Neural correlates of model-based behavior in internet gaming disorder and alcohol use disorder

MINA KWON[1†] , HANGNYOUNG CHOI[2,3†] ,
HARHIM PARK[1] , WOO-YOUNG AHN[1,4,5**] and
YOUNG-CHUL JUNG[2,3,6*]

[1] Department of Psychology, Seoul National University, Seoul, South Korea

[2] Department of Psychiatry, Yonsei University College of Medicine, Seoul, South Korea

[3] Institute of Behavioral Science in Medicine, Yonsei University College of Medicine, Seoul, South Korea

[4] Department of Brain and Cognitive Sciences, Seoul National University, Seoul, South Korea

[5] AI Institute, Seoul National University, Seoul, South Korea

[6] Institute for Innovation in Digital Healthcare, Yonsei University, Seoul, South Korea

FULL-LENGTH REPORT

ABSTRACT

Background: An imbalance between model-based and model-free decision-making systems is a common feature in addictive disorders. However, little is known about whether similar decision-making deficits appear in internet gaming disorder (IGD). This study compared neurocognitive features associated with model-based and model-free systems in IGD and alcohol use disorder (AUD). *Method:* Participants diagnosed with IGD ($n = 22$) and AUD ($n = 22$), and healthy controls ($n = 30$) performed the two-stage task inside the functional magnetic resonance imaging (fMRI) scanner. We used computational modeling and hierarchical Bayesian analysis to provide a mechanistic account of their choice behavior. Then, we performed a model-based fMRI analysis and functional connectivity analysis to identify neural correlates of the decision-making processes in each group. *Results:* The computational modeling results showed similar levels of model-based behavior in the IGD and AUD groups. However, we observed distinct neural correlates of the model-based reward prediction error (RPE) between the two groups. The IGD group exhibited insula-specific activation associated with model-based RPE, while the AUD group showed prefrontal activation, particularly in the orbitofrontal cortex and superior frontal gyrus. Furthermore, individuals with IGD demonstrated hyper-connectivity between the insula and brain regions in the salience network in the context of model-based RPE. *Discussion and Conclusions:* The findings suggest potential differences in the neurobiological mechanisms underlying model-based behavior in IGD and AUD, albeit shared cognitive features observed in computational modeling analysis. As the first neuroimaging study to compare IGD and AUD in terms of the model-based system, this study provides novel insights into distinct decision-making processes in IGD.

†Co-first authors with equal contribution.

*Corresponding author.
E-mail: eugenejung@yuhs.ac

**Corresponding author.
E-mail: wahn55@snu.ac.kr

INTRODUCTION

Internet gaming disorder (IGD) encompasses a dysfunctional pattern of gaming characterized by impaired control and elevated priority given to gaming, leading to interference with occupational, social, and academic functions. Previous studies have established similarities between IGD and substance use disorders (SUDs) or gambling disorder in terms of cognitive

functioning, neural activities, and other clinical features (for a review, see Vaccaro & Potenza, 2019). Consequently, the World Health Organization officially recognized gaming disorder as a medical condition in the International Classification of Disease-11 (World Health Organization, 2018). However, ongoing debate persists regarding the classification of IGD as an addictive disorder in the same category as SUDs (Kuss, Griffiths, & Pontes, 2017; Saunders et al., 2017; Starcevic, 2017). This debate arises from the distinct clinical aspects of IGD compared to SUDs (Dowling, 2014; Starcevic, 2017), such as less prominent physiological withdrawal and tolerance, which can be attributed to the absence of pharmacological effect in IGD (King, Herd, & Delfabbro, 2017; Yen, Lin, Wu, & Ko, 2022). Therefore, it is critical to identify shared and distinct neurobiological mechanisms of IGD and other substance-related addictive disorders.

According to the reinforcement learning model of addiction, addiction is a transition from goal-directed behaviors into stimulus-driven habitual behaviors (Everitt & Robbins, 2005; Lüscher, Robbins, & Everitt, 2020). Goal-directed behaviors are regulated by the "model-based" system, in which the agent computes and compares possible actions based on the outcomes of each action (Daw, Niv, & Dayan, 2005; Dayan & Niv, 2008). In contrast, habitual behaviors depend on "model-free" system, which relies on previously learned associations between reward and outcome (Daw et al., 2005; Dayan & Niv, 2008; Schultz, Dayan, & Montague, 1997). Thus, model-free decision-making is usually faster and more efficient compared to model-based decision-making but, at the same time, rigid and inflexible. While the model-based and model-free systems cooperatively and competitively arbitrate to make optimal decisions maximizing cumulative rewards (Daw et al., 2005; Drummond & Niv, 2020; Lee, Shimojo, & O'Doherty, 2014), numerous literatures demonstrated impairments in the model-based system and an overreliance on the model-free system in individuals with addiction (Groman, Massi, Mathias, Lee, & Taylor, 2019; Lucantonio, Caprioli, & Schoenbaum, 2014; Voon, Reiter, Sebold, & Groman, 2017). For example, individuals with alcohol use disorder (AUD) and binge drinkers exhibited deficits in model-based control (Chen et al., 2021; Doñamayor, Strelchuk, Baek, Banca, & Voon, 2018; Sebold et al., 2014, 2017; Voon et al., 2015), and severity of alcohol addiction was negatively correlated with model-based behavior in the general population (Gillan, Kosinski, Whelan, Phelps, & Daw, 2016). Furthermore, individuals with gambling disorder, a behavioral addiction without the confound of substance's neurotoxicity, also showed impaired model-based control (Wyckmans et al., 2019).

While these findings indicate impaired model-based systems as a common psychopathology in addiction, other studies did not report such impairments in individuals with AUD (Voon et al., 2015), high familial AUD risk groups (Reiter, Deserno, Wilbertz, Heinze, & Schlagenhauf, 2016), and young social drinkers (Nebe et al., 2018). Moreover, to the best of our knowledge, the model-based system related to

IGD has not been investigated yet. This represents a critical gap in understanding IGD, as applying reinforcement learning models in other addictive disorders have effectively captured the aberrant decision-making processes that could reveal essential features of addiction (Groman, Thompson, Lee, & Taylor, 2022; Gueguen, Schweitzer, & Konova, 2021). Therefore, to examine whether IGD has similar alterations in model-based systems, it is crucial to compare neuro-cognitive mechanisms of IGD with those of other substance-related addictive disorders in terms of model-based control. As model-based behavior could be influenced by various factors such as task structure, task instruction, and participants' cognitive capability (Schad et al., 2014; Silva & Hare, 2020), neuroimaging studies that simultaneously compare different types of addiction are required.

In this study, we investigated model-based and model-free systems in individuals with IGD, individuals with AUD, and healthy controls (HC) using a two-stage task that can assess the balance between the two systems. By applying computational modeling, we aimed to compare latent cognitive processes regarding model-based control in IGD and AUD. Additionally, we used model-based functional magnetic resonance imaging (fMRI) analysis to compare neural activities associated with model-based and model-free systems in IGD and AUD.

METHODS

Participants

Participants were recruited through community and university-based advertisements in Seoul, Korea, and classified by a board-certified psychiatrist into three distinct groups based on DSM-5 criteria for AUD and IGD. Following the inclusion process (Fig. 1), a total of 74 participants were included in the behavioral and computational modeling analyses (HC, $N = 30$; IGD, $N = 22$; AUD, $N = 22$), and fMRI analysis was conducted on 69 participants after data quality check (HC, $N = 28$; IGD, $N = 20$; AUD, $N = 21$).

Measures

Two-stage task. Participants completed the two-stage task (Fig. 2), which was developed to assess a balance between model-based and model-free behavior (Daw, Gershman, Seymour, Dayan, & Dolan, 2011), inside the fMRI scanner. The task consists of two stages. In the first-stage, participants made a choice between two options, which determined the subsequent second-stage room (blue or yellow). Transitions from the first to the second stage occurred with fixed probabilities: common transitions with a probability of 0.7 and rare transitions with a probability of 0.3. In the second stage, participants made another choice between two options, followed by a feedback stage where monetary rewards were given probabilistically. A hypothetical participant who only uses a model-based strategy (model-based agent) makes choices based on the task structure, specifically the transition probabilities. On the other hand, a participant
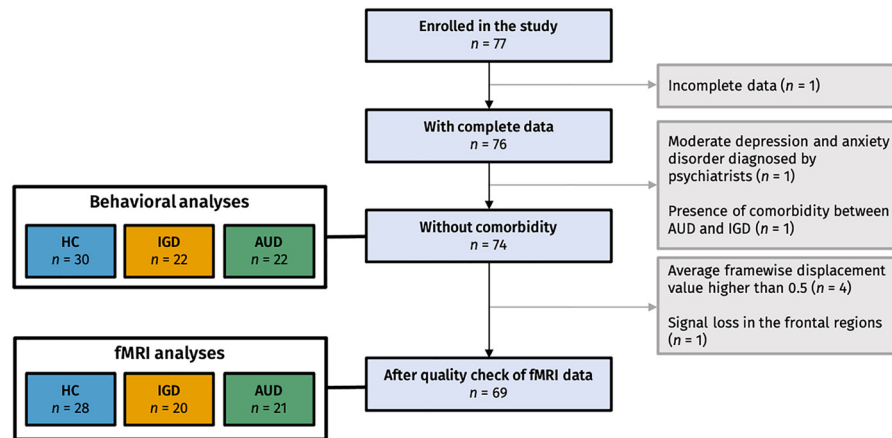
Enrolled in the study
n = 77

Incomplete data (n = 1)

With complete data
n = 76

Moderate depression and anxiety disorder diagnosed by psychiatrists (n = 1)

Presence of comorbidity between AUD and IGD (n = 1)

Behavioral analyses

HC	IGD	AUD
n = 30	n = 22	n = 22

Without comorbidity
n = 74

Average framewise displacement value higher than 0.5 (n = 4)

