perseveration following $5\text{-}HT_{2A}R$ blockade and improved reversal learning behaviour resulting from $5\text{-}HT_{2C}R$ antagonism. Computational modelling thus enables us to investigate the roles of the different $5\text{-}HT_2$ receptors more precisely in different aspects of RL behaviour.

Materials and methods

Animals

Subjects were male hooded Lister rats ($N = 36$; Charles River, UK) (Fig. 1) housed in groups of three or four throughout the experiments. The rats underwent two experiments. In the first experiment ($5\text{-}HT_{2A}R$ antagonism), all 36 rats were included. In the following $5\text{-}HT_{2C}R$ antagonist experiment, 35 rats were included, as one rat had to be euthanised due to seizures. The rats were housed under a reverse 12-h light/dark cycle with lights off at 0700 h. All training and testing was performed during the dark phase. To ensure sufficient motivation for task performance, the animals were food restricted with *ad libitum* access to water and fed once daily at random times after testing. Their body weights were maintained at 85% of their free-feeding weight. All experiments were subject to regulation by the United Kingdom Home Office (PPL 70/7548) in accordance with the Animals (Scientific Procedures) Act 1986.

Drugs

M100907 (R-(+)-α-(2,3-dimethoxyphenyl)-1-[2-(4-fluorophenylethyl)]-4-piperidinemethanol) (Sigma Aldrich, #M3324), a highly selective $5\text{-}HT_{2A}R$ antagonist (Kehne et al. 1996), was dissolved in 0.01 M phosphate-buffered saline (PBS) and 0.1 M hydrochloride, and adjusted with NaOH to pH 7. M100907 was administered at 0 (vehicle), 0.03 or 0.1 mg/kg.

SB-242084 (Eli Lilly, Indianapolis, IN, USA) was first dissolved in polyethene glycol 400 (PEG400) (Fisher Scientific, Loughborough, UK) at 20% of the final required volume, and then made up by 10% (w/v) hydroxypropyl-beta-cyclodextrin (Sigma-Aldrich, Poole, UK) in saline, and checked that the pH was 7. For systemic treatment, SB-242084 was administered intraperitoneally (i.p.) at doses of 0 (vehicle), 0.3 or 1.0 mg/kg in a volume of 1 ml/kg, 30 min prior to testing. Drugs were divided into the aliquots required for each test day and frozen at $-80\,°C$.
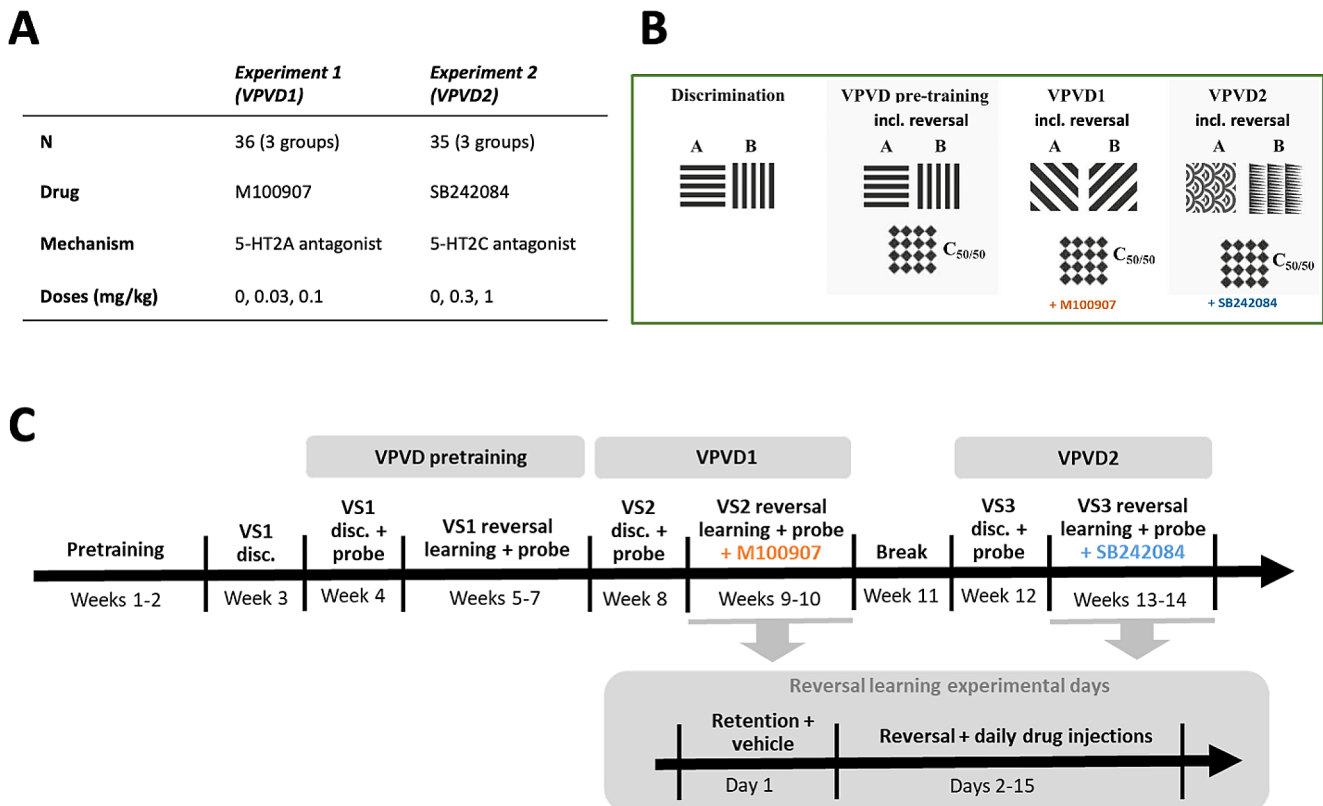

Fig. 1 Experimental design. (A) Table of groups and treatments. (N, number of subjects). (B) VPVD stages and stimuli in the M100907 and SB-242,084 experiments. A is the 100% reinforced stimulus, B is the 0% reinforced stimulus, C is reinforced on 50% of probe trials. (C) Experimental timeline, including pretraining and experimental weeks. (disc., discrimination; VPVD, valence-probe visual discrimination (1 and 2); VS, visual stimulus pair (1, 2 and 3); VS2, visual stimulus pair 2; VS3, visual stimulus pair 3

Valence-probe visual discrimination task with reversal

Behavioural training was performed as previously described in (Alsiö et al. 2019). The VPVD task can assess the effect of positive or negative feedback on learning through a neutral stimulus that is probabilistically reinforced (Phillips et al. 2018). For experimental timeline and design see Fig. 1 and for additional information on the apparatus, behavioural pre-training, and touchscreen visual discrimination and reversal, see Supplementary Materials.

After pre-training, the rats progressed to the VPVD task. The VPVD task was a three-stimulus task, during which responses to one stimulus (A+) were rewarded, whereas responding to the other stimulus (B−) was punished with a time-out. A third stimulus, probabilistically rewarded on average 50% of the time ($C_{50/50}$), was paired with either the A+ or B− on 'probe' trials (Fig. 1).

The trial structure was kept constant, but a tone was played every time a trial was rewarded, and the stimulus duration was unlimited to ensure that animals completed the probe trials. The probe stimulus and frequency of probe trials (every 4 or 5 trials) were determined based on a previous study (Alsiö et al. 2019). After optimization, each of the probe trials was presented once every 8 trials: randomized, but never on the first trial within any 8-trial bin. There was a maximum of 200 trials per session. Both the intertrial interval and time-out (on non-rewarded trials) were 5 s. Rats were initially tested for 5 days on the same A+ and B− as during the pre-training reversal (i.e., 'horizontal bars' vs. 'vertical bars'). The animals then completed a visual discrimination with a novel pair of stimuli ('slashes' vs. 'backslashes'; counterbalanced across rats). Training continued for a minimum of 5 sessions but could be extended to allow rats to reach 80% correct on the standard trials within the task. Once all rats had reached the criterion, all rats progressed to the 'reversal learning experiment'. On the day before reversal and start of drug treatment, the rats received a saline injection and were given a retention test session. The next day, rats were matched for stimulus–reward contingencies, performance on the probe trials before reversal and pre-training reversal performance, and accordingly allocated to a drug group. The stimulus–reward contingencies were reversed on the first day of reversal and then remained the same for the duration of the training sessions (i.e., there were only between-session reversals). The drug was administered before testing each day. The same stimulus ('diamonds') was used as the probe stimulus for all rats and across each of the phases, both during training and test trials. Training during the SB-242084 experiment followed the same procedure as above but rats were trained on a new pair of stimuli ('arcs' vs. 'triangles' counterbalanced across rats; the probe stimulus was kept the same) before reversal of the new stimulus − reward contingencies. In this case, the allocation into drug groups was also balanced based on previous drug exposure.

