

Compulsion is associated with impaired goal-directed and habitual learning and responding in obsessive-compulsive disorder

Quanhao Yu^{a,b,c,d,1}, Feng Gao^{a,b,c,1}, Chuting Li^{a,b,c},
Jie Xia^{a,b,c}, Yanyuan Cao^{a,b,c}, Xiang Wang^{a,b,c}, Chuman Xiao^{a,b,c},
Jingjie Lu^{a,b,c}, Qian Liu^{a,b,c}, Jie Fan^{a,b,c,*}, Xiongzhao Zhu^{a,b,c,*}

^a Medical Psychological Center, the Second Xiangya Hospital, Central South University, Changsha, Hunan, China

^b Medical Psychological Institute of Central South University, Changsha, Hunan, China

^c National Clinical Research Center for Mental Disorders, Changsha, Hunan, China

^d Shanghai Key Laboratory of Psychotic Disorders, Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, Shanghai, China

ARTICLE INFO

Keywords:

Obsessive-compulsive disorder
Goal-directed behavior
Habitual behavior
Reinforcement learning modeling

ABSTRACT

Background: Previous research has found that compulsions in obsessive-compulsive disorder (OCD) are associated with an imbalance between goal-directed and habitual responses. However, the cognitive mechanisms underlying how goal-directed and habitual behaviors are learned, and how these learning deficits affect the response process, remain unclear. The present study aimed to investigate these cognitive mechanisms and examine how they were involved in the mechanism of compulsions.

Methods: A total of 49 patients with OCD and 38 healthy controls (HCs) were recruited to perform the revised “slip of action test”. A reinforcement learning model was constructed, and model parameters including learning rates, reinforcement sensitivity, and perseveration were estimated using a hierarchical Bayesian approach. Comparisons of these parameters were made between the OCD group and HCs, and the associations with performance during the outcome devalued stage and clinical presentations were assessed.

Results: In the outcome devalued stage, patients with OCD exhibited greater responsiveness to the devalued outcome, indicating their impairment in flexible and goal-directed behavioral control. Computational modeling further revealed that, during the instrumental learning stage, patients with OCD showed reduced learning rates, decreased perseveration, and heightened reinforcement sensitivity as compared with HCs. The learning rate and perseveration during instrumental learning were significantly correlated with the performance in the outcome devalued stage and compulsive scores in OCD.

Conclusions: The results indicate that patients with OCD exhibit deficits in updating the associative strength based on prediction errors and are more likely to doubt established correct associations during goal-directed and habitual learning. These deficits may contribute to the inflexible goal-directed behavioral control and are involved in the mechanism of compulsion in OCD.

Introduction

Obsessive-compulsive disorder (OCD) is a common and chronic psychiatric disorder affecting 2%-3% of the population (Stein et al., 2019). Compulsion, manifested as repetitive, ritualistic behaviors or mental acts, stands as one of the core symptoms of OCD and significantly impair the social functioning of individuals with OCD (Amerio, Tonna, Odone, Stubbs, & Ghaemi, 2016; Robbins, Vaghi, & Banca, 2019). To

date, the neuropsychological mechanism of compulsion remains unclear, posing a substantial obstacle to the prevention and treatment of OCD.

Recent studies suggest that compulsions in OCD may be related to the imbalanced execution between the goal-directed and habitual actions (Gillan, Robbins, Sahakian, van den Heuvel, & van Wingen, 2016). Goal-directed behaviors are actions that are executed to achieve or avoid specific outcomes. These behaviors are highly adaptive but require more

* Corresponding authors at: Medical Psychological Center, the Second Xiangya Hospital, Central South University, Changsha, Hunan 410011, China.

E-mail addresses: fanjie1025@csu.edu.cn (J. Fan), xiongzhaozhu@csu.edu.cn (X. Zhu).

¹ Quanhao Yu and Feng Gao authors contributed equally to this study.

cognitive resources in a novel environments (Worbe, Savulich, de Wit, Fernandez-Egea, & Robbins, 2015). With repetition, such as daily routines, goal-directed actions can become habitual. Habitual behaviors are controlled by external stimulus-response associations and can be triggered automatically by certain stimulus (de Wit, Corlett, Aitken, Dickinson, & Fletcher, 2009). Habitual behaviors help individuals simplify the complex world, but can also lead to “slips of action” toward outcomes that are currently devalued (Gillan, 2021; Poldrack et al., 2005). The ability to flexibly shift between goal-directed and habitual responding is crucial for normal functioning in everyday life. Yet, this flexibility is compromised in patients with OCD and it is thought to be involved in the mechanism of compulsions.

The initial exploration of the imbalanced execution between goal-directed and habitual behavior in OCD was conducted using the outcome devaluation paradigm. This revealed that, compared to healthy controls (HCs), patients with OCD couldn't refrain from responding to the devalued stimuli (Gillan et al., 2011). Impairments in goal-directed behavior and an over-reliance on habitual behavior in OCD have also been observed in other studies using the contingency degradation paradigm and the two-step task and in the context of avoidance (Gillan et al., 2014).

While previous studies have indeed illuminated the impaired execution between goal-directed and excessive habitual behavioral control in patients with OCD, these investigations have primarily focused on behavioral responses during the outcome devaluation stage, neglecting the crucial process of instrumental learning of goal-directed and habitual behavior. Goal-directed and habitual learning are highly related and interdependent processes (de Wit & Dickinson, 2009; Dickinson, 1985; Valentin, Dickinson, & O'Doherty, 2007). During the initial stages of an individual's adaptation to a novel environment, goal-directed learning may predominate as individuals establish various stimulus-response-outcome associations and strive to optimize their actions (de Wit & Dickinson, 2009). As these associations become consistently reinforced through learning, the individual's behavioral system gradually shifts towards habitual behavior until a new outcome is required. However, the nature of learning processing in OCD, its effects on subsequently goal-directed and habitual responding process, and its association with compulsions remains unclear. This gap in understanding hinders our comprehension of OCD pathology.

The present study aimed to investigate the cognitive mechanisms underlying instrumental learning of goal-directed and habitual behaviors in patients with OCD and to examine how these mechanisms are involved in behavior execution associated with compulsion. To achieve this, we adopted a revised version of the “slip of action test”, which includes both the instrumental learning stage and the slip of action stage (Watson, van Wingen, & de Wit, 2018). A reinforcement learning model was applied to analyze the learning processes in both patients with OCD and healthy controls. The hypotheses were that 1) patients with OCD would exhibit deficits in the process of goal-directed and habitual learning, as reflected by the parameters of the reinforcement learning model; 2) the performance in the instrument learning phase would be strongly associated with behavior in devaluation phase in OCD, and would be specifically linked to the compulsion rather than obsession.