Signal loss in the frontal regions (n = 1)

fMRI analyses

HC	IGD	AUD
n = 28	n = 20	n = 21

After quality check of fMRI data
n = 69

Fig. 1. Participant recruitment and inclusion process

Participants were recruited from September 2018 to August 2019 through community and university-based advertisements in Seoul, Korea. Initially, 77 participants were enrolled, but three of them were subsequently excluded: one participant did not complete data collection, one participant was excluded due to moderate depression and anxiety disorder diagnosed by psychiatrists, and one participant was excluded due to the presence of comorbidity between alcohol use disorder (AUD) and internet gaming disorder (IGD). Consequently, a final sample of 74 participants were included in the behavioral analyses and computational modeling. Participants underwent evaluation by a board-certified psychiatrist using a semi-structured interview based on the Structured Clinical Interview from the DSM-IV (SCID; Kübler, 2013) to assess major psychiatric disorders. Additionally, two psychiatrists classified participants into three distinct groups based on DSM-5 criteria for AUD and IGD (HC, N = 30; IGD, N = 22; AUD, N = 22). The DSM-5's provisional diagnosis of IGD, outlined in the section recommending conditions for further research, was utilized in this study. For the functional magnetic resonance imaging (fMRI) analysis, participants with an average framewise displacement value higher than 0.5 (Power, Barnes, Snyder, Schlaggar, & Petersen, 2012) were excluded from the analysis (N = 4). Additionally, one participant was excluded from the analysis due to signal loss in the frontal regions. This resulted in a total of 69 participants included in the fMRI analysis (HC, **N** = 28; IGD, N = 20; AUD, N = 21).

using a model-free strategy (model-free agent) makes choices solely based on the outcome of the previous trial. In the following analyses, we removed non-response trials (Fig. S1). We performed one-way ANOVA with group factors (i.e., HC, IGD, AUD), revealing no significant group difference in the number of non-response trials (F(2, 66) = 0.980, p = 0.381).

Psychometric measures. Game addiction and alcohol addiction were estimated using modified versions of Young Internet Addiction Test (IAT; Young, 1998) and the Korean version of alcohol use disorder identification test (AUDIT; Saunders, Aasland, Babor, Fuente, & Grant, 1993), respectively. The severity of depressive symptoms and anxiety symptoms were measured using the Beck Depression Inventory (BDI; Beck, Steer, Ball, & Ranieri, 1996) and the Beck Anxiety Inventory (BAI; Beck, Epstein, Brown, & Steer, 1988). For details on other psychometric measures, group differences for each measure (Table S1), and correlations between each measure (Fig. S2), please refer to the Supplementary materials.

Procedure

Participants first completed assessments for the psychometric measures. Then they received instructions for the task and completed the two-stage task inside the fMRI scanner. Lastly, participants underwent an interview with a psychiatrist. Additional details on task instruction, fMRI data acquisition and preprocessing can be found in the Supplementary materials.

Statistical analysis

Behavioral analysis. We conducted a factorial analysis of choice behavior, as replicated from Daw et al. (2011), to calculate stay proportions of first-stage choices at the population level for each group. Linear mixed-effects logistic regression was performed using lme4 package (Bates, Mächler, Bolker, & Walker, 2015) in R to estimate the effects of preceding reward, transition probability, and their interaction on choice behavior. The model included random intercepts and random slopes for the effect of reward and transition probability, with participants as the random factor. We additionally examined the presence of group difference by adding the group factor as covariates in the regression model using the entire participants. Anxiety and depression scores were also included as covariates to control for the effects of psychiatric symptoms.

Computational modeling. We applied a computational model to the choice behavior based on the hybrid algorithm developed by Gläscher, Daw, Dayan, and O'Doherty (2010), referring to Daw et al. (2011). This model incorporated both model-based reinforcement learning and model-free temporal difference learning. The model-based weight parameter (ω) determined the weight given to the model-based learning. Learning rate parameter (α) determines how quickly an agent updates its expected values based on reward prediction errors (RPEs), with higher α indicating faster updates. The perseverance parameter (π) reflects an individual's tendency to persist with a previously chosen option, regardless of the
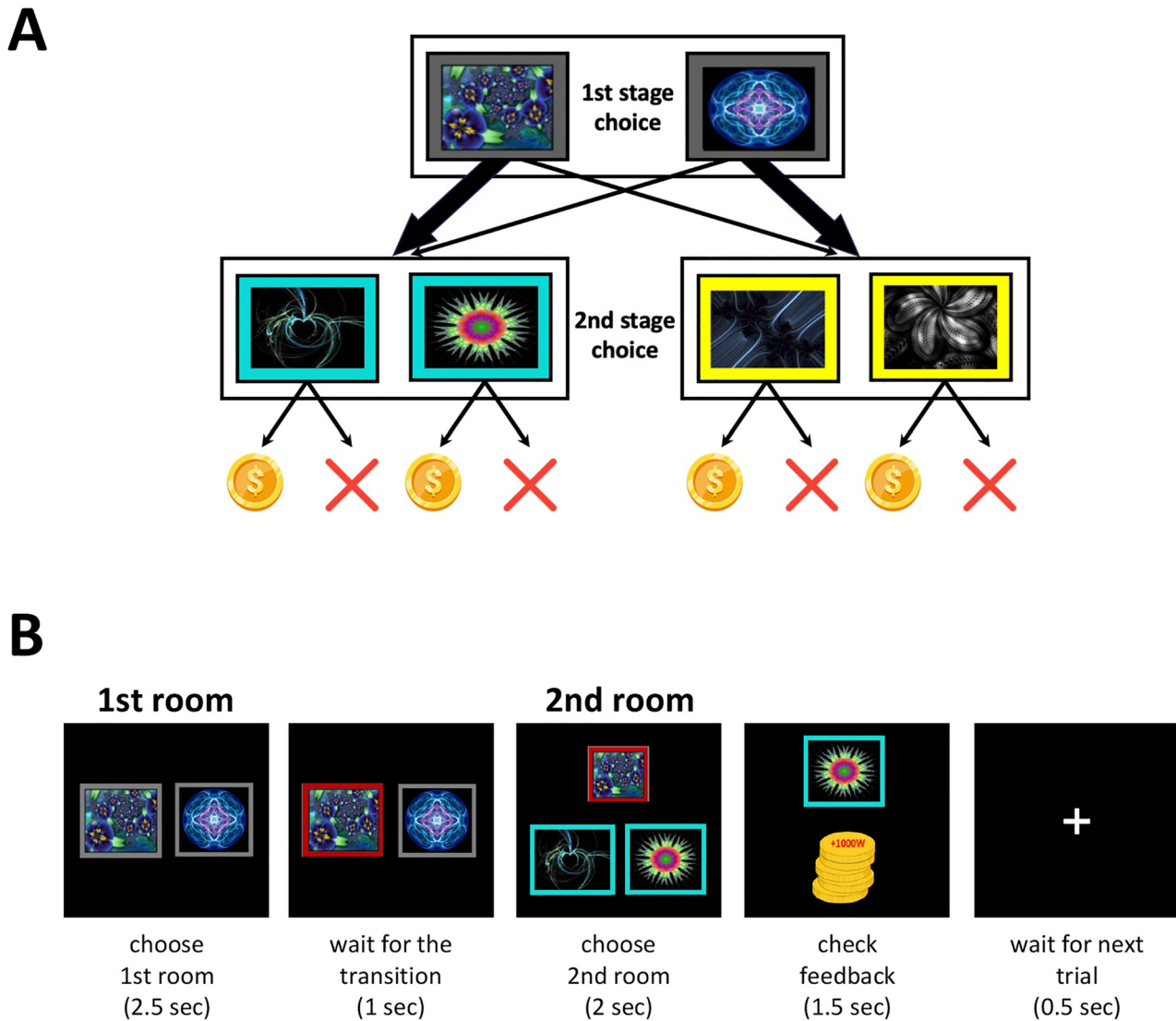
A

B
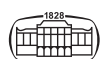
Fig. 2. Schematic representation of the Two-stage task

(A) Thick arrows represent common transitions (70%) and thin arrows indicate rare transitions (30%). For example, choosing the left option in the first stage has a 70% chance of transitioning to the blue room in the second stage, and a 30% chance of transitioning to the yellow room. The same probabilities apply in reverse for choosing the right option. After making a choice between the two options in one of the second stages, rewards were given based on slowly varying probabilities ranging from 0.25 to 0.75. Four distinct reward probability distributions were counterbalanced within each group. In the feedback stage, the chosen stimulus from the second stage remained on the screen as a reminder, and outcomes were represented by images of '+1000 W' on coins (indicating a reward) or a red 'X' (no reward). The value of 1000 won (Korean currency) is approximately 0.76 USD. A hypothetical participant who only uses a model-based strategy (model-based agent) makes choices based on the task structure, specifically the transition probabilities. If a choice in the first stage resulted in a reward through a common transition (70%), the model-based agent repeats the same choice in the next trial. However, if the reward was obtained through a rare transition (30%), the agent switches the choice. On the other hand, a participant using a model-free strategy (model-free agent) makes choices solely based on the outcome of the previous trial; selecting the same option if rewarded and selecting the opposite if not, irrespective of the transition structure. (B) Process of a single task trial with time points. Inside the fMRI scanner, participants completed 201 task trials, divided into three runs (7.5 seconds per trial, 8.38 minutes per run).

expected values, with higher π indicating a greater inclination to repeatedly choose the same option consecutively. For each trial, we computed model-free and model-based RPEs, where RPE represents the difference between the actual reward received and the expected reward (Schultz, 2016). We used hierarchical Bayesian analysis for estimating model parameters and group comparisons (Ahn, Krawitz, Kim, Busemeyer, & Brown, 2011; Kruschke, 2014). Additional

information regarding the complete equations and details on the computational model and group comparison method can be found in the Supplementary materials.

