Ph.D. Dissertation in Natural Sciences

Imbalanced arbitration between decision-making strategies in obsessive-compulsive disorder:

An integration of neuroscience and computational reinforcement learning

강박장애에서 의사결정 전략 사이의 조율 불균형: 강화학습 모델을 이용한 계산신경과학 연구

August 2021

Graduate School of Natural Sciences
Seoul National University
Brain and Cognitive Sciences Major

Taekwan Kim

Imbalanced arbitration between decision-making strategies in obsessive-compulsive disorder:

An integration of neuroscience and computational reinforcement learning

Jun Soo Kwon

Submitting a Ph.D. Dissertation in Natural Sciences

August 2021

Graduate School of Natural Sciences
Seoul National University
Brain and Cognitive Sciences Major

Taekwan Kim

Confirming the Ph.D. Dissertation written by
Taekwan Kim
August 2021

Chair	정 천 기
Vice Chair	권 준 수
Examiner	이 상 훈
Examiner	이 인 아
Examiner	이 상 완

Abstract

Introduction: Habit bias, resulted from imbalanced arbitration between goal-directed and habitual controls, is thought to underlie compulsive symptoms of patients with obsessive-compulsive disorder (OCD). A computational reinforcement learning (RL) model accounts for that, between the goal-directed (model-based; MB) and habitual (model-free; MF) RL systems, brain allocates weight to a controller with higher reliability in state or reward prediction. However, it remains unclear whether the impaired arbitration in OCD is attributed to faulty estimation of the reliability in the RLs and if inferior frontal gyrus (IFG) and/or frontopolar cortex (FPC), known to track the reliability signals, are grounded on this impairment.

Methods: The sequential two-choice Markov decision task was used to dissociate the MB and MF learning strategies. Thirty patients with OCD and thirty-one healthy controls (HCs) underwent a fMRI scan while performing the behavioral task. Behaviors of the arbitration process were estimated through a computational model based on RL algorithms. The model parameters and their neural estimates were compared between groups. Regression analyses were conducted to examine if neural differences explained faulty estimation of the reliability, in addition to compulsion severity, in OCD.

Results: Patients with OCD earned less reward and showed higher perseveration than HCs. During MB-favored trials, the uncertainty

of prediction based on the MF strategy was lower in patients, which led to higher maximum reliability of the RL systems arbitrating behaviors (i.e., stability of the arbitration) and higher probability to choose the MF strategy. The higher stability of the arbitration was associated with hyperactive signal of the lateral orbitofrontal cortex (OFC)/FPC in patients. Patients increased connectivity strength between the OFC/FPC and precuneus when choosing an action strategy. On the other hand, the hyperactive IFG signal was inversely associated with the lower stability of the arbitration and less severe compulsion in patients.

Conclusions: It was demonstrated that the hyperactive neural arbitrators encoding the excessively stable arbitration in which the MF reliability was predominant underlay the imbalanced arbitration in OCD. Therefore, the findings suggest the IFG and FPC as brain biomarkers useful to plan a neurocircuit-based treatment for habit biases and compulsions of OCD.

Keywords: obsessive-compulsive disorder; decision-making; goal-directed control; habitual control; reinforcement learning; neural arbitrator; task-based fMRI

Student ID: 2015-20480

Table of Contents

List of Figures

List of Tables

Background

Clinical characteristics of obsessive-compulsive disorder

Obsessive-compulsive disorder (OCD) is a psychiatric disorder that affects 1 to 3% of the world's population and provokes severe role impairments to 65% of patients.[1,2] OCD was classified as an anxiety disorder prior to the current diagnostic system using the Diagnostic and Statistical Manual of Mental Disorders (DSM) fifth edition because anxiety is commonly observed in the disorder.[3] However, OCD has been moved to a category of obsessive-compulsive related disorders as anxiety is not the most essential feature of this disorder.[4] Categorical diagnosis of OCD is made based on severity of obsessions and compulsions of patients. By definition, obsessions are persistent unwanted thoughts that are intrusive and causes distress, and compulsions are inflexible repetitive behaviors that are thought to reduce the distress following obsessions or form excessive habits and perseveration (Figure 1).[5,6]

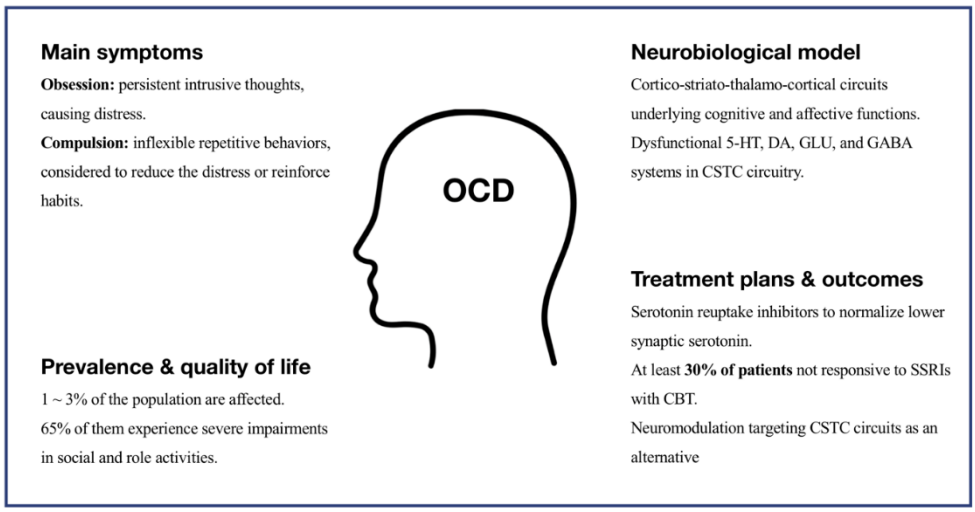

Figure 1. Characteristics of obsessive-compulsive disorder.

Although patients with OCD share common clinical features, they are not homogeneous group. For example, individual patients suffer different profiles of obsessive-compulsive (OC) symptoms across several symptom dimensions: aggressive/checking, contamination/cleaning, symmetry/ordering, and sexual/religious.[7,8] In addition, OCD patients are divided into at least two subtypes depending on age at illness onset.[9] Patients who had illness onset before puberty (early-onset subgroup) have difference of genetic loading, brain structures, and neurocognitive functions, compared to patients with illness onset after young adulthood (late-onset subgroup).[9,10] The aforementioned heterogeneity issue suggests that the current symptom-based classification of OCD does not satisfactorily capture the pathophysiology on an individual basis and

that the diagnostic classification should be complemented by using biological evidence.[10,11]

Theoretical models for OCD symptomatology

There are two main theories that account for symptomatology of OCD. A conventional theory views that obsessions arise first with distress in patients, and compulsions are performed as secondary phenomena to reduce the anxiety that follows obsession and provide relief.[5] This model is based on cognitive bias between values of repeating behaviors and those of switching to alternatives.[12] Patients with OCD may assign higher values to the cost of ceasing repetitive behaviors than the benefits thereof because relief resulting from the compulsions are rewarding to them.

However, there are two main limitations in the conventional view to account for some phenomena of OCD. First, in experimental settings of habit formation, patients with OCD often carry out compulsion-like behaviors in the absence of any prior obsession.[6,13,14] This evidence raises the possibility that compulsive behaviors of OCD could be independent of distress following obsessive thoughts.[15] Second, OCD is considered to be ego-dystonic as patients have difficulty to control urge to action even though they have insight that compulsive behaviors are irrational and want to cease these actions;[16] however, the theory on the basis of the cognitive bias is not enough to explain about the intact insight of patients.[15]

An alternative theory of OCD has been developed to address these concerns. This theory views that habit biases cooperate with trait anxiety (not necessarily) and foster compulsive urges and actions in patients with OCD.[15,17] Patients may experience cognitive dissonance and attempt to resolve this discrepancy between their cognitions and compulsive urges by developing irrational thoughts (i.e., obsessions) about threat or fear of the conflict.[15,18] Interacting with anxiety, the obsessive thoughts reinforce the compulsions through cognitive dissonance, forming a vicious cycle of sorts.[15] The idea that excessive habit formation serves as a key component in this alternative model suggests that disrupted balance between goal-directed and habitual controls might produce the habit biases and be a basis of OCD symptomatology.

Neurocircuitry mechanisms of OCD

Brain dysfunctions of OCD has been explained by the well-accepted cortico-striato-thalamo-cortical (CSTC) circuitry model, which is a neural basis of execution-inhibition balance for cognitive and affective processes.[19-22] In the CSTC circuitry, direct and indirect pathways counterbalance each other to regulate the execution and inhibition of actions.[23] Through the direct pathway, the CSTC loops project striatal Go signal through the internal globus pallidus to the thalamus, subsequently disinhibiting thalamo-cortical excitatory signal to execute actions. On the other hand, the indirect pathway projects striatal NoGo signal through the internal and external

segments of globus pallidus to the thalamus, suppressing the thalamo-cortical signal to inhibit actions.[23,24] In OCD, imbalance of the neurocircuitry is thought to underlie maladaptive shifting of attention to optimal choices.[25]

Multifaceted impairments of cognitive and affective controls in OCD can be explained by several parallel but often interactive CSTC circuits, each of which is connected with cortical regions involved in distinct functions.[21,22] Five parallel circuits are currently proposed for the CSTC model as follows: the fronto-limbic circuit (ventromedial prefrontal cortex [vmPFC]-amygdala) relevant for extinction of emotional response, dorsal cognitive circuit (dorsal prefrontal cortex-dorsal caudate) for working memory and planning, ventral cognitive circuit (ventrolateral prefrontal cortex [vlPFC]/anterolateral orbitofrontal cortex [OFC]-ventral caudate) for response inhibition and cognitive flexibility, ventral affective circuit (medial OFC/vmPFC-nucleus accumbens [NAcc]) for affective and reward processing, and sensorimotor circuit (supplementary motor area [SMA]-putamen) for habit formation.[22]

When generating and/or extinguishing emotional responses, such as fear or uncertain threat, patients have hyperactive vmPFC and amygdala and reduced functional connectivity between these regions,[26-28] suggestive of hyperactive fronto-limbic circuit in OCD.[29] While hypoactive dorsal cognitive circuit (failed recruitment of dlPFC and caudate, in addition to reduced dlPFC-striatal connectivity) is neural basis of impaired working memory and planning in OCD,[30,31] hypoactive ventral cognitive circuit (failed

recruitment of vlPFC/inferior frontal gyrus [IFG] and caudate, in addition to reduced vlPFC–caudate connectivity) underlie impairments in response inhibition and attention shifting.[31–33] However, the anterolateral OFC within the ventral cognitive circuit is hyperactive during symptom provocation and normalized after symptom improvement.[34] Also, the strength of vlPFC–pallidal subnetwork is strengthened as a compensatory neural mechanism for cognitive inflexibility.[25] Similar to the fronto–limbic circuit, the ventral affective circuit includes vmPFC; but, main areas of the affective circuit are reward–circuit regions, such as medial OFC and NAcc. When anticipating reward/punishment or making reward–based decision–making, patients fail to recruit the OFC and/or NAcc and have hyperconnectivity among reward–circuit regions.[35–38] In terms of the sensorimotor circuit, it has been hypothesized (but not empirically tested) to be related to habit–like compulsions of OCD.[15,19] The sensorimotor circuit is considered to become hyperactive in patients with multiple repetitions of habitual behaviors over time, which may play a role in the transition from goal–directed to habitual behaviors.[15]

In addition to the abnormalities within circuits, between–circuit interactions are considered to arise and integrate subprocesses of each circuit in order to execute complex behaviors.[22] This hypothesis was supported by studies showed that higher cognitive functions, such as attention shifting and emotion regulation, involved interactions between dorsal cognitive and ventral cognitive/affective circuits.[25,39]

Treatment strategies and unsatisfactory responses in patients with OCD

As described earlier in the conventional theory, the abnormalities in fear extinction has been recognized as a major pathological problem of OCD.[5,27] Thus, this phenotypic feature has been targeted to develop first-line treatments for OCD; for example, the exposure and response prevention, which is a type of cognitive behavioral therapy (CBT), is used to normalize the dysfunctional fear processing and has responsive up to 70% of patients.[40]

The other first-line treatment for OCD is a pharmacotherapy based on serotonergic dysfunction.[41] OCD is characterized by lower synaptic serotonin (5-HT) in the CSTC circuitry: diminished serotonin transporter (SERT) availability in the thalamus and raphe nuclei, increased 5-HT_{1A} receptor availability in the striatum, and increased 5-HT_{2A} autoreceptor availability in prefrontal areas.[41] The selective serotonin reuptake inhibitors (SSRIs) is, therefore, used to normalize the synaptic 5-HT concentration in the circuitry.[1]

However, the SSRIs are not effective to roughly half of patients.[1] Thus, it raises attention to neurotransmitter systems beyond the serotonin dysfunction in order to better design treatment options for OCD. Neuroimaging studies have revealed striatal dopaminergic (increased transporter density and diminished receptor binding) and glutamatergic (increased neurotransmitter concentration) hyperactivity in the CSTC circuitry of OCD.[41,42]

Additionally, patients also have diminished γ-aminobutyric acid (GABA) level in the medial prefrontal cortex.[43] The above literatures suggest a collective neurochemical model that the diminished inhibitory signals of serotonergic and GABAergic pathways provoke the striatal dopaminergic and glutamatergic hyperactivity, resulting in overactive CSTC circuitry in OCD.[44]

Current issues to be addressed in developing neurobiological evidence-based treatments for OCD

The unsatisfactory treatment outcomes in psychiatry diseases have been pointed out and thought to be attributed to symptom-based categorical diagnoses and phenotypic heterogeneity.[45] In case of OCD, at least 30 % of patients are not responsive to first-line treatments.[46,47] The Research Domain of Criteria (RDoC) initiative was introduced as an alternative to address the heterogeneity issue.[48] The RDoC uses dimensional approach to assess biotypes, both across and within diagnoses, in a basis of underlying biological features and cognitive/behavioral measures, pursuing evidence-based treatments. The RDoC framework consists of diverse domain matrices, such as positive or negative valence system and cognitive system. In particular, constructs in the positive valence system (e.g., learning and valuation of reward) and those in the cognitive system (goal selection and response inhibition) are relevant to pathophysiology of OCD.[49,50]

The dimensional classification based on biological evidence (i.e., biotyping) needs to build obvious biological bases, such as brain biomarkers, that capture clinical and neurocognitive features of the disorder.[48,51] Also, building brain biomarkers and determining biotypes of OCD are expected to help addressing the heterogeneity issue and developing neurocircuit-based treatments.[29,44] There was an effort to apply RDoC-based classification differentiating OCD patients responsive to treatments from treatment-resistant group by using functional brain organization.[52] However, this previous study is not sufficient to guide neurocircuit-based treatments for OCD because of lack of information about brain-behavior relationships.