Hierarchical bayesian reinforcement learning modelling

The VPVD data were modelled with RL models using a hierarchical Bayesian approach. In total, nine different models were implemented in Stan (version 2.26.1), containing different combinations of parameters. The methods and models tested are described in more detail in the Supplementary Materials.

Q-values were updated on each trial using the following equation:

$$Q_{t+1}(c_t) = Q_t(c_t) + \alpha \times (r_t - Q_t(c_t)) \tag{1}$$

where $Q_{t+1}(c_t)$ is the Q-value of the stimulus chosen on the current trial for the next, $Q_t(c_t)$ is the expected value of the stimulus selected on the current trial, α is the learning rate and r_t is the reinforcement on trial t (1 for reward and 0 for punishment). The learning rate reflects how much the Q-value is updated based on the prediction error $r_t - Q_t(c_t)$, with higher α driving faster learning.

Next, the softmax decision rule was used to calculate the probability of making one of two choices:

$$P(c_t = L|Q_t(L), Q_t(R)) = \frac{e^{Q_t(L)\beta}}{e^{Q_t(L)\beta} + e^{Q_t(R)\beta}} \tag{2}$$

$Q_t(L)$ and $Q_t(R)$ are the Q-values of the left and right stimuli, and β is the reinforcement sensitivity parameter, which determines to what extent the subject is driven by its reinforcement history (versus random choice). Lower values of β indicate greater exploration and lower sensitivity to reinforcement, whereas greater values represent increased exploitation and greater sensitivity to reinforcement.

The behavioural data were simulated with the posterior group mean parameters from the winning model, to ensure that the model could reproduce behavioural observations. The simulations were then analysed using a conventional approach as described below.

Statistical analyses

Data across days within one reversal were collapsed, and trial outcomes were coded as perseverative, random, or learning depending on performance over bins of 30 trials in a rolling window, as described in detail and illustrated previously (Hervig et al. 2020), and following binomial distribution probabilities (Jones and Mishkin 1972).

The main measures were percentage correct responses ('% correct') on the standard A−< B+ trials and '% optimal choice' for the negative and positive probe trials across sessions. The optimal choice percentage was defined as the percentage of trials where the highest reward-probability option was chosen. Only data up to (and including) the first block of 30 trials where a rat reached criterion (24/30 correct) were analysed.

We also analysed response and collection latencies. Drug effects on standard parameters were analysed using linear mixed-effects models with the lmer package in R as described previously (Phillips et al. 2018) and as recommended for such data (Wickham 2014). The model contained two fixed factors (dose and session or dose and phase) and one random factor (subject). When relevant, further analyses were performed by conducting separate multilevel models on 'dose' for each session or phase. These analyses were followed by *post hoc* Dunnett's corrected pairwise comparisons with the relevant vehicle condition. Significance was set at $\alpha = 0.05$.

Visualization and statistical tests were performed with R, version 4.1.2 (R Core Team 2021). Response frequencies were square-root transformed, latencies were log transformed and probabilities were arcsine transformed to ensure normality, as confirmed with a quantile–quantile plot of residuals.

Results

Experiment 1: effects of systemic 5-HT$_{2A}$R blockade on reversal learning and reinforcement learning parameters

Effects of systemic 5-HT$_{2A}$R blockade on reinforcement learning processes: computational modeling

After computational modeling of VPVD choice behaviour, Model 9 was the best-fitting model (Table 1).

This model included the following parameters: α_{rew} (reward learning rate), α_{pun}, (punishment learning rate), β (reinforcement sensitivity), κ_{stim} (stimulus stickiness), κ_{side} (side stickiness), and the discount factor ρ. Learning from negative feedback was decreased by both low (difference in parameter per-group mean, posterior 95% highest density interval (HDI) excluding zero (group difference, $0 \notin 95\%$ HDI)) and high (group difference, $0 \notin 75\%$ HDI) doses of M100907. There was some evidence that low, but not high, dose M100907, also decreased the reinforcement sensitivity parameter (reflecting decreased sensitivity to reinforcement) (group difference, $0 \notin 75\%$ HDI) and increased the stimulus stickiness parameter (group difference, $0 \notin 75\%$ HDI). The side (location) stickiness parameter was increased in the low dose group (group difference, $0 \notin 95\%$ HDI) and slightly increased in the high dose group (group difference, $0 \notin 75\%$ HDI). The reward learning rate and discount factor were unaffected by M100907 treatment (no group differences, $0 \in 75\%$ HDI) (Fig. 2 and Table 2. The mean and standard deviation of the novel discount factor ρ for each group can be found in Supplementary Table 2

Furthermore, we simulated the behavioural data using the extracted parameters from the winning model. The data modelled was separated into standard, positive and negative probe trials. The simulations were able to capture the dynamics of behaviour on the VPVD task, as can be seen in the Supplementary Materials (Figure SF.1)

Effects of 5-HT$_{2A}$R blockade on VPVD reversal: standard behavioural parameters

There was weak evidence that systemic M100907 impaired performance on the VPVD task. On the standard (A−< B+) trials, there was a trend towards a main effect of dose ($F_{2,35} = 2.93$, $p = 0.066$) and a trend towards a dose × session interaction ($F_{26,455} = 1.52$, $p = 0.051$) (Fig. 2A). As there were evident trending effects (although non-significant), we

Table 1 Model comparison summary. Models were assumed to be equiprobable *a priori*

| Model	Parameters	Rank (M100907)	Log marginal likelihood (M100907)	Log posterior P (M100907)	Rank (SB-248,420)	Log marginal likelihood (SB-248,420)	Log posterior P (SB-248,420)
1	α, β	9	-57308.08	-5267.60	9	-53630.42	-3381.41
2	α, β, κ_{stim}	8	-57299.67	-5259.07	8	-53593.89	-3344.88
3	α, β, κ_{side}	5	-52320.21	-279.61	4	-50351.08	-102.07
4	α, β, κ_{side}, κ_{stim}	4	-52270.79	-230.18	3	-50322.69	-73.68
5	α_{rew}, α_{pun}, β	6	-57182.04	-5141.43	6	-53561.99	-3312.98
6	α_{rew}, α_{pun}, β, κ_{stim}	7	-57243.58	-5202.97	7	-53588.28	-3339.27
7	α_{rew}, α_{pun}, β, κ_{side}	3	-52194.25	-153.64	1	-50249.01	0.000
8	α_{rew}, α_{pun}, β, κ_{side}, κ_{stim}	2	-52144.88	-104.27	5	-50386.28	-127.28
9	α_{rew}, α_{pun}, β, κ_{side}, κ_{stim}, ρ	1	-52040.60	0.000	2	-50255.83	-6.82

