Altered model-based prediction error signalling in the medial prefrontal cortex in patients with obsessive-compulsive disorder: An fMRI study

Pritha Sen [1,2], Kathrin Koch [1,2], Benita Schmitz-Koep [1,2], Deniz Gürsel [1,2], Franziska Knolle[1,2,3*]

Author affiliations:

1 School of Medicine and Health, Department of Diagnostic and Interventional Neuroradiology, Technical University of Munich, Munich, Germany

2 School of Medicine and Health, TUM-NIC Neuroimaging Center, Technical University of Munich, Munich, Germany

3 Department of Psychiatry, University of Cambridge, Cambridge, UK

Running title: Model-free and model-based decision-making in OCD

Keywords: two-step task; medial prefrontal cortex; goal-directed decisions; computational psychiatry; neuropsychiatric disorders

Number of words in abstract: 250

Number of words in text: 3875

Number of figures: 4

Number of tables: 2

*corresponding author:

Dr. Franziska Knolle

Department of Diagnostic and Interventional Neuroradiology, School of Medicine and Health, Klinikum rechts der Isar, Technical University Munich, 81675 Munich, Germany

Phone: +49 89 4041 7976 Fax: +49 89 4140 4653

Email: franziska.knolle@tum.de ORCID: 0000-0002-9542-613X

Abstract

Background: Compared to healthy individuals, patients with obsessive-compulsive disorder (OCD) are found to rely more on model-free decision-making strategies which may underlie symptom expression. It is, however, unclear whether these behavioural differences are represented in neural alterations of model-free and model-based decision-making when tested simultaneously.

Methods: We investigated the neural signatures of 22 OCD patients and 22 matched controls who completed a two-step Markov decision-making task during functional MRI scanning. We used hierarchical Bayesian modelling and Bayesian statistics to examine model-based and model-free decision-making behaviours. Parametric regressors were employed for model-free and model-based reward prediction errors to inform neural reward presentation, which we analysed using a Bayesian Multilevel Modeling (BML) approach. Associations between significant activations and symptoms as well as cognitive scores were explored using Bayesian linear regression.

Results: While controls received significantly more rewards and were significantly less stochastic compared to patients, both groups similarly relied on model-free decision-making strategies. Importantly, our group comparison of neural reward prediction error responses showed greater activation for model-based reward prediction error in the medial prefrontal cortex (mPFC) in OCD patients compared to controls, but no differences for model-free reward prediction error processing. Increased mPFC activity was associated with lower depressive symptoms but not with other clinical or cognitive scores.

Conclusion: These findings support the notion that OCD is associated with a bias towards habitual behaviour, which may be expressed through altered mPFC activity underlying goal-directed behaviour. Tackling the hyperactivity on the mPFC may provide a new target for optimised treatment and interventions.

Introduction

One of the key impairments of obsessive-compulsive disorder (OCD) is the inability to make effective and advantageous decisions (1,2). Decision-making allows successful interaction with the environment through planning, monitoring, feedback integration and feedback-based updating. This process can be categorized into two distinct systems – model-free and model-based decision-making, which operate in parallel to choose the most suitable action (3–5). Model-free decision-making is reinforcement-learning based, and therefore automatic, guided by past successful choices, and inflexible to changes (6–8). Model-free decision-making may be described as habit learning, which requires direct responses to environmental stimuli without following a goal or explicit value. Model-based decisions, on the other hand, require an understanding of action-outcome contingencies, and with this an understanding of a given task. They are, therefore, deliberate, flexible, and adaptive, aiming for optimal outcomes (6,9,10), which relates to goal-directed behaviour. They may be described as a form of instrumental conditioning, where outcomes are induced by explicit actions. Distinct neural systems represent these goal-directed and model-based actions (6,11). A recent meta-analysis (7) revealed overlapping but distinct regions, showing that while the goal-directed system has been linked to activity in the ventral striatum, the medial prefrontal cortex (mPFC) and the orbitofrontal cortex (OFC), the habit system has been associated with activity in ventral striatum as well, but also in the globus pallidus and caudate head.

Individuals with OCD are usually characterized by a relatively high inflexibility and tendency towards sticking to habits and routines. Thus, it does not come as a surprise that they often exhibit an imbalance in these two systems, favouring rigid and repetitive decisions (1,10,12), which lead to inflexibility and resistance to change (10,13–15), potentially due to an overreliance on model-free decision-making. Additionally, compulsions in OCD are also shown to arise from this core difficulty of regulating actions and thoughts in a goal-directed manner (16–19). Neuroimaging studies looking at model-based and model-free decision-making separately, suggest that this shift towards habitual decisions in OCD is associated with alterations in the cortico-striato-thalamo-cortical (CSTC) circuits. Activity within these loops, involving – amongst others – OFC,, anterior cingulate cortex (ACC) and caudate nucleus, which are responsible for governing habits and automated behaviours (7,20–22), have been proposed to become hyperactive or unresponsive to inhibitory signals in OCD, potentially establishing the neuronal basis for the incapacity to inhibit obsessive or compulsive thoughts or behaviours (23,24). Additionally, reduced grey matter volume and decreased activity in the medial prefrontal cortex (mPFC), a region responsible for constructing, updating and

maintaining an internal model of the environment to guide goal-directed behaviour (25–27), has been linked to over-reliance on habitual control in OCD patients (25,28). Furthermore, disruption of activity in various regions of the basal ganglia, such as, subthalamic nucleus, nucleus accumbens, has been found to be associated with impaired behavioural inhibition, resulting in heightened impulsive choices (29,30).

Few studies investigate model-based and model-free decision-making simultaneously in OCD (8,31). The two-step Markov decision-making task (9) distinguishes between model-free and model-based learning behaviourally, and when combined with fMRI, revealed that the striatal reward prediction error (RPE) signal reflected both model-free and model-based decisions in healthy individuals (9). In a behavioural set-up of this two-step task, Voon and colleagues (32) found that individuals with different compulsive behaviours (binge eating disorders, methamphetamine addiction, and OCD patients), resorted to more model-free decisions. A follow-up behavioural study comparing the influence of reward and loss outcomes on decisions in this task demonstrated that, following rewarded trials, OCD patients leaned more towards model-free decisions, while the reverse was observed in loss trials (8). Notably, patients displayed a heightened tendency to maintain the same first-level choice in consecutive trials, irrespective of win or loss, compared to healthy controls. Furthermore, the severity of compulsion scores correlated with habitual learning in rewarded trials whereas obsession scores correlated with increased choice switching following loss trials. In a very recent study, Kim and colleagues (31), for the first time, explored neurocomputational alteration during model-based and model-free decision-making in OCD patients using the two-step task and applying a computational model which was designed to account for an uncertainty-based arbitration process of model-based and model-free strategies. They found, first, that patients showed hypoactivity in the inferior frontal gyrus when tracking the reliability of goal-directed behaviour; and second, that, patients exhibited weaker right ipsilateral ventrolateral prefronto-putamen coupling to reach specific goals, which was correlated with more severe compulsivity. More evidence, however, is needed to understand the neural mechanisms of altered model-based and model-free decision-making and their relevance for clinical symptoms in OCD.