Model-based fMRI analysis. In the first-level analysis, we performed a model-based fMRI analysis (O'Doherty, Hampton, & Kim, 2007) to identify voxels of which blood oxygenation level dependent (BOLD) activity is correlated

with model-free and model-based RPEs. Time series of standard RPE estimates were extracted as a model-free regressor. For a model-based regressor, we computed difference regressor representing the residual prediction error not accounted for by model-free RPE values (Daw et al., 2011). The difference regressor was computed by subtracting the model-free RPE values from the hypothetical RPE values that would have occurred if the participant solely used a model-based strategy during the task. Both the model-free and model-based regressors were included as parametric modulators in the design matrix at the onset of second stage and feedback stage. We did not orthogonalize the regressors to control the effects of each other, and also to enhance the interpretability of the activation maps (Mumford, Poline, & Poldrack, 2015). Additional regressors include timepoints of choice response in the first and second stage, as well as the onset of fixation and first stage. Six motion regressors estimated by fMRIPrep were also added as nuisance regressors to account for head movement. In summary, there were twelve regressors in the GLM: two parametric regressors (model-free RPE and model-based RPE), four nuisance regressors indicating timepoints to be regressed out (choice time in the first and second stage, fixation onset, and first stage onset), and six motions regressors. By using the two parametric regressors of interest, two contrasts were obtained for each participant: one for identifying voxels showing BOLD activity correlated with the model-free regressor, and the other with the model-based regressor.

In the second-level analysis, first-level contrast images were entered into random effects analysis to generate second-level contrasts. Depression and anxiety scores were included as covariates to control for the effect of psychiatric symptoms. The results were thresholded at $p < 0.001$ (uncorrected for multiple comparisons) with an extent threshold of $k \geq 10$ voxels. This criterion was adopted to discern potentially meaningful neural activations between the groups, given the exploratory nature of the investigation.

We then examined correlations between the neural correlates showing significant group differences and the model-based weight parameter (ω). This analysis aimed to explore the relationship between model-based behavior and the strength of the neural correlates associated with the model-based and model-free systems, capturing how the strength of each neural correlate varied based on the degree of model-based control. Mean beta values of peak voxels (3 mm sphere) in each significant brain region were extracted for each individual, and Pearson correlations were computed with individual estimates of ω.

Functional connectivity analysis - psychophysiological interaction analysis. In the model-based fMRI analysis, we found that the insula played a critical role in model-based learning, particularly in the IGD group. To examine how the insula interacts with other brain regions in the context of model-based learning, we conducted a psychophysiological interaction (PPI) analysis (Friston et al., 1997). As insular cortices are known as cortical hubs of the salience network (Seeley, 2019; Seeley et al., 2007), we examined functional connectivity between the bilateral insula and other brain regions related to the salience network (see Supplementary materials). In the first-level analysis, the left and right insula (i.e., 3 mm sphere centered on peak voxels from the second-level model-based fMRI analysis) were used as seed regions, and the model-free and model-based RPEs were included as parametric modulators. Four nuisance regressors and six motion regressors used in the model-based fMRI were also included. In the second-level analysis, one-sample t-tests were conducted within each group to examine connectivity patterns. We also conducted two-sample t-tests (HC vs. IGD; IGD vs. AUD; HC vs. AUD) to examine group differences. To mitigate the influence of psychiatric symptoms, depression and anxiety scores were integrated as covariates. The results were thresholded at $p < 0.001$ (uncorrected) with an extent threshold of $k \geq 10$ voxels and were small-volume-corrected using the brain mask of the salience network.

Ethics

The study was conducted in accordance with the Institutional Review Board at Severance Hospital, Seoul, Korea (IRB No. 40-2014-0745). All participants were provided with detailed information about the study protocol and provided written informed consent before participating.

RESULTS

Behavioral results

The observed stay probability (Fig. 3) demonstrates a combination of model-based and model-free learning in all three groups, which is further supported by the results of linear mixed effects logistic regression (Table S3). The main effects of reward (HC, $p < 1e-4$; IGD, $p < 1e-10$; AUD, $p < 1e-06$) and transition probability (HC, $p < 1e-9$; IGD, $p < 9e-10$; AUD, $p = 0.001$) were significant in all three groups, indicating that all groups accounted for both rewards (i.e., model-free learning) and the transition structure (i.e., model-based learning) when making decisions. Additionally, the interaction between reward and transition probability also showed a significant effect, supporting the presence of a mixture of model-based and model-free strategies across the groups (HC, $p < 2e-16$; IGD, $p < 2e-16$; AUD, $p < 2e-16$). However, the regression model did not reveal a significant main effect of the group, even with the inclusion of anxiety and depression as covariates ($p > 0.1$)."

Modeling results

We found significant group differences in the second-stage learning rate parameter α_2, and perseverance parameter π (Fig. 4). Both the IGD and AUD groups exhibited higher α_2 ($\alpha_{2,IGD} - \alpha_{2,HC}$ 95% highest density interval (HDI) = [0.072, 0.359]; $\alpha_{2,AUD} - \alpha_{2,HC}$ 95% HDI = [0.011, 0.313]) and higher π estimates ($\pi_{IGD} - \pi_{HC}$ 95% HDI = [0.100, 1.059]; $\pi_{AUD} - \pi_{HC}$ 95% HDI = [0.010, 1.021]) compared to the HC group. However, no credible group differences were
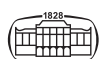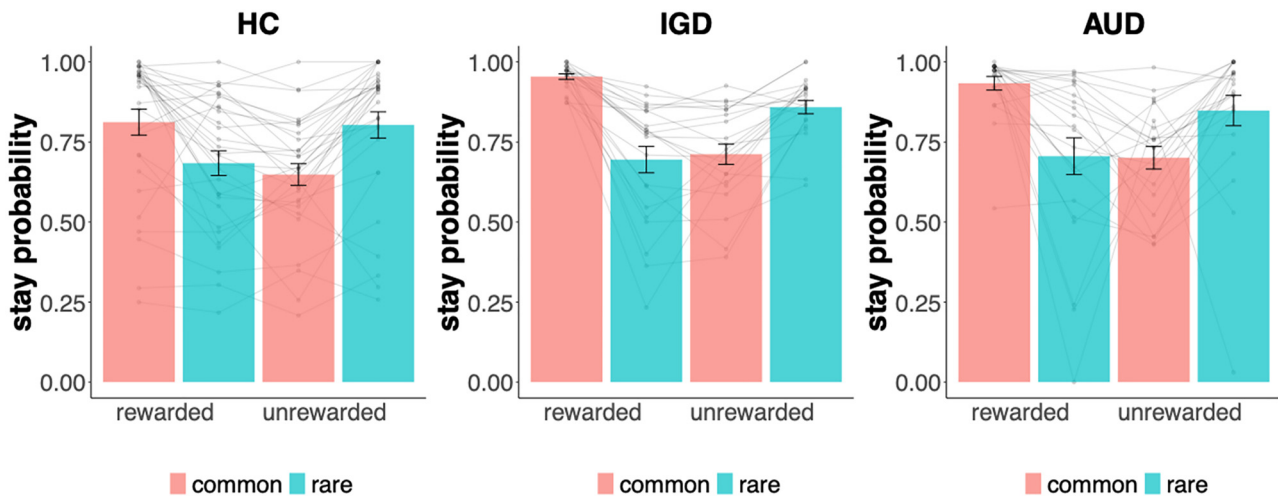
Fig. 3. Stay probability of each group
The figure displays the actual stay proportion of each group (HC, healthy control; IGD, internet gaming disorder; AUD, alcohol use disorder) based on reward and transition probability. Each data point represents the individual participant's data, connected by lines. The error bars represent the standard error of the mean.

observed in the other four parameters including the model-based weight parameter ω. See Fig. S4 for the distributions of group differences in the model parameter estimates.

Model-based fMRI results

The model-based fMRI analysis revealed significant group difference between the HC, IGD, and AUD groups, when the model-based RPE was employed as a parametric modulator (Fig. 5A and B). Specifically, the activation of the right orbitofrontal cortex (OFC) in the AUD group was greatly correlated with the model-based RPE, compared to the IGD group ($t = 4.22$, $k = 13$, $p < 0.001$). Furthermore, the bilateral insular activation of the IGD group showed greater correlation with the model-based RPE, compared to the HC group (left: $t = 4.20$, $k = 17$, $p < 0.001$; right: $t = 4.71$, $k = 33$, $p < 0.001$). Lastly, activation of the left superior frontal gyrus (SFG) of the AUD group was greatly correlated with the model-based RPE, compared to the HC group ($t = 4.18$, $k = 26$, $p < 0.001$). However, no significant group differences were observed when utilizing the model-free RPE as a parametric modulator. See Table S4 for the second level results of each group.

Further analysis examined correlation between the significant results of the model-based fMRI analysis (i.e., correlation of each region with model-based RPE) and the level of model-based control (e.g., individual estimates of ω). We found a negative correlation between the beta value of the right insula and ω in the IGD group ($r = -0.45$, $p < 0.05$), but not in the AUD group ($r = -0.09$, $p = 0.709$) or the HC group ($r = -0.19$, $p = 0.334$) (Fig. 5C). In other words, the coupling between the right insula and the model-based RPE was stronger in individuals with lower model-based weight parameter estimates, only in the IGD group. No significant correlations were found for the left insula.