Method

Participants

A total of 49 patients with OCD and 38 HCs participated in the present study. Patients with OCD were recruited from the psychology clinic at the Second Xiangya Hospital of Central South University, Changsha, Hunan, China. Inclusion criteria for OCD patients include a DSM-5 diagnosis of OCD, confirmed by two experienced psychiatrists, an age range of 16 and 45 years, and right-handedness. Exclusion criteria included primary diagnoses of other DSM-5 disorders such as schizophrenia, bipolar disorder, or anxiety disorders, as well as a

documented history of any significant medical or neurological conditions. Of the 49 OCD patients, 20 were unmedicated. Among the 29 medicated patients, 28 were taking selective serotonin reuptake inhibitors (SSRIs), including sertraline, escitalopram oxalate, paroxetine, fluvoxamine, and fluoxetine. One patient was taking flupentixol-melitracen tablets. The HCs were recruited from local communities and universities. Inclusion criteria for HCs included no history of meeting the diagnostic criteria for any psychiatric or mood disorder under DSM-5, an age range of 16 and 45 years, and right-handedness. All participants provided written informed consent before completing the measures. The study was approved by the Ethics Committee of the Second Xiangya Hospital of Central South University.

Questionnaires

The Yale-Brown Obsessive-Compulsive Scale (Y-BOCS) was used to assess the severity of obsessive and compulsive (OC) symptoms in patients (Goodman et al., 1989). All participants completed the Obsessive-compulsive Inventory-Revised (OCI-R), the Beck Depression Inventory (BDI), and the State-Trait Anxiety Inventory (STAI) to evaluate the OC symptoms, depression, and anxiety level. Verbal intelligence (IQ) was measured by the Chinese version of the Wechsler Intelligence Test III (Foa et al., 2002; Spielberger, Gonzalez-Reigosa, Martinez-Urrutia, Natalicio, & Natalicio, 1971).

Slip of action test

We used a revised version of the “slip of action test” (Harold & Sellers, 2018), programmed in E-prime 2.0, to evaluate performance in the learning process and dual-system execution (see Fig. 1). This test comprised two stages: the instrumental learning stage (Fig. 1a) and the outcome devalued stage (Fig. 1b). During the instrumental learning stage, participants were instructed to establish stimulus-response-outcome associations through trials and errors based on feedback. Each trial began with a black fixation “+” presented at the center of a white screen for a random duration of 2–4 seconds. Subsequently, a closed box labeled with a fruit (stimulus fruit) appeared in the center of the screen, and participants were instructed to respond by pressing either the right key or left key within a 2-second response window. One of the keys would trigger the appearance of another fruit (outcome fruit) inside the box and award a point. Faster and correct responses would result in more points (ranging from 1 to 5), while an incorrect key press would lead to an empty box with no points awarded. If no response was recorded within the time limit, “Too late” appeared on screen. Feedback was displayed for 1 second. Participants were instructed that their goal was to earn as many points as possible and remember the associations between stimulus fruits, responses, and outcome fruits. The stage consisted of 12 blocks with 12 trials in each, for a total 144 trials. There were 12 fruit images, forming six stimulus-response-outcome associations in total.

The outcome devaluation stage allowed for a direct assessment of relative habitual and goal-directed behavior execution. At the beginning of each trial of a block, all 6 outcome fruits were displayed on the screen for 5 seconds, with two of them devalued (indicated by two red crosses). Participants were instructed to continue responding to the still-valuable outcomes and to stop responding to the devalued outcomes, as doing so would lead to the subtraction of points. The outcome devaluation stage began only after participants had correctly completed the recollections test, where they were asked to identify the devalued outcome fruits. Following this, a fixation cross was presented for a random duration of 2.5–4.5 seconds. The stimulus fruit would appear for 1.5 seconds, during which participants had to decide whether to respond with the correct key or refrain from responding. Feedback was provided at the end of each block. In the devalued phase, nine blocks included all possible combinations of right and left responses paired with outcomes that were devalued. Each block consisted of 24 trials, where the six stimulus fruits

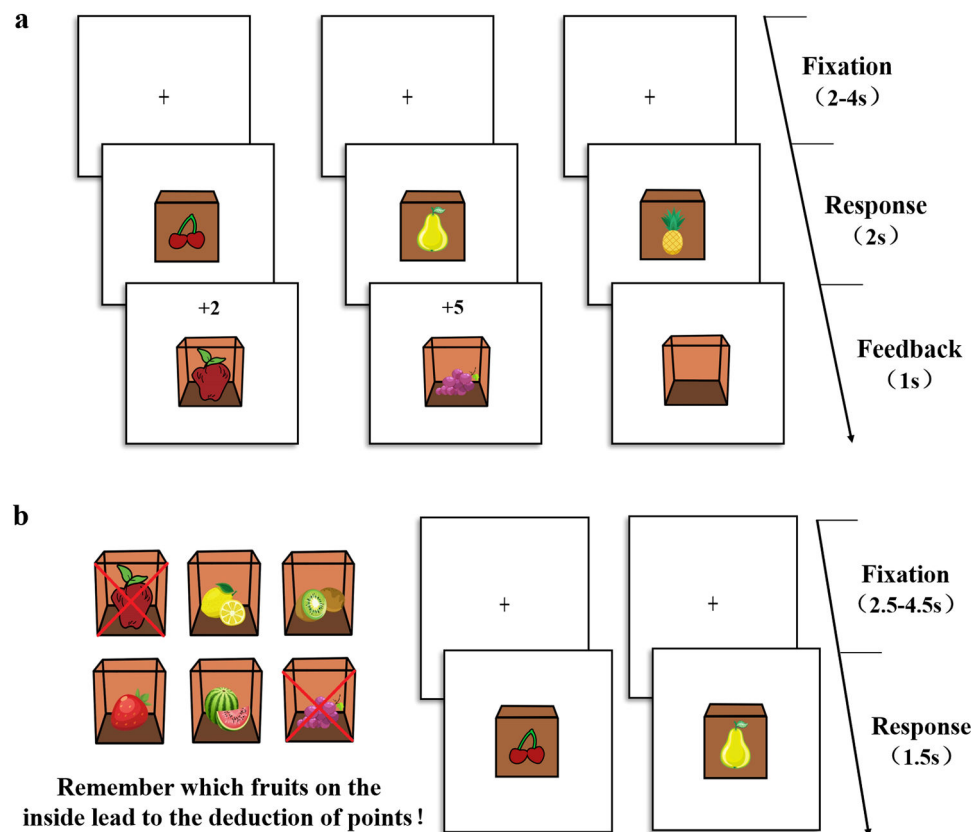


Fig. 1. The process of “slip of action test”. a, During the instrumental training stage, participants saw fruits on the outside of the box first, and then had to learn whether press left or right key to collect the fruit inside the box to establish and reinforce the stimulus-action-outcome association. Faster and correct responses led to higher point gains. In this example, the cherry stimulus is paired with the apple outcome, and the pear stimulus paired with the grape. b, During the outcome devaluation stage, at the beginning of each block of trials, two certain fruit outcomes were devalued (indicated by two red crosses), signifying that participants should no longer respond to those outcomes, as doing so would lead to subtraction of points. Participants were instructed to respond to the stimulus when the associated fruit outcome was not devalued and withhold respond to the devalued fruit (e.g., the apple is devalued so no response should be made to the cherry).

were shown four times in random order. Additionally, there were three blocks, which did not contain devalued outcome fruit, and participants were instructed to press the correct key to earn points like the instrumental learning stage. Each filler blocks consisted of 12 trials, with each of the six stimuli fruits was shown twice in random order.