A

B
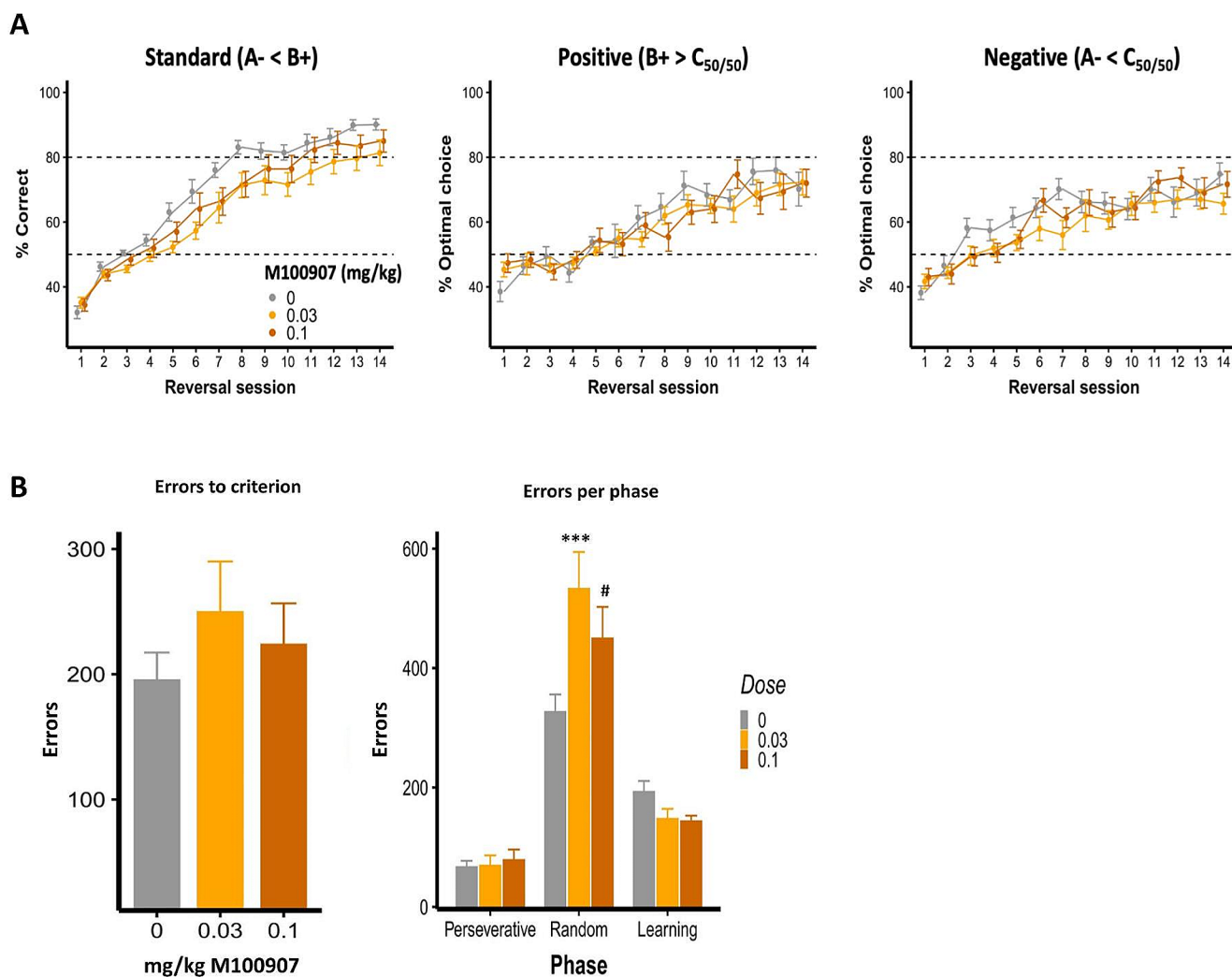
Fig. 2 Effects of M100907 on VPVD parameters. **(A)** Percent correct and percent optimal choice across sessions. **(B)** Errors to criterion and errors per phase. Results are represented as mean ± standard error of the mean (SEM); *** $p < 0.01$, # $p < 0.1$

Table 2 Summary of the effects of low and high dose M100907 and SB-242084 on reinforcement learning parameters (↑/↓, increase/decrease; - indicates no change at those levels; blank cells for parameters not tested with a given data set, as described in the Supplementary Methods). Red indicates $0 \notin 95\%$ HDI; orange indicates $0 \notin 75\%$ HDI

Parameter	M100907–0.03 mg/kg	M100907–0.1 mg/kg	SB-242,084–0.3 mg/kg	SB-242,084–1.0 mg/kg
α_{rew}	-	-	-	-
α_{pun}	↓	↓	-	-
β	↓	-	-	↓
κ_{side}	↑	↑	↓	↓
κ_{stim}	↑	-		
ρ	-	-		

performed further *post hoc* analyses within each session. *Post hoc* comparisons following correction for multiple comparisons revealed that the 0.03 mg/kg dose significantly reduced

correct responding on sessions 6 ($t_{112} = -2.50$, $p = 0.027$), 8 ($t_{112} = -2.63$, $p = 0.019$), 13 ($t_{112} = -2.79$, $p = 0.012$) and 14 ($t_{112} = -2.37$, $p = 0.036$). On positive and negative probe trials, we found no dose × session interactions (positive: $F_{26,455} = 1.30$, $p = 0.15$; negative: $F_{26,455} = 1.12$, $p = 0.31$) or main effect of dose (positive: $F_{2,35} = 0.30$, $p = 0.74$; negative: $F_{2,35} = 1.52$, $p = 0.23$) on % optimal choice.

For errors to criterion, there was a significant drug × phase interaction ($F_{4,105} = 3.85$, $p = 0.0058$), but no effect of M100907 overall ($F_{2,105} = 0.21$, $p = 0.81$). Further analysis based on planned pairwise comparisons showed that 0.03 mg/kg M100907 significantly increased errors in the random phase ($t_{115} = 3.59$, $p = 0.0010$), while there was a trend of 1 mg/kg M100907 towards increasing errors ($t_{115} = 2.18$, $p = 0.060$) in this phase.

Experiment 2: effects of systemic 5-HT$_{2C}$R blockade on reversal learning and reinforcement learning parameters

Effects of systemic 5-HT$_{2C}$R blockade on reinforcement learning processes: computational modeling

Model 7 was the winning model for this dataset (including parameters α_{rew}, α_{pun}, β and κ_{side}) (Model 9 did not converge; see **Supplementary Material**). It showed that learning from positive and negative feedback were unaffected by SB-242084 (no group differences, $0 \in 75\%$ HDI) (Fig. 3 and Table 2. High-dose SB-242084 decreased the reinforcement sensitivity parameter (i.e., reducing sensitivity to feedback) (group difference, $0 \notin 75\%$ HDI). The side stickiness parameter was decreased by low-dose (group difference, $0 \notin 95\%$ HDI) and high-dose (group difference, $0 \notin 75\%$ HDI) SB-242,084. We also simulated

the data for this experiment using the extracted parameters FigureSF.2

Effects of 5-HT$_{2C}$R blockade on VPVD reversal: standard behavioural parameters