Therefore, the aim of this study was to explore the neural mechanisms of model-based and model-free decision-making in OCD patients using the two-step task with fMRI, incorporating hierarchical Bayesian modelling and Bayesian Multilevel Modeling (BML) which performs Bayesian analyses based on pre-specified regions of interest. All other statistical analyses were also performed using Bayesian statistical methods, following a timely debate of

misinterpretations of p-values and the replication crisis (33). Bayesian statistical approaches incorporate prior information into the analysis allowing for regularisation of models and preventing overfitting, which increases the precision of the findings, making them more generalisable and reliable (34,35). Based on the literature, our regions of interest included the caudate, putamen, nucleus accumbens, mPFC and ACC for both model-free and model-based RPE processing. Timeseries of the model-free and model-based decision-making were used to explore which strategy was reflected in each region and whether this differed across groups. We hypothesized that the OCD patients would exhibit predominantly model-free behaviour with caudate, putamen, nucleus accumbens and ACC activations reflecting more model-free RPE, while the healthy controls would display a mixture of model-free and model-based behaviour, with the mPFC and the nucleus accumbens reflecting more model-based RPE. Furthermore, we expected that neural alterations would link to symptom strength. We also expected that neural alterations in model-based prediction-errors would be associated with stronger compulsivity.

Methods

Participants

Twenty-five adults with OCD and 23 healthy controls were recruited for this study out of which one healthy control and three OCD patients had to be excluded from the final analysis due to corrupted imaging data. The final data for the analysis, therefore, consisted of 22 OCD patients and 22 healthy controls. Demographic information, clinical and cognitive scores are presented in Table 1. Patients met Diagnostic and Statistical Manual of Mental Disorders (DSM)-IV criteria for OCD (36). Exclusion criteria for both groups were a history of severe head injuries, seizures, neurological diseases, schizophrenia, autism, substance and alcohol abuse/dependency, mental retardation, severe medical conditions, and pregnancy. 16 patients were under medication while the other six were medication-naïve or had stopped medication at least one week before scanning. The study was approved by the Technical University of Munich ethics committee, and all participants gave informed consent after receiving a complete description of the study.

Table 1. Demographic characteristics, clinical characteristics and cognitive functioning scores of healthy controls and patients with obsessive-compulsive disorder.

	OCD	Healthy Controls	Group Differences
Demographic characteristics			
Age (SD)	33.4 (11.8)	32.2 (8.33)	B=-0.66, 95% CI [-6.5, 5.05]
Gender	17F, 5M	13F, 9M	B=-5.7, 95% CI [-15.22, -2.3]
Clinical scores			
Years since onset	14.32 (9.41)		
Age of onset	19.04 (10.06)		
Y-BOCS Obsessions	11.2 (4.16)		
Y-BOCS Compulsions	10 (4.23)		
Y-BOCS total	21.2 (7.04)		
OCI-R	26.9 (9.47)		
HAMD	12.5 (4.6)		
Cognitive scores			
TMT-A (s)	29 (13.2)	22 (6.10)	B=-5.69, 95% CI [-11.79, 0.34]
TMT-B (s)	66.5 (31.5)	56.1 (18.9)	B=-9.48, 95% CI [-26.4, 7]
TMT (B-A) (s)	37.5 (23)	34.1 (17.4)	B=-3.18, 95% CI [-15.45, 9.65]

Note: Data are presented as mean (SD) except for Gender. Abbreviations: M=Male; F=Female; SD=Standard Deviation; TMT=Trail-making test; Y-BOCS=Yale-Brown Obsessive-Compulsive Scale; OCI-R=Obsessive-Compulsive Inventory – Revised; HAMD=Hamilton rating scale for depression; B = Bayesian estimate; CI not containing zero indicates substantial effect.

Task description

All the participants performed the two-step task by Daw and colleagues (9) inside an fMRI scanner, which is a reinforcement learning task to measure the relative degree of model-free against model-based decision-making. A detailed description of the task has been described elsewhere (37,38), in the supplementary material, and in Figure 1.