Before the instrument learning stage, participants completed a short test phase with four associations consisting of eight pictures of different drinks to ensure that the participants fully understood the rules of the task. After completing the instrumental learning stage, participants also completed paper-and-pencil questionnaires of contingency knowledge to evaluate whether they remembered the associations.

Statistical analysis and computational modeling

Comparison of demographic and clinical features

Two-sample *t*-tests and Chi-square tests were used to evaluate the demographic and clinical differences between patients with OCD and HCs.

Analyses of standard outcome measures for the slip of action task

The standard outcome measures for the slip of action task were accuracy rate (ACC) and reaction time (RT) in the instrument learning stage, as well as response rates (%) on valued and devalued trials during the outcome devaluation stage. The devaluation sensitivity index (DSI), reflecting the percentages of responses on valuable outcomes minus devalued outcomes, was also calculated as an indication of the sensitivity of the outcome value in the devaluation stage. A higher DSI suggests a greater tendency toward goal-directed behavior responding. The

number of consecutive incorrect responses for each stimulus was also calculated.

A repeated measures ANCOVA was utilized to examine group differences in accuracy and response time during the instrument learning stage. Two-sample *t*-tests were used to evaluate the differences between the two groups in response rates (%) on valued and devalued trials during the outcome devaluation stage. The Mann-Whitney U test was employed to compare the number of consecutive incorrect responses in each stimulus for the first time between patients with OCD and HCs. Partial correlations were conducted to evaluate the relationship between task performance and clinical measures. For the comparison and correlation analyses, the covariates of verbal IQ and medication status were controlled. All analyses were conducted in SPSS version 22 and R version 4.0.2.

Computational modeling

This study employs the classical Q-Learning algorithm to model the process of action selection during the trial-by-trial instrumental learning stage (Schaaf, Jepma, Visser, & Huizenga, 2019). During each trial, there are two possible responses for each stimulus (press right key or left key) and participants assigns an expected value to each response: $V_t(R)$ and $V_t(L)$. These values are initialized to 0.5, and V_t is updated on each trial as following algorithm:

$$V_{t+1} = V_t + \alpha(R_t - V_t) \quad (1)$$

During each trial, the expected value of a specific response linked to a specific stimulus could also be viewed as the associative strength, which increases when a response is reinforced. Associative strength was

updated trial by trial based on prediction errors, which represent the discrepancy between the expected outcome V_t and actual outcome R_t . Larger prediction errors lead to greater changes in associative strength. Additionally, the learning rate parameter α regulates the impact of prediction errors on updating the associative strength. Higher learning rates (close to 1) indicate greater sensitivity to prediction errors and fast adaptation of associative strength, whereas lower learning rates (near 0) lead to slower adaptation.

The instrumental learning stage consists of six specific stimulus-response-outcome associations. Using only one learning rate parameter, α , to describe this task may overlook the learning differences between these distinct associations. To better capture these variations, this study optimizes the classical Q-learning model by establishes separate learning rates for each type of stimulus-response association. Thus, Six different $V_{type,t+1}$ were updated separately on each trial according to the following algorithm:

$$V_{type,t+1} = V_{type,t} + \alpha_{type}(R_t - V_{type,t}) \quad (2)$$

We defined a perseveration parameter τ to represent the tendency to repeat the previous response of the same stimulus. For individuals, the probability of making either response should be equal upon the first appearance of each stimulus type. Each type of stimulus-response association is independent of the others. Therefore, the response and outcome for a particular stimulus type will only influence subsequent occurrences of the same stimulus-response association. For trial t with specific stimulus and response k , we defined $C_{type,t}^k$ to be 1 if the subject chose response k on the previous the same stimulus trial, and 0 otherwise.

In sum, there were three parameters in this model: learning rate α , reinforcement sensitivity β and perseveration τ . The probability of a particular choice i for a specific type of stimulus on a given trial t followed a softmax rule, as described by the following equation.

$$p(i, type, t) = \frac{e^{\beta V_{type,t}^i + \tau C_{type,t}^i}}{\sum_{k=1}^n e^{\beta V_{type,t}^k + \tau C_{type,t}^k}} \quad (3)$$

The parameter β , referred to as reinforcement sensitivity or inverse temperature, reflects the degree of randomness or noise in the decision-making process (Gershman, 2016; Schaaf, Jepma, Visser, & Huizenga, 2019). Lower β values indicate greater randomness in choices and reduced sensitivity to expected reward values, while higher β values reflect a stronger tendency to choose stimuli with higher expected rewards. The parameter τ determined the degree of the tendency to perseverate the same choice to a certain type of stimulus.

Parameter estimation and statistical analyses

The model was estimated using a hierarchical Bayesian framework implemented in RStan (version 2.21.2), which employs Hamiltonian Markov Chain Monte Carlo sampling. Priors for the means of the group-level hyperparameters were assigned separately. For learning rates (α), we provided a prior beta (1.1, 1.1) distribution with range [0, 1]. The prior for reinforcement sensitivity (β) was a prior gamma (4.82, 0.88) distribution (Gershman, 2016). For perseveration (τ), we used a prior normal (0, 1) distribution (Gershman, 2016). Each subject-specific parameter was drawn from the distribution of its group-level parameter (Lim et al., 2019). The standard deviation of α and τ was given a prior half-normal (0, 0.17) distribution, and the standard deviation of β was drawn from a prior half-normal (0, 2) distribution.

We used 4 independent Markov chain Monte Carlo (MCMC) chains, each with 2000 burn-in samples. The convergence of Markov chains is assessed using the potential scale reduction factor $\hat{R}$, with values less than 1.1 indicating sufficient convergence (Brooks & Gelman, 1998; Gelman, Carlin, Stern, & Rubin, 1995). To examine the differences in the three parameters: learning rate (α), reinforcement sensitivity (β) and perseveration (τ) between patients with OCD and HCs, we calculated the

posterior distribution of the group difference. The 95% HDI of the group difference that did not overlap with zero indicated credible group differences (Kruschke, 2011). We also explored the correlations between the model parameters and the standard measures in the slip of action task as well as the clinical characteristics of OCD with verbal IQ and medication status controlled as covariates.

Results

Demographic and clinical characteristics

As shown in Table 1, the two groups did not differ in terms of age and gender distribution. Verbal intelligence of patients with OCD was lower than that of HCs. Moreover, as expected, patients with OCD exhibited a higher level of OC, depressive, and anxiety symptoms than HCs (all $p < 0.05$).

Standard outcome measures during the task

Instrumental learning stage

The standard results of the instrumental learning stage are presented in Figs. 2a-c. In both groups, ACC during the instrumental learning stage gradually increased across blocks and then leveled off (Fig. 2a), while RT gradually decreased across blocks and eventually stabilized as well (Fig. 2b).