Systemic SB-242084 impaired performance in the VPVD reversal learning task. On the standard (A−< B+) trials, there was a trend towards a main effect of dose ($F_{2,35} = 3.15$, $p = 0.055$) but no dose × session interaction ($F_{26,455} = 0.81$, $p = 0.74$) (Fig. 3). On positive probe trials, there was a significant main effect of dose on % optimal choice($F_{2,35} = 7.38$, $p = 0.0021$) but no dose × session interaction ($F_{26,455} = 1.04$, $p = 0.41$). As there were evident trending effects (although non-significant), we performed further *post hoc* analyses within each session for the standard (A−< B+) trials. *Post hoc* comparisons revealed that the 1.0 mg/kg SB-242084 significantly reduced % correct on sessions 7
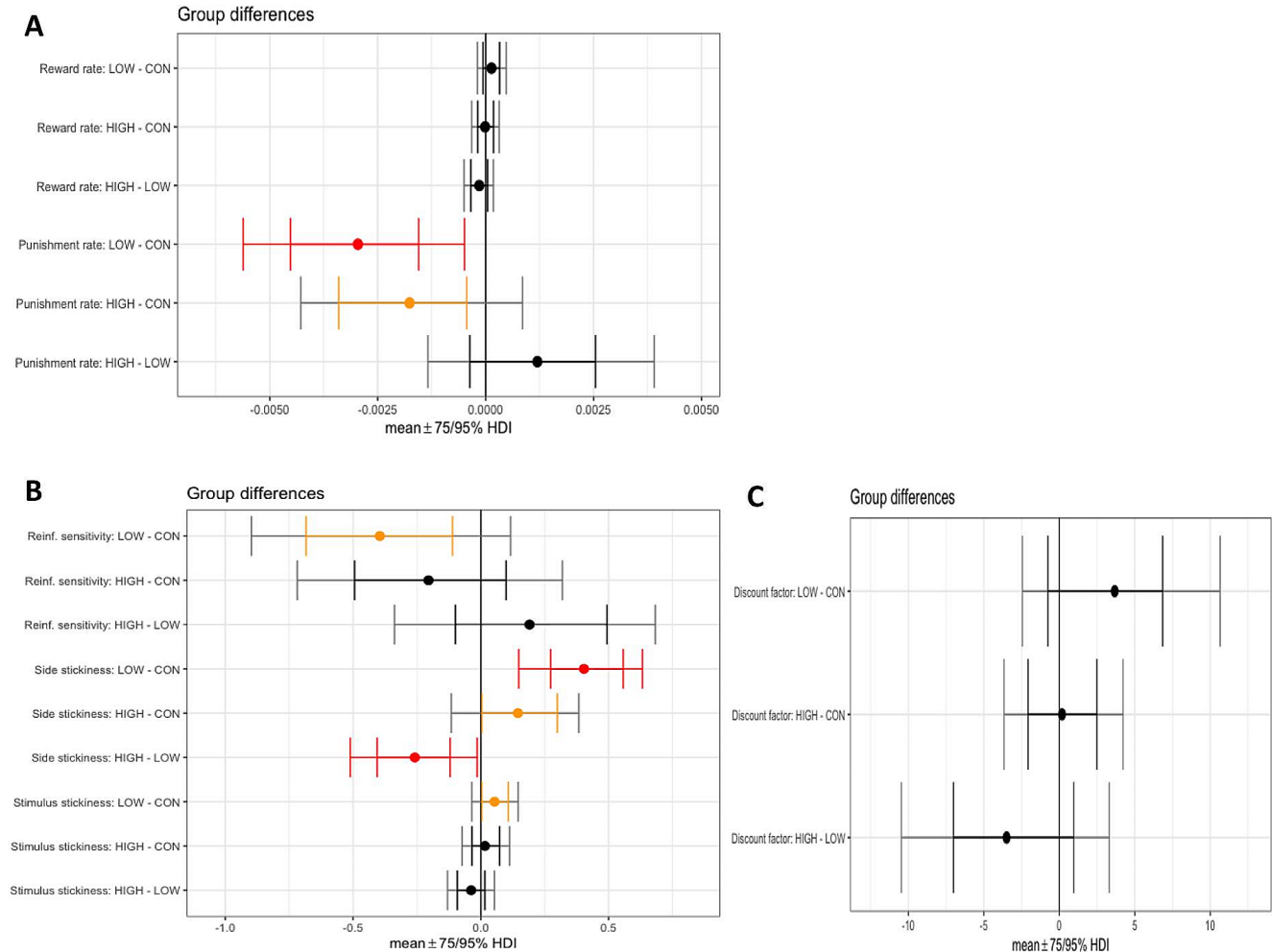
Fig. 3 Results from the hierarchical Bayesian winning RL model 9, showing differences in group mean parameters following M100907 administration. **(A)** Reward and punishment learning rate parameters. **(B)** Reinforcement sensitivity, side and stimulus stickiness parameters.

(C) Discount factor. (LOW, low dose; HIGH, high dose; CON, vehicle; Reinf., reinforcement; HDI, highest posterior density interval. Red indicates $0 \notin 95\%$ HDI; orange indicates $0 \notin 75\%$ HDI)

$(t_{91.8} = -2.63, p = 0.020)$ and 8 $(t_{91.8} = -2.35, p = 0.040)$. On positive probe trials, *post hoc* analyses showed that % optimal choice was significantly decreased on sessions 8 $(t_{423} = -2.48, p = 0.026)$, 9 $(t_{423} = -2.61, p = 0.018)$, 11 $(t_{423} = -2.39, p = 0.034)$ and 12 $(t_{423} = -2.24, p = 0.049)$.

For errors to criterion, we found no effect of SB-242,084 overall $(F_{2,105} = 1.80, p = 0.17)$. When analysing the effect of SB-242084 on errors per phase, we found a trend towards a main effect of dose $(F_{2,35} = 3.15, p = 0.055)$ and significant effect of phase $(F_{2,70} = 53.15, p < 0.0001)$, but no dose × phase interaction $(F_{4,70} = 0.50, p = 0.73)$.

Win-stay/lose-shift and latency analyses for both experiments can be found in the **Supplementary Materials**.

DISCUSSION

These findings indicated contrasting, as well as common, effects of 5-HT$_{2A}$ and 5-HT$_{2C}$ R antagonists on measures of RL and cognitive flexibility in the rat. We used a computational modelling approach to visual discrimination reversal that characterized novel drug effects not seen previously using standard behavioural measures. The RL parameters enabled us to gain a deeper insight into the latent mechanisms underlying behaviour on the VPVD task.

Effects of 5-HT$_{2A}$R antagonism on reinforcement learning and cognitive flexibility

Selective blockade of 5-HT$_{2A}$Rs using M100907 impaired reversal learning as reflected by reductions in % correct on standard trials and an increasing frequency of errors after the initial perseverative phase at the random choice and learning phases. This impairment was not associated with changes in response or collection latencies, showing that it was unlikely to be caused by motivational or sensorimotor deficits. Computational analyses revealed that 5-HT$_{2A}$R antagonism impaired learning from negative feedback, decreased the reinforcement sensitivity parameter and increased both side and stimulus 'stickiness', suggesting differential effects of 5-HT$_{2A}$R blockade on value-dependent (reinforcement sensitivity) compared to value-independent (stickiness) choices, which may reflect distinct facets of the cognitive flexibility construct.

Previous studies using systemic (Boulougouris et al. 2008) or intra-lateral OFC (Hervig et al. 2020) M100907 have also shown impaired reversal learning performance, consistent with the present findings. Moreover, lower 5-HT$_{2A}$R binding in the rat OFC is associated with more perseveration during spatial reversal (Barlow et al. 2015). Our findings may seem inconsistent with studies showing that the 5-HT$_{2A}$R antagonist ketanserin normalizes impairments in flexibility resulting from lysergic acid diethylamide (LSD), which is a

partial 5-HT$_{2A}$R agonist, as well as general improvements in set-shifting following ketanserin administration in rats (Baker et al. 2011; Pokorny et al. 2020; Torrado Pacheco et al. 2023). However, such apparent inconsistencies may have resulted from the use of different paradigms to assess flexibility, such as set-shifting, which may involve distinct neural and 5-HT dependent substrates than reversal learning (Clarke et al. 2005; Dias et al. 1996).

Dose may also be a relevant factor. The lower dose of 0.03 mg/kg M100907 affected reversal learning more than the 0.1 mg/kg dose, possibly reflecting an inverted U-curve effect, as previously reported for 5-HT$_{2A}$R antagonists (Marek et al. 2005). Dose-response studies have shown that moderate systemic doses of M100907 are more effective than low and high doses on a response-inhibition task and that intra-lOFC infusions with moderate M100907 doses induce the most detrimental effects on reversal learning (Furr et al. 2012; Marek et al. 2005). The high-dose of the 5-HT$_{2A}$R antagonist may have induced receptor internalization, an established mechanism for the 5-HT$_{2A}$R which produces such apparently paradoxical effects (Roth 2011) (Fig. 4).

The findings align with our initial hypothesis of increased stickiness following 5-HT$_{2A}$R blockade. Selective depletions of 5-HT in the marmoset OFC and amygdala using 5,7-DHT also results in increased side stickiness rates, similar to our findings following 5-HT$_{2A}$R antagonism (Rygula et al. 2015), suggesting that 5-HT$_{2A}$Rs in these areas may modulate the stickiness parameter, i.e., repeating responses regardless of previous outcomes. This accords with the demonstration that side stickiness is correlated with functional connectivity between the amygdala and medial OFC in rats (Zühlsdorff et al. 2023).