According to the alternative theory, excessive habit formation is a key component in the vicious cycle of compulsions and obsessions. Importantly, imbalance between goal-directed and habitual controls is thought to produce the habit biases and underlies the symptomatology. However, it remains unclear how the arbitration of decision-making strategies becomes imbalanced in OCD. Therefore, this research aimed to reveal deficient cognitive component in the arbitration process, build brain biomarkers underlying the imbalanced arbitration between two strategies of decision-making, and provide biological evidence to guide a neurocircuit-based treatment for OCD impairments, especially habit bias. Considering the neurocircuitry model of OCD, it is expected that the imbalanced arbitration may be related to imbalanced competition between the direct and indirect pathways of CSTC

circuitry. The impaired arbitration in OCD may also involve the ventral cognitive circuit that inhibits repetitive habitual behaviors and the dorsal cognitive circuit that controls goal-directed planning.

Chapter 1. Reliability-based competition between model-based and model-free learning strategies in OCD

Introduction

Compulsive symptoms of OCD are translated into inflexible and habitual behaviors that result in biases toward habits.[13] In addition to the excessive habits, patients have difficulty to use forward models of action–outcome scenarios to perform goal–directed behaviors.[6,14] As optimal decision–making can be achieved by adaptive integration of goal–directed and habitual learning processes, the deficits in two action control systems should be discussed along the same lines in OCD.[53,54] In a recent study, it has been demonstrated that patients with OCD are more likely to employ habitual learning system, at expense of goal–directed planning, when facing decision problems.[55] The above literatures have established a theory of imbalanced arbitration between goal–directed and habitual decision–making strategies, which explains compulsive behaviors of OCD. In addition, the imbalanced arbitration of decision–making is considered as a useful source for dimensional biotyping of OCD because the arbitration process comprises the positive valence (e.g., reward–based learning) and cognitive control (inhibition and planning) systems of the RDoC matrix[56,57].

Although the theory of imbalanced arbitration provides insight into symptomatology of OCD,[13,54,55,58] there is lack of understanding about which cognitive component in the arbitration process provokes the impaired arbitration. This issue can be dealt with using a computational model of the arbitration process. In the computational framework, goal–directed and habitual behaviors are

modeled as two different classes of reinforcement learning (RL) that updates action values from prediction errors.[57,59] Goal-directed behaviors can be expressed by a model-based (MB) RL algorithm in which agents compute action values using an internal model about state-transition and outcome structure of the environment. On the other hand, a model-free (MF) RL algorithm that learns cached action values directly from trial and error experience without building an explicit model of the environment describes habitual behaviors.[57] A theoretical study of the dual-action choice systems suggests that a competition between the two learning systems is arbitrated according to uncertainty of each controller; each controller is deployed when its learning system is most accurate.[57] Furthermore, a recent computational model has empirically demonstrated that human brain evaluates reliability of each learning strategy and uses this information to proportionately allocate degree of behavioral control of each system.[60] For instance, if the reliability of MF learning strategy is more predominant than that of MB learning strategy, brain is likely to produce greater reliance on habitual system than goal-directed system. In addition, stability of the arbitration between the two RL systems can be measured by the reliability of whichever strategy that provides more accurate predictions (i.e., maximum reliability).[60,61] Therefore, this computational model, which exhibits details of the arbitration process, is considered to be useful to test whether the reliability estimation process is a cognitive basis of the imbalanced arbitration that is biased toward habitual system in OCD.

In the first chapter of dissertation, it was aimed to test whether the impaired arbitration in OCD is attributed to faulty estimation of the reliability in the RL systems. It was hypothesized that the arbitration of dual-action systems is imbalanced in OCD patients because their arbitration is excessively stable (i.e., higher maximum reliability) in which MF learning strategy is predominant (i.e., higher MF reliability).

Methods

Participants

Thirty-one patients with OCD were recruited via the OCD clinic in Seoul National University Hospital (SNUH). Licensed psychiatrists made the diagnosis of OCD using the Structured Clinical Interview for DSM-IV Axis-I Disorder (SCID-I), patient edition.[62] OC symptoms and accompanying anxiety and depression of patients were assessed using the Yale-Brown Obsessive Compulsive Scale (Y-BOCS) and Hamilton Rating Scales for Anxiety and Depression (HAM-A/D).[63-65] Clinicians in the OCD clinic introduced this study to outpatients who had no current comorbid psychotic or movement disorders. Patients who suffered moderate or more severe OC symptom (Y-BOCS total ≥ 15) at the time of study and had no CBT history in recent 1 year were asked about their willingness to participate in this study. Patients taking more than two antidepressants, high dose of antipsychotics, or a mood stabilizer at the time of study were not included. The procedures for recruitment and experiment of OCD patients are presented in Figure 2.

The study also recruited thirty-one healthy controls (HCs) through internet advertisements, and healthy participants were screened using the SCID-I, non-patient edition.[66] Exclusion criteria was set as follows: a lifetime history of psychotic or neurological disorders, substance use disorder, mental retardation, or incomplete

data collection. After excluding one sample with missing behavioral data, 30 patients and 31 HCs from the original dataset remained. All participants were informed of a complete description of the experimental procedures and provided written informed consent. This study was approved by SNUH Institutional Review Board (IRB No. H-1908-208-1066) and performed in accordance with ethical guidelines of the Declaration of Helsinki.

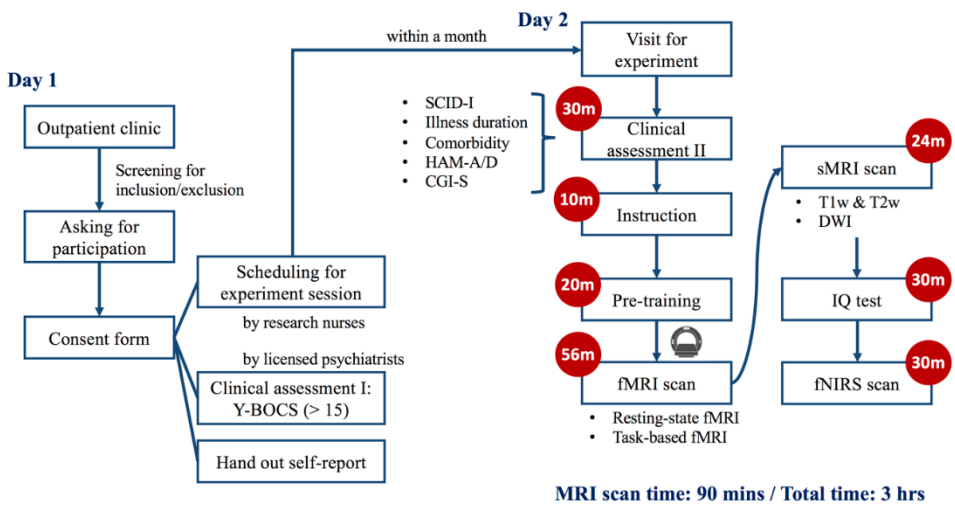

Figure 2. Flowchart of the recruitment and experimental procedures for patients with obsessive-compulsive disorder.

Behavioral task

Participants performed the sequential two-choice Markov decision task, which was designed to examine behavioral arbitration between MB (goal-directed) and MF (habitual) learning strategies (Figure 3).[60] Participants made two sequential choices (pressing left or

right button) to gain a coin valued 40, 20, or 10 scores (10 scores equals to 200 KRW) at the reward stage. If making no choice in four seconds, a random choice was made followed by a monetary penalty for this trial. Participants were instructed to collect as many coins as possible and learn about associations between states and outcomes to find optimal choices. Each trial of the task started at the same starting state. A choice at the initial stage led to one of four states (two states exist per a choice option; one of the two states is selected by a certain state–transition probability) at the second stage. In the same manner, a choice from a state of the second stage led to one of four reward states (total 16 reward states). Participants were informed that the contingencies might change during the experiment without notice about the state–transition probability. The state–transition probability changed between high uncertainty (P = (0.5, 0.5) when MF system is favored) and low uncertainty (P = (0.9, 0.1) when MB system is favored) conditions, which dissociates the two behavioral control systems.

The task design also incorporated specific and flexible goal conditions to observe MB and MF behaviors, respectively.[60] In the specific goal condition, participants needed to guide their behaviors to obtain a specific coin whose color was matched to a presented collecting box (red, yellow, or blue) at the trial. If a color of coin did not match with the collecting box color, participants failed to obtain a monetary reward of that coin. By changing the goal–state (collecting box color) every few trials, the specific goal condition

made participants actively consider about currently valuable goal. In contrast, participants in the flexible goal condition were presented with a white collecting box and allowed to collect any color of coin. In association with the both goal and uncertainty conditions, the task dissociates the goal-directed system (MB-favored block; specific goal with low uncertainty) from the habitual system (MF-favored block; flexible goal with high uncertainty). The task randomized the sequence of total four blocks, including two other blocks with arbitrary conditions (specific goal with high uncertainty or flexible goal with low uncertainty).

There was a uniform-distributed temporal interval between 1 to 4 seconds when preceding to next stage or trial, and the duration of reward stage was two seconds. The experimental task comprises six sessions (40 trials on average per session). Prior to the experiment sessions, participants performed one pre-training session of 100 trials, which have been demonstrated as enough time for participants to learn about the rules of the two-choice decision task.[67] The pre-training session was designed to help learning about the task with free decision-making strategy; the state-transition probability was fixed at high uncertainty condition with flexible goal condition for first 80 trials and specific goal condition for the rest of the session.

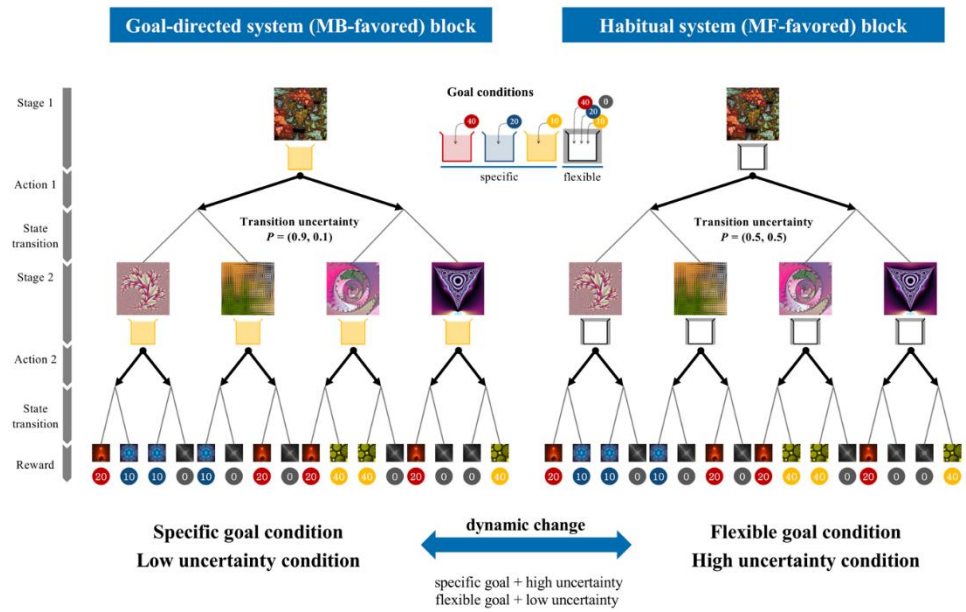

Figure 3. The paradigm of sequential two-choice Markov decision task. In the sequential two-choice decision task, participants needed to perform goal-directed behaviors in MB-favored block (specific goal with low uncertainty) while it was preferred to perform habitual behaviors in MF-favored block (flexible goal with high uncertainty). In every few trials, block types were dynamically changed to make participants arbitrate their behavioral strategy for the purpose of optimal decision outcomes.

Computational model of decision arbitration system

Lee et al. (2014) developed a computational model that arbitrates MB and MF learning strategies depending on reliability of each controller.[60] This arbitration model updates reliability of each MB and MF learning by Bayesian estimation and Pearce-Hall

associability rule respectively and incorporates model bias toward MF system by programming biophysical two-state transition; the model was named "mixedArb-dynamic" and was better fitted to explain the arbitration process than other versions, such as dual Bayesian model.[60] In this study, the "mixedArb-dynamic" model was modified to update the MF reliability differently depending on sign of reward prediction error (RPE) signal (please see below for details). The computational model contains six free parameters: the threshold for defining zero state prediction error (SPE), learning rate of the model estimating absolute RPE, the amplitude of a transition rate function (MB → MF), the amplitude of a transition rate function (MF → MB), the level of stochastic decision-making (inverse softmax temperature), and learning rate of the MB and MF systems.[60] To optimize the free parameters, model fitting process was iterated one hundred times with randomly generated seed parameters in each run.

Following the previously established computational framework,[60] the arbitration model estimated parameters in following steps (Figure 4): reinforcement learning (state-action values corresponding to SPE and RPE), reliability estimation (reliability of learning strategies and stability of the arbitration), reliability competition (model-choice probability), and value integration (state-action value of the arbitration).