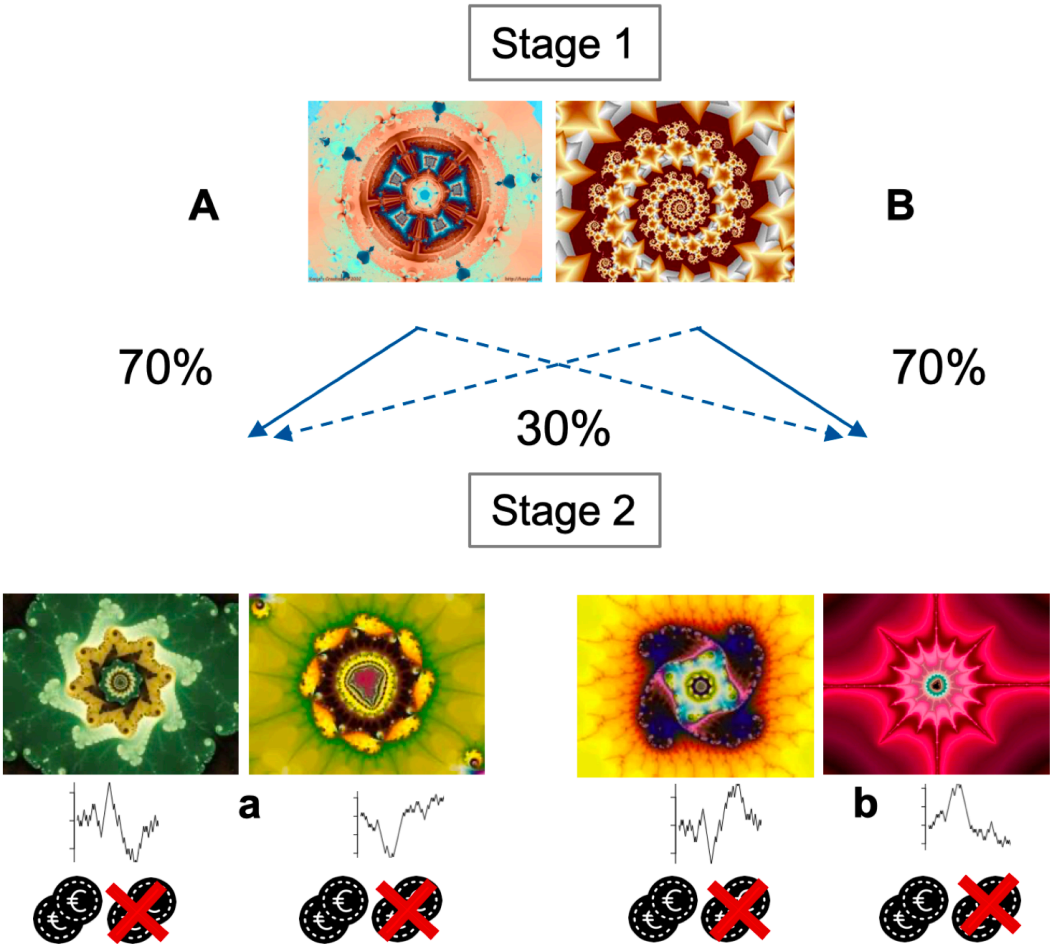

Figure 1 Two-step Markov decision task. On each trial, a first stage choice between two images (here, A or B) leads probabilistically to a second stage choice (here, images in set 'a' or images in set 'b'). Importantly, each first-stage stimulus is more strongly (70% vs 30%) associated with a particular second-stage state throughout the experiment, imposing a task structure that can be exploited in a model-based choice. All stimuli in stage 2 are associated with a probabilistic reward changing slowly and independently based on Gaussian random walks, forcing subjects to continuously learn and explore the second stage choices.

Analysis

Behavioural Analysis

Performance analysis

To measure performance in the task, reward percentages for both groups were calculated as the ratio of number of rewarded trials to the total number of trials per subject and compared using robust Bayesian ANOVA, followed by a Bayes factors (BF_{10}) analysis, quantifying the more likely hypothesis based on the observed data. Evidence for alternative hypothesis over null hypothesis was identified if $BF_{10} > 1$ (BF_{10} [1-3]: anecdotal evidence; BF_{10} [3-10]: moderate

evidence; BF_{10} [10-30]: strong evidence; BF_{10} [30-100]: very strong evidence; $BF_{10} > 100$: extreme evidence).

Analysis of stay probability

To identify differences between patients and controls towards a decision bias, we assessed the probability for staying with the same stage 1 choice in the subsequent trial which is deterministic of model-free and model-based choices. See supplementary material for details. Stay probabilities for ideal model-free and model-based performances were also simulated along and compared visually with actual stay probability in the task.

We used Bayesian regression analysis, assuming normal distribution, to investigate whether the stay probability varied based on participant groups (healthy controls, OCD patients), previous trial transition type (common or uncommon), previous trial outcome (rewarded or unrewarded), as well as their interactions. As we had multiple observations per person, we incorporated a group level intercept to account for the resulting dependency in the data. A substantial effect in the interaction between transition type and outcome would reflect model-based decisions, and lack of a substantial effect of interaction would indicate model-free decisions that is purely guided by reward without considering the transition structure. The equation for the model that we fit is as follows:

stay probability ~ previous trial reward*previous trial transition type*participants + (1|Subject)

Computational modelling of task behaviour

Hierarchical Bayesian modelling was employed on the behavioural data (39) to understand the underlying mechanisms of decision-making in OCD patients compared to healthy controls and inform the imaging analysis. We fitted four models, each with 4 chains, 1000 burn-in samples and 3000 samples, after which model selection for further analysis was determined based on model-convergence and lowest leave one out information criterion (LOOIC) values. The 6-parameter model was found to be the winning model and was used for further analysis. See supplementary material for description of models, fitting procedures, and model comparisons. The 6-parameter model estimates the following parameters:

Learning rate ('alpha 1 and alpha 2'): showing the efficiency of learning over the trials in stage 1 and stage 2, takes values from 0 to 1. The higher the score, the better the subjects understand and perform in the task.

Inverse-temperature/choice stochasticity ('beta1 and beta2'): referring to the proportions of stochastic/random choices made during the task in stage 1 and stage 2, with beta=0 for completely random responding and beta=∞ for stochastically choosing the highest value option.
Perseverance ('pi'): Perseverance determines how strongly the subject(s) stick to their decisions. Higher the score on the perseverance scale, lower the chance of switching to a different image.
Model-weights ('w'): referring to the degree of model-based influence on choices, takes values from 0 to 1, '0' indicating more model-free decisions and '1' being more model-based.

Trial-by-trial model-regressors were extracted from the winning computational model for both model-free and model-based prediction errors to be used as parametric modulators in the imaging analysis ((9), supplementary material).

To investigate group differences for the model parameters, we employed Bayesian robust one-way ANOVAs with each parameter as the dependent variable and group as the predictor variable, followed by a Bayes factors (BF_{10}) analysis.

fMRI Analysis

Details of image acquisition and pre-processing are described in the supplementary material. After preprocessing, subject-specific design matrices were defined using general linear modelling as implemented in SPM12 (40). The fMRI analysis was centred around analysing the time-series of model-free and model-based RPEs. Four conditions were entered to calculate the average BOLD responses across all trials at each of the four time points in the task (i.e., onset first stage images, button-press first stage, onset second stage images, and button-press with reward presentation on the second stage). The computationally derived reward prediction error signals (mf_RPE and mfb_RPE) were entered as parametric regressors in the fourth condition, i.e. the button-press second stage with reward presentation. For the second-level models, the contrasts of interest defined were 'model-free reward prediction error' and 'model-based reward prediction error'. The per-subject estimate of the model-based effect (*'w'* obtained from the computational model) was also included as a second-level covariate for the model-based regressor to test the correspondence between behavioural and neural estimates of the model-based effect. These contrasts were taken to a second-level one-sample t-test across all controls and patients.

The second level analysis was performed on specific regions of interest that have been implicated in model-free and model-based RL in past studies – caudate, putamen, nucleus

accumbens, medial prefrontal cortex (mPFC) and anterior cingulate cortex (ACC) (Figure 3A). Anatomical masks of the caudate, putamen and nucleus accumbens were derived from the Harvard-Oxford Subcortical Atlas and created and combined using FSL (41); anatomic masks for the ACC was created combining Brodmann areas 24, 32, 33 as implemented in the WFU PickAtlas Toolbox in SPM, and mPFC was created by constructing a sphere of 5mm radius around the mPFC coordinates (x=-4, y=60, z=14) identified by Daw and colleagues (9) for model-free and model-based decisions. We used a sphere to define the mPFC to precisely target the specific coordinates since the mPFC is a functionally heterogeneous area and atlas-based regions can be quite large, potentially encompassing areas not specifically relevant to our study's hypotheses. Parameter estimates for these ROIs were extracted from the second level contrast images using MarsBaR in SPM and entered into the Region-Based Analysis Program through BML (42), in AFNI (43).