Effects of 5-HT$_{2C}$R antagonism on reinforcement learning and cognitive flexibility

Antagonism of 5-HT$_{2C}$Rs with SB-242084 decreased % correct and % optimal choice on the VPVD task at high doses. Previous data have shown that this agent can improve serial reversal performance in the initial perseverative phases due to reduced perseveration but that there is an overall decremental effect on performance, possibly due to impaired (re-) learning of associations after perseveration has been overcome (Alsiö et al. 2015). This interpretation is supported by differential roles of 5-HT in lateral orbitofrontal and medial prefrontal cortex (Alsiö et al. 2019). In probabilistic reversal tasks, where there is already a high baseline of response shifting, further increases are unlikely to improve performance and may impair it (e.g., human data in (Kanen et al. 2019). Using RL models, we found here that 5-HT$_{2C}$R blockade decreased the reinforcement sensitivity parameter at a higher dose and decreased side stickiness at low and high doses. In both the
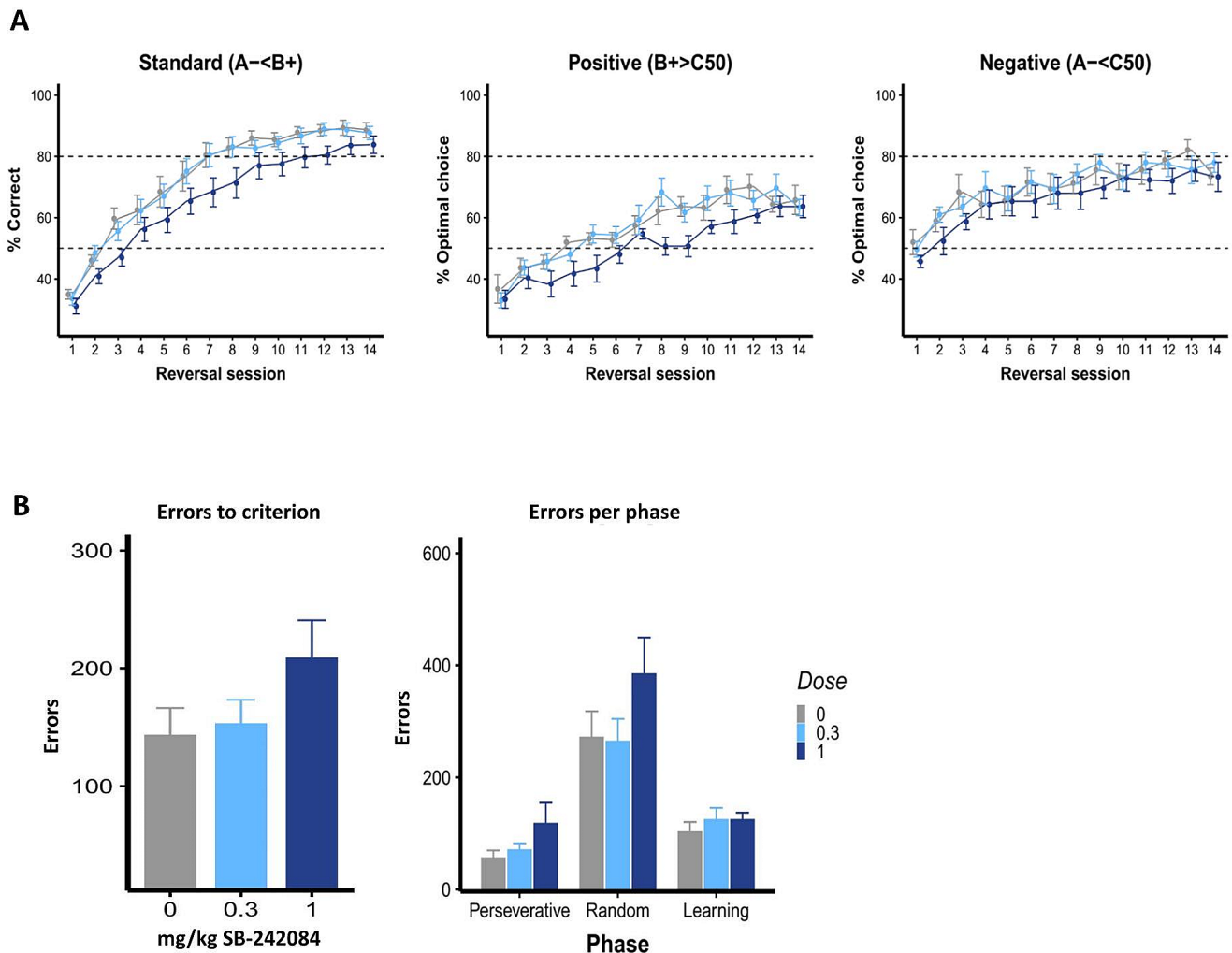
A

B

Fig. 4 Effects of SB-242048 on VPVD parameters. **(A)** Percent correct and percent optimal choice across sessions. **(B)** Errors to criterion and errors per phase. Results are represented as mean ± SEM; *** $p < 0.01$, # $p < 0.1$

present study and in Phillips et al. (2018), SB-242084 impaired performance and reduced reinforcement sensitivity. This drug therefore appeared to enhance flexible responding as reflected by the reinforcement sensitivity and side stickiness parameters and (Fig. 5) this may account for the initial positive effects on serial reversal. This observation is in accordance with studies showing SB-242084 to improve performance during perseverative phases of serial visual reversal learning (Boulougouris et al. 2008). Our findings indicate that this improvement may be due to decreased side stickiness following SB-242084 administration. However, the reduction in reinforcement sensitivity may lead to an overall deficit in performance.

Implications for mechanisms of action of SSRIs and psychedelics in psychiatric disorders

In a recent analysis, lower doses of the SSRI citalopram increases the reward learning rate and decreases side stickiness, whilst decreasing reward rate and increasing reinforcement sensitivity at a higher dose (Luo et al. 2023). Acute escitalopram in healthy human participants reduces the reward learning rate, decreases reinforcement sensitivity, and decreases stimulus stickiness (Luo et al. 2023), partially aligning with our findings following 5-HT$_{2C}$R blockade. Our findings using selective 5-HT$_{2A}$R and 5-HT$_{2C}$R antagonists may thus aid our understanding of mechanisms underlying cognitive flexibility and RL.