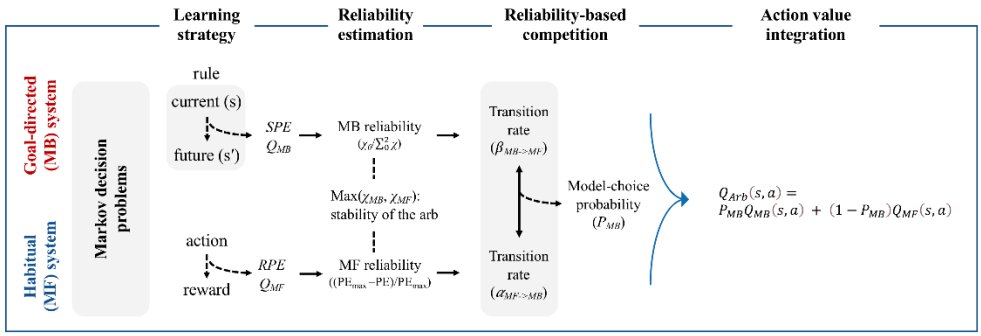

Figure 4. The computational framework estimating model parameters that are used to account for the reliability-based arbitration of RL systems. Between MB and MF RL systems, brain allocates weight to a controller with higher reliability in state or reward prediction. The computational model estimated model parameters in a sequence of four steps. At RL step, prediction errors were estimated while updating state-action values. Based on the prediction errors, reliability of the predictions was estimated in each RL system. In a dynamical two-state transition model, the reliability values of MB and MF RL systems calculated model-choice probability. This computational model finally estimated state-action value of the arbitration integrating the action values of the two RL systems in a weighted average manner by the model-choice probability.

First, the arbitrator computed state-action values of the MB and MF systems by using prediction errors of each strategy. By employing the MB learning algorithm with the FORWARD learning and BACKWARD planning components, a state-action-state transition probability (T(s, a, s')) and the corresponding state-action value (Q_{MB}) were updated based on the state prediction error

(SPE, δ_{SPE}).[67] The BACKWARD planning was applied whenever an agent was presented with an explicit goal (e.g., change in a specific goal condition or transition from the flexible to specific goal condition). The MF learning was expressed by the SARSA algorithm that computed reward prediction error (RPE) and updated the corresponding state–action value (Q_{MF}).[59]

in case of MB reinforcement learning:

FORWARD learning,

$$\delta_{SPE} = 1 - T(s,a,s'),$$

$$\Delta T(s,a,s') = \eta\delta_{SPE},$$

$$Q_{MB}(s,a) = \sum_{s'} T(s,a,s')\{r(s') + max_{a'} Q_{MB}(s',a')\}.$$

BACKWARD planning,

for $i = 3, 2$

for $s \in S_{i-1}$

$Q_{MB}(s,a) = \sum_{s'} T(s,a,s')\{r(s_i) + max_{a'} Q_{MB}(s',a')\}.$

end

end

in case of MF reinforcement learning:

$$\delta_{RPE} = r(s') + \gamma Q_{MF}(s',a') - Q_{MF}(s,a),$$

$$\Delta Q_{MF}(s,a) = \alpha\delta_{RPE}.$$

where a and a' refer to the action in the current (s) and the next state (s'), r(s') indicates the earned reward value at s' state, η and α denote the learning rate of MB and MF, and γ refers to the temporal discount factor.

Second, the arbitrator estimated reliability of each learning strategy by determining the extent to which the prediction errors are high or low. A hierarchical Bayesian model with inverse Fano factor was

used to estimate the reliability of MB learning strategy given the SPE history.[60] Because the SPE is always positive, less erroneous prediction produced higher reliability of the system.

$$E(\theta_j|D) = \frac{1 + \#PE_j}{3 + |D|}, j = 0, 1, 2,$$

$$Var(\theta_j|D) = \frac{(1 + \#PE_j)(2 + \sum_{j \neq i} \#PE_i)}{(3 + |D|)^2(4 + |D|)}, j = 0, 1, 2,$$

$$\chi_j = \frac{E(\theta_j|D)}{Var(\theta_j|D)}, j = 0, 1, 2,$$

$$\chi_{MB} = \chi_0 / \sum_{i=0}^{2} \chi_j.$$

where θ_0, θ_1, θ_2 represents the probability of making zero, negative, or positive prediction error (PE), $\#PE_j$ refers to the number of events that cause PE_j, $|D| = \sum_{j=0}^{2} \#PE_j$, the cardinality of a set (D) of events causing prediction errors, E and Var denote the expectation and variance of the posterior $\theta|D$, and χ_{MB} is the reliability of MB learning strategy.

For the MF learning, the absolute RPE estimate was computed in the same manner with the previous study.[60] The Pearce–Hall associability rule was used to estimate the reliability;[68] however, it depended on positive or negative sign of RPE. In case of positive RPE, it was assumed that agents would decrease the reliability of current learning strategy and facilitate explorative learning (i.e., transition to MB learning strategy) because of an unexpected positive reward following an erroneous prediction based on the current strategy. In case of negative RPE, the reliability of current learning strategy would be increased, which hampers the explorative learning, because 1) expected negative rewards

following erroneous predictions do not mean that the current learning strategy is unreliable and because 2) this prediction–reward pair rather may give hints about correct choices.[69]

$$\Delta\Omega = \eta(|RPE| - \Omega),$$
$$\chi_{MF} = \begin{cases} (RPE_{max} - \Omega)/(2 \times RPE_{max}), & RPE > 0 \\ (RPE_{max} + \Omega)/(2 \times RPE_{max}), & RPE \leq 0 \end{cases}.$$

where Ω is the absolute RPE estimator, η refers to the learning rate of the model estimating Ω, RPE_{max}, the upper bound of RPE, equals to 40, and χ_{MF} is the reliability of MF learning strategy. Stability of the arbitration between MB and MF systems can be measured by the reliability of whichever strategy that provides more accurate predictions (i.e., maximum reliability, $max(\chi_{MB}, \chi_{MF})$).[60,61]

Third, the reliability signals of learnings were used as transition rates to compute the probability of choosing MB (P_{MB}) or MF strategy ($1-P_{MB}$) in a dynamical two–state transition model, inspired by biophysical neuronal models.[70]

$$\alpha(\chi_{MF}) = A_\alpha/(1 + e^{B_\alpha \chi_{MF}}),$$
$$\beta(\chi_{MB}) = A_\beta/(1 + e^{B_\beta \chi_{MB}}),$$
$$\frac{dP_{MB}}{dt} = \alpha(1 - P_{MB}) - \beta P_{MB}.$$

where α is the MF → MB transition rate, β is the MB → MF transition rate, A, B represents the maximum transition rate and the steepness of each learning model, and P_{MB} is the probability of choosing MB model, which is equals to $1-P_{MF}$.

Finally, the computational model estimated the state−action value of arbitration (Q_{Arb}) integrating Q_{MB} and Q_{MF} signals in a weighted average manner by the arbitrator weight, P_{MB}.[67]

$$Q_{Arb}(s,a) = P_{MB}Q_{MB}(s,a) + (1 - P_{MB})Q_{MF}(s,a).$$

Given the value signal of the arbitrator, this model stochastically selects an action according to the softmax function.[67]

Statistical analyses

Demographic data and the model parameters of behavioral arbitration were compared by using the chi−squared test or two−sample *t*−test between OCD patients and HCs, while the Mann−Whitney *U* test was used for the observed behaviors to address non−normality issue. Behavioral variables were separately tested in each MB−favored (specific goal with low uncertainty) or MF−favored (flexible goal with high uncertainty) block. Because of insufficiency of the experimental blocks to reveal model differences on choice consistency, this behavioral measure was separated depending on whether the arbitrator predicts predominantly MB ($P_{MB} > 0.5$) or MF strategy ($P_{MB} < 0.5$).[60] All the above analyses were performed using the statistical functions in the SciPy library (www.scipy.org/).

Correlations between the observed behaviors and the estimated model parameters were examined in patients to demonstrate how altered model parameters would account for impaired decision−making performance.

Results

Demographic and clinical characteristics

Demographic variables, including intelligence quotient (IQ) and education year, were comparable between patients and HCs. Patients suffered moderate severity of OC symptom (Y-BOCS$_{Total}$ = 22.03 ± 5.28). Details of demographic and clinical information are presented in Table 1.

Table 1. Demographic, clinical, and behavioral characteristics

Variables	OCD (N = 30)	HCs (N = 31)	Statistic	*p* value
Demographic information				
Age, year	27.50 ± 6.53	24.97 ± 4.64	t = 1.74	0.088
Male/female	18/12	13/18	$\chi 2$ = 1.99	0.158
Handedness (left/right)	2/28	4/27	$\chi 2$ = 0.67	0.414
IQ	113.00 ± 13.55	112.77 ± 8.72	t = 0.08	0.938
Education, year	14.52 ± 1.78	15.00 ± 1.95	t = −0.98	0.333
Clinical outcomes				
Y-BOCS score				
Total	22.03 ± 5.28	–	–	–
Obsession	11.27 ± 3.17	–	–	–
Compulsion	10.77 ± 2.85	–	–	–
HAM-A score	6.59 ± 4.79	–	–	–
HAM-D score	7.34 ± 6.01	–	–	–
Observed behaviors				
Mean reward				

	MB-favored**	12.41 ± 4.67	15.28 ± 3.37	U = 291.0	0.006
	MF-favored	17.79 ± 3.42	17.37 ± 2.98	U = 430.0	0.309
Hit rate					
	MB-favored**	0.51 ± 0.19	0.65 ± 0.13	U = 271.0	0.003
	MF-favored	0.63 ± 0.08	0.59 ± 0.06	U = 376.0	0.101
Choice consistency					
	MB-favored**	0.64 ± 0.07	0.59 ± 0.07	U = 285.0	0.005
	MF-favored	0.90 ± 0.13	0.87 ± 0.13	U = 412.0	0.224

OCD: obsessive-compulsive disorder; HCs: healthy controls; IQ: intelligence quotient; Y-BOCS, Yale-Brown Obsessive Compulsive Scale; HAM-A, Hamilton Anxiety Rating Scale; HAM-D, Hamilton Depression Rating Scale; MB-favored: within trials favoring model-based system; MF-favored: within trials favoring model-free system

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Behaviors biased toward habits and inefficient decision-making

Patients had smaller amounts of mean reward (U = 291.0, p = 0.006) and hit rate (U = 271.0, p = 0.003) and higher choice consistency at the initial stage (U = 285.0, p = 0.005) than HCs when MB system was favored while these performances were comparable between groups during MF-favored trials (Table 1).

Regarding free parameters of the arbitration model (Table 2), patients had greater amplitude of the MB → MF transition rate function (t = 3.18, p = 0.003) and lower level of the stochastic decision-making (t = −2.45, p = 0.017) than HCs. Group

differences of the computed model parameters are presented in Table 3, which showed how the arbitration behaviors are altered compared to HCs. Patients had greater negative RPE (t = −3.08, p = 0.003) and higher reliability of MF learning strategy (t = 2.67, p = 0.010) than HCs during MB−favored trials while SPE and reliability of MB learning strategy were comparable between groups (Figure 5). Also, patients showed more stable arbitration between learning strategies (t = 2.63, p = 0.011) than HCs when MB system was favored (Figure 6). During MF−favored trials, all the computed parameters except lower MF learning reliability in patients (t = −2.81, p = 0.007) were comparable between groups. As a result of the reliability−based competition, patients chose MF learning strategy more often than HCs when MB system was preferred (lower P_{MB}; t = −2.63, p = 0.012) but similarly when MF system is favored (t = −1.81, p = 0.075). When combining the state−action values with the arbitration weight, patients showed lower Q_{Arb} than HCs (t = −2.26, p = 0.028) during MB−favored trials (Figure 7).

As shown in Figures 3, 4, & 5, the lower hit rate of patients was correlated with higher stability of the arbitration (r = −0.52, p = 0.003), higher reliability of MF strategy (r = −0.95, p < 0.001), and lower action value of the arbitration (r = 0.82, p < 0.001), when MB system was favored. Likewise, the hit rate of HCs was correlated with MF reliability (r = −0.91, p < 0.001), arbitration stability (r = −0.42, p = 0.020), and action value of the arbitration (r = 0.70, p < 0.001) during MB−favored trials.