The 12 blocks were divided into three equally phases: the beginning, middle, and end. A two-way repeated measure ANCOVA was conducted to examine the differences in ACC and RT between OCD and HCs. For ACC, results revealed a significant main effect of phase ($F_{(2,81)} = 0.185$, $p < 0.001$). The main effect of group ($F_{(1,82)} = 0.185$, $p > 0.05$) and the interaction effect between group and phase ($F_{(2,81)} = 1.188$, $p > 0.05$) were not significant. For RT, the main effect of group, the main effect of phase, and the interaction effect between group and phase were all not significant ($F_{(1,82)} = 0.185$, $p > 0.05$; $F_{(2,81)} = 2.475$, $p > 0.05$; $F_{(2,81)} = 0.113$, $p > 0.05$). As shown in Fig. 2c, after the first response error for a specific type of stimulus, there were no significant differences in the number of consecutive incorrect responses for each stimulus type between patients with OCD and HCs.

Outcome devaluation stage

The results of the outcome devaluation stage are depicted in Fig. 2d. Three HCs who did not complete the outcome devaluation stage were excluded from comparison. Compared to HCs, the OCD group demonstrated a significantly lower response rate to valuable outcomes ($t = -2.82$, $p < 0.01$), a significantly higher response rate to devalued

Table 1

Demographic and clinical characteristics of patients with OCD and HCs.

	OCD (N=49)	HC (N=38)	t/χ^2	p	ϕ /cohen's d
Age (year)	22.88±6.14	21.39±2.42	1.543	0.128	0.32
Gender (female, %)	25 (51.02)	25 (65.8)	1.910	0.167	0.15
Verbal intelligence	108.86±9.77	115.55±6.65	-3.795	<0.001	-0.80
Y-BOCS	20.33±6.50	–			
Y-BOCS obsession	10.35±3.90	–			
Y-BOCS compulsion	9.98±4.05	–			
OCI-R	31.02±15.28	17.26±9.94	5.000	<0.001	1.07
BDI	21.44±12.89	6.08±7.14	7.009	<0.001	1.47
STAI-S	52.57±12.41	36.42±9.05	6.932	<0.001	1.49

Note: Y-BOCS, Yale-Brown Obsessive-Compulsive Scale; OCI-R, Obsessive-compulsive Inventory Revised; BDI, Beck Depression Inventory; STAI-S, State-Trait Anxiety Inventory-State Form.

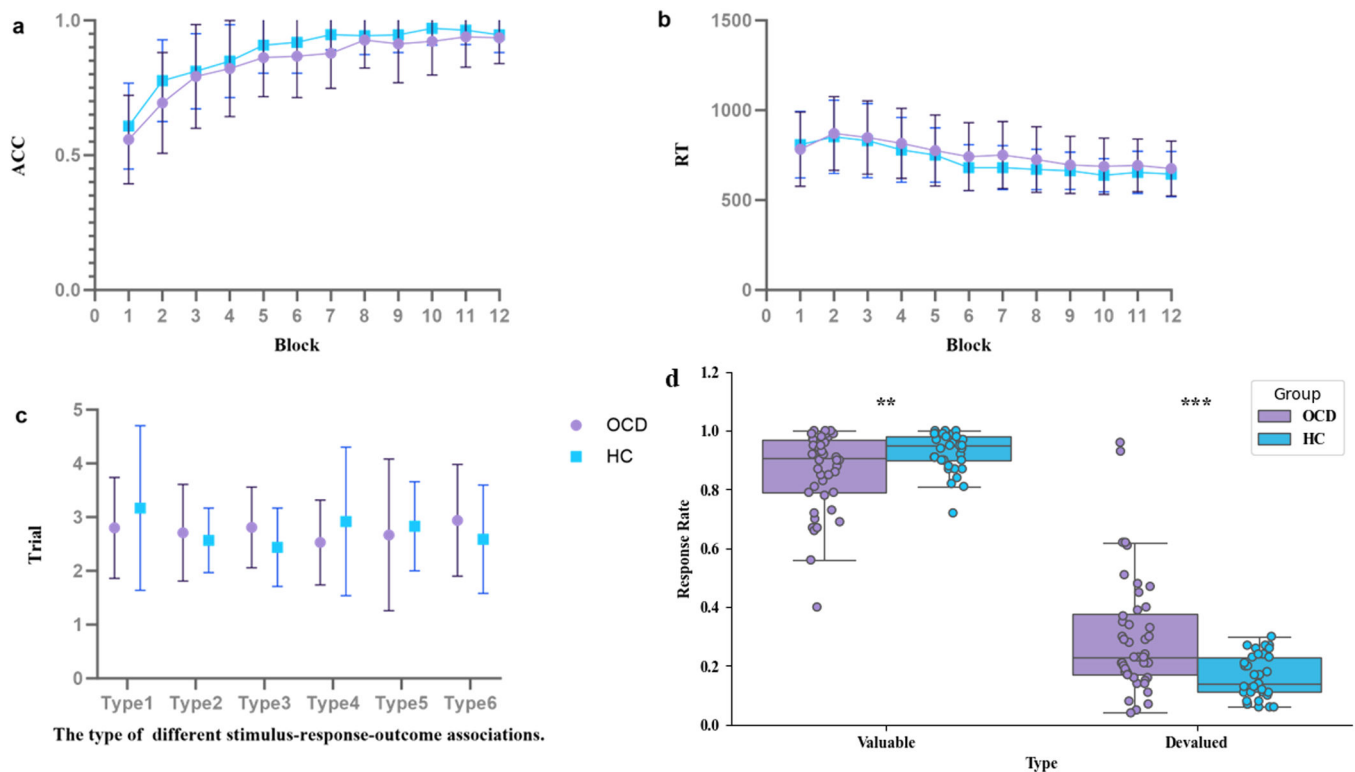


Fig. 2. Results from analyses of standard outcome measures. a and b, Accuracy and reaction times in each block of instrumental learning. Both groups had learned the correct response to make (press left or right key) in the presence of each stimulus through trial and error. c, There were no significance differences in the number of continuous incorrect responses in each type of stimulus between patients with OCD and HCs. d, The results of outcome devalued stage. Patients with OCD showed significantly lower response rate to valuable outcomes, a significantly higher response rate to devalued outcomes. OCD, obsessive-compulsive disorder; HCs, healthy controls; **, $p < 0.01$; ***, $p < 0.001$.

outcomes ($t = 3.85$, $p < 0.001$), and consequently, a significant decrease in the DSI ($t = -3.93$, $p < 0.001$).

Computation modeling during instrumental learning

Gelman-Rubin testing demonstrated that the model achieved good convergence, as all $\hat{R}$ values were less than 1.1. The learning rates for each association in OCD and HCs was depicted in Fig S1 a, b (see Supplementary materials). Results from computational modeling parameters showed that patients with OCD had significantly higher learning rates for one type of association (95% HDI [0.0135, 0.0255]), and significantly lower learning rates for four types of associations (95% HDI [-0.0145, -0.0417]; [-0.0719, -0.1886]; [-0.0012, -0.0150]; [-0.0093, -0.0188]) compared to HCs (see Fig. 3a, 3b, 3c, 3e, 3f). The learning rate for one type of association did not differ between groups (Fig. 3d). Moreover, compared to HCs, patients with OCD exhibited higher reinforcement sensitivity (95% HDI [0.45, 1.01], Fig. 3g) and lower perseveration (95% HDI [-0.22, -0.25], Fig. 3h), particularly in the first half of the instrumental learning phase (95% HDI [-0.23, -0.26], Fig S1 c, d).