Psilocybin and other psychedelics are receiving increased attention for their therapeutic potential in treating neuropsychiatric disorders such as MDD and anxiety (Carhart-Harris et al. 2016, 2021; Goldberg et al. 2020). Even though their mechanisms are poorly understood, one hypothesis is that psilocybin improves cognitive flexibility (Baker et al. 2011; Torrado Pacheco et al. 2023). Psilocybin, which primarily exerts its psychoactive effects through 5-HT$_{2A}$R agonism (Madsen et al. 2019), has been shown to increase cognitive
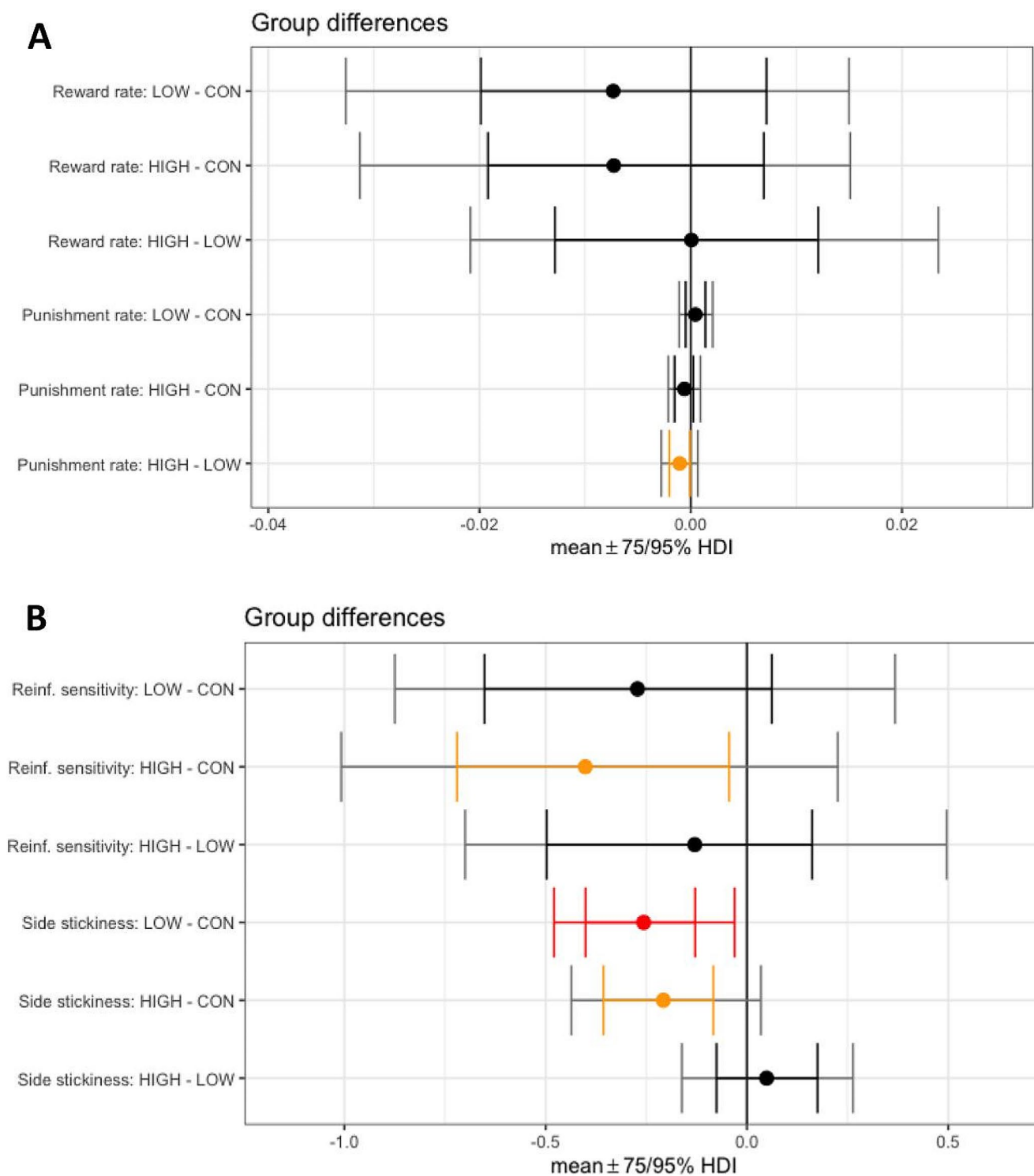
Fig. 5 Results from the hierarchical Bayesian winning RL model 7, showing differences in group mean parameters following SB-242048 administration. **(A)** Reward and punishment learning rate parameters. **(B)** Reinforcement sensitivity and side stickiness parameters. (LOW, low dose; HIGH, high dose; CON, vehicle; Reinf., reinforcement; HDI, highest posterior interval. Red indicates $0 \notin 95\%$ HDI; orange indicates $0 \notin 75\%$ HDI)

flexibility in individuals with MDD for at least 4 weeks (Doss et al. 2021). Ayahuasca, which contains the $5\text{-HT}_{2A}\text{R}$ agonist dimethyltryptamine, similarly increases cognitive flexibility in healthy volunteers (Kuypers et al. 2016; Murphy-Beiner and Soar 2020). In contrast, 2,5-dimethoxy-4-iodoamphetamine, a $5\text{-HT2}_{A/C}\text{R}$ agonist, impairs flexible strategy choice, highlighting different mechanisms of actions of hallucinogenic substances (Torrado Pacheco et al. 2023). Finally, a recent study investigating the effects on RL parameters of the psychedelic LSD, a partial $5\text{-HT}_{2A}\text{R}$ agonist, has reported increased reward and punishment learning, and reduced stimulus stickiness (Kanen et al. 2022). Overall, these results suggest that $5\text{-HT}_{2A}\text{R}$ agonism can improve flexibility. In the present study, we show that antagonism of this receptor decreases the punishment learning rate and increases stickiness, mirroring these hypothetical effects of $5\text{-HT}_{2A}\text{R}$ agonism. A limitation of our study is the fact that only male animals were included; therefore, sex-dependent effects could not be investigated.

In summary, we report that both $5\text{-HT}_{2A}\text{R}$ and $5\text{-HT}_{2C}\text{R}$ antagonism altered performance on a visual reversal task. We characterized this impairment using RL models, finding that $5\text{-HT}_{2A}\text{R}$ blockade reduced both learning from punishment and reinforcement sensitivity, but increased stickiness. $5\text{-HT}_{2C}\text{R}$ blockade impaired learning from positive feedback as assessed using conventional measures, suggesting a dissociation between the two receptors: the $5\text{-HT}_{2C}\text{R}$ is essential for learning from positive feedback and the $5\text{-HT}_{2A}\text{R}$ is important for learning from negative feedback. Additionally, $5\text{-HT}_{2C}\text{R}$ antagonism reduced reinforcement sensitivity and side stickiness parameters, indicating increased flexibility. These results provide novel insights into the mechanisms of 5-HT and the involvement of different 5-HT receptors in cognitive flexibility. This may be important for our understanding of neuropsychiatric conditions such as MDD and OCD, as well as for research into future treatments such as psychedelic agents that act as $5\text{-HT}_{2A}\text{R}$ agonists.

Declarations

References

Aghajanian GK, Marek GJ (1999) Serotonin and hallucinogens. Neuropsychopharmacology 21. https://doi.org/10.1016/S0893-133X(98)00135-3

Alsiö J, Nilsson SRO, Gastambide F, Wang RAH, Dam SA, Mar AC, Tricklebank M, Robbins TW (2015) The role of 5-HT2C receptors in touchscreen visual reversal learning in the rat: a cross-site study. https://doi.org/10.1007/s00213-015-3963-5. Psychopharmacology

Alsiö J, Phillips BU, Sala-Bayo J, Nilsson SRO, Calafat-Pla TC, Rizwand A, Plumbridge JM, López-Cruz L, Dalley JW, Cardinal RN, Mar AC, Robbins TW (2019) Dopamine D2-like receptor stimulation blocks negative feedback in visual and spatial reversal learning in the rat: behavioural and computational evidence. Psychopharmacology 236(8):2307–2323. https://doi.org/10.1007/s00213-019-05296-y

Alsiö J, Lehmann O, Mckenzie C, Theobald DE, Searle L, Xia J, Dalley JW, Robbins TW (2021) Serotonergic innervations of the Orbitofrontal and Medial-prefrontal cortices are differentially involved in visual discrimination and reversal learning in rats, vol 31. Cerebral Cortex, p 1090. (New York, NY)2https://doi.org/10.1093/CERCOR/BHAA277

Alvarez BD, Morales CA, Amodeo DA (2021) Impact of specific serotonin receptor modulation on behavioral flexibility. In

Pharmacology Biochemistry and Behavior (Vol. 209). https://doi.org/10.1016/j.pbb.2021.173243

Amargós-Bosch M, Bortolozzi A, Puig MV, Serrats J, Adell A, Celada P, Toth M, Mengod G, Artigas F (2004) Co-expression and in vivo Interaction of Serotonin1A and Serotonin2A receptors in pyramidal neurons of pre-frontal cortex. Cereb Cortex 14(3). https://doi.org/10.1093/cercor/bhg128

Andersen KAA, Carhart-Harris R, Nutt DJ, Erritzoe D (2021) Therapeutic effects of classic serotonergic psychedelics: a systematic review of modern-era clinical studies. Acta Psychiatrica Scand (Vol 143(2). https://doi.org/10.1111/acps.13249

APA (2010) Practice guideline for the treatment of patients with major depressive disorder (third edition). American Psychiatric Association

Baker PM, Thompson JL, Sweeney JA, Ragozzino ME (2011) Differential effects of 5-HT2A and 5-HT2C receptor blockade on strategy-switching. Behav Brain Res 219(1):123. https://doi.org/10.1016/J.BBR.2010.12.031

Bari A, Theobald DE, Caprioli D, Mar AC, Aidoo-Micah A, Dalley JW, Robbins TW (2010) Serotonin modulates sensitivity to reward and negative feedback in a probabilistic reversal learning task in rats. Neuropsychopharmacology 35(6):1290–1301. https://doi.org/10.1038/npp.2009.233

Barlow RL, Alsiö J, Jupp B, Rabinovich R, Shrestha S, Roberts AC, Robbins TW, Dalley JW (2015) Markers of serotonergic function in the orbitofrontal cortex and dorsal raphé nucleus predict individual variation in spatial-discrimination serial reversal learning. Neuropsychopharmacology: Official Publication Am Coll Neuropsychopharmacol 40(7):1619–1630. https://doi.org/10.1038/NPP.2014.335