As the BML is a recent neuroimaging technique, a concise overview is presented here (for a deeper understanding, refer to Chen and colleagues (42)). BML involves estimating the likelihood of a hypothesis based on observed data, represented as a probability density function termed the posterior distribution (abbreviated as P+). This distribution is derived by combining empirical data with a model and prior expectations. BML offers the advantage of consolidating data across subjects and Regions of Interest (ROIs), facilitating the integration of shared information among these regions. The more the posterior distribution deviates from 0, the stronger the evidence for an effect. The evidence for effects with a probability (P+) of <.10 or >.90 as categorized as weak, <.05 or >.95 as moderate, and <.025 or >.975 as strong. A probability (P+) of >.10 or <.90 is considered no evidence. Effects with a P+ of <.10 or >.90 are discussed in the main text, while comprehensive effects are presented in the probability distribution plots.

Associations between neural responses and clinical and cognitive scores

To examine the relationships between clinical and cognitive symptoms, and brain activations in patients that were significantly different from those in controls, we conducted Bayesian linear regressions for each of the symptom (i.e., Y-BOCS obsession and compulsion, OCI-R scores, HAMD) and cognitive scores (i.e., TMT-A, TMT-B, TMT-BA) as the outcome variable and significant brain activations as predictors covarying for age of illness onset. The response variables were assumed to follow a Gaussian distribution for the residuals, and priors for the regression coefficients were specified using a normal distribution with a mean of 0 and standard deviation of 10. The posterior distributions were sampled using the No-U-Turn Sampler with 4

chains, each containing 2000 iterations, including a 1000-iteration warmup phase. In Bayesian regression analysis, an effect was identified by inspecting the credibility intervals. If 0 is not contained in the credibility interval, a significant effect was detected.

General statistical implementations

All statistical analyses were performed using R Statistical Software (version 4.3.1) (44); hierarchical Bayesian modelling was performed using hBayesDM (39); Bayesian linear regression and ANOVA analysis was carried out using the brms package (version 2.20.4) (45); imaging analysis was performed using Statistical Parametric Mapping software, SPM12 (40); masks for region of interest analysis was made using FMRIB Software Library (FSL) (version 6.0.5.2) (41); and BML analysis for the regions of interest was performed using the Region Based Analysis (RBA) Program (42) in the Analysis of Functional NeuroImages (AFNI) Software (version 23.2.12) (43).

Results

Performance differences across groups

The robust Bayesian ANOVA revealed a substantial group difference in reward percentage for healthy controls (Mean: 60.3, SD: 3.43) and OCD patients (Mean: 54.6, SD: 3.5) (B=5.61, EE=1, 95% CI [3.64, 7.57]) with Bayes Factor analysis revealing extreme evidence (BF_{10}=4779.379), indicating that controls performed better than OCD patients.

Stay probability differences across groups

Results of the Bayesian regression analysis revealed no main effect for group, but a substantial main effect of reward (rewarded vs. unrewarded: B=0.14, EE=0.03, 95% CI [0.07, 0.21]), indicating that all participants stayed on the same first level choice in the next trial if the previous trial was rewarded, suggesting a model-free behaviour. We found no evidence for interaction effects indicating that the subjects followed a model-free decision-making approach. Stay probabilities for ideal model-free and model-based behaviour are displayed in Figure 2a and actual stay probabilities in the current task are displayed in Figure 2b.

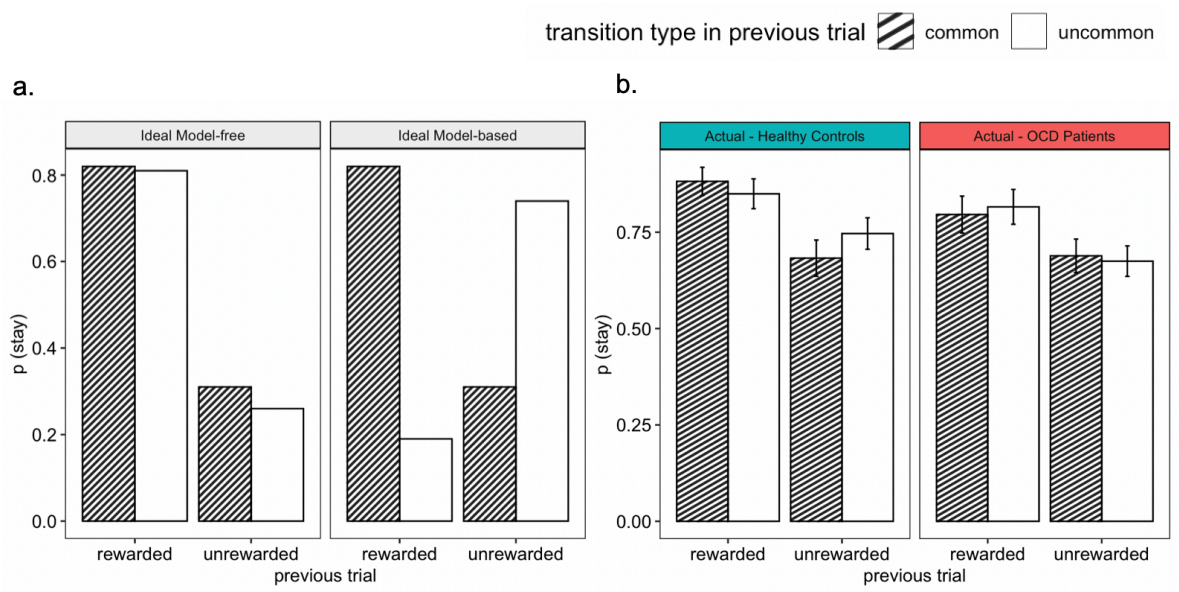

Figure 2 Ideal and actual stay-switch behaviours. a. Ideal Model-free and Ideal Model-based decision-making behaviour: Model-free RL predicts that a first-stage choice yielding a reward is likely to be repeated on the upcoming trial, regardless of a common or an uncommon transition, Model-based RL predicts that an uncommon transition should affect the value of the next first stage option, leading to a predicted interaction between reward and transition probability. **b.** Actual stay proportions, averaged across subjects, displaying a characteristic of using both model-free and model-based decision-making strategies.

Modelling parameters differences across groups

The robust Bayesian ANOVA revealed group differences in stochasticity at stage 1 'beta 1' (healthy controls vs. OCD patients: B=1.57, EE=0.7, 95% CI [0.21, 2.94]). The Bayes Factor analysis for stochasticity at stage 1 'beta 1' scores indicated anecdotal evidence (BF_{10}=2.65), suggesting that patients were more random in their choices compared to controls. None of the other parameters revealed substantial effects.

Differences in the neural signature of model-free and model-based prediction errors

Group comparisons between healthy controls and OCD patients in the model-based RPE signal revealed that the OCD patients had a higher activation in the mPFC compared to healthy controls (strong evidence; P+=0.02). We did not find any group differences in the model-free RPE signal in patients and controls. Parameter estimates for group comparisons in model-based and model-free RPE signal are displayed in Figure 3B.