Associations between goal-directed and habitual learning and responding

For the partial correlation analyses, six patients with OCD who did not memorize the association's contingency and three healthy controls who did not complete the outcome devaluation stage were excluded. Results were presented in Fig. 4.

Specifically, in patients with OCD, the learning rate was found to be significantly positively correlated with the rate of response to valuable outcomes (mean overall learning rate: $r = 0.372$, $p < 0.05$, Fig. 4a; mean 4 lower learning rate: $r = 0.411$, $p < 0.05$). Additionally, significant

negative correlations were observed between reinforcement sensitivity and the rate of response to devalued outcomes (OCD: $r = -0.354$, $p < 0.05$, Fig. 4b; HCs: $r = -0.401$, $p < 0.05$), and significant positive correlations were found between reinforcement sensitivity and the DSI (OCD: $r = 0.409$, $p < 0.01$, Fig. 4c; HCs: $r = 0.385$, $p < 0.05$) in both groups. Perseveration was significantly negatively correlated with the rate of response to devalued outcomes in OCD patients ($r = -0.348$, $p < 0.05$, Fig. 4d). Correlation results for HCs were shown in Fig S2 (see Supplementary materials).

Correlation between behavioral performance and clinical presentations

The partial correlation analyses revealed that patients with OCD who had higher Y-BOCS compulsive scores showed lower learning rates (mean overall learning rate: $r = -0.372$, $p < 0.05$; mean 4 lower learning rate: $r = -0.426$, $p < 0.05$, Fig. 4e, Fig S3) and lower perseveration parameters ($r = -0.376$, $p < 0.05$, Fig. 4f). No other significant correlations were observed.

Discussion

The present study utilized computational modeling to explore the cognitive mechanisms underlying goal-directed and habitual learning in patients with OCD, with a particular focus on how these mechanisms influence goal-directed and habitual responses and compulsive behaviors. The results showed that OCD patients exhibited stronger responses to devalued outcomes. The reinforcement learning model further revealed that OCD patients had a lower learning rate, higher reinforcement sensitivity, and reduced perseveration when learning new stimulus-response-outcome associations. These learning deficits were linked to impaired goal-directed and habitual behavior execution, which

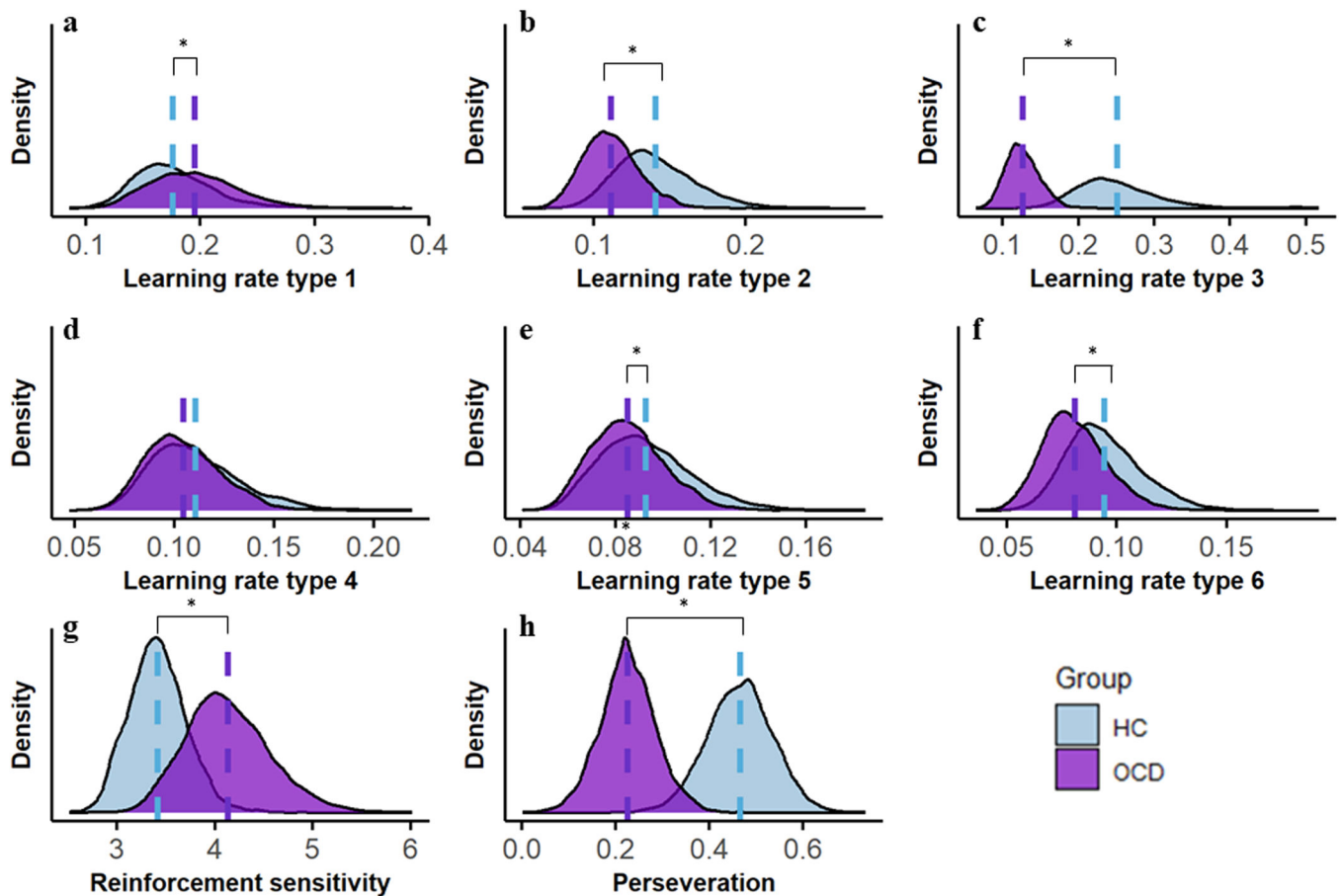


Fig. 3. Results from reinforcement learning model. a-f, group differences in the learning rate of 6 types of associations. g, OCD patients had higher reinforcement sensitivity as compared with HCs in the instrumental learning. h, OCD patients had reduced perseveration as compared with HCs. OCD, obsessive-compulsive disorder; HCs, healthy controls; *, The difference between the two groups was significant.

was associated with the severity of their compulsions. These findings support our hypothesis that OCD patients exhibit abnormalities in behavior learning processes, which strongly influence subsequent dual-system behavior execution. Notably, these abnormalities were specifically linked to compulsions rather than obsessions, providing further insight into the psychological mechanisms underlying compulsive behavior in OCD.