Berlin GS, Hollander E (2014) Compulsivity, impulsivity, and the DSM-5 process. CNS Spectr 19(1):62–68. https://doi.org/10.1017/S1092852913000722

Boulougouris V, Glennon JC, Robbins TW (2008) Dissociable effects of selective 5-HT2A and 5-HT2C receptor antagonists on serial spatial reversal learning in rats. Neuropsychopharmacology 33(8). https://doi.org/10.1038/sj.npp.1301584

Carhart-Harris R, Friston KJ (2019) REBUS and the anarchic brain: toward a Unified Model of the Brain Action of Psychedelics. Pharmacol Rev 71(3):316–344. https://doi.org/10.1124/PR.118.017160

Carhart-Harris R, Bolstridge M, Rucker J, Day CMJ, Erritzoe D, Kaelen M, Bloomfield M, Rickard JA, Forbes B, Feilding A, Taylor D, Pilling S, Curran VH, Nutt DJ (2016) Psilocybin with psychological support for treatment-resistant depression: an open-label feasibility study. Lancet Psychiatry 3(7):619–627. https://doi.org/10.1016/S2215-0366(16)30065-7

Carhart-Harris R, Giribaldi B, Watts R, Baker-Jones M, Murphy-Beiner A, Murphy R, Martell J, Blemings A, Erritzoe D, Nutt DJ (2021) Trial of Psilocybin versus Escitalopram for Depression. N Engl J Med 384(15):1402–1411. https://doi.org/10.1056/NEJMOA2032994/SUPPL_FILE/NEJMOA2032994_DATA-SHARING.PDF

Chamberlain SR, Fineberg NA, Blackwell AD, Robbins TW, Sahakian BJ (2006) Motor inhibition and cognitive flexibility in obsessive-compulsive disorder and trichotillomania. Am J Psychiatry 163(7). https://doi.org/10.1176/ajp.2006.163.7.1282

Clarke HF, Dalley JW, Crofts HS, Robbins TW, Roberts AC (2004) Cognitive inflexibility after Prefrontal Serotonin Depletion. Science 304(5672):878–880. https://doi.org/10.1126/science.1094987

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 Neuroscience: Official J Soc Neurosci 25(2):532–538. https://doi.org/10.1523/JNEUROSCI.3690-04.2005

Clevenger SS, Malhotra D, Dang J, Vanle B, IsHak WW (2018) The role of selective serotonin reuptake inhibitors in preventing relapse of major depressive disorder. Therapeutic Adv Psychopharmacol 8(1). https://doi.org/10.1177/2045125317737264

R Core Team (2021) R: A Language and Environment for Statistical Computing. In R Foundation for Statistical Computing

Daw ND (2009) Trial-by-trial data analysis using computational models

Den Ouden HEM, Daw ND, Fernandez G, Elshout JA, Rijpkema M, Hoogman M, Franke B, Cools R (2013) Dissociable effects of dopamine and serotonin on reversal learning. Neuron 80(4):1090–1100. https://doi.org/10.1016/J.NEURON.2013.08.030

Dias R, Robbins TW, Roberts AC (1996) Dissociation in prefrontal cortex of affective and attentional shifts. Nature 380(6569). https://doi.org/10.1038/380069a0

Doss MK, Považan M, Rosenberg MD, Sepeda ND, Davis AK, Finan PH, Smith GS, Pekar JJ, Barker PB, Griffiths RR, Barrett FS (2021) Psilocybin therapy increases cognitive and neural flexibility in patients with major depressive disorder. Translational Psychiatry 11(1). https://doi.org/10.1038/s41398-021-01706-y

Fineberg NA, Hollander E, Pallanti S, Walitza S, Grünblatt E, Dell'Osso BM, Albert U, Geller DA, Brakoulias V, Reddy J, Arumugham YC, Shavitt SS, Drummond RG, Grancini L, De Carlo B, Cinosi V, Chamberlain E, Ioannidis SR, Rodriguez K, Menchon CI (2020) J. M. Clinical advances in obsessive-compulsive disorder: A position statement by the International College of Obsessive-Compulsive Spectrum Disorders. In International Clinical Psychopharmacology. https://doi.org/10.1097/YIC.0000000000000314

Furr A, Danet Lapiz-Bluhm M, Morilak DA (2012) 5-HT2A receptors in the orbitofrontal cortex facilitate reversal learning and contribute to the beneficial cognitive effects of chronic citalopram treatment in rats. Int J Neuropsychopharmacol 15(9). https://doi.org/10.1017/S1461145711001441

Goldberg SB, Pace BT, Nicholas CR, Raison CL, Hutson PR (2020) The experimental effects of psilocybin on symptoms of anxiety and depression: a meta-analysis. Psychiatry Res 284. https://doi.org/10.1016/J.PSYCHRES.2020.112749

Hervig M, Piilgaard L, Božic T, Alsiö J, Robbins TW (2020) Glutamatergic and serotonergic modulation of rat medial and lateral orbitofrontal cortex in visual serial reversal learning. Psychol Neurosci 13(3):438. https://doi.org/10.1037/PNE0000221

Hervig M, Fiddian L, Piilgaard L, Bozič T, Blanco-Pozo M, Knudsen C, Olesen SF, Alsiö J, Robbins TW (2020a) Dissociable and paradoxical roles of rat medial and lateral Orbitofrontal cortex in Visual serial reversal learning. https://doi.org/10.1093/cercor/bhz144. Cerebral Cortex

Herzallah MM, Moustafa AA, Natsheh JY, Abdellatif SM, Taha MB, Tayem YI, Sehwail MA, Amleh I, Petrides G, Myers CE, Gluck MA (2013) Learning from negative feedback in patients with major depressive disorder is attenuated by SSRI antidepressants. Front Integr Nuerosci 7(SEP). https://doi.org/10.3389/fnint.2013.00067

Iigaya K, Fonseca MS, Murakami M, Mainen ZF, Dayan P (2018) An effect of serotonergic stimulation on learning rates for rewards apparent after long intertrial intervals. Nat Commun 2018 9:1(1):1–10. https://doi.org/10.1038/s41467-018-04840-2

Izquierdo A, Carlos K, Ostrander S, Rodriguez D, McCall-Craddolph A, Yagnik G, Zhou F (2012) Impaired reward learning and intact motivation after serotonin depletion in rats. Behav Brain Res 233(2):494–499. https://doi.org/10.1016/J.BBR.2012.05.032

Jentsch JD, Taylor JR (2001) Impaired inhibition of conditioned responses produced by Subchronic Administration of Phencyclidine to rats. Neuropsychopharmacol 2000 24(1):66–74. https://doi.org/10.1016/s0893-133x(00)00174-3

Jones B, Mishkin M (1972) Limbic lesions and the problem of stimulus-reinforcement associations. Exp Neurol 36(2). https://doi.org/10.1016/0014-4886(72)90030-1

Kanen JW, Ersche KD, Fineberg NA, Robbins TW, Cardinal RN (2019) 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 236(8):2337–2358. https://doi.org/10.1007/s00213-019-05325-w

Kanen JW, Apergis-Schoute AM, Yellowlees R, Arntz FE, van der Flier FE, Price A, Cardinal RN, Christmas DM, Clark L, Sahakian BJ, Crockett MJ, Robbins TW (2021) Serotonin depletion impairs both pavlovian and instrumental reversal learning in healthy humans. Mol Psychiatry 2021 26(12):7200–7210. https://doi.org/10.1038/s41380-021-01240-9

Kanen JW, Luo Q, Kandroodi MR, Cardinal RN, Robbins TW, Nutt DJ, Carhart-Harris RL, Ouden HE (2022) Effect of lysergic acid diethylamide (LSD) on reinforcement learning in humans. Psychol Med 1–12 M. den. https://doi.org/10.1017/S0033291722002963

Kehne JH, Baron BM, Carr AA, Chaney SF, Elands J, Feldman DJ, Frank RA, Van Giersbergen PLM, Mccloskey TC, Johnson MP, Mccarty DR, Poirot M, Senyah Y, Siegel BW, Widmaier C (1996) Preclinical characterization of the potential of the putative atypical antipsychotic MDL 100,907 as a potent 5-HT2A antagonist with a favorable CNS safety profile. Journal of Pharmacology and Experimental Therapeutics