Table 2. Estimated free parameters of the arbitration model

Subjects	Free parameters[a]					
	1	2	3	4	5	6
OCD 01	0.331	0.019	10.000	4.183	0.053	0.200
OCD 02	0.795	0.308	0.102	9.310	0.164	0.011
OCD 03	0.394	0.099	0.122	2.936	0.149	0.111
OCD 04	0.337	0.032	0.416	1.353	0.010	0.034
OCD 05	0.779	0.219	0.122	2.636	0.166	0.010
OCD 06	0.662	0.310	9.583	9.419	0.132	0.094
OCD 07	0.621	0.010	0.100	4.902	0.152	0.196
OCD 08	0.785	0.182	0.124	0.894	0.137	0.197
OCD 09	0.358	0.132	3.746	0.114	0.037	0.198
OCD 10	0.300	0.049	9.998	0.116	0.141	0.141
OCD 11	0.790	0.347	9.754	9.769	0.023	0.194
OCD 12	0.628	0.056	0.115	2.394	0.099	0.130
OCD 13	0.433	0.035	0.101	0.145	0.123	0.076
OCD 14	0.441	0.324	0.108	9.595	0.097	0.200
OCD 15	0.630	0.233	0.419	8.907	0.080	0.013
OCD 16	0.793	0.089	7.433	6.045	0.010	0.010
OCD 17	0.300	0.021	0.157	8.641	0.118	0.187
OCD 18	0.677	0.350	9.972	9.882	0.047	0.200
OCD 19	0.300	0.287	9.995	0.100	0.034	0.026
OCD 20	0.364	0.185	0.137	3.496	0.105	0.145
OCD 21	0.378	0.075	8.269	0.425	0.060	0.179
OCD 22	0.587	0.021	0.117	0.320	0.056	0.199
OCD 23	0.406	0.145	0.107	0.106	0.193	0.200
OCD 24	0.431	0.191	0.114	5.080	0.063	0.088
OCD 25	0.702	0.200	9.984	9.326	0.024	0.046
OCD 26	0.650	0.275	0.118	7.687	0.075	0.200
OCD 27	0.371	0.055	0.112	8.476	0.155	0.200
OCD 28	0.302	0.052	0.158	5.891	0.010	0.199
OCD 29	0.603	0.195	0.135	6.332	0.121	0.165

OCD 30	0.363	0.199	0.125	8.430	0.065	0.199
HC 01	0.325	0.039	0.102	0.112	0.163	0.200
HC 02	0.399	0.032	0.138	0.771	0.120	0.045
HC 03	0.342	0.295	0.104	9.615	0.056	0.198
HC 04	0.309	0.015	0.117	1.709	0.108	0.195
HC 05	0.657	0.298	0.107	7.663	0.154	0.030
HC 06	0.517	0.175	0.107	4.643	0.217	0.067
HC 07	0.413	0.032	0.217	0.226	0.045	0.200
HC 08	0.492	0.103	4.581	2.708	0.024	0.049
HC 09	0.405	0.078	0.108	9.721	0.052	0.119
HC 10	0.335	0.059	5.734	3.024	0.131	0.092
HC 11	0.749	0.045	0.102	7.702	0.076	0.093
HC 12	0.340	0.271	0.101	6.818	0.042	0.175
HC 13	0.797	0.024	0.113	6.020	0.083	0.195
HC 14	0.604	0.066	0.122	5.228	0.116	0.200
HC 15	0.458	0.204	0.106	2.876	0.178	0.172
HC 16	0.449	0.193	0.112	3.133	0.181	0.157
HC 17	0.796	0.327	0.104	7.454	0.032	0.198
HC 18	0.799	0.341	0.118	10.000	0.323	0.142
HC 19	0.443	0.281	0.129	9.995	0.154	0.200
HC 20	0.357	0.047	0.112	2.996	0.285	0.081
HC 21	0.552	0.152	0.111	3.747	0.295	0.038
HC 22	0.506	0.057	0.109	1.435	0.130	0.134
HC 23	0.783	0.041	0.150	6.590	0.125	0.010
HC 24	0.418	0.152	0.120	0.489	0.013	0.199
HC 25	0.309	0.066	0.115	9.033	0.140	0.200
HC 26	0.637	0.222	0.108	0.100	0.220	0.071
HC 27	0.795	0.074	0.103	9.157	0.111	0.200
HC 28	0.300	0.014	0.110	5.209	0.252	0.019
HC 29	0.362	0.342	0.105	8.852	0.115	0.195
HC 30	0.591	0.261	0.111	7.310	0.109	0.199
HC 31	0.354	0.171	0.143	0.107	0.079	0.200

OCD: obsessive–compulsive disorder; HC: healthy control

[a] Parameter 1: the threshold for defining zero state prediction error, (2) learning rate of the model estimating absolute reward prediction error, (3) the amplitude of a transition rate function (MB → MF), (4) the amplitude of a transition rate function (MF → MB), (5) the level of stochastic decision-making (inverse softmax temperature), and (6) learning rate of the MB and MF systems

Table 3. Computational parameters accounting for behavioral arbitration between model-based and model-free systems

Model parameters	OCD (N = 30)	HCs (N = 31)	*t* statistic	*p* value
Prediction error				
SPE				
MB-favored	0.38 ± 0.06	0.38 ± 0.04	0.21	0.832
MF-favored	0.52 ± 0.02	0.51 ± 0.02	0.31	0.758
RPE				
MB-favored**	−1.00 ± 3.86	1.58 ± 2.51	−3.08	0.003
MF-favored†	1.06 ± 2.83	−0.29 ± 2.76	1.88	0.065
Reliability of learning strategies				
Reliability of MB learning				
MB-favored	0.23 ± 0.09	0.21 ± 0.10	0.92	0.362
MF-favored	0.13 ± 0.09	0.13 ± 0.07	0.07	0.948
Reliability of MF learning				
MB-favored*	0.23 ± 0.06	0.20 ± 0.04	2.67	0.010
MF-favored**	0.21 ± 0.04	0.24 ± 0.03	−2.81	0.007
Stability of arbitration (maximum reliability)				
MB-favored*	0.36 ± 0.05	0.32 ± 0.06	2.63	0.011
MF-favored	0.31 ± 0.04	0.32 ± 0.04	−1.41	0.165

Arbitration weight and action value				
Model−choice probability (P_{MB})				
MB−favored*	0.76 ± 0.06	0.79 ± 0.03	−2.63	0.012
MF−favored	0.30 ± 0.13	0.35 ± 0.10	−1.81	0.075
State−action value of arbitration (Q_{Arb})				
MB−favored*	4.96 ± 3.50	6.76 ± 2.66	−2.26	0.028
MF−favored	8.85 ± 5.27	9.67 ± 5.23	−0.61	0.547

MB−favored: within trials favoring model−based system; MF−favored: within trials favoring model−free system; SPE: state prediction error; RPE: reward prediction error

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; † marginally significant level

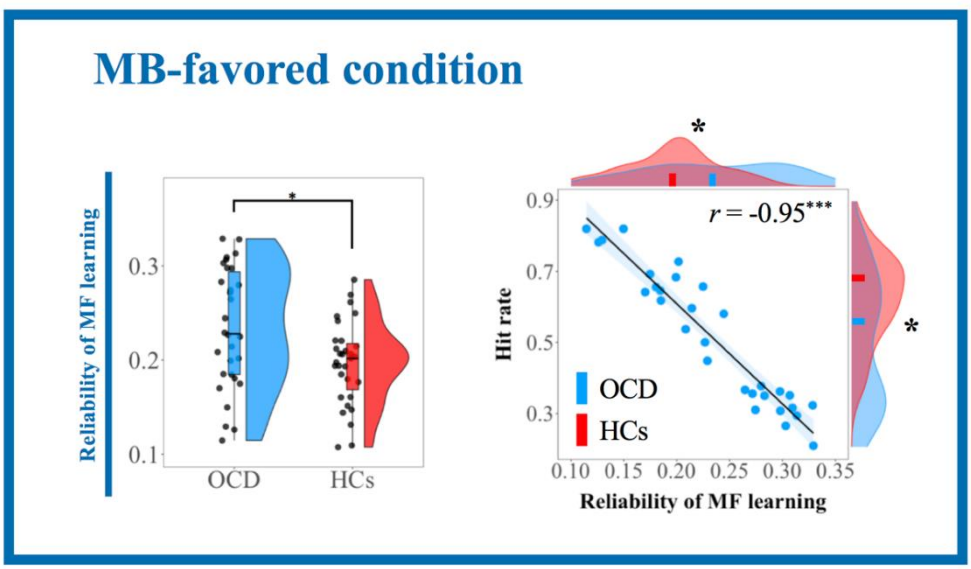

Figure 5. Higher reliability of MF learning in OCD patients. Compared to HCs, patients with OCD had higher reliability of MF learning during MB−favored trials, which was associated with lower hit rate in the sequential two−choice decision task.

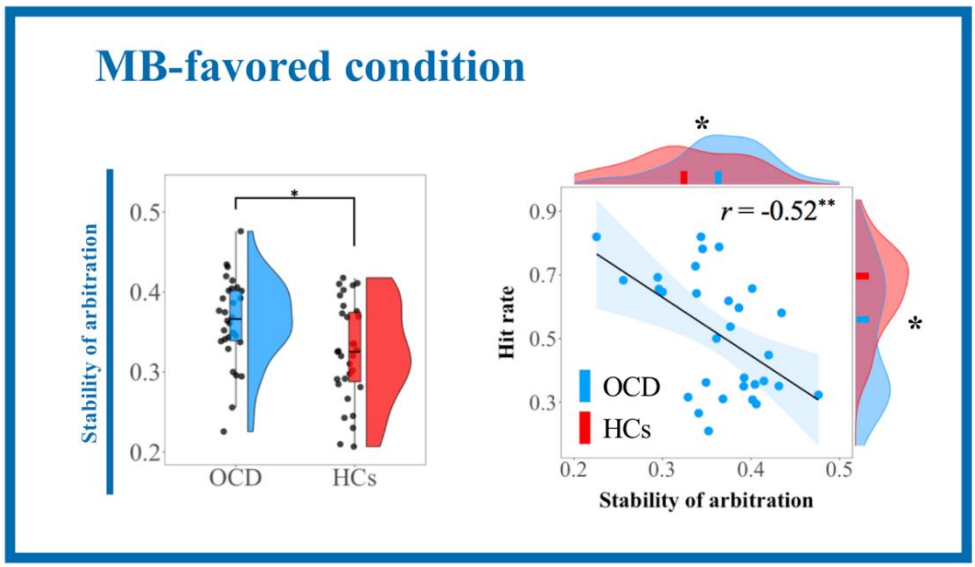

Figure 6. Higher stability of the arbitration in OCD patients.

Compared to HCs, patients with OCD had higher stability of the arbitration during MB-favored trials, which was associated with lower hit rate in the sequential two-choice decision task.

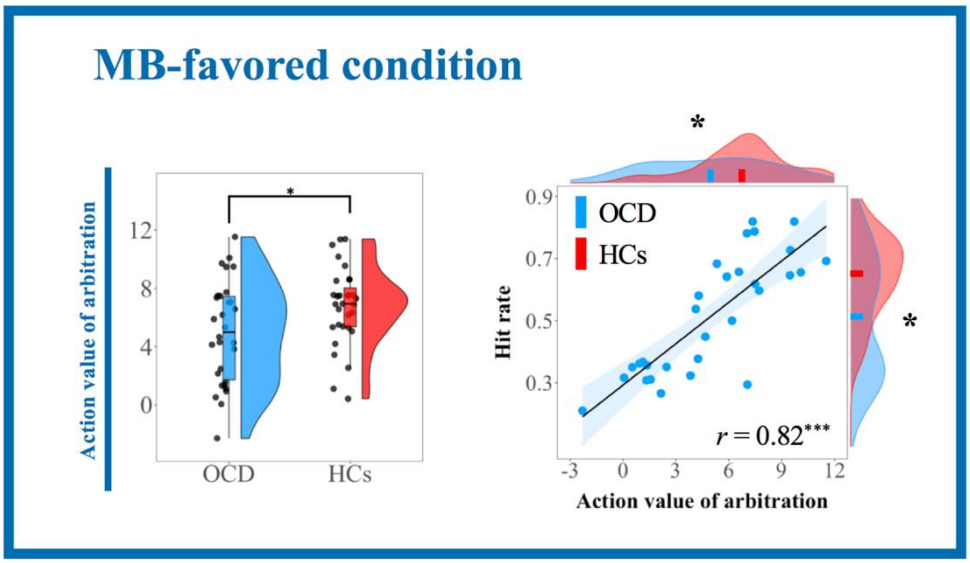

Figure 7. Lower state−action value of the arbitration in OCD patients.

Compared to HCs, patients with OCD had lower state−action value of the arbitration during MB−favored trials, which was associated with lower hit rate in the sequential two−choice decision task.

Discussion

Imbalance between goal-directed and habitual controls is considered to produce habit biases, which is a key component in vicious cycle of OC symptoms. To reveal deficient cognitive component underlying the imbalanced decision-making in OCD, this study applied the computational model of reliability-based arbitration between MB (goal-directed) and MF (habitual) learning strategies. Consistent with the hypothesis, OCD patients had higher reliability of MF learning and more stably arbitrated the learning systems during MB-favored trials. In addition, the atypical arbitration and MF learning reliability was associated with lower performance of decision-making in patients.

According to the dual-action choice theory, human brain computes prediction error signals from the dichotomous MB and MF systems and chooses a system having less prediction uncertainty and higher reliability.[57,60] From this view, the higher reliability of MF strategy accounts for biases toward habitual system in OCD. The higher stability of the arbitration then explains whether the decision arbitration is inflexibly fixed at a certain behavioral strategy.[60,61] Therefore, the excessively stable arbitration that heavily relies on MF system is considered to be the key cognitive factor underlying the imbalance between goal-directed and habitual actions in OCD. Because the reliability differences were observed during MB-favored trials, it is worth to note that patients have the

difficulty to achieve balanced arbitration when goal-directed planning is required.

Chapter 2. Aberrant neural arbitrators underlying the imbalanced arbitration between decision-making strategies in OCD

Introduction

The final aim of this doctoral research was to build brain biomarkers underlying the imbalanced arbitration and provide biological evidence to guide a neurocircuit-based treatment for difficulties (especially, habit bias and compulsions) in OCD. Through the behavioral study described in Chapter 1, the excessively stable arbitration with predominant MF learning strategy was suggested as a crucial cognitive component underlying the imbalanced arbitration in OCD. In this second chapter of dissertation, it was aimed to understand how the impaired arbitration process is mapped in OCD brain and provide it as brain biomarkers to be targeted in future neurocircuit-based treatments.

From a computational neuroscience study of the reliability-based arbitration, the inferior frontal gyrus (IFG) and frontopolar cortex (FPC) were identified as the neural arbitrators that encode reliability signals of MB and MF learnings and stability signal of the arbitration (i.e., maximum reliability).[60] The neural arbitrators are involved in implementing the arbitration process between the MB and MF learning systems and transmit reliability signals of the controllers to other brain areas encoding action values, including dorsal striatum, to guide behavioral choices.[60] The brain regions participating in the reliability-based arbitration process have been frequently observed as neural substrates of excessive habit formation, cognitive inflexibility, and clinical symptoms of OCD.[25,34,55,71] Especially, FPC area at the lateral OFC is abnormally

hyperactive during symptom provocation and has reduced global connectivity in OCD.[34,72] The IFG is hypoactive during response inhibition and coupled with the hyperactive caudate when forming excessive habits in OCD.[32,71] However, it has not been tested if the imbalanced arbitration in OCD is attributed to dysfunctions of the neural substrates involved in the reliability-based arbitration process. Based on evidence of the neural dysfunctions in OCD, it was hypothesized that hyperactive lateral OFC/FPC and hypoactive IFG are impaired neural arbitrators and underlie the excessively stable arbitration with predominant MF learning strategy in OCD.