At the outcome devaluation stage, our study found that patients with OCD exhibited greater responses to devalued outcomes and had higher DSI compared with HCs. These findings are consistent with previous research and indicate that after goal-directed and habitual behavior learning, patients with OCD were less sensitive to the outcome value change, and showed over-reliance on habitual response in the behavior execution stage (Gillan et al., 2011). This response pattern aligns with the hallmark features of compulsions, which are characterized by insensitivity to outcomes, even when these outcomes become excessive or seemingly irrational (Gillan, 2021).

While much research has focused on the imbalanced dual-system behavior execution, few studies have explored the learning process in OCD, overlooking the relationship between dual-system behavior learning and subsequent execution. In the present study, we combined the standard indices and computational modeling parameters to reveal the specific instrumental learning process of OCD. Results from standard indices showed no significant differences in ACC and RT during the instrumental learning stage between the two groups. Additionally, patients with OCD did not exhibit impairment in promptly adjusting their behavior in response to negative feedback, as evidenced by the number of continuous incorrect responses in each stimulus. These findings

suggested that both groups may have the capacity to learn the stimulus-response-outcome associations through trial and error. However, these explicit behavioral indices only demonstrate that both groups successfully learned these associations, without capturing how individuals learn during the instrumental learning phase.

Further insights from computational modeling revealed that OCD patients exhibited a lower learning rate, higher reinforcement sensitivity, and reduced perseveration during behavior learning. Among the six types of associations, OCD patients demonstrated a lower learning rate for four types, but a higher learning rate for one (Type 1 associations). The relatively higher learning rate for Type 1 may be attributed to physical similarities between stimuli, such as color or shape. For instance, Type 1, which pairs a red apple with a red cherry, is likely easier to learn due to these shared visual characteristics. OCD patients had a higher learning rate for this association, while showing lower learning rates for the others. This may indicate that, unlike HCs, whose learning rates display a more gradual, gradient-like pattern across multiple associations, suggesting a balanced approach to learning, OCD patients may exhibit a disproportionate focus on certain associations. This could reflect an overemphasis on specific connections, such as the easier ones, at the expense of learning others (details illustrated in Fig S1). Overall, despite association 1, OCD patients tended to have lower learning rate than HCs. The lower learning rates may suggest that OCD patients struggle with updating associative strength based on prediction errors when learning multiple associations. This finding aligns with previous studies, which have shown that patients with OCD exhibit reduced learning efficiency, potentially influenced by beta-gamma activity in the medial orbitofrontal cortex (Grover, Nguyen, Viswanathan,

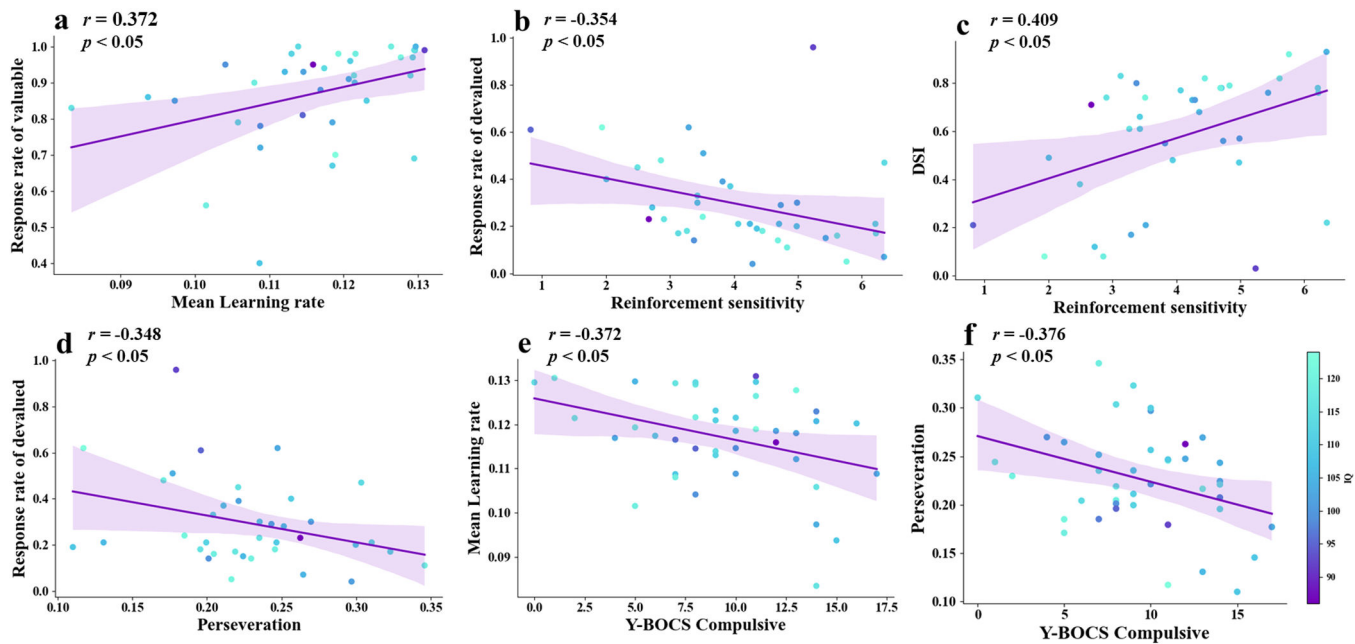


Fig. 4. The relationship between parameters in instrumental learning, performance in the outcome stage, and compulsive severity in patients with OCD, was analyzed with verbal IQ and medication status as covariates. a, the mean learning rate exhibited a positive correlation with the rate of response to valuable outcomes. b, the reinforcement sensitivity parameter exhibited a significant negative correlation with the rate of response to devalued outcomes. c, reinforcement sensitivity parameter exhibited a significant positive correlation with the index of devaluation sensitivity (DSI). d, the perseveration showed a significant negative correlation with the rate of response to devalued outcomes. e, perseveration showed significant negative correlations between Y-BOCS compulsive scores. f, mean of all type of learning rate showed significant negative correlations between Y-BOCS compulsive scores. OCD, obsessive-compulsive disorder; HCs, healthy controls.

& Reinhart, 2021; Hiebert et al., 2020). Patients with OCD also presented a lower tendency to persist with the same responses within the same type of association. Considering that both groups achieved an ACC of over 90% at the end of the instrumental learning stage, it may be inferred that perseveration is approximately equivalent to the tendency to persist in correct responses within the same type of association. Also our results show that the difference in perseveration primarily be attributed to the performance at the beginning of the instrumental learning stage. Thus, lower perseveration are less likely to maintain "stickiness" to correct responses and are more sensitive to negative feedback, meaning they tend to doubt previously established associations. This reduced "stickiness" (i.e., increased switching between responses) when learning optimal behaviors has also been observed in other studies (Fradkin, Ludwig, Eldar, & Huppert, 2020; Kanen, Ersche, Fineberg, Robbins, & Cardinal, 2019; Ruan et al., 2023), collectively suggesting that OCD patients may place less weight on prior experiences and engage more in over-exploratory behaviors. Additionally, patients with OCD showed heightened sensitivity to negative feedback, aligning with the clinical characteristics of compulsions. OCD patients often exhibit excessive concern with avoiding negative consequences rather than seeking rewards, and their compulsions frequently involve repetitive behaviors that rigidly adhere to rules, initially aimed at alleviating or preventing unpleasant or undesirable outcomes (Chamberlain & Menzies, 2009; Voon et al., 2015). In sum, the analyses of standard outcome measures did not show group differences in the learning phase. The alterations in OCD patients were only apparent when exploring parameters derived from the computational modeling analysis. The findings from reinforcement learning modeling, which support our hypothesis 1, revealed the cognitive mechanisms in the goal-directed and habitual learning in OCD, and provide a more expressive understanding of instrumental learning deficits in OCD.