PPI results

The PPI analysis revealed significant correlations between the insula and other brain regions in the salience network, only in the IGD group (Fig. 6 & Table S5). The right insula was correlated with the right putamen ($t = 5.23$, $p < 0.001$), left insula ($t = 5.75$, $p < 0.001$), and occipital lobe ($t = 4.23–4.37$, $p < 0.001$) in the context of model-based RPE. Similarly, the left insula was correlated with the anterior cingulate cortex (ACC; $t = 4.24$, $p < 0.001$), right superior temporal gyrus (STG; $t = 3.67$, $p < 0.001$), and occipital lobe ($t = 3.72–5.14$, $p \leq 0.001$). Group comparisons indicated distinctive connectivity patterns in the IGD group (Table S6). Connectivity between the right insula and the left insula was greater in the IGD group compared to HC ($t = 3.89$, $p < 0.001$). Furthermore, connectivity between the insula and occipital lobe was greater in the IGD group compared to the AUD (right insula as seed: $t = 3.52–4.41$, $p < 0.001$; left insula as seed: $t = 3.52–3.57$, $p < 0.001$) and HC (left insula as seed: $t = 3.41–3.79$, $p < 0.001$). Overall, our findings demonstrate that IGD is associated with distinctive patterns of hyper-connectivity in the insula and its interactions with the salience network, suggesting a potential link to ongoing reward processes unique to the IGD group during model-based learning.

DISCUSSION

In this study, we compared neurocognitive features related to model-based learning in individuals with IGD and AUD. The main findings are as follows: (1) Both the IGD and AUD groups showed higher learning rate and perseverance parameter estimates compared to the HC group, contrary to our initial hypothesis. (2) We found an IGD-specific role of the insula in processing model-based behavior. Model-based
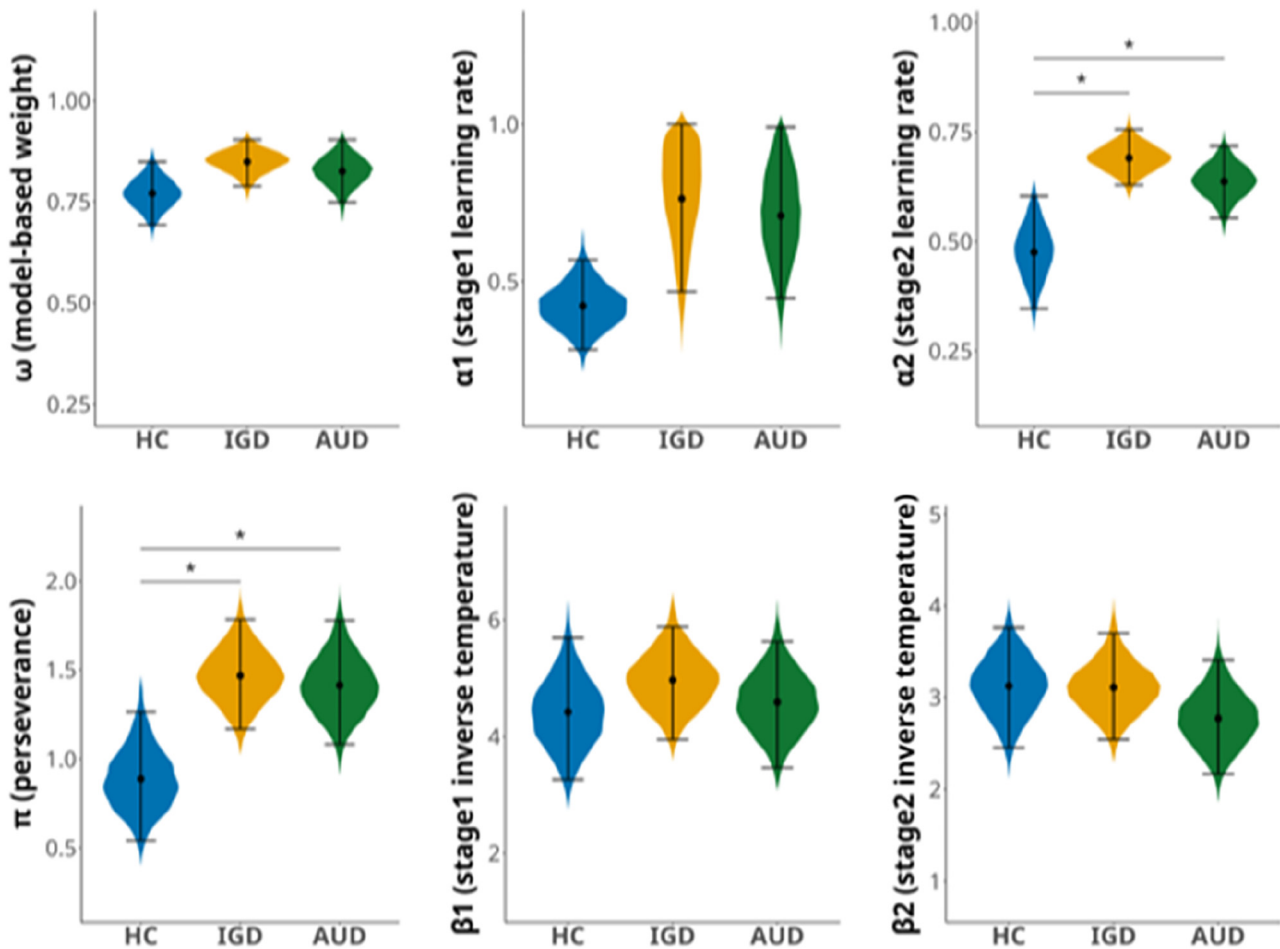
Fig. 4. Group comparison of parameter estimates
Each point is a group-wise mean of the estimates of six model parameter (HC, healthy control; IGD, internet gaming disorder; AUD, alcohol use disorder). The model-based weight parameter ω represents the degree of model-based control, with higher ω values indicating greater emphasis on model-based learning. Learning rate parameters (α_1, α_2) indicate how quickly reward prediction errors are updated in the temporal difference learning, independently for stage1 and stage2, with higher values indicating faster updates. The perseverance parameter (π) represents whether the participant has a tendency to choose the same option as on the previous trial, with higher π indicating a greater tendency to repeat choices. Inverse temperature parameters (β_1, β_2) determine the level of stochasticity in the participant's choice, with higher values indicating more deterministic choices. Each error bar represents the 95% highest density interval (HDI) of the parameter estimates. Asterisks indicate significant group differences. See Figure S4 for the distributions of the group differences.

RPE was strongly correlated with bilateral insula activation in the IGD group, while it showed correlation with the activation of the frontal regions (i.e., right OFC and left SFG) in the AUD group. Notably, only the IGD group exhibited hyper-connectivity between the bilateral insula and other brain regions in the salience network, including the putamen, ACC, STG, and occipital lobe, in the context of model-based RPE. In addition, we observed that the correlation between the right insula and model-based RPE was particularly stronger in individuals who showed lower model-based behavior (i.e., lower ω).

The finding of similarly high model-based behavior across the groups was consistently observed in both behavior and computational modeling analyses. While previous studies have yielded mixed findings regarding the impairment of model-based behavior in individuals with AUD (Chen et al., 2021; Doñamayor et al., 2018; Gillan et al., 2011; Reiter et al., 2016; Sebold et al., 2014; Voon et al., 2015), a recent study by Silva and Hare (2020) suggested that humans predominantly rely on model-based inference when they have an accurate conception of the task. In our study, we implemented several procedures recommended by Silva and Hare (2020) to ensure a clear understanding of the two-stage task (see Methods). Additionally, participants in our study exhibited above-normal working memory and processing speed, surpassing those reported in previous studies. These higher cognitive capabilities may have contributed to the enhanced model-based behavior observed in our participants (Schad et al., 2014). In essence, accurate understanding of the task along with higher cognitive capabilities of the participants could account for the preserved model-based behavior of IGD and AUD when compared to HC. Nevertheless, it is imperative to note that these speculative interpretations require validation through further
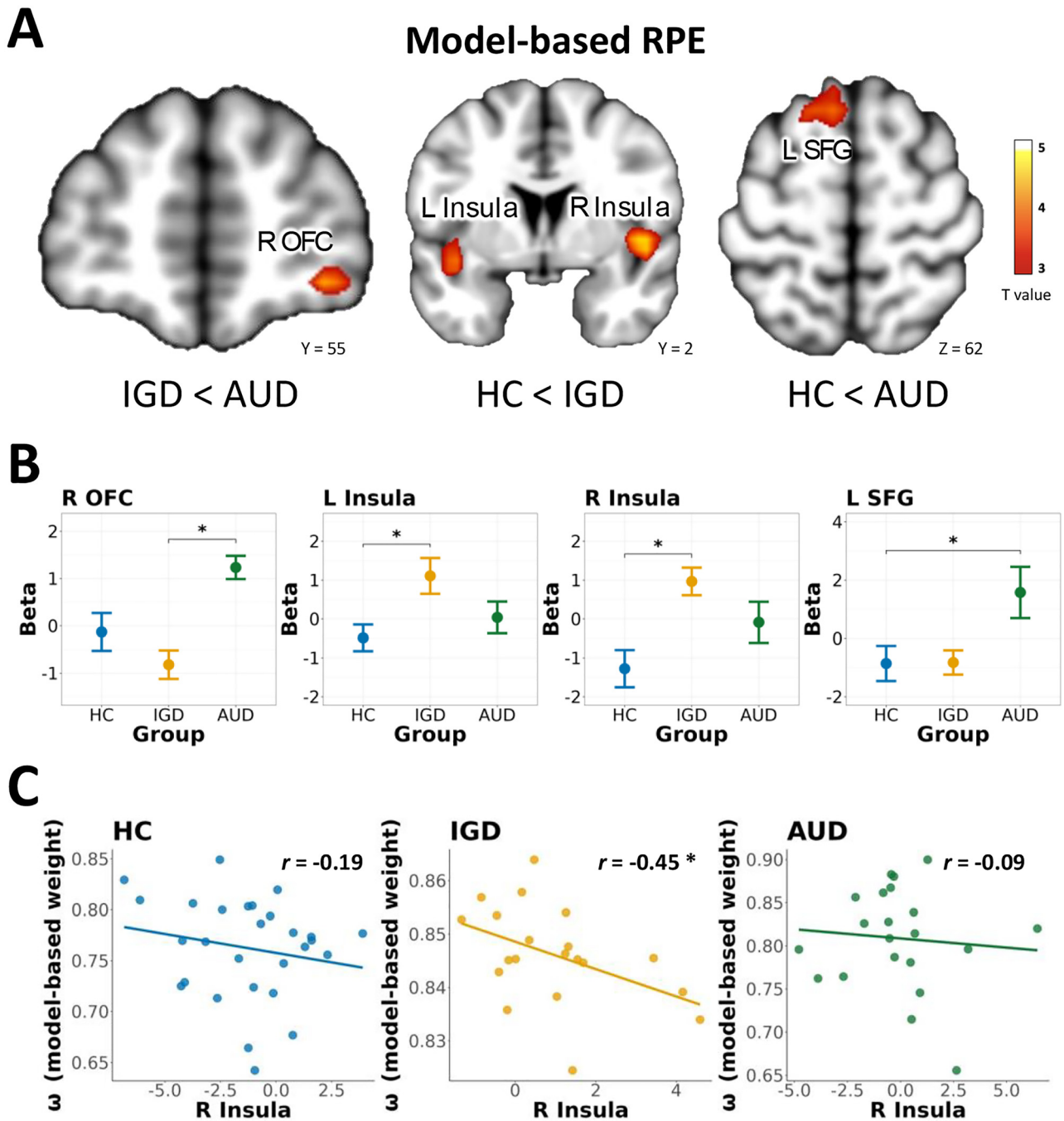
Fig. 5. Group differences in model-based fMRI analysis