Koob GF, Volkow ND (2016) Neurobiology of addiction: a neurocircuitry analysis. Lancet Psychiatry 3(8):760–773. https://doi.org/10.1016/S2215-0366(16)00104-8

Kuypers KPC, Riba J, de la Fuente Revenga M, Barker S, Theunissen EL, Ramaekers JG (2016) Ayahuasca enhances creative divergent thinking while decreasing conventional convergent thinking. https://doi.org/10.1007/s00213-016-4377-8. Psychopharmacology

Liu FY, Xing GG, Qu XX, Xu IS, Han JS, Wan Y (2007) Roles of 5-hydroxytryptamine (5-HT) receptor subtypes in the inhibitory effects of 5-HT on C-fiber responses of spinal wide dynamic range neurons in rats. J Pharmacol Exp Ther 321(3). https://doi.org/10.1124/jpet.106.115204

Luo Q, Kanen JW, Bari A, Skandali N, Langley C, Knudsen GM, Alsiö J, Phillips BU, Sahakian BJ, Cardinal RN, Robbins TW (2023) Comparable roles for serotonin in rats and humans for computations underlying flexible decision-making. Neuropsychopharmacol 2023 49(3):600–608. https://doi.org/10.1038/s41386-023-01762-6

Madsen MK, Fisher PM, Burmester D, Dyssegaard A, Stenbæk DS, Kristiansen S, Johansen SS, Lehel S, Linnet K, Svarer C, Erritzoe D, Ozenne B, Knudsen GM (2019) Psychedelic effects of psilocybin correlate with serotonin 2A receptor occupancy and plasma psilocin levels. Neuropsychopharmacol 2019 44(7):7. https://doi.org/10.1038/s41386-019-0324-0

Marek GJ, Martin-Ruiz R, Abo A, Artigas F (2005) The selective 5-HT2A receptor antagonist M100907 enhances antidepressant-like behavioral effects of the SSRI Fluoxetine. Neuropsychopharmacol 2005 30(12):2205–2215. https://doi.org/10.1038/sj.npp.1300762

Michely J, Eldar E, Erdman A, Martin IM, Dolan RJ (2022) Serotonin modulates asymmetric learning from reward and punishment in healthy human volunteers. Commun Biology 5(1). https://doi.org/10.1038/s42003-022-03690-5

Murphy-Beiner A, Soar K (2020) Ayahuasca's 'afterglow': improved mindfulness and cognitive flexibility in Ayahuasca drinkers. https://doi.org/10.1007/s00213-019-05445-3. Psychopharmacology

Perani D, Garibotto V, Gorini A, Moresco RM, Henin M, Panzacchi A, Matarrese M, Carpinelli A, Bellodi L, Fazio F (2008) In vivo PET study of 5HT(2A) serotonin and D(2) dopamine dysfunction in drug-naive obsessive-compulsive disorder. NeuroImage 42(1):306–314. https://doi.org/10.1016/J.NEUROIMAGE.2008.04.233

Phillips BU, Dewan S, Nilsson SRO, Robbins TW, Heath CJ, Saksida LM, Bussey TJ, Alsiö J (2018) Selective effects of 5-HT2C receptor modulation on performance of a novel valence-probe visual discrimination task and probabilistic reversal learning in mice. Psychopharmacology. https://doi.org/10.1007/s00213-018-4907-7

Pokorny T, Duerler P, Seifritz E, Vollenweider FX, Preller KH (2020) LSD acutely impairs working memory, executive functions, and cognitive flexibility, but not risk-based decision-making. Psychol Med 50(13):2255–2264. https://doi.org/10.1017/S0033291719002393

Roth BL (2011) Irving Page lecture: 5-HT(2A) serotonin receptor biology: interacting proteins, kinases and paradoxical regulation. Neuropharmacology 61(3):348–354. https://doi.org/10.1016/J.NEUROPHARM.2011.01.012

Rygula R, Clarke HF, Cardinal RN, Cockcroft GJ, Xia J, Dalley JW, Robbins TW, Roberts AC (2015) Role of central serotonin in anticipation of rewarding and punishing outcomes: effects of selective amygdala or orbitofrontal 5-HT depletion. Cereb Cortex 25(9):3064–3076. https://doi.org/10.1093/cercor/bhu102

Santana N, Bortolozzi A, Serrats J, Mengod G, Artigas F (2004) Expression of serotonin1A and serotonin2A receptors in pyramidal and GABAergic neurons of the rat prefrontal cortex. Cereb Cortex 14(10). https://doi.org/10.1093/cercor/bhh070

Seymour B, Daw ND, Roiser JP, Dayan P, Dolan R (2012) Serotonin selectively modulates reward value in human decision-making. J Neurosci 32(17). https://doi.org/10.1523/jneurosci.0053-12.2012

Stroud JB, Freeman TP, Leech R, Hindocha C, Lawn W, Nutt DJ, Curran HV, Carhart-Harris R (2018) Psilocybin with psychological support improves emotional face recognition in treatment-resistant depression. Psychopharmacology 235(2):459. https://doi.org/10.1007/S00213-017-4754-Y

Sutton RS, Barto AG (1998) Reinforcement learning: an introduction. IEEE Trans Neural Networks 9(5):1054–1054. https://doi.org/10.1109/tnn.1998.712192

Torrado Pacheco A, Olson RJ, Garza G, Moghaddam B (2023) Acute psilocybin enhances cognitive flexibility in rats. https://doi.org/10.1038/s41386-023-01545-z. Neuropsychopharmacology

Uddin LQ (2021) Cognitive and behavioural flexibility: neural mechanisms and clinical considerations. In Nature Reviews Neuroscience (Vol. 22, Issue 3, pp. 167–179). Nature Research. https://doi.org/10.1038/s41583-021-00428-w

Wickham H (2014) Tidy data. J Stat Softw 59(10):1–23. https://doi.org/10.18637/JSS.V059.I10

Zhu C, Kwok NT, kit, Chan TC, wan, Chan GH, kei, So SH (2021) wai. Inflexibility in Reasoning: Comparisons of Cognitive Flexibility, Explanatory Flexibility, and Belief Flexibility Between Schizophrenia and Major Depressive Disorder. Frontiers in Psychiatry, 11, 609569. https://doi.org/10.3389/FPSYT.2020.609569/BIBTEX

Zühlsdorff K, López-Cruz L, Dutcher EG, Jones JA, Pama C, Sawiak S, Khan S, Milton AL, Robbins TW, Bullmore ET, Dalley JW (2023) Sex-dependent effects of early life stress on reinforcement learning and limbic cortico-striatal functional connectivity. Neurobiol Stress 22:100507. https://doi.org/10.1016/J.YNSTR.2022.100507

Authors and Affiliations

Mona El- Sayed Hervig[1,2,3] · Katharina Zühlsdorff[1,2,4] · Sarah F. Olesen[5] · Benjamin Phillips[1] · Tadej Božič[5] · Jeffrey W. Dalley[1,2,6] · Rudolf N. Cardinal[2,6,7] · Johan Alsiö[1,2,8] · Trevor W. Robbins[1,2]

✉ Katharina Zühlsdorff
kz294@cam.ac.uk

[1] Department of Psychology, University of Cambridge, Cambridge CB2 3EB, UK

[2] Behavioural and Clinical Neuroscience Institute, University of Cambridge, Cambridge CB2 3EB, UK

[3] Department of Neuroscience, University of Copenhagen, Copenhagen DK-2200, Denmark

[4] The Alan Turing Institute, British Library, London NW1 2DVB, UK

[5] UCL Sainsbury Wellcome Centre for Neural Circuits and Behaviour, London W1T 4JG, UK

[6] Department of Psychiatry, Herchel Smith Building, Cambridge CB2 0SZ, UK

[7] Liaison Psychiatry Service, Cambridgeshire and Peterborough NHS Foundation Trust, Cambridge Biomedical Campus, Box 190, Cambridge CB2 0QQ, UK

[8] School of Physiology, Pharmacology & Neuroscience, University of Bristol, Bristol BS8 1TD, UK