Further correlation analysis in the present study revealed that the learning features in OCD was significantly correlated with their subsequent performance in dual-system behavior execution. Patients with OCD exhibited a higher sensitivity to expected rewards in behavior

learning process but respond more to devalued outcomes. This suggests that OCD patients may overly rely on expected value during behavior learning, leading to the formation of excessively strong stimulus-response associations. Such over formed habitual behaviors makes it challenging for them to adjust to new goals or adapt to changing environments. Also, in the present study, we found that the performance in the instrumental learning stage was significantly correlated with the severity of compulsions rather than the obsessions. These findings supported our hypothesis 2, clarifying the potential connection between the goal-directed learning and the subsequent impaired goal-directed and habitual responding in patients with OCD. They also strengthen our insight into the mechanism of compulsions (Gillan, Kosinski, Whelan, Phelps, & Daw, 2016; Peng et al., 2022; Zainal et al., 2023). During the past decades, cognitive-behavioral theory has been the dominant framework for explaining the underlying psychological mechanism of OCD. According to this theory, patients with OCD often engage in compulsive behaviors as a way to avoid potential negative consequences or discomfort and to neutralize anxiety or distress stemming from particular painful obsessions (Salkovskis, 1985). However, recent studies have proposed alternative hypotheses to clarify the relationship between obsessions and compulsions in OCD. Some studies have proposed that compulsive behaviors may not always arise as a direct consequence of obsessions, suggesting that such behaviors (e.g., ritualizing) can manifest independently (Robbins, Gillan, Smith, de Wit, & Ersche, 2012). This observation highlighted that the possibility that compulsions may function as an independent factor, further suggesting that compulsions are phenomenologically distinct from obsessions. The present study specifically characterizes the cognitive underpinnings of goal-directed and habitual learning and responding as being associated with compulsions in patients with OCD, providing some support for these emerging theories.

The current study has several limitations. First, the results of the slip of action test cannot definitively determine whether the bias toward habits is due to excessive reliance on habits, weak goal-directed control, or a combination of both. Further research could use different paradigms

or neuroimaging methods to clarify this hypothesis. Second, more than half of the patients with OCD were on pharmacotherapy, mainly SSRIs, which may have reduced the observed differences between groups (Voon et al., 2020).

Conclusion

Patients with OCD exhibit deficiency in goal-directed and habitual learning, characterized by an inability to update the association strength in response to prediction errors, and are more likely to doubt the correct associations that have been established. These instrumental learning deficits influenced the subsequently impaired goal-directed and habitual control associated with compulsion in OCD. These findings provide important insight into the pathophysiology of OCD.

Ethics approval

The authors affirm that all procedures contributing to this work adhere to the ethical standards set by the Institutional Ethics Board of the Second Hospital of Xiangya, Central South University.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This work was supported by the National Natural Science Foundation of China [grant number 82201673], and Hunan Provincial Innovation Foundation for Postgraduate [grant number CX20220356].

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.ijchp.2024.100531](https://doi.org/10.1016/j.ijchp.2024.100531).