(A) Brain regions exhibiting significant group differences in the second-level two-sample comparisons using model-based reward prediction error (RPE) as parametric modulators ($p < 0.001$, uncorrected; cluster size, $k \geq 10$). Significant group differences indicate that one group's brain activation is significantly correlated with the model-based RPE compared to the other group. The color bar indicates t-statistics from the results of two-sample t-tests. (B) Group-wise average beta value extracted from 3 mm spheres at peak MNI coordinates (R OFC: 35, 55, −8; L Insula: −40, 2, −8; R Insula, 46, 2, 2; L SFG: −6, 25, 62) for the brain regions showing significant group differences. Each dot represents the group-wise mean of the beta value for each region, and each error bar indicates the group-wise standard error. Asterisks denote significance based on two-sample t-test ($p < 0.001$). (C) Correlation between beta value of the left insula and the model-based weight parameter (ω) estimates. Each dot represents the beta value of the right insula (extracted from 3 mm spheres at peak MNI coordinates: 46, 2, 2) from the first-level analysis on the x-axis, and the individual estimates of the model-based weight parameter on the y-axis. The regression line represents the Pearson correlation between the beta value of the right insula and the estimates of ω. The asterisk denotes a significant correlation ($p < 0.05$). OFC = orbitofrontal cortex; SFG = superior frontal gyrus.
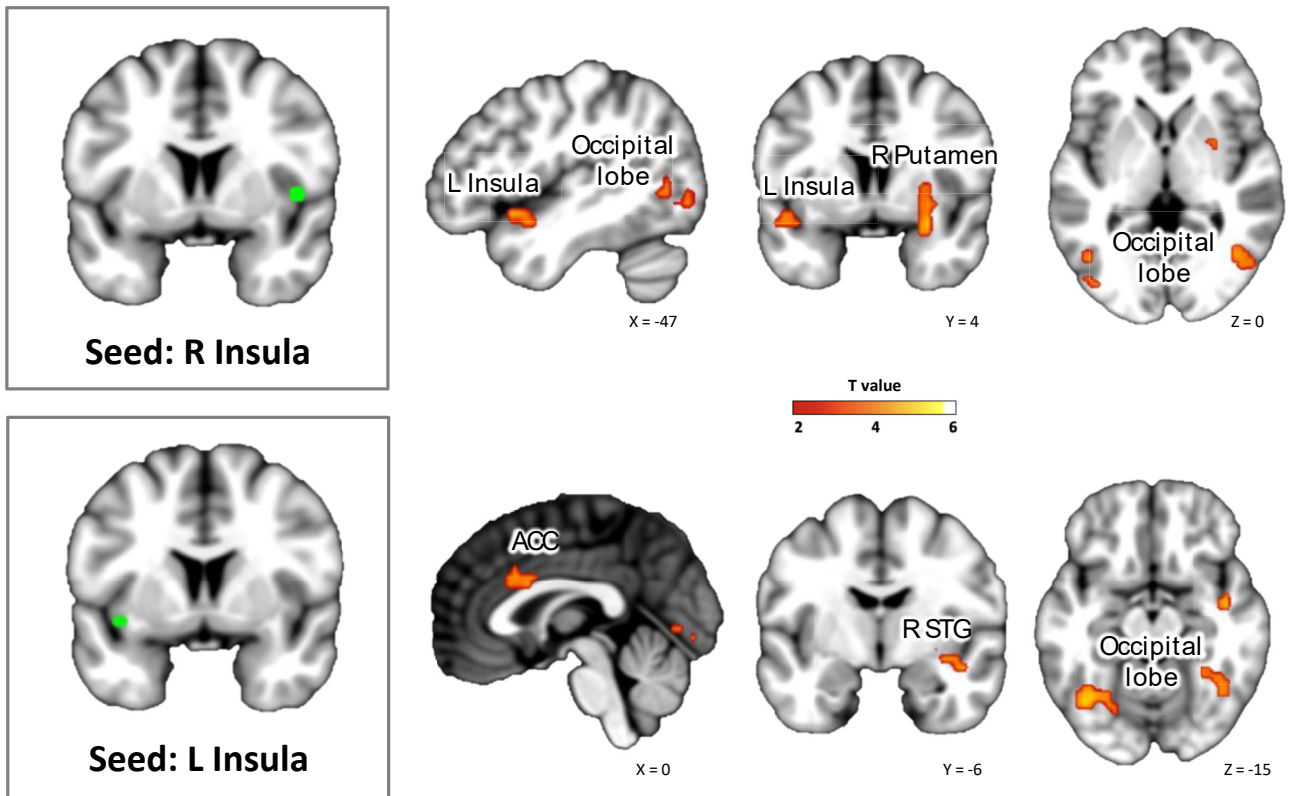
Fig. 6. Results of psychophysiological interaction (PPI) analysis results of the internet gaming disorder group
Using the right insula (MNI-coordinates: −40, 2, −8) and left insula (MNI-coordinates: 46, 2, 2) as seed regions (3 mm spheres), a PPI analysis was conducted with model-based reward prediction error (RPE) as the psychological variable. The brain maps display the effects of model-based RPE on each insula to the whole brain of the IGD group ($p < 0.001$, uncorrected; cluster size, $k \geq 10$). The color bar represents t-statistics from the results of one-sample *t*-tests. See Table S5 for the MNI coordinates of the peak voxels.

investigation. Future studies should encompass samples with diverse cognitive capabilities and manipulate participants' understanding of the task to substantiate and generalize the observed associations.

While the degree of model-based control (i.e., ω) was similar across groups, the IGD and AUD groups exhibited higher α_2 and π estimates, indicating shared cognitive aspects between alcohol and gaming addiction. Impulsivity, a well-known personality trait related to addiction (Bickel & Marsch, 2001; Koob & Volkow, 2016; Kozak et al., 2019; Poulton & Hester, 2019), might account for the higher learning rate (α_2) estimates in terms of reward sensitivity and cognitive control. The faster updates of immediate reward found in the addicted individuals may indicate enhanced salience of immediate rewards or impaired ability to inhibit responses to immediate rewards (Wu et al., 2017). Additionally, while a higher learning rate can be advantageous in a rapidly changing and uncertain environment, it might be detrimental in a setting like the two-stage task, where reward probability changes gradually (Eckstein & Collins, 2020).

Higher perseverance parameter (π) estimates in the IGD and AUD groups suggest a greater tendency to repeat previous choices regardless of reward, which could be related to habitual behavior or compulsivity, key characteristics of addictive disorders (Everitt & Robbins, 2015; Lucantonio et al., 2014; Lüscher et al., 2020; Ostlund & Balleine, 2009). Previous literature also reported elevated perseverance in addictive individuals, supporting the idea that high perseveration could be a detrimental feature related to addiction (Doñamayor et al., 2022; Ersche et al., 2011). Therefore, the shared pattern of higher learning rate and perseverance found in the IGD and AUD groups might reflect the neurocognitive characteristics of addictive behaviors, though further investigation is warranted.

Using fMRI, we found distinctive neural correlates of model-based behavior in the IGD compared to the AUD groups. Specifically, we observed that the insula plays a specific role in model-based behavior in IGD. The insula is a key component of the brain's salience network (Menon & Uddin, 2010; Uddin, 2015), involved in representing prediction errors related to reward variance (Preuschoff, Quartz, & Bossaerts, 2008) and processing salient stimuli (Jensen et al., 2007; Seeley et al., 2007). Previous literature reported hyper-connectivity of the salience network in individuals with behavioral addiction during resting state (Tolomeo & Yu, 2022), and also specifically in individuals with IGD during resting state, risky decision making, and executive control tasks (Hong et al., 2015; Lee, Lee, Lee, & Jung, 2017; Sun et al., 2012; Zheng et al., 2019). Consistent

with these findings, our results suggest that hyper-sensitivity of the insula, a key component of the salience network, might serve as a distinct neural marker of IGD.

To comprehend the hyperactivation of the insula specifically in the IGD group, it is crucial to consider its involvement in drug craving and addiction (Droutman, Read, & Bechara, 2015). Brain lesions in the insular cortex are known to disrupt addictive behavior, emphasizing the insula's significant role in addiction (Naqvi, Gaznick, Tranel, & Bechara, 2014). However, neuroimaging studies on SUDs have generally reported hypoactivation of insula during decision-making tasks (Nestor, Hester, & Garavan, 2010; Stewart, Connolly, et al., 2014; Stewart, May, et al., 2014). This discrepancy may be attributed to the differential processing of drug rewards and non-drug rewards in individuals with addiction (Madden, Petry, Badger, & Bickel, 1997). Drugs of addiction are highly salient reward, leading to hyperactivation of the salience network, while non-drug rewards are associated with decreased activity in the salience network compared to non-drug users (Cushnie, Tang, & Heilbronner, 2023). Therefore, the previous findings of hypoactivation of the insula during decision-making tasks in SUDs may be due to the use of non-salient rewards. Correspondingly, neuroimaging studies of IGD using gaming-related cues have indicated heightened activation of the salience network (Ko et al., 2009, 2013; Turel, He, Wei, & Bechara, 2021). Considering the task dynamics and reward structure in our study, which may resemble gaming behavior at least more than substance use, the excessive insula activation observed in the IGD group may reflect that the individuals treated the two-stage task like a game.