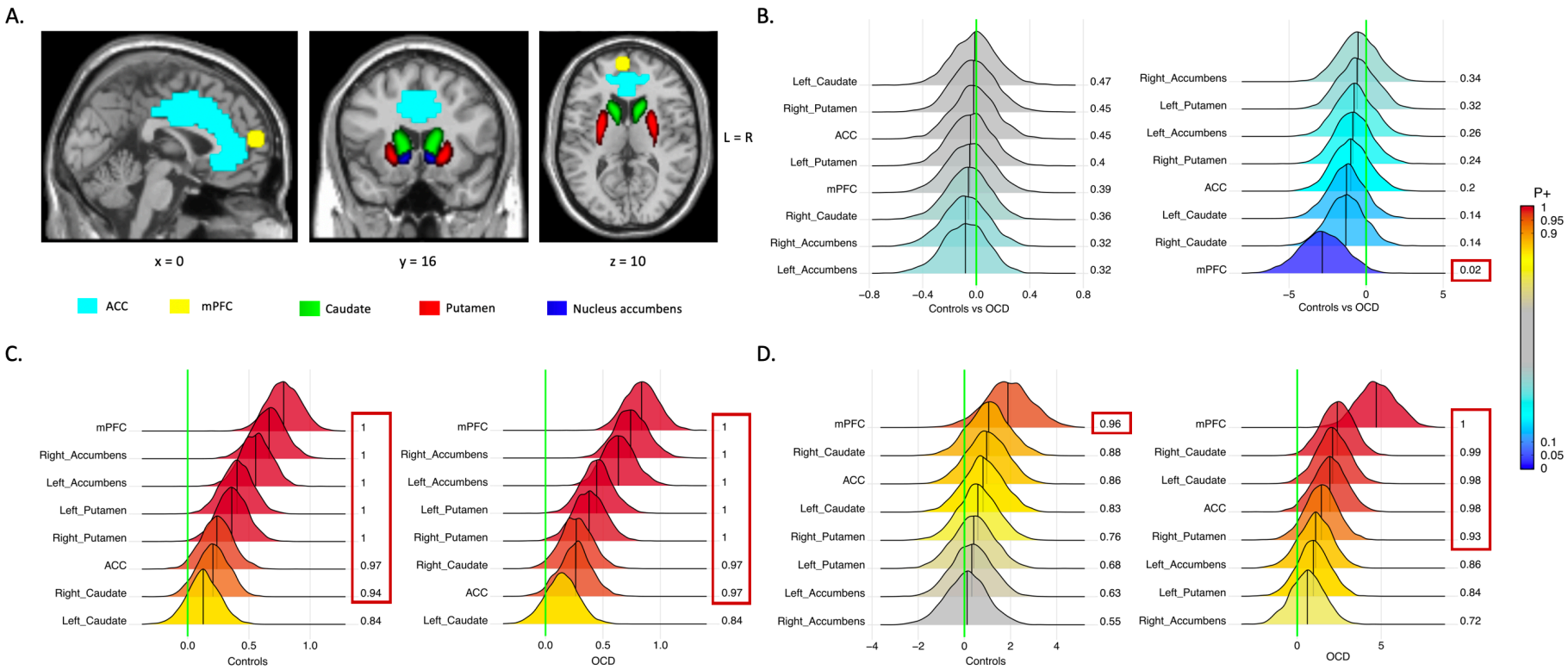

Figure 3 ROI analysis. A. Regions included in the ROI analysis. ROIs for model-free and model-based RPE; Masks were derived from the Harvard-Oxford Subcortical Atlas using FSL. **B. ROI analysis showing posterior distributions for OCD patients compared with healthy controls (left:** Model-free RPE; **right:** Model-based RPE). Values >.90 represent higher activation in healthy controls, and <.10 represent lower activation in the healthy controls. **C. ROI analysis showing posterior distributions within each group for model-free RPE (left:** Healthy controls; **right:** OCD patients). **D. ROI analysis showing posterior distributions within each group for model-based RPE (left:** Healthy controls; **right:** OCD patients). Note: Posterior distributions: strong evidence: (P+ > 0.975 or <0.025), moderate evidence (P+ > 0.95 or <0.05), weak evidence (P+ >0.90 or <0.10) of an effect. Boxes outlined in red highlight regions with substantial effects in each of these plots.

Within group effects for model-based RPE revealed activations in the mPFC (moderate evidence; P+=0.96) in healthy controls; and activations in the mPFC (strong evidence; P+=1), right (strong evidence; P+=0.99) and left (strong evidence; P+=0.98) caudate, ACC (strong evidence; P+=0.98) and right putamen (weak evidence; P+=0.93) in the OCD patients. Within group effects of model-free RPE revealed activations in the mPFC (strong evidence; Controls: P+=1, OCD: P+=1), right accumbens (strong evidence; Controls: P+=1, OCD: P+=1), left accumbens (strong evidence; Controls: P+=1, OCD: P+=1), left putamen (strong evidence; Controls: P+=1, OCD: P+=1), right putamen (strong evidence; Controls: P+=1, OCD: P+=1), ACC (moderate evidence; Controls: P+=0.97, OCD: P+=0.97) and right caudate (weak evidence in Controls; P+=0.94, moderate evidence in OCD; P+=0.97) for both healthy controls and OCD patients. Parameter estimates for group comparisons in model-based and model-free RPE signal for healthy controls and OCD patients are displayed in Figures 3C and 3D respectively.

Bayesian Linear Regressions results

First, we examined the relationships between model-based RPE signal in the mPFC, OCD symptoms and cognitive scores in OCD patients (Table 2). The Bayesian regression revealed that mPFC activity significantly and negatively predicted HAMD (B=-0.15, EE=0.08, 95% CI [-0.31, -0.00]), controlling for age of onset. No other regression analyses revealed significant results.

Table 2. Model-based activations in the mPFC in OCD predicting clinical and cognitive scores using Bayesian linear regression analysis.

	Estimate	Est. Error	95% CI [LL,UL]	R-hat	Bulk ESS	Tail ESS
Bayesian regression predicting Y-BOCS Obsessions						
Intercept	12.6	2.03	[8.7, 16.62]	1	3511	2398
mPFC	-0.08	0.07	[-0.22, 0.07]	1	3604	2561
age of onset	-0.06	0.1	[-0.25, 0.13]	1	3362	2407
sigma	4.25	0.68	[3.17, 5.79]	1	3394	3242