Methods

Participants

This neuroimaging study used the data from the same participants (30 patients with OCD and 31 HCs) of the behavioral study (see participants section of Chapter 1). This second study was approved by SNUH Institutional Review Board (IRB No. H-1908-208-1066) and performed in accordance with ethical guidelines of the Declaration of Helsinki.

fMRI data acquisition and preprocessing

Functional and structural neuroimaging data were acquired at Seoul National University Hospital using a 3T magnetic resonance imaging (MRI) scanner (Siemens Magnetom Trio), equipped with a 32-channel head coil. Both T1-weighted (T1w) and T2-weighted (T2w) MRI images with submillimeter voxel dimension were obtained for the purpose of higher accuracy of brain segmentation and registration. Each T1w and T2w MRI data was acquired with repetition time [TR]/echo time [TE] = 2,400/2.19 or 3,200/565 ms respectively, while applying common parameters of 0.85 mm isotropic voxel dimension and 320 slices. Task-based functional MRI (fMRI) data were scanned for six sessions using a gradient echo-planar imaging (EPI) pulse sequence (TR/TE = 1,500/30 ms, flip angle = 85° , phase encoding [PE] direction = left-right, 2.3

mm isotropic voxels, 66 slices, multiband x3) for max 9 min per session. To correct EPI distortion of fMRI data, a pair of blip images with PE directions in left−right and right−left was acquired.

The neuroimaging data were preprocessed following the HCP pipeline.[73] Non−brain tissues were extracted using a FNIRT−based masking method after aligning T1w and T2w images to the AC−PC line. To reduce a magnetic susceptibility−induced distortion, bias field was corrected for the structural images, which were used for a structural−functional coregistration step. For preprocessing the functional images, the fMRI timeseries data were realigned for head motion correction using rigid body transformation to the first volume reference. The motion parameters were provided as well for nuisance regression in first−level analysis following the preprocessing step. Because EPI images are susceptible to magnetic field inhomogeneity effect in a phase encoding direction, the EPI distortion was corrected by using the pair of blip images with reversed PE directions in the FSL topup.[74] The distortion−corrected functional images were coregistered to the T1w images using the boundary−based registration method for fine tuning.[75] Subsequently, a nonlinear warping was carried out to register the functional images to the standard MNI space with 2 mm isotropic voxel dimension. At last, spatial smoothing with a full width at half maximum Gaussian kernel of 6 mm was applied to improve signal−to−noise ratio.

Statistical analyses

The fMRI data was analyzed using the Statistical Parametric Mapping toolbox version 12 (SPM12; www.fil.ion.ucl.ac.uk/spm/). For a first-level analysis, general linear model (GLM) design matrices were constructed, each of which estimated neural correlates of the following regressors: (1) the SPE and RPE, (2) reliability of MB and MF learning strategies, (3) maximum reliability and reliability difference, or (4) the Q_{MB}, Q_{MF}, Q_{Arb}, and updated value signal after BACKWARD planning. All design matrices included regressors containing average neural signals at choice and outcome states, two prediction errors, and the six motion parameters. The GLM models were separated as above without applying serial orthogonalization in order to avoid impacts of the orthogonalization on result inferences and collinearity issues among the regressors.[76] For a second-level analysis, one-sample *t*-test was conducted across all participants to test replicability of the previous findings of neural arbitrator signals.[60] Then, independent two samples *t*-test of the parametric estimates was performed between patients and HCs. A brain region with significant difference was defined if a cluster survived the cluster-extent threshold of false discovery rate corrected *p* (*p*FDR) < 0.05 with peak-level threshold of uncorrected $p < 0.005$.[77] When the previously defined neural arbitrator regions (i.e., IFG and FPC) were included in a large cluster, small volume correction (10 mm sphere) on a local minimum region of the a priori ROIs was applied to increase spatial specificity. For the one-sample *t*-test, clusters were formed at

peak-level threshold of $p < 0.001$ to follow the test criterion of the previous study.[60] The survived brain regions were labeled based on the Automated Anatomical Labelling atlas 3 and NeuroSynth systems.[78,79]

To understand how neural interaction processes for the arbitration are altered in patients, it needed to investigate coupling between neural arbitrator regions encoding stability of the arbitration and other regions encoding state-action value of the arbitration. Thus, differences of psychophysiological interaction (PPI) between dysfunctional neural arbitrators and model-choice probability (P_{MB}) were additionally tested between groups. Based on two samples *t*-test results of neural activations encoding the arbitration stability (Table 5), the right anterolateral OFC and bilateral IFG regions were chosen. First eigenvariate of these neural signals were extracted from 5 mm spherical ROIs centered on their peak coordinates. The GLM for the PPI analysis included the interaction term as a covariate of interest while controlling for the physiological and psychological terms.

To explore neural mechanisms underlying dysfunctional behavioral arbitration and OC symptoms, multiple linear regression analyses were performed using stepwise method. Regression models explaining stability of the arbitration included activities of the neural arbitrators that encoded the maximum reliability and had significant group differences, while controlling for effects of IQ and a reliability parameter relevant to the neural arbitrators (Table 5). As exploratory analyses, partial correlation coefficients between

stability of the arbitration and activities of the neural arbitrator were also calculated while controlling for the same covariates. In addition, other regression models predicting OC symptom severity included activities of the neural arbitrators or effective connectivity thereof as covariates of interest, while controlling for an effect of anxiety symptom.

Results

Demographic and clinical characteristics

Demographic data were comparable between groups, and patients with OCD suffered moderate severity of OC symptoms. Demographic variables and severity of clinical symptoms are presented in Table 1.

Neural substrates of the arbitration between MB and MF strategies

Consistent with the previous study,[60] the IFG and FPC regions were found to encode stability of the arbitration; thus, these regions were defined as the neural arbitrators. The neural arbitrators estimated reliability signals of MB and MF systems. Regarding neural correlates of state–action values, the dorsal striatum was found to encode the Q_{MF} and Q_{Arb} signals. In addition to the insula and supplementary motor area (SMA),[60] neural signals of the inferior parietal lobe (IPL) and precuneus contained the Q_{Arb}. Details of neural correlates with the model parameters, including prediction errors, are presented in Table 4.

Table 4. Neural correlates of the arbitration model parameters

Brain region	MNI coordinates (mm)	Cluster size (*k*)	Statistics	
			T	Cluster *p*FDR
SPE[a]				
R Insula	38, 20, −02	2897	14.60	< 0.001
L Insula	−36, 18, −02	608	14.57	< 0.001
R Caudate/GP	10, 06, 0	1458	11.91	< 0.001
L dmPFC	02, 30, 50	1144	11.35	< 0.001
L Fusiform gyrus	−28, −72, −08	4469	10.34	< 0.001
R IPL	42, −50, 46	733	8.93	< 0.001
L Precentral gyrus	−54, 10, 44	766	8.22	< 0.001
L IFG	−44, 44, −04	228	8.17	< 0.001
L Substantia nigra	−06, −12, −16	101	7.34	< 0.001
R MFG	36, 54, 02	96	7.31	< 0.001
RPE[a]				
L Lingual gyrus	−18, −80, −08	619	7.94	< 0.001
L NAcc	−06, 06, −10	104	7.79	< 0.001
R NAcc/Putamen	14, 08, −10	111	7.68	< 0.001
R Cuneus	16, −92, 18	340	7.22	< 0.001
R Insula	28, 14, −16	53	7.15	< 0.001
Reliability of MB learning				
R Cuneus	14, −92, 18	3370	8.17	< 0.001
R IFG (FPC)	52, 40, −06	2527	7.18	< 0.001
R Fusiform gryus	28, −76, −06	355	6.34	< 0.001
L IFG	−54, 26, 02	315	5.75	< 0.001
R dmPFC	08, 36, 54	1616	5.64	< 0.001
L Insula	−30, 14, −18	232	5.12	< 0.001
L dmPFC (FPC)	−16, 54, 28	96	4.75	< 0.001
Reliability of MF learning				
R dmPFC	04, 32, 54	767	5.66	< 0.001
L Caudate	−10, −08, 26	85	5.20	0.019

R STG	50, −24, −04	135	4.55	0.005
R IFG	52, 24 , 06	133	4.25	0.005
R MFG	42, 10, 46	252	4.14	< 0.001
R vlPFC (FPC)	38, 56, 0	131	3.98	0.005
Stability of arbitration (maximum reliability)				
R dmPFC	06, 36, 56	1223	6.54	< 0.001
R IFG (FPC)	52, 40, −06	1726	6.22	< 0.001
R Cuneus	14, −90, 18	804	5.66	< 0.001
L Insula/IFG	−28, 18, −14	321	5.22	< 0.001
L Cerebellar Crus I	−18, −80, −30	513	4.90	< 0.001
R MTG	58, −30, −04	364	4.41	< 0.001
State−action value of MB learning (Q_{MB})				
R Precentral gyrus	40, −20, 54	10197	11.27	< 0.001
L Postcentral gyrus	−56, −18, 26	1403	7.41	< 0.001
L Insula/IFG	−34, 20, 10	843	6.58	< 0.001
R Insula	32, 24, 10	151	5.57	< 0.001
R Precentral gyrus	60, 06, 28	149	5.47	< 0.001
L NAcc/Caudate	0, 18, −04	72	4.56	< 0.001
R Calcarine sulcus	24, −50, 04	88	4.36	< 0.001
L Putamen	−20, 12, −06	91	4.33	< 0.001
State−action value of MF learning (Q_{MF})				
L dlPFC	−26, 0, 56	35318	14.40	< 0.001
R Insula/IFG	32, 24, 06	437	8.46	< 0.001
L Thalamus	−12, −18, 02	1345	7.18	< 0.001
L Caudate/GP	−14, 10, 08	480	6.99	< 0.001
R Caudate/GP	14, 06, 14	313	6.12	< 0.001
R IFG	60, 10, 22	370	5.53	< 0.001
L ITG	−58, −48, −14	237	4.96	< 0.001
State−action value of arbitration (Q_{Arb}), chosen−unchosen				
R Cuneus	14, −92, 18	8141	10.51	< 0.001
R Cerebellar Crus II	20, −76, −40	765	8.34	< 0.001
R Hippocampus	28, −06, −18	113	7.36	0.006
L Cerebellar Crus II	−40, −78, −36	1230	6.77	< 0.001

L Caudate	−18, −02, 28	225	6.63	< 0.001
L IFG	−52, 26, 02	575	6.51	< 0.001
L Amygdala	−24, −06, −14	192	5.78	0.001
R IFG	44, 28, −12	1269	5.65	< 0.001
R dmPFC	08, 52, 26	3380	5.60	< 0.001
L Postcentral gyrus	−38, −26, 52	1010	5.42	< 0.001
L Putamen	−30, −08, 04	160	5.07	0.001
State−action value of arbitration (Q_{Arb}), unchosen−chosen[a]				
L SMA	−04, 14, 50	5703	14.91	< 0.001
L SFG	−28, 0, 58	1187	11.77	< 0.001
L Insula	−30, 20, 08	372	11.75	< 0.001
R Insula	32, 22, 08	239	10.81	< 0.001
L IPL	−34, −46, 40	1154	9.20	< 0.001
L IFG	−58, 06, 18	594	8.53	< 0.001
R SPL/Precuneus	14, −68, 56	277	8.45	< 0.001
L Precuneus	−10, −70, 52	362	8.37	< 0.001
R Precuneus	16, −56, 20	94	7.88	< 0.001
R Precentral gyrus/IFG	60, 08, 22	94	7.82	< 0.001
L Caudate	−10, 10, 08	85	6.91	< 0.001

dmPFC: dorsomedial prefrontal cortex; dlPFC: dorsolateral prefrontal cortex; vlPFC: ventrolateral prefrontal cortex; FPC: frontopolar cortex; SFG: superior frontal gyrus; MFG: middle frontal gyrus; IFG: inferior frontal gyrus; STG: superior temporal gyrus; MTG: middle temporal gyrus; ITG: inferior temporal gyrus; SPL: superior parietal lobe; IPL: inferior parietal lobe; SMA: supplementary motor area; NAcc: nucleus accumbens; GP: globus pallidus

[a] cluster−forming peak threshold of family−wise error corrected $p < 0.05$

Dysfunctional neural signals of the arbitration between MB and MF strategies

When neural estimates of the model parameters were compared between groups, patients with OCD showed hyperactive right anterolateral OFC/FPC (MNI [32, 54, −16], cluster *p*FDR = 0.039) and bilateral IFG (MNI [50, 24, −12], cluster *p*FDR = 0.039; MNI [−56, 22, −4], cluster *p*FDR = 0.046) signals encoding stability of the arbitration than HCs (Figure 8). The hyperactivations of the same model parameter were also observed in the left OFC (MNI [−38, 28, −22], cluster *p*FDR = 0.030, k = 214) and right superior temporal gyrus (MNI [46, 0, −20], cluster *p*FDR < 0.001, k = 694) clusters that extended to local maxima of the IFG regions (Table 5). The aforementioned neural arbitrators of the right OFC/FPC (MNI [34, 48, −14], cluster *p*FDR = 0.006) and bilateral IFG (MNI [46, 34, −10], cluster *p*FDR = 0.026; MNI [−56, 24, −4], cluster *p*FDR = 0.004) were also hyperactive in patients than HCs when these regions contained MF reliability signal. Right IPL signal (MNI [34, −48, 36], cluster *p*FDR = 0.002) encoding action value of the arbitration was hyperactive in patients than HCs. When choosing a model between MB and MF strategies, patients had stronger effective connectivity from the right anterolateral OFC/FPC (MNI [32, 54, −16]), which contained arbitration stability signal, to the right precuneus (MNI [4, −54, 26], cluster *p*FDR = 0.018) found to encode state−action value of the arbitration than HCs. Detailed results of the brain functional differences are presented in Table 5.