Another notable finding is that individuals with AUD exhibited hyperactivation in the prefrontal regions, specifically the OFC and SFG, in the context of model-based behavior. The OFC has been identified as a key regulator of goal-directed, or model-based action planning and execution (Gremel & Costa, 2013; Jones et al., 2012; Kahnt, 2023; McDannald, Lucantonio, Burke, Niv, & Schoenbaum, 2011), encoding the value of stimuli and the relationship between stimuli and their expected outcomes (Howard, Gottfried, Tobler, & Kahnt, 2015; Lopatina et al., 2015). Similarly, the SFG is involved in higher cognitive functions and working memory associated with executive processing (Boisgueheneuc et al., 2006). Interestingly, while previous literature suggested decreased prefrontal activation in individuals with AUD during decision-making tasks and resting state, the individuals with AUD in our study showed increased activation of these regions. This discrepancy could be explained by differences in model-based performance, as we did not detect impairment of model-based control in AUD, contrary to previous findings of AUD or SUDs (Dom, Sabbe, Hulstijn, & Brink, 2005; Moorman, 2018). Notably, these findings align with previous studies that suggesting compensatory hyperactivation in prefrontal regions in cases of AUD, aiming to maintain task performance (López-Caneda et al., 2012). Building upon this prior finding, the hyperactivation of the prefrontal regions in the AUD group may be interpreted as efforts to counteract impaired model-

based systems, thus enabling model-based behavior. Importantly, compensatory hyperactivation in the frontal regions is specific to AUD and not observed in IGD. This distinction may be attributed to the absence of chemical intoxication in IGD, unlike in AUD and other SUDs (Grant, Potenza, Weinstein, & Gorelick, 2010; Han et al., 2015; Oscar-Berman & Marinković, 2007; Weinstein & Lejoyeux, 2020).

To the best of our knowledge, this is the first neuroimaging study that compared IGD and AUD focusing on the model-based behavior. Our findings suggest that IGD and AUD have shared mechanisms regarding the model-based behavior, indicating common cognitive characteristic related to reward sensitivity and compulsivity. However, the underlying neural mechanisms may differ due to various factors such as differences in reward of salience and the distinct brain changes related to alcohol and gaming behavior. Our findings provide valuable insights into the neurocognitive mechanisms underlying addictive disorders and highlight the need for further research to explore the complex role of insula and salience network in IGD. Nonetheless, this study has several limitations. First, we did not jitter the duration of fixation between trials, using 0.5-second-fixation for all inter-trial intervals. This methodological limitation may result in trial-preparatory activity and also overlap in the BOLD responses across trials (Friston, Zarahn, Josephs, Henson, & Dale, 1999). Thus, we recommend future studies to use jittering for inter-trial intervals, as well as sufficient intervals between the trials. Second, as this is the first study to compare neural correlates of model-based behavior between IGD and AUD, we acknowledge the broadness of our statistical methods and lack of specificity in exploring the common and specific neural correlates of each group. Thus, further investigation with more specificity is needed to identify the group difference in the neural correlates. Lastly, the generalizability of the findings is limited not only due to the relatively modest sample sizes of each group, but also the characteristics of the samples. Although participants were diagnosed with IGD or AUD by psychiatrists, they were not clinical patients seeking for treatment. Instead, they were recruited from community sites, which could potentially lead to variations in clinical characteristics compared to patients with more severe symptoms. Furthermore, the sample is restricted to young men in terms of gender and age, and the majority of individuals in the IGD group was predominantly engaged in a specific game genre, multiplayer online battle arena (MOBA), which may be influenced by Korean culture. Therefore, future studies should consider the influence of culture, and also use larger sample sizes and more stringent statistical thresholds.

CONCLUSIONS

Our study provides novel insights into the neurocognitive features and neural correlates of model-based learning in individuals with IGD and AUD. Despite both IGD and AUD exhibiting similar levels of model-based behavior, distinct

neural signatures were observed in the insula for IGD and prefrontal regions for AUD. These findings suggest potential differences in the neurobiological mechanisms underlying addictive behaviors in IGD and AUD, contributing to the growing body of evidence highlighting shared and distinct features of IGD and substance-related addictive disorders. Further prospective research is needed to better understand pathophysiology of IGD, and to address the controversy surrounding the diagnostic criteria and treatment approaches specific to IGD.

Authors' contribution: MK and HC are equal contributors to this work and designated as co-first authors. Y-CJ and W-YA are designated as co-corresponding authors. MK conducted statistical analysis, discussed and interpreted the results, and wrote the manuscript. HC designed and performed the experiments, discussed and interpreted the results, and wrote the manuscript. HP conducted statistical analysis for computational modeling. W-YA and Y-CJ obtained funding, supervised the project, and edited the paper. All authors had full access to all data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Conflicts of interest: The authors declare no conflict of interest.

SUPPLEMENTARY DATA

Supplementary data to this article can be found online at https://doi.org/10.1556/2006.2024.00006.

REFERENCES

Ahn, W.-Y., Krawitz, A., Kim, W., Busemeyer, J. R., & Brown, J. W. (2011). A model-based fMRI analysis with hierarchical bayesian parameter estimation. *Journal of Neuroscience, Psychology, and Economics*, 4(2), 95–110. https://doi.org/10.1037/a0020684.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1). https://doi.org/10.18637/jss.v067.i01.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology*, 56(6), 893–897. https://doi.org/10.1037/0022-006x.56.6.893.

Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. F. (1996). Comparison of beck depression inventories-IA and-II in psychiatric outpatients. *Journal of Personality Assessment*, 67(3), 588–597. https://doi.org/10.1207/s15327752jpa6703_13.

Bickel, W. K., & Marsch, L. A. (2001). Toward a behavioral economic understanding of drug dependence: Delay discounting processes. *Addiction*, 96(1), 73–86. https://doi.org/10.1046/j.1360-0443.2001.961736.x.

Boisgueheneuc, F. du, Levy, R., Volle, E., Seassau, M., Duffau, H., Kinkingnehun, S., … Dubois, B. (2006). Functions of the left superior frontal gyrus in humans: A lesion study. *Brain*, 129(12), 3315–3328. https://doi.org/10.1093/brain/awl244.

Chen, H., Mojtahedzadeh, N., Belanger, M. J., Nebe, S., Kuitunen-Paul, S., Sebold, M., … Smolka, M. N. (2021). Model-based and model-free control predicts alcohol consumption developmental trajectory in young adults: A 3-year prospective study. *Biological Psychiatry*, 89(10), 980–989. https://doi.org/10.1016/j.biopsych.2021.01.009.

Cushnie, A. K., Tang, W., & Heilbronner, S. R. (2023). Connecting circuits with networks in addiction neuroscience: A salience network perspective. *International Journal of Molecular Sciences*, 24(10), 9083. https://doi.org/10.3390/ijms24109083.

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, 69(6), 1204–1215. https://doi.org/10.1016/j.neuron.2011.02.027.

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, 8(12), 1704–1711. https://doi.org/10.1038/nn1560.

Dayan, P., & Niv, Y. (2008). Reinforcement learning: The good, the bad and the ugly. *Current Opinion in Neurobiology*, 18(2), 185–196. https://doi.org/10.1016/j.conb.2008.08.003.

Doñamayor, N., Ebrahimi, C., Arndt, V. A., Weiss, F., Schlagenhauf, F., & Endrass, T. (2022). Goal-directed and habitual control in human substance use: State of the art and future directions. *Neuropsychobiology*, 81(5), 403–417. https://doi.org/10.1159/000527663.

Doñamayor, N., Strelchuk, D., Baek, K., Banca, P., & Voon, V. (2018). The involuntary nature of binge drinking: Goal directedness and awareness of intention. *Addiction Biology*, 23(1), 515–526. https://doi.org/10.1111/adb.12505.

Dom, G., Sabbe, B., Hulstijn, W., & Brink, W. van D. (2005). Substance use disorders and the orbitofrontal cortex. *British Journal of Psychiatry*, 187(3), 209–220. https://doi.org/10.1192/bjp.187.3.209.

Dowling, N. A. (2014). Issues raised by the DSM-5 internet gaming disorder classification and proposed diagnostic criteria. *Addiction*, 109(9), 1408–1409. https://doi.org/10.1111/add.12554.

Droutman, V., Read, S. J., & Bechara, A. (2015). Revisiting the role of the insula in addiction. *Trends in Cognitive Sciences*, 19(7), 414–420. https://doi.org/10.1016/j.tics.2015.05.005.

Drummond, N., & Niv, Y. (2020). Model-based decision making and model-free learning. *Current Biology*, *30*(15), R860–R865. https://doi.org/10.1016/j.cub.2020.06.051.

Eckstein, M. K., & Collins, A. G. E. (2020). Computational evidence for hierarchically structured reinforcement learning in humans. *Proceedings of the National Academy of Sciences*, *117*(47), 29381–29389. https://doi.org/10.1073/pnas.1912330117.

Ersche, K. D., Roiser, J. P., Abbott, S., Craig, K. J., Müller, U., Suckling, J., … Bullmore, E. T. (2011). Response perseveration in stimulant dependence is associated with striatal dysfunction and can be ameliorated by a D2/3 receptor agonist. *Biological Psychiatry*, *70*(8), 754–762. https://doi.org/10.1016/j.biopsych.2011.06.033.