Bayesian regression predicting Y-BOCS Compulsions						
Intercept	9.9	2.16	[5.68, 14.23]	1	4613	2529
mPFC	0	0.08	[-0.16, 0.16]	1	4163	2406
age of onset	0.01	0.1	[-0.19, 0.21]	1	4314	2637
sigma	4.61	0.1	[-0.19, 0.21]	1	3199	2615
Bayesian regression predicting Y-BOCS total						
Intercept	22.46	3.62	[15.52, 29.57]	1	4211	2738
mPFC	-0.07	0.13	[-0.34, 0.18]	1	4421	2755
age of onset	-0.05	0.17	[-0.37, 0.29]	1	3667	2899
sigma	7.6	1.3	[5.56, 10.52]	1	3659	2762
Bayesian regression predicting OCI-R						
Intercept	25.99	4.8	[16.79, 35.83]	1	4202	2642
mPFC	-0.05	0.18	[-0.4, 0.3]	1	3536	2255
age of onset	0.06	0.22	[-0.38, 0.5]	1	4471	2870
sigma	10.33	1.73	[7.58, 14.32]	1	3464	2695
Bayesian regression predicting HAMD						
Intercept	12.18	2.08	[8.06, 16.25]	1	3869	2877
mPFC	-0.15	0.08	[-0.31, -0.00]	1	4314	2900
age of onset	0.05	0.10	[-0.14, 0.24]	1	3848	2675
sigma	4.53	0.76	[3.32, 6.37]	1	3090	2801
Bayesian regression predicting TMT-A						
Intercept	22.42	5.97	[10.69, 33.98]	1	3820	2595
mPFC	-0.27	0.22	[-0.69, 0.15]	1	4115	2702
age of onset	0.38	0.28	[-0.18, 0.93]	1	3786	2664

sigma	12.94	2.08	[9.57, 17.52]	1	3296	2735
Bayesian regression predicting TMT-B						
Intercept	49.31	13.6	[22.06, 76.69]	1	3907	2693
mPFC	-0.87	0.5	[-1.84, 0.14]	1	4459	2913
age of onset	1.08	0.64	[-0.16, 2.32]	1	3886	2410
sigma	29.7	4.88	[21.98, 40.81]	1	3675	2890
Bayesian regression predicting TMT(B-A)						
Intercept	26.84	10.94	[5.24, 48.35]	1	4357	2523
mPFC	-0.6	0.38	[-1.35, 0.15]	1	4175	2468
age of onset	0.7	0.51	[-0.31, 1.7]	1	4316	2640
sigma	22.57	3.77	[16.52, 31.23]	1	3365	2611

Note: All Bayesian regressions are conducted assuming a Gaussian distribution and are controlled for age of onset. Y-BOCS=Yale-Brown Obsessive-Compulsive Scale; OCI-R=Obsessive-Compulsive Inventory – Revised; HAMD=Hamilton rating scale for depression; Estimate=Bayesian estimate for the predictors; TMT=Trail-making test; CI=Credible interval; LL=lower limit; UL=upper limit; R-hat=potential scale reduction factor on split chains (at convergence, R-hat=1); Bulk ESS=general effective sample size; Tail ESS=effective sample size for the tails of the distribution; sigma=standard deviation of the residuals in the Gaussian model. CI not containing zero indicates substantial effect.

Discussion

We investigated neural correlates of model-free and model-based prediction errors in OCD patients versus healthy controls using an fMRI-version of the two-step Markov decision task and computational modelling. Behaviourally, the model parameters revealed that the patients were more random and received fewer rewards than healthy individuals, and relied more strongly on model-free strategies compared to model-based, which confirms previous findings (8,38,46). Importantly, we found higher activation in the mPFC in OCD compared to controls for model-based RPE processing, but no differences for model-free RPE processing. These

findings support the notion that OCD is associated with a bias towards habitual behaviour, which may be expressed through altered mPFC activity underlying goal-directed behaviour (11,32,47).

Activations in the mPFC for model-based RPE align with literature highlighting the region's role in updating and maintaining an internal model of the environment to guide goal-directed behaviour (25–27). Supporting our findings, previous studies have observed relative hyperactivation of the mPFC, including the supplementary motor area in OCD patients during task performances involving inhibitory control and conflict processing compared to healthy controls (10,48). In a study combining volume-localised proton magnetic resonance spectroscopy and fMRI in an inhibitory control task, Yücel and colleagues (48) argued that while having primary relevance to the pathogenesis of the disorder, increased activation in the mPFC reflects a compensatory mechanism to achieve adequate task performance, potentially due to biochemical abnormalities in the dorsal ACC. Together these results suggest that hyperactivity in the mPFC might be a potential neural correlate of altered goal-directed behaviour in OCD. Additionally, it has been proposed that the mPFC plays a central role in resolving conflicts between the brain's habitual and goal-directed system (49,50), either by deliberately overriding the habitual system when required (49), or by mediating the two systems to naturally interact (50). This supports our findings as both groups, when analysed separately, show model-based and model-free prediction error processing in the mPFC.

OCD patients showing higher activations for model-based RPE in the mPFC compared to controls, and also showing overlapping region activations during both model-free and model-based decisions, further suggests that the patients may over-recruit regions responsible for habitual decisions when they are unable to resolve conflict between habitual and goal-directed strategies. This may support previous findings of a disrupted arbitration mechanism in OCD patients (28,51) and might potentially be responsible for the dominance of habitual behaviours in OCD. In line with our findings, past studies have also found structural and functional alterations within the fronto-striatal circuits, particularly the mPFC, putamen and caudate responsible for the imbalance between model-free and model-based decision-making in OCD (28,52). Additionally, the mPFC is a crucial part of the dorsal cognitive CSTC circuit which contributes to habit formation and compulsivity in OCD (53). Since the nucleus accumbens and the left putamen were activated only during model-free RPE, our results suggest activity in these regions as potential neural correlates for model-free decisions in OCD.

Furthermore, our results revealed that the same brain regions reflected model-free RPE in the controls and patients. These regions included activations in the bilateral nucleus accumbens and putamen, the ACC, the mPFC, and the right caudate nucleus, which are in accordance with our hypothesis and the literature (7,20,23,24,46,54–57), confirming the contributions of these regions to model-free RPE processing. Left caudate was not associated with model-free RPE in our study, as also reported by Guida and colleagues (58), indicating fewer and less strong activations in the left caudate for habitual activities in everyday-life.

Moreover, in OCD patients mPFC RPE signalling negatively predicted depressive symptoms measured using HAMD. However, we did not find any other symptom associations. Given the hyperactivity of the mPFC in patients it is surprising that depressive symptoms are declining with increasing activity. Although mPFC signalling alterations during reward and reward prediction error processing have been reported in patients with major depressive disorder (59–61), future studies should further explore these associations.

Limitations

This study has several limitations. First, our moderate sample size did not allow to explore subgroups of patients with specific symptoms. Subsequent studies should aim to replicate the results in a larger sample allowing for a differentiation in habitual and goal-directed decision-making between several symptom subtypes (e.g., washers versus checkers). Second, we were unable to elicit model-based decision-making in healthy controls, potentially due to a combination of task instruction, reward presentation and task duration (62,63). Future studies should replicate our results with more trials, potentially in a longitudinal framework to provide a more thorough understanding. Third, our analysis included both medicated and unmedicated patients. Future studies should aim for larger sample sizes to verify results in separate analyses for medicated and unmedicated individuals.