References

- Amerio, A., Tonna, M., Odone, A., Stubbs, B., & Ghaemi, S. (2016). Psychiatric comorbidities in comorbid bipolar disorder and obsessive-compulsive disorder patients. *Asian Journal of Psychiatry*, 21, 23–24. <https://doi.org/10.1016/j.ajp.2016.02.009>
- Brooks, S. P., & Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7(4), 434–455. <https://doi.org/10.1080/10618600.1998.10474787>
- Chamberlain, S. R., & Menzies, L. (2009). Endophenotypes of obsessive-compulsive disorder: rationale, evidence and future potential. *Expert review of neurotherapeutics*, 9(8), 1133–1146. <https://doi.org/10.1586/ern.09.36>
- de Wit, S., Corlett, P. R., Aitken, M. R., Dickinson, A., & Fletcher, P. C. (2009). Differential engagement of the ventromedial prefrontal cortex by goal-directed and habitual behavior toward food pictures in humans. *Journal of Neuroscience*, 29(36), 11330–11338. <https://doi.org/10.1523/JNEUROSCI.1639-09.2009>
- de Wit, S., & Dickinson, A. (2009). Associative theories of goal-directed behaviour: a case for animal-human translational models. *Psychological Research PRPF*, 73(4), 463–476. <https://doi.org/10.1007/s00426-009-0230-6>
- Dickinson, A. (1985). Actions and habits: the development of behavioural autonomy. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 308 (1135), 67–78. <https://doi.org/10.1098/rstb.1985.0010>
- Foa, E. B., Huppert, J. D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G., & Salkovskis, P. M. (2002). The Obsessive-Compulsive Inventory: development and validation of a short version. *Psychological Assessment*, 14(4), 485. <https://doi.org/10.1037/1040-3590.14.4.485>
- Fradkin, I., Ludwig, C., Eldar, E., & Huppert, J. D. (2020). Doubting what you already know: Uncertainty regarding state transitions is associated with obsessive compulsive symptoms. *PLoS computational biology*, 16(2), Article e1007634. <https://doi.org/10.1371/journal.pcbi.1007634>
- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (1995). *Bayesian data analysis*. Chapman and Hall/CRC.
- Gershman, S. J. (2016). Empirical priors for reinforcement learning models. *Journal of Mathematical Psychology*, 71, 1–6. <https://doi.org/10.1016/j.jmp.2016.01.006>
- Gillan, C. M. (2021). Recent developments in the habit hypothesis of OCD and compulsive disorders. *The neurobiology and treatment of OCD: accelerating progress* (pp. 147–167). https://doi.org/10.1007/7854_2020_199
- Gillan, C. M., Kosinski, M., Whelan, R., Phelps, E. A., & Daw, N. D. (2016). Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *Elife*, 5, e11305. <https://doi.org/10.7554/eLife.11305>
- Gillan, C. M., Morein-Zamir, S., Urcelay, G. P., Sule, A., Voon, V., Apergis-Schoute, A. M., & Robbins, T. W. (2014). Enhanced avoidance habits in obsessive-compulsive disorder. *Biological psychiatry*, 75(8), 631–638. <https://doi.org/10.1016/j.biopsych.2013.02.002>
- Gillan, C. M., Pappmeyer, M., Morein-Zamir, S., Sahakian, B. J., Fineberg, N. A., Robbins, T. W., & de Wit, S. (2011). Disruption in the balance between goal-directed behavior and habit learning in obsessive-compulsive disorder. *American Journal of Psychiatry*, 168(7), 718–726. <https://doi.org/10.1176/appi.ajp.2011.10071062>
- Gillan, C. M., Robbins, T. W., Sahakian, B. J., van den Heuvel, O. A., & van Wingen, G. (2016). The role of habit in compulsivity. *European Neuropsychopharmacology*, 26(5), 828–840. <https://doi.org/10.1016/j.euroneuro.2015.12.033>
- Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Fleischmann, R. L., Hill, C. L., & Charney, D. S. (1989). The Yale-Brown obsessive compulsive scale: I. Development, use, and reliability. *Archives of general psychiatry*, 46(11), 1006–1011. <https://doi.org/10.1001/archpsyc.1989.01810110048007>
- Grover, S., Nguyen, J. A., Viswanathan, V., & Reinhart, R. M. (2021). High-frequency neuromodulation improves obsessive-compulsive behavior. *Nature Medicine (New York, NY, United States)*, 27(2), 232–238. <https://doi.org/10.1038/s41591-020-01173-w>
- Harold, G. T., & Sellers, R. (2018). Annual Research Review: Interparental conflict and youth psychopathology: an evidence review and practice focused update. *Journal of Child Psychology & Psychiatry & Allied Disciplines*, 59(4), 374. <https://doi.org/10.1111/jcpp.12893>
- Hiebert, N. M., Lawrence, M. R., Ganjavi, H., Watling, M., Owen, A. M., Seergobin, K. N., & MacDonald, P. A. (2020). Striatum-mediated deficits in stimulus-response learning and decision-making in OCD. *Frontiers in Psychiatry*, 11, 13. <https://doi.org/10.3389/fpsyt.2020.00013>
- Kanen, J. W., Ersche, K. D., Fineberg, N. A., Robbins, T. W., & Cardinal, R. N. (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, 2337–2358. <https://doi.org/10.1007/s00213-019-05325-w>
- Kruschke, J. K. (2011). Bayesian assessment of null values via parameter estimation and model comparison. *Perspectives on Psychological Science*, 6(3), 299–312. <https://doi.org/10.1177/1745691611406925>
- Lim, T. V., Cardinal, R. N., Savulich, G., Jones, P. S., Moustafa, A. A., Robbins, T., & Ersche, K. D. (2019). Impairments in reinforcement learning do not explain enhanced habit formation in cocaine use disorder. *Psychopharmacology*, 236(8), 2359–2371. <https://doi.org/10.1007/s00213-019-05330-z>
- Peng, Z., He, L., Wen, R., Verguts, T., Seger, C. A., & Chen, Q. (2022). Obsessive-compulsive disorder is characterized by decreased Pavlovian influence on instrumental behavior. *PLOS Computational Biology*, 18(10), Article e1009945. <https://doi.org/10.1371/journal.pcbi.1009945>
- Poldrack, R. A., Sabb, F. W., Foerster, K., Tom, S. M., Asarnow, R. F., Bookheimer, S. Y., & Knowlton, B. J. (2005). The neural correlates of motor skill automaticity. *Journal of Neuroscience*, 25(22), 5356–5364. <https://doi.org/10.1523/JNEUROSCI.3880-04.2005>
- Robbins, T. W., Gillan, C. M., Smith, D. G., de Wit, S., & Ersche, K. D. (2012). Neurocognitive endophenotypes of impulsivity and compulsivity: towards dimensional psychiatry. *Trends in Cognitive Sciences*, 16(1), 81–91. <https://doi.org/10.1016/j.tics.2011.11.009>
- Robbins, T. W., Vaghi, M. M., & Banca, P. (2019). Obsessive-compulsive disorder: puzzles and prospects. *Neuron*, 102(1), 27–47. <https://doi.org/10.1016/j.neuron.2019.01.046>
- Ruan, Z., Seger, C. A., Yang, Q., Kim, D., Lee, S. W., Chen, Q., & Peng, Z. (2023). Impairment of arbitration between model-based and model-free reinforcement learning in obsessive-compulsive disorder. *Frontiers in Psychiatry*, 14, Article 1162800. <https://doi.org/10.3389/fpsyt.2023.1162800>
- Salkovskis, P. M. (1985). Obsessional-compulsive problems: A cognitive-behavioural analysis. *Behaviour Research and Therapy*, 23(5), 571–583. [https://doi.org/10.1016/0005-7967\(85\)90105-6](https://doi.org/10.1016/0005-7967(85)90105-6)
- Schaaf, J. V., Jepma, M., Visser, I., & Huizenga, H. M. (2019). A hierarchical Bayesian approach to assess learning and guessing strategies in reinforcement learning. *Journal of Mathematical Psychology*, 93, 102276. <https://doi.org/10.1016/j.jmp.2019.102276>
- Spielberger, C. D., Gonzalez-Reigosa, F., Martinez-Urrutia, A., Natalicio, L. F., & Natalicio, D. S. (1971). The state-trait anxiety inventory. *Revista Interamericana de Psicologia/Interamerican journal of psychology*, 5(3 & 4). <https://doi.org/10.30849/rjp.v5i3%20%204.620>
- Stein, D. J., Costa, D. L., Lochner, C., Miguel, E. C., Reddy, Y. J., Shavitt, R. G., & Simpson, H. B. (2019). Obsessive-compulsive disorder. *Nature Reviews Disease Primers*, 5(1), 52. <https://doi.org/10.1038/s41572-019-0102-3>
- Valentin, V. V., Dickinson, A., & O'Doherty, J. P. (2007). Determining the neural substrates of goal-directed learning in the human brain. *Journal of Neuroscience*, 27 (15), 4019–4026. <https://doi.org/10.1523/JNEUROSCI.0564-07.2007>
- Voon, V., Baek, K., Enander, J., Worbe, Y., Morris, L., Harrison, N. A., & Daw, N. (2015). Motivation and value influences in the relative balance of goal-directed and habitual behaviours in obsessive-compulsive disorder. *Translational Psychiatry*, 5(11), e670. <https://doi.org/10.1038/tp.2015.165>

- Voon, V., Joutsa, J., Majuri, J., Baek, K., Nord, C. L., Arponen, E., & Kaasinen, V. (2020). The neurochemical substrates of habitual and goal-directed control. *Translational Psychiatry*, 10(1), 1–9. <https://doi.org/10.1038/s41398-020-0762-5>
- Watson, P., van Wingen, G., & de Wit, S. (2018). Conflicted between goal-directed and habitual control, an fMRI investigation. *eneuro*, 5(4). <https://doi.org/10.1523/ENEURO.0240-18.2018>
- Worbe, Y., Savulich, G., de Wit, S., Fernandez-Egea, E., & Robbins, T. W. (2015). Tryptophan depletion promotes habitual over goal-directed control of appetitive responding in humans. *International Journal of Neuropsychopharmacology*, 18(10), pyv013. <https://doi.org/10.1093/ijnp/pyv013>
- Zainal, N. H., Camprodon, J. A., Greenberg, J. L., Hurtado, A. M., Curtiss, J. E., Berger-Gutierrez, R. M., & Wilhelm, S. (2023). Goal-Directed Learning Deficits in Patients with OCD: A Bayesian Analysis. *Cognitive Therapy and Research*, 1–12. <https://doi.org/10.1007/s10608-022-10348-3>