Table 5. Dysfunctional neural signals of the arbitration system in patients with OCD

Brain region	MNI coordinates (mm)	Cluster size (*k*)	Statistics	
			T	Cluster *p*FDR
OCD > HCs				
Reliability of MF learning				
L Paracentral gyrus	−06, −20, 64	1301	5.01	< 0.001
R STG	62, −06, 0	885	4.37	< 0.001
L dmPFC	−14, 38, 22	434	4.15	0.002
R OFC (FPC)	34, 48, −14	321	4.07	0.006
R IFG[a]	46, 34, −10	117	3.76	0.026
L IFG	−56, 24, −04	352	3.81	0.004
L dlPFC	−50, 18, 40	363	3.62	0.004
Stability of arbitration (maximum reliability)				
L ACC	−06, −02, 34	318	5.18	0.006
R OFC (FPC)	32, 54, −16	186	4.74	0.039
R Insula	46, −04, −02	316	4.69	0.006
R STG/IFG	46, 0, −20	694	4.60	< 0.001
R IFG[a]	50, 24, −12	104	3.39	0.015
L OFC/IFG	−38, 28, −22	214	4.46	0.030
L IFG[a]	−56, 22, −04	65	3.65	0.046
L Postcentral gyrus	−32, −40, 66	320	4.08	0.006
R Postcentral gyrus	22, −34, 74	245	3.86	0.019
State−action value of arbitration (Q_{Arb})				
R IPL	34, −48, 36	426	4.42	0.002
Effective connectivity of the R OFC (FPC)$_{\text{maximum reliability}}$[b] when choosing a model system (× P_{MB})				
R Precuneus	04, −54, 26	327	0.018	4.10

dmPFC: dorsomedial prefrontal cortex; dlPFC: dorsolateral prefrontal cortex; ACC: anterior cingulate cortex; OFC: orbitofrontal

cortex; FPC: frontopolar cortex; IFG: inferior frontal gyrus; STG: superior temporal gyrus; IPL: inferior parietal lobe

[a] small volume correction (10 mm sphere) on a local minimum region of IFG or FPC, the neural arbitrators[60]

[b] first eigenvariate of the right OFC (FPC) signal (5 mm spherical ROI on MNI [32, 54, −16]) encoding the arbitration stability

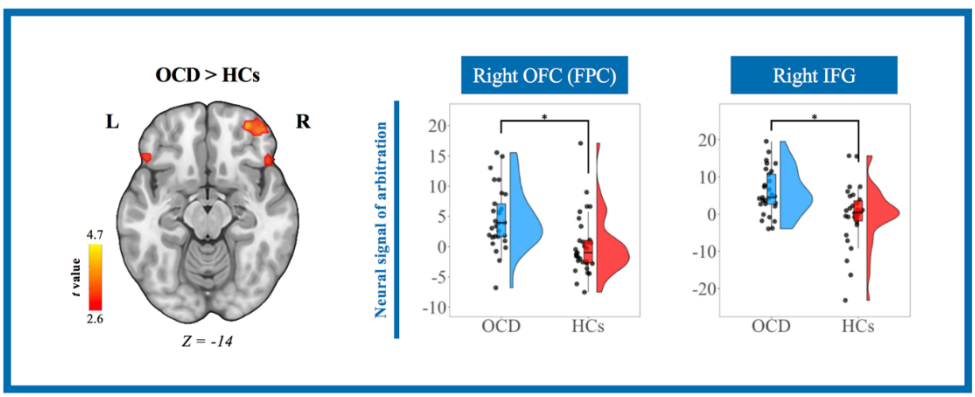

Figure 8. Dysfunctional neural arbitrators in OCD. Compared to HCs, patients with OCD had hyperactive signals in the right anterolateral orbitofrontal cortex (OFC)/frontopolar cortex (FPC) and bilateral inferior frontal gyrus (IFG) regions that encoded stability of the arbitration.

Neural arbitrators accounting for the impaired arbitration and OC symptoms

47 % of the higher stability of the arbitration in patients during MB−favored trials was explained (F = 9.46, p < 0.001) by effects of the right OFC/FPC signal encoding arbitration stability (β = 0.42, p =

0.007), MF reliability (β = 0.65, p < 0.001), and IQ (β = 0.27, p = 0.06) in the model predicting the behavioral impairment (Figure 9). The right OFC/FPC hyperactivation was also positively correlated with the higher stability of the arbitration during MF−favored trials (r = 0.47, p = 0.011) while the right IFG hyperactivation relevant to the arbitration stability had opposite correlation with the same measure (r = −0.70, p < 0.001) in patients (Figure 9).

Severity of OC symptoms was explained by brain functional differences (Figure 10). The right IFG signal containing the arbitration stability (β = −0.37, p = 0.026), together with HAM−A score (β = 0.53, p = 0.002), explained compulsion score of patients (F = 7.99, p = 0.002, adjusted R^2 = 0.33). In addition, obsession score of patients was explained (F = 14.29, p < 0.001, adjusted R^2 = 0.49) by effects of the connectivity between right OFC/FPC and precuneus (β = −0.29, p = 0.046) and HAM−A score (β = 0.72, p < 0.001).

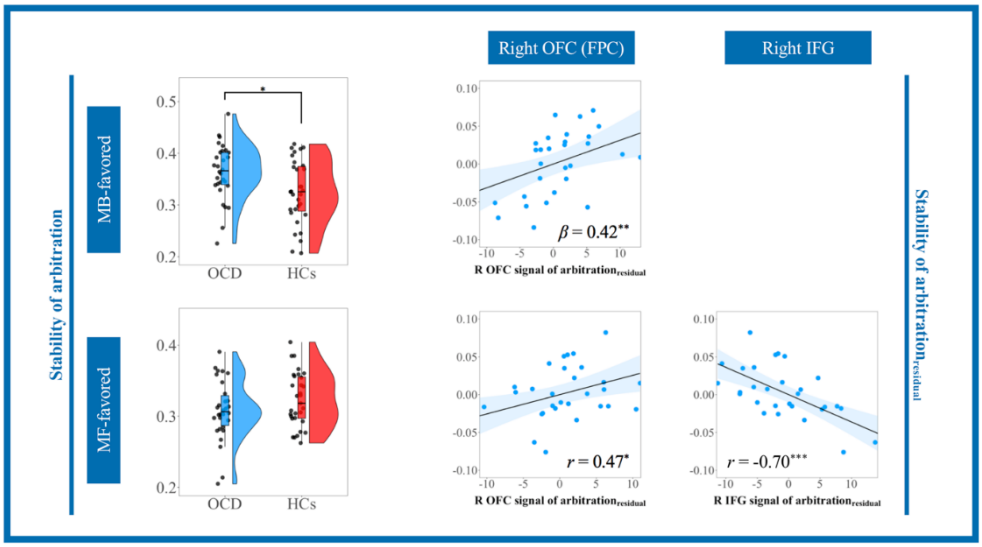

Figure 9. Neural arbitrator signals underlying the impaired arbitration in OCD. The impaired arbitration in patients was explained by their hyperactive OFG/FPC signal encoding stability of the arbitration while the hyperactive right IFG signal in patients compensated the arbitration dysfunction during MF−favored trials.

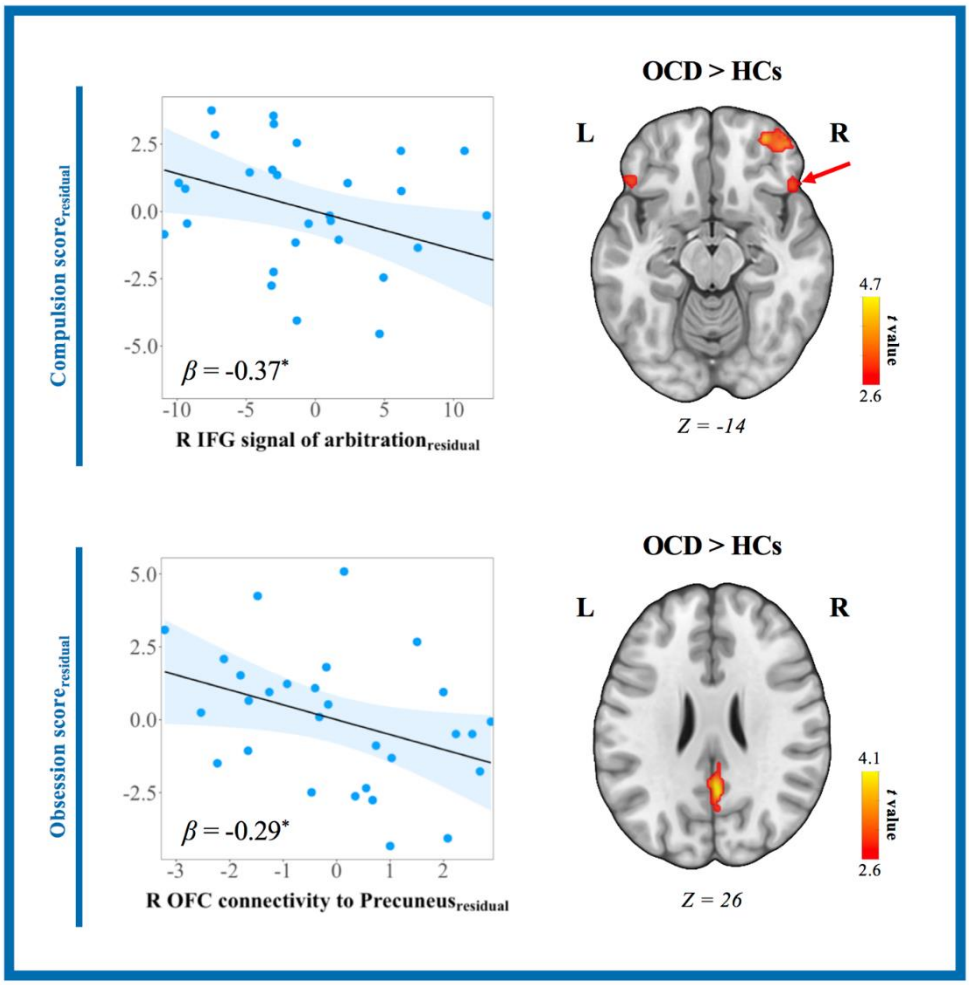

Figure 10. Neural arbitrator signals underlying clinical symptom severity of OCD. The increased right IFG signal encoding stability of the arbitration and the strengthened functional connectivity of the OFC/FPC with the precuneus containing action value of the arbitration were compensatorily related to less severe compulsion and obsession, respectively, in patients with OCD.

Discussion

Based on the finding of the excessively stable arbitration in which MF RL strategy was predominant in OCD, the subsequent neuroimaging study expected that neural arbitrators estimating the reliability of learning strategies would be dysfunctional in patients. Consistent with the hypothesis, hyperactive lateral OFC/FPC signal in OCD underlay the imbalanced arbitration between MB and MF RL systems. Interestingly, the right IFG, the other neural arbitrator, was also hyperactive in patients but relevant to a compensating mechanism for the impaired arbitration and compulsivity.

The main purpose of this research project was to build brain biomarkers underlying the imbalanced arbitration between goal-directed and habitual systems in OCD for developing a neurocircuit-based treatment.[29,44] The hyperactive anterolateral OFC/FPC and IFG regions were identified to be relevant to the impaired arbitration in OCD. These two regions are known as the neural arbitrators that compute reliability of each learning strategy and arbitrate decision-making by choosing a more reliable controller.[60] In OCD, these neural substrates are parts of the ventral cognitive CSTC circuit, which is relevant for impaired response inhibition and inflexible attention switching.[21,22,25,31,80] This study provided a new evidence of pathological role of these two regions. The neural arbitrators were demonstrated to influence the imbalanced arbitration of OCD but in different ways. While the hyperactive lateral OFC/FPC was related to the excessively stable

arbitration that was biased toward MF system, the right IFG region compensatorily increased its activation to normalize the impaired arbitration. Therefore, it is suggested that the hyperactive lateral OFC/FPC underlies the impaired arbitration and habit bias in OCD.

To provide the biomarkers as biological evidence for neurotherapeutics, it needs to confirm if the neural substrates of the impaired arbitration would be further related to severity of clinical symptoms. Through the regression analyses, it was figured out that the hyperactive right IFG signal was linked to less severe compulsion. Considering its compensatory mechanism on the decision arbitration and compulsivity, the right IFG and its role in ventral cognitive circuit are thought to be neuroscientifically-informed treatment targets for compulsions of OCD. On the other hand, the precuneus relevant for action value of the arbitration was more strongly connected with the dysfunctional OFC/FPC in patients with less severe obsessions. Thus, the OFC/FPC-precuneus connectivity may be an additional target for treating obsessive symptoms. Because the disrupted goal-directed planning and excessive habits are transdiagnostically related to compulsivity, not confined to OCD diagnosis,[81] the neurobiological evidence provided by this study would be applicable to other compulsive disorders as well.

General Discussion

Current symptom-based classification of OCD does not satisfactorily capture the pathophysiology on an individual basis.[10,11] In case of OCD, at least 30 % of patients are not responsive to first-line treatments.[46,47] Dimensional classification based on biological evidence is suggested as an alternative. Building brain biomarkers and determining biotypes of OCD are expected to help addressing the heterogeneity issue and developing neurocircuit-based treatments.[29,44]

The unsatisfactory outcome of first-line treatments in OCD draws attention to develop alternative treatments that are based on well-accepted CSTC circuitry model (Figure 11).[44,46,47,82,83] According to the collective neurobiological model, OCD is characterized by diminished inhibitory signals of serotonergic and GABAergic pathways, which further provoke striatal dopaminergic and glutamatergic hyperactivity (i.e., hyperactive CSTC circuitry).[44] From this extended neurobiological model, poor performances of decision-making in OCD could be linked to the dopaminergic hyperactivity beyond the serotonergic dysfunction, which is more relevant to emotion regulation and fear extinction.[41] CSTC circuitry comprises direct and indirect pathways that are divided by distinct dopaminergic system but interacting each other. Striatal neurons in the direct pathway express mostly D1 receptor and project Go signal through the internal globus pallidus to thalamus, consequently executing actions. On the other hand, other types of striatal neurons

expressing mostly D2 receptor occupy the indirect pathway and project Nogo signal through the internal and external segments of globus pallidus to the thalamus, consequently inhibiting actions.[23,24,84] These two counterbalancing dopaminergic systems further participate in goal-directed and habitual behaviors.[85] Therefore, the balance between striatal dopaminergic systems is a key in adaptive behavioral controls, and dysfunctional communication between the striatal regions and the prefrontal controllers is considered to provoke maladaptive behaviors, such as compulsivity and habit bias in psychiatric diseases.