Conclusion

Taken together, this study revealed an overreliance on model-free behaviour and a hyperactivity in the mPFC in the OCD patients compared to controls in response to model-based prediction error processing, suggesting a disruption of the arbitration system responsible for resolving conflict between habitual and goal-directed decisions. Thus, the mPFC may

provide a new target for optimised treatment and intervention, such as non-invasive stimulations using transcranial direct current stimulation.

Acknowledgements

We would like to thank the patients and their families as well as the healthy controls for their time and effort. Furthermore, we would like to thank Dr. Felix Brandl and Dr. Christian Sorg for their insights and support during data acquisition.

Competing interests

The authors report no competing interests.

References

1. Robbins TW, Vaghi MM, Banca P (2019): Obsessive-Compulsive Disorder: Puzzles and Prospects. *Neuron* 102: 27–47.
2. Sachdev PS, Malhi GS (2005): Obsessive–Compulsive Behaviour: A Disorder of Decision-Making. *Aust N Z J Psychiatry* 39: 757–763.
3. Daw ND, Niv Y, Dayan P (2005): Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control [no. 12]. *Nat Neurosci* 8: 1704–1711.
4. Gläscher J, Daw N, Dayan P, O'Doherty JP (2010): States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron* 66: 585–595.
5. Grosskurth ED, Bach DR, Economides M, Huys QJM, Holper L (2019): No substantial change in the balance between model-free and model-based control via training on the two-step task. *PLOS Computational Biology* 15: e1007443.
6. Balleine BW, O'Doherty JP (2010): Human and Rodent Homologies in Action Control: Corticostriatal Determinants of Goal-Directed and Habitual Action [no. 1]. *Neuropsychopharmacol* 35: 48–69.

7. Huang Y, Yaple ZA, Yu R (2020): Goal-oriented and habitual decisions: Neural signatures of model-based and model-free learning. *NeuroImage* 215: 116834.

8. Voon V, Baek K, Enander J, Worbe Y, Morris LS, Harrison NA, *et al.* (2015): Motivation and value influences in the relative balance of goal-directed and habitual behaviours in obsessive-compulsive disorder [no. 11]. *Transl Psychiatry* 5: e670–e670.

9. Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ (2011): Model-Based Influences on Humans' Choices and Striatal Prediction Errors. *Neuron* 69: 1204–1215.

10. Gillan CM, Robbins TW (2014): Goal-directed learning and obsessive–compulsive disorder. *Philosophical Transactions of the Royal Society B: Biological Sciences* 369: 20130475.

11. Robbins TW, Banca P, Belin D (2024): From compulsivity to compulsion: the neural basis of compulsive disorders. *Nature Reviews Neuroscience* 25: 313–333.

12. Gillan CM, Papmeyer M, Morein-Zamir S, Sahakian BJ, Fineberg NA, Robbins TW, de Wit S (2011): Disruption in the Balance Between Goal-Directed Behavior and Habit Learning in Obsessive-Compulsive Disorder. *AJP* 168: 718–726.

13. Morein-Zamir S, Papmeyer M, Gillan CM, Crockett MJ, Fineberg NA, Sahakian BJ, Robbins TW (2013): Punishment promotes response control deficits in obsessive-compulsive disorder: evidence from a motivational go/no-go task. *Psychological medicine* 43: 391–400.

14. Murray GK, Knolle F, Ersche KD, Craig KJ, Abbott S, Shabbir SS, *et al.* (2019): Dopaminergic drug treatment remediates exaggerated cingulate prediction error responses in obsessive-compulsive disorder. *Psychopharmacology (Berl)* 236: 2325–2336.

15. Wheaton MG, Gillan CM, Simpson HB (2019): Does cognitive-behavioral therapy affect goal-directed planning in obsessive-compulsive disorder? *Psychiatry Research* 273: 94–99.

16. Gillan CM (2017): Habits and goals in OCD.

17. Nestadt G, Kamath V, Maher BS, Krasnow J, Nestadt P, Wang Y, *et al.* (2016): Doubt and the Decision-making Process in Obsessive-Compulsive Disorder. *Med Hypotheses* 96: 1–4.

18. Pushkarskaya H, Tolin D, Ruderman L, Kirshenbaum A, Kelly JM, Pittenger C, Levy I (2015): Decision-Making Under Uncertainty in Obsessive-Compulsive Disorder. *J Psychiatr Res* 69: 166–173.

19. Robbins TW (2022): Cognitive flexibility, OCD and the brain. *Brain* 145: 814–815.

20. Burton AC, Nakamura K, Roesch MR (2015): From ventral-medial to dorsal-lateral striatum: Neural correlates of reward-guided decision-making. *Neurobiology of Learning and Memory* 117: 51–59.

21. Miller KJ, Ludvig EA, Pezzulo G, Shenhav A (2018): Chapter 18 - Realigning Models of Habitual and Goal-Directed Decision-Making. In: Morris R, Bornstein A, Shenhav A, editors. *Goal-Directed Decision Making*. Academic Press, pp 407–428.

22. Simmler LD, Ozawa T (2019): Neural circuits in goal-directed and habitual behavior: Implications for circuit dysfunction in obsessive-compulsive disorder. *Neurochemistry International* 129: 104464.

23. Burguière E, Monteiro P, Mallet L, Feng G, Graybiel AM (2015): Striatal circuits, habits, and implications for obsessive–compulsive disorder. *Current Opinion in Neurobiology* 30: 59–65.

24. Graybiel AM, Rauch SL (2000): Toward a Neurobiology of Obsessive-Compulsive Disorder. *Neuron* 28: 343–347.

25. Ahmari SE, Rauch SL (2022): The prefrontal cortex and OCD. *Neuropsychopharmacology* 47: 211–224.

26. Seow TXF, Benoit E, Dempsey C, Jennings M, Maxwell A, O'Connell R, Gillan CM (2021): Model-Based Planning Deficits in Compulsivity Are Linked to Faulty Neural Representations of Task Structure. *J Neurosci* 41: 6539–6550.

27. Szalisznyó K, Silverstein DN (2021): Computational predictions for OCD pathophysiology and treatment: a review. *Frontiers in Psychiatry* 12: 687062.

28. Ruan Z, Seger CA, Yang Q, Kim D, Lee SW, Chen Q, Peng Z (2023): Impairment of arbitration between model-based and model-free reinforcement learning in obsessive–compulsive disorder. *Front Psychiatry* 14: 1162800.

29. Christakou A, Robbins TW, Everitt BJ (2004): Prefrontal cortical-ventral striatal interactions involved in affective modulation of attentional performance: implications for corticostriatal circuit function. *J Neurosci* 24: 773–780.

30. Jahanshahi M, Obeso I, Rothwell JC, Obeso JA (2015): A fronto–striato–subthalamic–pallidal network for goal-directed and habitual inhibition [no. 12]. *Nat Rev Neurosci* 16: 719–732.