Everitt, B. J., & Robbins, T. W. (2005). Neural systems of reinforcement for drug addiction: From actions to habits to compulsion. *Nature Neuroscience*, *8*(11), 1481–1489. https://doi.org/10.1038/nn1579.

Everitt, B. J., & Robbins, T. W. (2015). Drug addiction: Updating actions to habits to compulsions ten years on. *Annual Review of Psychology*, *67*(1), 1–28. https://doi.org/10.1146/annurev-psych-122414-033457.

Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *NeuroImage*, *6*(3), 218–229. https://doi.org/10.1006/nimg.1997.0291.

Friston, K. J., Zarahn, E., Josephs, O., Henson, R. N. A., & Dale, A. M. (1999). Stochastic designs in event-related fMRI. *NeuroImage*, *10*(5), 607–619. https://doi.org/10.1006/nimg.1999.0498.

Gillan, C. M., Kosinski, M., Whelan, R., Phelps, E. A., & Daw, N. D. (2016). Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *ELife*, *5*, e11305. https://doi.org/10.7554/elife.11305.

Gillan, C. M., Papmeyer, M., Morein-Zamir, S., Sahakian, B. J., Fineberg, N. A., Robbins, T. W., & Wit, S. de. (2011). Disruption in the balance between goal-directed behavior and habit learning in obsessive-compulsive disorder. *The American Journal of Psychiatry*, *168*(7), 718–726. https://doi.org/10.1176/appi.ajp.2011.10071062.

Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, *66*(4), 585–595. https://doi.org/10.1016/j.neuron.2010.04.016.

Grant, J. E., Potenza, M. N., Weinstein, A., & Gorelick, D. A. (2010). Introduction to behavioral addictions. *The American Journal of Drug and Alcohol Abuse*, *36*(5), 233–241. https://doi.org/10.3109/00952990.2010.491884.

Gremel, C. M., & Costa, R. M. (2013). Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nature Communications*, *4*(1), 2264. https://doi.org/10.1038/ncomms3264.

Groman, S. M., Massi, B., Mathias, S. R., Lee, D., & Taylor, J. R. (2019). Model-free and model-based influences in addiction-related behaviors. *Biological Psychiatry*, *85*(11), 936–945. https://doi.org/10.1016/j.biopsych.2018.12.017.

Groman, S. M., Thompson, S. L., Lee, D., & Taylor, J. R. (2022). Reinforcement learning detuned in addiction: Integrative and translational approaches. *Trends in Neurosciences*, *45*(2), 96–105. https://doi.org/10.1016/j.tins.2021.11.007.

Gueguen, M. C., Schweitzer, E. M., & Konova, A. B. (2021). Computational theory-driven studies of reinforcement learning and decision-making in addiction: What have we learned? *Current Opinion in Behavioral Sciences*, *38*, 40–48. https://doi.org/10.1016/j.cobeha.2020.08.007.

Han, J. W., Han, D. H., Bolo, N., Kim, B., Kim, B. N., & Renshaw, P. F. (2015). Differences in functional connectivity between alcohol dependence and internet gaming disorder. *Addictive Behaviors*, *41*, 12–19. https://doi.org/10.1016/j.addbeh.2014.09.006.

Hong, S.-B., Harrison, B. J., Dandash, O., Choi, E.-J., Kim, S.-C., Kim, H.-H., … Yi, S.-H. (2015). A selective involvement of putamen functional connectivity in youth with internet gaming disorder. *Brain Research*, *1602*, 85–95. https://doi.org/10.1016/j.brainres.2014.12.042.

Howard, J. D., Gottfried, J. A., Tobler, P. N., & Kahnt, T. (2015). Identity-specific coding of future rewards in the human orbitofrontal cortex. *Proceedings of the National Academy of Sciences*, *112*(16), 5195–5200. https://doi.org/10.1073/pnas.1503550112.

Jensen, J., Smith, A. J., Willeit, M., Crawley, A. P., Mikulis, D. J., Vitcu, I., & Kapur, S. (2007). Separate brain regions code for salience vs. valence during reward prediction in humans. *Human Brain Mapping*, *28*(4), 294–302. https://doi.org/10.1002/hbm.20274.

Jones, J. L., Esber, G. R., McDannald, M. A., Gruber, A. J., Hernandez, A., Mirenzi, A., & Schoenbaum, G. (2012). Orbitofrontal cortex supports behavior and learning using inferred but not cached values. *Science*, *338*(6109), 953–956. https://doi.org/10.1126/science.1227489.

Kahnt, T. (2023). Computationally informed interventions for targeting compulsive behaviors. *Biological Psychiatry*, *93*(8), 729–738. https://doi.org/10.1016/j.biopsych.2022.08.028.

King, D. L., Herd, M. C. E., & Delfabbro, P. H. (2017). Tolerance in internet gaming disorder: A need for increasing gaming time or something else? *Journal of Behavioral Addictions*, *6*(4), 525–533. https://doi.org/10.1556/2006.6.2017.072.

Ko, C.-H., Liu, G.-C., Hsiao, S., Yen, J.-Y., Yang, M.-J., Lin, W.-C., … Chen, C.-S. (2009). Brain activities associated with gaming urge of online gaming addiction. *Journal of Psychiatric Research*, *43*(7), 739–747. https://doi.org/10.1016/j.jpsychires.2008.09.012.

Ko, C.-H., Liu, G.-C., Yen, J.-Y., Yen, C.-F., Chen, C.-S., & Lin, W.-C. (2013). The brain activations for both cue-induced gaming urge and smoking craving among subjects comorbid with Internet gaming addiction and nicotine dependence. *Journal of Psychiatric Research*, *47*(4), 486–493. https://doi.org/10.1016/j.jpsychires.2012.11.008.

Koob, G. F., & Volkow, N. D. (2016). Neurobiology of addiction: A neurocircuitry analysis. *The Lancet Psychiatry*, *3*(8), 760–773. https://doi.org/10.1016/s2215-0366(16)00104-8.

Kozak, K., Lucatch, A. M., Lowe, D. J. E., Balodis, I. M., MacKillop, J., & George, T. P. (2019). The neurobiology of impulsivity and substance use disorders: Implications for treatment. *Annals of the New York Academy of Sciences*, *1451*(1), 71–91. https://doi.org/10.1111/nyas.13977.

Kruschke, J. (2014). *Doing bayesian data analysis: A tutorial with R, JAGS, and Stan* (2nd ed.). Cambridge, Massachusetts: Academic Press.

Kübler, U. (2013). Structured clinical interview for DSM-IV (SCID). In Gellman, M. D., & Turner, J. R. (Eds.), *Encyclopedia of behavioral medicine* (pp. 1919–1920). Springer. https://link.springer.com/referenceworkentry/10.1007/978-1-4419-1005-9_66.

Kuss, D. J., Griffiths, M. D., & Pontes, H. M. (2017). Chaos and confusion in DSM-5 diagnosis of Internet Gaming Disorder: Issues, concerns, and recommendations for clarity in the field. *Journal of Behavioral Addictions*, 6(2), 103–109. https://doi.org/10.1556/2006.5.2016.062.

López-Caneda, E., Cadaveira, F., Crego, A., Gómez-Suárez, A., Corral, M., Parada, M., … Holguín, S. R. (2012). Hyperactivation of right inferior frontal cortex in young binge drinkers during response inhibition: A follow-up study. *Addiction*, 107(10), 1796–1808. https://doi.org/10.1111/j.1360-0443.2012.03908.x.

Lüscher, C., Robbins, T. W., & Everitt, B. J. (2020). The transition to compulsion in addiction. *Nature Reviews Neuroscience*, 21(5), 247–263. https://doi.org/10.1038/s41583-020-0289-z.

Lee, D., Lee, J., Lee, J. E., & Jung, Y.-C. (2017). Altered functional connectivity in default mode network in Internet gaming disorder: Influence of childhood ADHD. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 75, 135–141. https://doi.org/10.1016/j.pnpbp.2017.02.005.

Lee, S. W., Shimojo, S., & O'Doherty, J. P. (2014). Neural Computations underlying arbitration between model-based and model-free learning. *Neuron*, 81(3), 687–699. https://doi.org/10.1016/j.neuron.2013.11.028.

Lopatina, N., McDannald, M. A., Styer, C. V., Sadacca, B. F., Cheer, J. F., & Schoenbaum, G. (2015). Lateral orbitofrontal neurons acquire responses to upshifted, downshifted, or blocked cues during unblocking. *ELife*, 4, e11299. https://doi.org/10.7554/elife.11299.

Lucantonio, F., Caprioli, D., & Schoenbaum, G. (2014). Transition from 'model-based' to 'model-free' behavioral control in addiction: Involvement of the orbitofrontal cortex and dorsolateral striatum. *Neuropharmacology*, 76, 407–415. https://doi.org/10.1016/j.neuropharm.2013.05.033.

Madden, G. J., Petry, N. M., Badger, G. J., & Bickel, W. K. (1997). Impulsive and self-control choices in opioid-dependent patients and non-drug-using control patients: Drug and monetary rewards. *Experimental and Clinical Psychopharmacology*, 5(3), 256. https://doi.org/10.1037/1064-1297.5.3.256.

McDannald, M. A., Lucantonio, F., Burke, K. A., Niv, Y., & Schoenbaum, G. (2011). Ventral striatum and orbitofrontal cortex are both required for model-based, but not model-free, reinforcement learning. *The Journal of Neuroscience*, 31(7), 2700–2705. https://doi.org/10.1523/jneurosci.5499-10.2011.

Menon, V., & Uddin, L. Q. (2010). Saliency, switching, attention and control: A network model of insula function. *Brain Structure and Function*, 214(5–6), 655–667. https://doi.org/10.1007/s00429-010-0262-0.

Moorman, D. E. (2018). The role of the orbitofrontal cortex in alcohol use, abuse, and dependence. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 87(Pt A), 85–107. https://doi.org/10.1016/j.pnpbp.2018.01.010.