A theory-driven approach suggests that habit bias and imbalance between goal-directed and habitual controls underlie symptomatology of OCD.[15] Among several CSTC circuits, the imbalanced arbitration and habit bias of OCD are thought to involve disrupted ventral cognitive circuit, which is known to underlie response inhibition and adaptive attention shifting (Figure 11).[31-34] Non-invasive neuromodulation approach, such as transcranial magnetic stimulation, can modulates prefrontal and striatal dopamine releases and be applicable to enhance the performance of decision-making.[86-88] Therefore, it is expected that a neurocircuit-based treatment that decreases the lateral OFC/FPC activity or increases the IFG activity would normalize the dopaminergic dysfunctions within the ventral cognitive CSTC circuit and improve the imbalanced arbitration and compulsion of OCD.

This study has limitations to be addressed. First, CBT is considered to improve goal-directed planning in OCD although this

effect is controversial.[89,90] Because the long-term efficacy is not clearly known and becomes less effective at one year after the treatment,[91,92] it was planned to recruit patients who did not experience the CBT session in recent one year so as to minimize the CBT effect on the arbitration performance. Second, a large number of free parameters were used to explain the complex arbitration process; however, this high parameter count renders concerns about overfitting of the model parameters.[93] To address this issue and optimize the parameters, model fitting was iterated large enough times with random seeds to make it as reproducible as possible.

In conclusion, it was identified that the hyperactive neural arbitrators underlying the excessively stable arbitration in which MF learning strategy was predominant were brain biomarkers for the imbalanced arbitration between goal-directed and habitual controls in OCD. Based on their associations with compulsion severity, this study suggests that the hyperactive IFG and its connections within the ventral cognitive CSTC circuit can guide a neurocircuit-based treatment for habit bias and compulsivity in OCD.

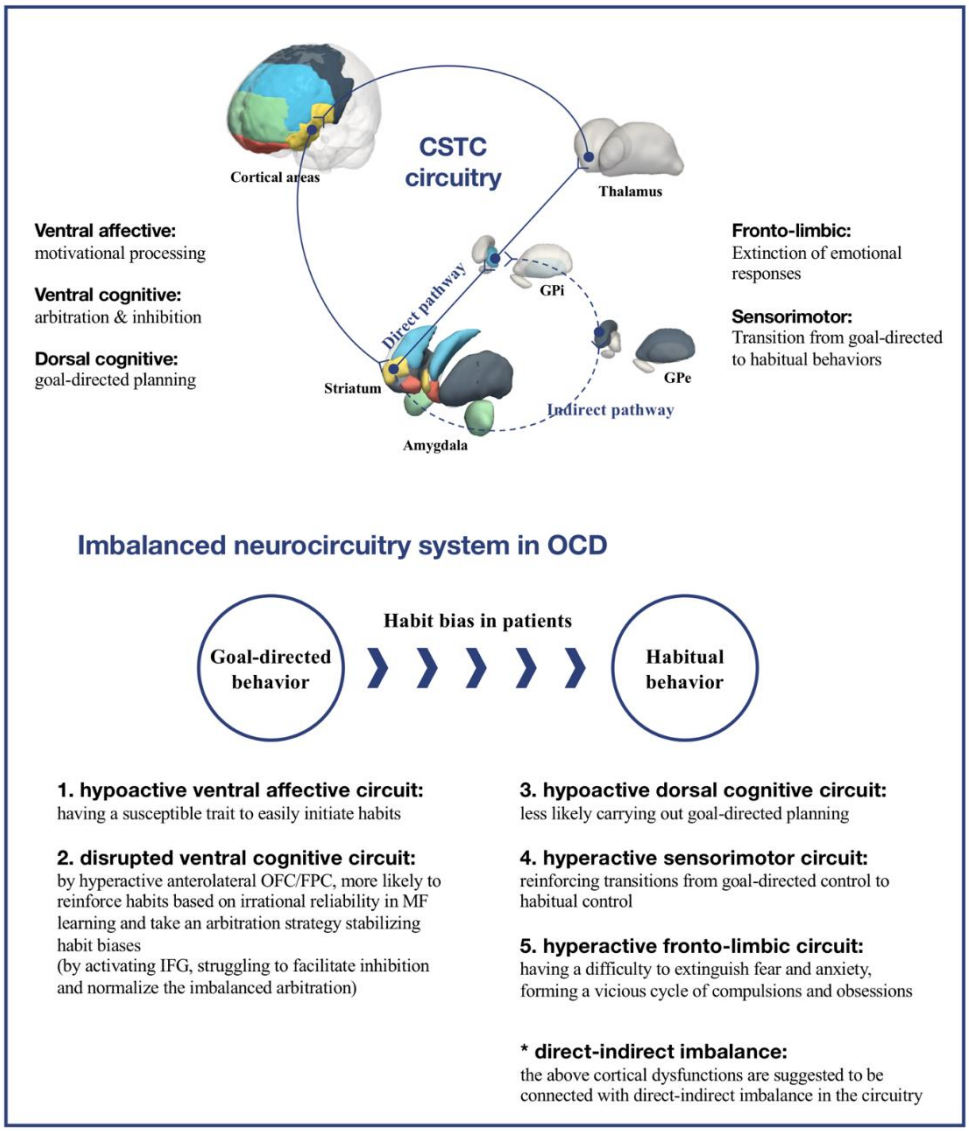

Figure 11. A model of neurocircuitry system implicated in imbalanced arbitration between goal-directed and habitual behaviors in OCD.

Hypoactive ventral affective circuit in OCD reflects that patients are sensitive to initiate habits. Because of their irrational belief about that model-free learning is the most reliable strategy for decision-making, patients are more likely to reinforce the habits and take an

arbitration strategy that stabilizes the habit biases. Hyperactive anterolateral OFC/FPC underlies this imbalance in the arbitration between goal-directed and habitual systems, while the IFG, the other neural arbitrator, is hyperactivated to struggle to facilitate inhibitions for the habitual behaviors and normalize the imbalanced arbitration. The disrupted arbitration system also influences other CSTC circuits as follows. The dorsal cognitive circuit becomes hypoactive, making patients not to carry out goal-directed planning. The reinforced transitions from goal-directed to habitual behaviors make the sensorimotor circuit hyperactive in patients. Lastly, patients become to have maladaptive behaviors (i.e., impaired extinction of emotional responses) because the excessive habits make anxiety and fear blunt, which is relevant to hyperactive fronto-limbic circuit.

References

1 Jenike, M. A. Obsessive-compulsive disorder. *N Engl J Med* **350**, 259-265 (2004).

2 Ruscio, A. M., Stein, D. J., Chiu, W. T. & Kessler, R. C. The epidemiology of obsessive-compulsive disorder in the National Comorbidity Survey Replication. *Mol Psychiatry* **15**, 53-63 (2010).

3 Stein, D. J. *et al.* Should OCD be classified as an anxiety disorder in DSM-V? *Depress Anxiety* **27**, 495-506 (2010).

4 Phillips, K. A. *et al.* Should an obsessive-compulsive spectrum grouping of disorders be included in DSM-V? *Depress Anxiety* **27**, 528-555 (2010).

5 Salkovskis, P. M., Forrester, E. & Richards, C. Cognitive-behavioural approach to understanding obsessional thinking. *Br J Psychiatry* **173**, 53-63 (1998).

6 Gillan, C. M. *et al.* Enhanced avoidance habits in obsessive-compulsive disorder. *Biol Psychiatry* **75**, 631-638 (2014).

7 Rosario-Campos, M. C. *et al.* The Dimensional Yale-Brown Obsessive-Compulsive Scale (DY-BOCS): an instrument for assessing obsessive-compulsive symptom dimensions. *Mol Psychiatry* **11**, 495-504 (2006).

8 van den Heuvel, O. A. *et al.* The major symptom dimensions of obsessive-compulsive disorder are mediated by partially distinct neural systems. *Brain* **132**, 853-868 (2009).

9 Taylor, S. Early versus late onset obsessive–compulsive disorder: evidence for distinct subtypes. *Clin Psychol Rev* **31**, 1083–1100 (2011).

10 Kim, T. *et al.* Neural bases of the clinical and neurocognitive differences between earlyand late–onset obsessive–compulsive disorder. *J Psychiatry Neurosci* **45**, 234–242 (2020).

11 Lack, C. W. Obsessive–compulsive disorder: evidence–based treatments and future directions for research. *World J Psychiatry* **2**, 86–90 (2012).

12 Salkovskis, P. M. *et al.* Responsibility attitudes and interpretations are characteristic of obsessive compulsive disorder. *Behav Res Ther* **38**, 347–372 (2000).

13 Gillan, C. M. *et al.* Disruption in the balance between goal–directed behavior and habit learning in obsessive–compulsive disorder. *Am J Psychiatry* **168**, 718–726 (2011).

14 Gillan, C. M. *et al.* Counterfactual processing of economic action–outcome alternatives in obsessive–compulsive disorder: further evidence of impaired goal–directed behavior. *Biol Psychiatry* **75**, 639–646 (2014).

15 Gillan, C. M. & Robbins, T. W. Goal–directed learning and obsessive–compulsive disorder. *Philos Trans R Soc B* **369**, 20130475 (2014).

16 Clark, D. A. & O'Connor, K. Thinking Is Believing: Ego–Dystonic Intrusive Thoughts in Obsessive–Compulsive Disorder. In: *Intrusive Thoughts in Clinical Disorders: Theory, Research, and Treatment*. 145–174 (Guilford Press, 2005).

17 Robbins, T. W., Gillan, C. M., Smith, D. G., de Wit, S. & Ersche, K. D. Neurocognitive endophenotypes of impulsivity and compulsivity: towards dimensional psychiatry. *Trends Cogn Sci* **16**, 81-91 (2012).

18 Festinger, L. *A Theory of Cognitive Dissonance*. (Stanford University Press, 1957).

19 Graybiel, A. M. & Rauch, S. L. Toward a neurobiology of obsessive-compulsive disorder. *Neuron* **28**, 343-347 (2000).

20 Pauls, D. L., Abramovitch, A., Rauch, S. L. & Geller, D. A. Obsessive-compulsive disorder: an integrative genetic and neurobiological perspective. *Nat Rev Neurosci* **15**, 410-424 (2014).

21 Milad, M. R. & Rauch, S. L. Obsessive-compulsive disorder: beyond segregated cortico-striatal pathways. *Trends Cogn Sci* **16**, 43-51 (2012).

22 van den Heuvel, O. A. *et al.* Brain circuitry of compulsivity. *Eur Neuropsychopharmacol* **26**, 810-827 (2016).

23 Jahanshahi, M., Obeso, I., Rothwell, J. C. & Obeso, J. A. A fronto-striato-subthalamic-pallidal network for goal-directed and habitual inhibition. *Nat Rev Neurosci* **16**, 719-732 (2015).

24 Maia, T. V. & Frank, M. J. From reinforcement learning models to psychiatric and neurological disorders. *Nat Neurosci* **14**, 154-162 (2011).

25 Kim, T., Jung, W. H., Yoon, B. Y., Kim, M. & Kwon, J. S. Cortico-striato-thalamo-cortical network for cognitive flexibility

in obsessive-compulsive disorder in *OHBM 2020 Annual Meeting* (2020).

26 Thorsen, A. L. *et al.* Emotional processing in obsessive-compulsive disorder: a systematic review and meta-analysis of 25 functional neuroimaging studies. *Biol Psychiatry Cogn Neurosci Neuroimaging* **3**, 563-571 (2018).

27 Apergis-Schoute, A. M. *et al.* Neural basis of impaired safety signaling in obsessive compulsive disorder. *Proc Natl Acad Sci USA* **114**, 3216-3221 (2017).

28 Paul, S. *et al.* Amygdala-prefrontal connectivity during appraisal of symptom-related stimuli in obsessive-compulsive disorder. *Psychol Med* **49**, 278-286 (2019).

29 Shephard, E. *et al.* Toward a neurocircuit-based taxonomy to guide treatment of obsessive-compulsive disorder. *Mol Psychiatry* (2021).

30 van den Heuvel, O. A. *et al.* Frontal-striatal dysfunction during planning in obsessive-compulsive disorder. *Arch Gen Psychiatry* **62**, 301-309 (2005).

31 Vaghi, M. M. *et al.* Specific frontostriatal circuits for impaired cognitive flexibility and goal-directed planning in obsessive-compulsive disorder: evidence from resting-state functional connectivity. *Biol Psychiatry* **81**, 708-717 (2017).

32 de Wit, S. J. *et al.* Presupplementary motor area hyperactivity during response inhibition: a candidate endophenotype of obsessive-compulsive disorder. *Am J Psychiatry* **169**, 1100-1108 (2012).

33 Kang, D. H. *et al.* Neural correlates of altered response inhibition and dysfunctional connectivity at rest in obsessive-compulsive disorder. *Prog Neuro-Psychopharmacol Biol Psychiatry* **40**, 340-346 (2013).

34 Nakao, T. *et al.* Brain activation of patients with obsessive-compulsive disorder during neuropsychological and symptom provocation tasks before and after symptom improvement: a functional magnetic resonance imaging study. *Biol Psychiatry* **57**, 901-910 (2005).

35 Jung, W. H. *et al.* Abnormal corticostriatal-limbic functional connectivity in obsessive-compulsive disorder during reward processing and resting-state. *Neuroimage Clin* **3**, 27-38 (2013).

36 Figee, M. *et al.* Dysfunctional reward circuitry in obsessive-compulsive disorder. *Biol Psychiatry* **69**, 867-874 (2011).

37 Norman, L. J. *et al.* Frontostriatal dysfunction during decision making in attention-deficit/hyperactivity disorder and obsessive-compulsive disorder. *Biol Psychiatry Cogn Neurosci Neuroimaging* **3**, 694-703 (2018).