31. Kim T, Lee SW, Lho SK, Moon S-Y, Kim M, Kwon JS (2024): Neurocomputational model of compulsivity: deviating from an uncertain goal-directed system. *Brain* 147: 2230–2244.

32. Voon V, Derbyshire K, Rück C, Irvine MA, Worbe Y, Enander J, *et al.* (2015): Disorders of compulsivity: a common bias towards learning habits. *Mol Psychiatry* 20: 345–352.

33. Anderson SF (2020): Misinterpreting p: The discrepancy between p values and the probability the null hypothesis is true, the influence of multiple testing, and implications for the replication crisis. *Psychological methods* 25: 596.

34. Edwards W, Lindman H, Savage LJ (1963): Bayesian statistical inference for psychological research. *Psychological Review* 70: 193–242.

35. Van De Schoot R, Winter SD, Ryan O, Zondervan-Zwijnenburg M, Depaoli S (2017): A systematic review of Bayesian articles in psychology: The last 25 years. *Psychological Methods* 22: 217–239.

36. First M, Spitzer R, Gibbon M, Williams J (2002): *Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition (SCID-I/P).* Biometrics Research, New York State Psychiatric Institute: New York.

37. Brandl F, Knolle F, Avram M, Leucht C, Yakushev I, Priller J, *et al.* (2023): Negative symptoms, striatal dopamine and model-free reward decision-making in schizophrenia. *Brain* 146: 767–777.

38. Knolle F, Sen P, Culbreth A, Koch K, Schmitz-Koep B, Gürsel DA, *et al.* (2023): Investigating disorder-specific and transdiagnostic alterations in model-based and model-free decision-making. *medRxiv* 2023–12.

39. Ahn W-Y, Haines N, Zhang L (2017): Revealing Neurocomputational Mechanisms of Reinforcement Learning and Decision-Making With the hBayesDM Package. *Computational Psychiatry* 1: 24–57.

40. Ashburner J, Barnes G, Chen C-C, Daunizeau J, Flandin G, Friston K, *et al.* (2021): SPM12 Manual.

41. Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM (2012): FSL. *NeuroImage* 62: 782–790.

42. Chen G, Xiao Y, Taylor PA, Rajendra JK, Riggins T, Geng F, *et al.* (2019): Handling Multiplicity in Neuroimaging Through Bayesian Lenses with Multilevel Modeling. *Neuroinformatics* 17. https://doi.org/10.1007/s12021-018-9409-6

43. Cox RW (1996): AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical research* 29: 162–173.

44. R Core Team (2021): *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

45. Bürkner P-C (2017): brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software* 80: 1–28.

46. Banca P, Voon V, Vestergaard MD, Philipiak G, Almeida I, Pocinho F, *et al.* (2015): Imbalance in habitual versus goal directed neural systems during symptom provocation in obsessive-compulsive disorder. *Brain* 138: 798–811.

47. Robbins TW (2024): The brain's arbitration system and obsessive-compulsive disorder. *Brain* 147: 1929–1930.

48. Yücel M, Harrison BJ, Wood SJ, Fornito A, Wellard RM, Pujol J, *et al.* (2007): Functional and biochemical alterations of the medial frontal cortex in obsessive-compulsive disorder. *Archives of general psychiatry* 64: 946–955.

49. Bogdanov M, Timmermann JE, Gläscher J, Hummel FC, Schwabe L (2018): Causal role of the inferolateral prefrontal cortex in balancing goal-directed and habitual control of behavior. *Sci Rep* 8: 9382.

50. McLaughlin AE, Diehl GW, Redish AD (2021): Potential roles of the rodent medial prefrontal cortex in conflict resolution between multiple decision-making systems. *Int Rev Neurobiol* 158: 249–281.

51. Gruner P, Anticevic A, Lee D, Pittenger C (2016): Arbitration between Action Strategies in Obsessive-Compulsive Disorder. *The Neuroscientist* 22: 188–198.

52. Kasal MI, Besiroglu L, Zorlu N, Dikmeer N, Bilge A, Durmaz E, *et al.* (2021): Fronto-striatal structures related with model-based control as an endophenotype for obsessive–compulsive disorder. *Scientific Reports* 11: 11951.

53. Stein DJ, Allen A, Bobes J, Eisen JL, Figuera ML, Iikura Y, *et al.* (2000): Quality of Life in Obsessive-Compulsive Disorder. *CNS Spectrums* 5: 37–39.

54. Calzà J, Gürsel DA, Schmitz-Koep B, Bremer B, Reinholz L, Berberich G, Koch K (2019): Altered cortico–striatal functional connectivity during resting state in obsessive–compulsive disorder. *Frontiers in Psychiatry* 10: 319.

55. Gillan CM, Apergis-Schoute AM, Morein-Zamir S, Urcelay GP, Sule A, Fineberg NA, *et al.* (2015): Functional Neuroimaging of Avoidance Habits in Obsessive-Compulsive Disorder. *AJP* 172: 284–293.

56. Harrison BJ, Soriano-Mas C, Pujol J, Ortiz H, López-Solà M, Hernández-Ribas R, *et al.* (2009): Altered Corticostriatal Functional Connectivity in Obsessive-compulsive Disorder. *Archives of General Psychiatry* 66: 1189–1200.

57. Milad MR, Rauch SL (2012): Obsessive-compulsive disorder: beyond segregated cortico-striatal pathways. *Trends in Cognitive Sciences* 16: 43–51.

58. Guida P, Michiels M, Redgrave P, Luque D, Obeso I (2022): An fMRI meta-analysis of the role of the striatum in everyday-life vs laboratory-developed habits. *Neuroscience & Biobehavioral Reviews* 141: 104826.

59. Downar J, Geraci J, Salomons TV, Dunlop K, Wheeler S, McAndrews MP, *et al.* (2014): Anhedonia and reward-circuit connectivity distinguish nonresponders from responders to dorsomedial prefrontal repetitive transcranial magnetic stimulation in major depression. *Biological psychiatry* 76: 176–185.

60. Rupprechter S, Romaniuk L, Series P, Hirose Y, Hawkins E, Sandu A-L, *et al.* (2020): Blunted medial prefrontal cortico-limbic reward-related effective connectivity and depression. *Brain* 143: 1946–1956.

61. Yaple ZA, Tolomeo S, Yu R (2021): Abnormal prediction error processing in schizophrenia and depression. *Human Brain Mapping*.

62. Akam T, Costa R, Dayan P (2015): Simple Plans or Sophisticated Habits? State, Transition and Learning Interactions in the Two-Step Task ((J. Daunizeau, editor)). *PLoS Comput Biol* 11: e1004648.

63. Feher da Silva C, Hare TA (2020): Humans primarily use model-based inference in the two-stage task [no. 10]. *Nat Hum Behav* 4: 1053–1066.