Mumford, J. A., Poline, J.-B., & Poldrack, R. A. (2015). Orthogonalization of regressors in fMRI models. *Plos One*, 10(4), e0126255. https://doi.org/10.1371/journal.pone.0126255.

Naqvi, N. H., Gaznick, N., Tranel, D., & Bechara, A. (2014). The insula: A critical neural substrate for craving and drug seeking under conflict and risk. *Annals of the New York Academy of Sciences*, 1316(1), 53–70. https://doi.org/10.1111/nyas.12415.

Nebe, S., Kroemer, N. B., Schad, D. J., Bernhardt, N., Sebold, M., Müller, D. K., … Smolka, M. N. (2018). No association of goal-directed and habitual control with alcohol consumption in young adults. *Addiction Biology*, 23(1), 379–393. https://doi.org/10.1111/adb.12490.

Nestor, L., Hester, R., & Garavan, H. (2010). Increased ventral striatal BOLD activity during non-drug reward anticipation in cannabis users. *NeuroImage*, 49(1), 1133–1143. https://doi.org/10.1016/j.neuroimage.2009.07.022.

O'Doherty, J. P., Hampton, A., & Kim, H. (2007). Model-based fMRI and its application to reward learning and decision making. *Annals of the New York Academy of Sciences*, 1104(1), 35–53. https://doi.org/10.1196/annals.1390.022.

Oscar-Berman, M., & Marinković, K. (2007). Alcohol: Effects on neurobehavioral functions and the brain. *Neuropsychology Review*, 17(3), 239–257. https://doi.org/10.1007/s11065-007-9038-6.

Ostlund, S. B., & Balleine, B. W. (2009). On habits and addiction: An associative analysis of compulsive drug seeking. *Drug Discovery Today: Disease Models*, 5(4), 235–245. https://doi.org/10.1016/j.ddmod.2009.07.004.

Poulton, A., & Hester, R. (2019). Transition to substance use disorders: Impulsivity for reward and learning from reward. *Social Cognitive and Affective Neuroscience*, 15(10), 1182–1191. https://doi.org/10.1093/scan/nsz077.

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage*, 59(3), 2142–2154. https://doi.org/10.1016/j.neuroimage.2011.10.018.

Preuschoff, K., Quartz, S. R., & Bossaerts, P. (2008). Human insula activation reflects risk prediction errors as well as risk. *The Journal of Neuroscience*, 28(11), 2745–2752. https://doi.org/10.1523/jneurosci.4286-07.2008.

Reiter, A. M. F., Deserno, L., Wilbertz, T., Heinze, H.-J., & Schlagenhauf, F. (2016). Risk factors for addiction and their association with model-based behavioral control. *Frontiers in Behavioral Neuroscience*, 10, 26. https://doi.org/10.3389/fnbeh.2016.00026.

Saunders, J. B., Aasland, O. G., Babor, T. F., Fuente, J. R. D. L., & Grant, M. (1993). Development of the alcohol use disorders identification test (AUDIT : WHO collaborative project on early detection of persons with harmful alcohol consumption-II. *Addiction*, 6(88), 791–804. https://doi.org/10.1111/j.1360-0443.1993.tb02093.x.

Saunders, J. B., Hao, W., Long, J., King, D. L., Mann, K., Fauth-Bühler, M., … Poznyak, V. (2017). Gaming disorder: Its delineation as an important condition for diagnosis,
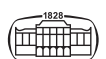
management, and prevention. *Journal of Behavioral Addictions*, 6(3), 271–279. https://doi.org/10.1556/2006.6.2017.039.

Schad, D. J., Jünger, E., Sebold, M., Garbusow, M., Bernhardt, N., Javadi, A.-H., … Huys, Q. J. M. (2014). Processing speed enhances model-based over model-free reinforcement learning in the presence of high working memory functioning. *Frontiers in Psychology*, 5, 1450. https://doi.org/10.3389/fpsyg.2014.01450.

Schultz, W. (2016). Dopamine reward prediction error coding. *Dialogues in Clinical Neuroscience*, 18(1), 23–32. https://doi.org/10.31887/dcns.2016.18.1/wschultz.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593.

Sebold, M., Deserno, L., Nebe, S., Nebe, S., Schad, D. J., Garbusow, M., … Huys, Q. J. M. (2014). Model-based and model-free decisions in alcohol dependence. *Neuropsychobiology*, 70(2), 122–131. https://doi.org/10.1159/000362840.

Sebold, M., Nebe, S., Garbusow, M., Guggenmos, M., Schad, D. J., Beck, A., … Heinz, A. (2017). When habits are dangerous: Alcohol expectancies and habitual decision making predict relapse in alcohol dependence. *Biological Psychiatry*, 82(11), 847–856. https://doi.org/10.1016/j.biopsych.2017.04.019.

Seeley, W. W. (2019). The salience network: A neural system for perceiving and responding to homeostatic demands. *The Journal of Neuroscience*, 39(50), 9878–9882. https://doi.org/10.1523/jneurosci.1138-17.2019.

Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., … Greicius, M. D. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *The Journal of Neuroscience*, 27(9), 2349–2356. https://doi.org/10.1523/jneurosci.5587-06.2007.

Silva, C. F. da, & Hare, T. A. (2020). Humans primarily use model-based inference in the two-stage task. *Nature Human Behaviour*, 1–14. https://doi.org/10.1038/s41562-020-0905-y.

Starcevic, V. (2017). Internet gaming disorder: Inadequate diagnostic criteria wrapped in a constraining conceptual model. *Journal of Behavioral Addictions*, 6(2), 110–113. https://doi.org/10.1556/2006.6.2017.012.

Stewart, J. L., Connolly, C. G., May, A. C., Tapert, S. F., Wittmann, M., & Paulus, M. P. (2014). Striatum and insula dysfunction during reinforcement learning differentiates abstinent and relapsed methamphetamine-dependent individuals. *Addiction*, 109(3), 460–471. https://doi.org/10.1111/add.12403.

Stewart, J. L., May, A. C., Poppa, T., Davenport, P. W., Tapert, S. F., & Paulus, M. P. (2014). You are the danger: Attenuated insula response in methamphetamine users during aversive interoceptive decision-making. *Drug and Alcohol Dependence*, 142, 110–119. https://doi.org/10.1016/j.drugalcdep.2014.06.003.

Sun, Y., Ying, H., Seetohul, R. M., Xuemei, W., Ya, Z., Qian, L., … Ye, S. (2012). Brain fMRI study of crave induced by cue pictures in online game addicts (male adolescents). *Behavioural Brain Research*, 233(2), 563–576. https://doi.org/10.1016/j.bbr.2012.05.005.

Tolomeo, S., & Yu, R. (2022). Brain network dysfunctions in addiction: A meta-analysis of resting-state functional connectivity. *Translational Psychiatry*, 12(1), 41. https://doi.org/10.1038/s41398-022-01792-6.

Turel, O., He, Q., Wei, L., & Bechara, A. (2021). The role of the insula in internet gaming disorder. *Addiction Biology*, 26(2), e12894. https://doi.org/10.1111/adb.12894.

Uddin, L. Q. (2015). Salience processing and insular cortical function and dysfunction. *Nature Reviews Neuroscience*, 16(1), 55–61. https://doi.org/10.1038/nrn3857.

Vaccaro, A. G., & Potenza, M. N. (2019). Diagnostic and classification considerations regarding gaming disorder: Neurocognitive and neurobiological features. *Frontiers in Psychiatry*, 10, 405. https://doi.org/10.3389/fpsyt.2019.00405.

Voon, V., Derbyshire, K., Rück, C., Irvine, M. A., Worbe, Y., Enander, J., … Bullmore, E. T. (2015). Disorders of compulsivity: A common bias towards learning habits. *Molecular Psychiatry*, 20(3), 345–352. https://doi.org/10.1038/mp.2014.44.

Voon, V., Reiter, A., Sebold, M., & Groman, S. (2017). Model-based control in dimensional Psychiatry. *Biological Psychiatry*, 82(6), 391–400. https://doi.org/10.1016/j.biopsych.2017.04.006.

Weinstein, A., & Lejoyeux, M. (2020). Neurobiological mechanisms underlying internet gaming disorder. *Dialogues in Clinical Neuroscience*, 22(2), 113–126. https://doi.org/10.31887/dcns.2020.22.2/aweinstein.

World Health Organization (2018). *ICD-11 for mortality and morbidity statistics 2018*. Retrieved from https://icd.who.int/browse11/l-m/en.

Wu, X., Wang, T., Liu, C., Wu, T., Jiang, J., Zhou, D., & Zhou, J. (2017). Functions of learning rate in adaptive reward learning. *Frontiers in Human Neuroscience*, 11, 592. https://doi.org/10.3389/fnhum.2017.00592.

Wyckmans, F., Otto, A. R., Sebold, M., Daw, N., Bechara, A., Saeremans, M., … Noël, X. (2019). Reduced model-based decision-making in gambling disorder. *Scientific Reports*, 9(1), 19625. https://doi.org/10.1038/s41598-019-56161-z.

Yen, J.-Y., Lin, P.-C., Wu, H.-C., & Ko, C.-H. (2022). The withdrawal-related affective, gaming urge, and anhedonia symptoms of internet gaming disorder during abstinence. *Journal of Behavioral Addictions*, 11(2), 481–491. https://doi.org/10.1556/2006.2022.00008.

Young, K. S. (1998). Internet addiction: The emergence of a new clinical disorder. *Cyber Psychology & Behavior*. https://doi.org/10.1089/cpb.1998.1.237.

Zheng, H., Hu, Y., Wang, Z., Wang, M., Du, X., & Dong, G. (2019). Meta-analyses of the functional neural alterations in subjects with Internet gaming disorder: Similarities and differences across different paradigms. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 94, 109656. https://doi.org/10.1016/j.pnpbp.2019.109656.