38 Moreira, P. S. *et al.* Altered response to risky decisions and reward in patients with obsessive-compulsive disorder. *J Psychiatry Neurosci* **45**, 98-107 (2020).

39 Han, H. J. *et al.* Disruption of effective connectivity from the dorsolateral prefrontal cortex to the orbitofrontal cortex by negative emotional distraction in obsessive-compulsive disorder. *Psychol Med* **46**, 921-932 (2016).

40 Eddy, K. T., Dutra, L., Bradley, R. & Westen, D. A multidimensional meta-analysis of psychotherapy and pharmacotherapy for obsessive-compulsive disorder. *Clin Psychol Rev* **24**, 1011-1030 (2004).

41 Graat, I., Figee, M. & Denys, D. Neurotransmitter Dysregulation in OCD. In: *Obsessive-Compulsive Disorder: Phenomenology, Pathophysiology, and Treatment*. 271-288 (Oxford University Press, 2017).

42 Rosenberg, D. R. *et al.* Decrease in caudate glutamatergic concentrations in pediatric obsessive-compulsive disorder patients taking paroxetine. *J Am Acad Child Adolesc Psychiatry* **39**, 1096-1103 (2000).

43 Simpson, H. B. *et al.* Investigation of cortical glutamate-glutamine and gamma-aminobutyric acid in obsessive-compulsive disorder by proton magnetic resonance spectroscopy. *Neuropsychopharmacology* **37**, 2684-2692 (2012).

44 Dougherty, D. D. *et al.* Neuroscientifically informed formulation and treatment planning for patients with obsessive-compulsive disorder: a review. *JAMA Psychiatry* **75**, 1081-1087 (2018).

45 Hyman, S. E. The diagnosis of mental disorders: the problem of reification. *Annu Rev Clin Psychol* **6**, 155-179 (2010).

46 Ost, L. G., Havnen, A., Hansen, B. & Kvale, G. Cognitive behavioral treatments of obsessive-compulsive disorder. a systematic review and meta-analysis of studies published 1993-2014. *Clin Psychol Rev* **40**, 156-169 (2015).

47 Bloch, M. H., McGuire, J., Landeros-Weisenberger, A., Leckman, J. F. & Pittenger, C. Meta-analysis of the dose-response relationship of SSRI in obsessive-compulsive disorder. *Mol Psychiatry* **15**, 850-855 (2010).

48 Insel, T. R. The NIMH Research Domain Criteria (RDoC) project: precision medicine for psychiatry. *Am J Psychiatry* **171**, 395-397 (2014).

49 Hauser, T. U. *et al.* Increased fronto-striatal reward prediction errors moderate decision making in obsessive-compulsive disorder. *Psychol Med* **47**, 1246-1258 (2017).

50 Chamberlain, S. R., Blackwell, A. D., Fineberg, N. A., Robbins, T. W. & Sahakian, B. J. The neuropsychology of obsessive compulsive disorder: the importance of failures in cognitive and behavioural inhibition as candidate endophenotypic markers. *Neurosci Biobehav Rev* **29**, 399-419 (2005).

51 Woo, C. W., Chang, L. J., Lindquist, M. A. & Wager, T. D. Building better biomarkers: brain models in translational neuroimaging. *Nat Neurosci* **20**, 365-377 (2017).

52 Kwak, S. *et al.* Defining data-driven subgroups of obsessive-compulsive disorder with different treatment responses based on resting-state functional connectivity. *Transl Psychiatry* **10**, 359 (2020).

53 Balleine, B. W. & O'Doherty, J. P. Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology* **35**, 48-69 (2010).

54 Gruner, P., Anticevic, A., Lee, D. & Pittenger, C. Arbitration between action strategies in obsessive-compulsive disorder. *Neuroscientist* **22**, 188-198 (2016).
55 Voon, V. *et al.* Disorders of compulsivity: a common bias towards learning habits. *Mol Psychiatry* **20**, 345-352 (2015).
56 Yucel, M. *et al.* A transdiagnostic dimensional approach towards a neuropsychological assessment for addiction: an international Delphi consensus study. *Addiction* **114**, 1095-1109 (2019).
57 Daw, N. D., Niv, Y. & Dayan, P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci* **8**, 1704-1711 (2005).
58 Banca, P. *et al.* Imbalance in habitual versus goal directed neural systems during symptom provocation in obsessive-compulsive disorder. *Brain* **138**, 798-811 (2015).
59 Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An Introduction*. (MIT Press, 1998).
60 Lee, S. W., Shimojo, S. & O'Doherty, J. P. Neural computations underlying arbitration between model-based and model-free learning. *Neuron* **81**, 687-699 (2014).
61 Heo, S., Sung, Y. & Lee, S. W. Effects of subclinical depression on prefrontal-striatal model-based and model-free learning. *PLoS Comput Biol* **17**, e1009003 (2021).
62 First, M. B., Spitzer, R. L., Gibbon, M. & Williams, J. B. W. *Structured Clinical Interview for DSM-IV-TR Axis I Disorders,*

Research Version, Patient Edition. (Biometrics Research, New York State Psychiatric Institute, 2002).

63 Goodman, W. K. *et al.* The Yale-Brown Obsessive Compulsive Scale. I. development, use, and reliability. *Arch Gen Psychiatry* **46**, 1006-1011 (1989).

64 Hamilton, M. The assessment of anxiety states by rating. *Br J Med Psychol* **32**, 50-55 (1959).

65 Hamilton, M. Development of a rating scale for primary depressive illness. *Br J Soc Clin Psychol* **6**, 278-296 (1967).

66 First, M. B., Spitzer, R. L., Gibbon, M. & Williams, J. B. W. *Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-Patient Edition*. (Biometrics Research, New York State Psychiatric Institute, 2002).

67 Glascher, J., Daw, N., Dayan, P. & O'Doherty, J. P. States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron* **66**, 585-595 (2010).

68 Pearce, J. M. & Hall, G. A model for pavlovian learning - variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol Rev* **87**, 532-552 (1980).

69 Schultz, W. Reward prediction error. *Curr Biol* **27**, R369-R371 (2017).

70 Dayan, P. & Abbott, L. F. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. (MIT Press, 2001).

71 Gillan, C. M. *et al.* Functional neuroimaging of avoidance habits in obsessive–compulsive disorder. *Am J Psychiatry* **172**, 284–293 (2015).

72 Anticevic, A. *et al.* Global resting–state functional magnetic resonance imaging analysis identifies frontal cortex, striatal, and cerebellar dysconnectivity in obsessive–compulsive disorder. *Biol Psychiatry* **75**, 595–605 (2014).

73 Glasser, M. F. *et al.* The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage* **80**, 105–124 (2013).

74 Andersson, J. L., Skare, S. & Ashburner, J. How to correct susceptibility distortions in spin–echo echo–planar images: application to diffusion tensor imaging. *NeuroImage* **20**, 870–888 (2003).

75 Greve, D. N. & Fischl, B. Accurate and robust brain image alignment using boundary–based registration. *NeuroImage* **48**, 63–72 (2009).

76 Mumford, J. A., Poline, J. B. & Poldrack, R. A. Orthogonalization of regressors in fMRI models. *PLoS ONE* **10**, e0126255 (2015).

77 Woo, C. W., Krishnan, A. & Wager, T. D. Cluster–extent based thresholding in fMRI analyses: pitfalls and recommendations. *NeuroImage* **91**, 412–419 (2014).

78 Rolls, E. T., Huang, C. C., Lin, C. P., Feng, J. & Joliot, M. Automated anatomical labelling atlas 3. *NeuroImage* **206**, 116189 (2020).

79 Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C. & Wager, T. D. Large-scale automated synthesis of human functional neuroimaging data. *Nat Methods* **8**, 665-670 (2011).

80 Gu, B. M. *et al.* Neural correlates of cognitive inflexibility during task-switching in obsessive-compulsive disorder. *Brain* **131**, 155-164 (2008).

81 Gillan, C. M. *et al.* Comparison of the association between goal-directed planning and self-reported compulsivity vs obsessive-compulsive disorder diagnosis. *JAMA Psychiatry* **77**, 77-85 (2020).

82 Rapinesi, C. *et al.* Brain stimulation in obsessive-compulsive disorder (OCD): a systematic review. *Curr Neuropharmacol* **17**, 787-807 (2019).

83 Thorsen, A. L. *et al.* Stable inhibition-related inferior frontal hypoactivation and fronto-limbic hyperconnectivity in obsessive-compulsive disorder after concentrated exposure therapy. *Neuroimage Clin* **28**, 102460 (2020).

84 Gerfen, C. R. Molecular effects of dopamine on striatal-projection pathways. *Trends Neurosci* **23**, S64-70 (2000).

85 Wunderlich, K., Smittenaar, P. & Dolan, R. J. Dopamine enhances model-based over model-free choice behavior. *Neuron* **75**, 418-424 (2012).

86 Strafella, A. P., Paus, T., Barrett, J. & Dagher, A. Repetitive transcranial magnetic stimulation of the human prefrontal cortex induces dopamine release in the caudate nucleus. *J Neurosci* **21**, RC157 (2001).

87 Strafella, A. P., Paus, T., Fraraccio, M. & Dagher, A. Striatal dopamine release induced by repetitive transcranial magnetic stimulation of the human motor cortex. *Brain* **126**, 2609–2615 (2003).

88 Weissengruber, S., Lee, S. W., O'Doherty, J. P. & Ruff, C. C. Neurostimulation reveals context–dependent arbitration between model–based and model–free reinforcement learning. *Cereb Cortex* **29**, 4850–4862 (2019).

89 Huyser, C., Veltman, D. J., Wolters, L. H., de Haan, E. & Boer, F. Functional magnetic resonance imaging during planning before and after cognitive–behavioral therapy in pediatric obsessive–compulsive disorder. *J Am Acad Child Adolesc Psychiatry* **49**, 1238–1248 (2010).

90 Wheaton, M. G., Gillan, C. M. & Simpson, H. B. Does cognitive–behavioral therapy affect goal–directed planning in obsessive–compulsive disorder? *Psychiatry Res* **273**, 94–99 (2019).

91 van Dis, E. A. M. *et al.* Long–term outcomes of cognitive behavioral therapy for anxiety–related disorders: a systematic review and meta–analysis. *JAMA Psychiatry* **77**, 265–273 (2020).

92 Andersson, E. *et al.* Long–term efficacy of internet–based cognitive behavior therapy for obsessive–compulsive disorder with or without booster: a randomized controlled trial. *Psychol Med* **44**, 2877–2887 (2014).

93 Kriegeskorte, N. & Douglas, P. K. Cognitive computational neuroscience. *Nat Neurosci* **21**, 1148–1160 (2018).

국문 초록

서론: 목적-지향적 행동전략과 습관적 행동전략 사이의 조율 불균형으로 발생하는 습관 편향은 강박장애(OCD) 주증상인 강박행동의 기저를 이룬다. 강화학습 인공지능 알고리즘에 기반한 계산신경과학 모델은 이러한 두 행동전략 사이의 조율 기전을 설명할 수 있다. 사람의 뇌는 목적-지향적(모델-기반) 학습 시스템과 습관적(모델-자유) 학습 시스템의 상태/보상 예측 신뢰도를 계산하고, 신뢰도가 높은 학습 시스템을 선택하여 의사결정을 조율한다. 하지만, 강박장애 환자에서 나타나는 의사결정 조율 불균형이 잘못된 학습전략 신뢰도 추정에 원인을 둔 것인지 아직 불분명하다. 또한, 학습전략 신뢰도 계산을 담당하는 하전두회(IFG)와 전두극피질(FPC)의 기능 손상이 이러한 조율 불균형의 신경생물학적 기저인지 연구가 필요하다.

방법: 연구참여자들의 모델-기반 학습전략과 모델-자유 학습전략 행동을 분리해 관찰하기 위해 마르코프 의사결정 과제(sequential two-choice Markov decision task)를 사용했다. 30 명의 강박장애 환자와 31 명의 건강 대조군이 연구에 참여했으며, 의사결정 과제를 수행함과 동시에 기능적 뇌 자기공명영상(fMRI)을 촬영했다. 강화학습 알고리즘에 기반한 계산모델을 이용해 의사결정 조율 과정 동안의 행동을 추정했다. 모델 행동변수 및 관련 뇌 기능에 대해 환자군과 대조군 사이의 차이를 통계적으로 검증했으며, 해당 뇌 기능 차이가 신뢰도 추정 오류 및 강박행동 증상을 설명하는지 회귀분석을 통해 확인했다.

결과: 강박장애 환자들은 대조군에 비해 의사결정 과제 수행 시 보상 획득에 더 큰 어려움을 겪고 더 보속적으로 행동했다. 모델-기반 학습전략이 필요한 상황에서, 환자들은 오히려 모델-자유 학습전략을 과도히 신뢰했다. 그 결과, 환자들에서 두 학습전략 사이의 조율 안정성이

더 높았으며, 모델-자유 학습전략으로의 편향이 확인되었다. 환자에서 과도히 높은 조율 안정성은 전두극피질 영역 중 전외측 안와전두피질(anterolateral OFC)의 과활성화와 관련있었으며, 신뢰도 정보를 바탕으로 학습전략을 선택할 때 전외측 안와전두피질과 쐐기앞소엽 사이의 기능적 연결성이 비정상적으로 강화되었다. 반면, 환자에서 과활성화된 하전두회는 조율 안정성 및 강박행동 중증도와 부적 상관관계를 보였다.

결론: 본 연구는 강박장애의 의사결정 조율 불균형이 모델-자유 학습전략에 편향된 조율을 야기하는 뇌 기능 이상에 원인이 있음을 밝혔다. 나아가, 예측 신뢰도를 추정하는 하전두회 및 전두극피질을 강박행동 및 습관 편향에 대한 신경회로-기반 치료의 뇌 생물지표로 제안한다.

주요어: 강박장애, 의사결정, 목적-지향적 행동 조절, 습관적 행동 조절, 강화학습, 신경조율자, 과제-기반 기능적 뇌 자기공명영상

학번: 2015-20480