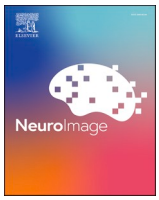
NeuroImage 320 (2025) 121479

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg



Individual differences in policy precision: Links to suicidal risk and network dynamics

Dayoung Yoon [a,*], Jaejoong Kim [b], Do Hyun Kim [c], Dong Woo Shin [d], Su Hyun Bong [a], Jaewon Kim [a], Hae-Jeong Park [e,f,g], Hong Jin Jeon [h], Bumseok Jeong [a,i,*]

[a] Graduate School of Medical Science and Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Republic of Korea
[b] Department of Psychiatry and Behavioural Sciences, University of Minnesota, MN 55454, USA
[c] Department of Psychiatry, Dankook University College of Medicine, Cheonan, Republic of Korea
[d] Department of Neurology, Ewha Womans University Mokdong Hospital, Seoul, Republic of Korea
[e] Graduate School of Medical Science, Brain Korea 21 Project, Department of Nuclear Medicine, Department of Psychiatry, Yonsei University College of Medicine, Seoul, Republic of Korea
[f] Center for Systems and Translational Brain Science, Institute of Human Complexity and Systems Science, Yonsei University, Seoul, Republic of Korea
[g] Department of Cognitive Science, Yonsei University, Seoul, Republic of Korea
[h] Department of Psychiatry, Depression Center, Samsung Medical Center, Sungkyunkwan University School of Medicine, 81 Irwon-Ro Gangnam-gu. Seoul 06351, Republic of Korea
[i] KAIST Institute for Health Science and Technology and Graduate School of Data Science, KAIST, Daejeon, Republic of Korea

ARTICLE INFO

Keywords:
Decision-making
Active inference
fMRI
Behavioural modelling

ABSTRACT

The Behavioural modelling of decision-making processes has advanced our understanding of impairments associated with various psychiatric conditions. While many studies have focused on models that best fit behavioural data, the extent to which such models reflect biologically plausible mechanisms remains underexplored. To bridge this gap, we developed a probabilistic two-armed bandit task model grounded in the active inference framework and evaluated its performance against established reinforcement learning (RL) models. Our model not only matched but outperformed conventional RL models in explaining individual variability in choice behaviour. A central feature of our model is the optimisation of policy precision based on previous outcomes. This process captures the dynamic balance between model-based predictions derived from the internal generative model and the influence of immediate past observations. Importantly, incorporating the temporal dynamics of policy precision significantly improved the model's capacity to explain large-scale brain network activity and inter-subject variability. We found that increases in policy precision were positively associated with default mode network dominance and negatively associated with states dominated by dorsal attention and frontoparietal networks. These opposing associations suggest functional coordination between these systems, as supported by the correlations between brain state transitions and behavioural parameters. Furthermore, prolonged dominance of another brain state, characterised by elevated ventral attention network activity and stronger inter-network connectivity, appeared to disrupt this coordination. Finally, we found that heightened sensitivity to negative outcomes in a loss-related context was associated with high suicidal risk among individuals with major depressive disorder.

1. Introduction

Computational psychiatry has taken on the challenge of bridging the gap between mental phenomena and their assumed origins in brain function (Huys et al., 2016; Montague et al., 2012). The development of mathematical models has been central to this effort. Owing to the intricate cognitive processes involved in decision-making, such as evaluating evidence for potential outcomes and option selection, decision-making has become a means to explain psychiatric symptoms (Chen et al., 2015; Maia and Frank, 2011). Reinforcement learning (RL)

* Corresponding authors.
E-mail addresses: skfjhd@kaist.ac.kr (D. Yoon), kjj11033@gmail.com (J. Kim), dohyun.kim@dankook.ac.kr (D.H. Kim), dongwooshin.stroke@gmail.com (D.W. Shin), npdrbong@kaist.ac.kr (S.H. Bong), jwk1921@kaist.ac.kr (J. Kim), parkhj@yonsei.ac.kr (H.-J. Park), jeonhj@skku.edu (H.J. Jeon), bs.jeong@kaist.ac.kr, leesangwon@kaist.ac.kr (B. Jeong).

https://doi.org/10.1016/j.neuroimage.2025.121479
Received 14 January 2025; Received in revised form 19 September 2025; Accepted 19 September 2025
Available online 22 September 2025

has facilitated decision-making, as well as the development of various models augmented with additional parameters, to more accurately explain the agent's decision (Ahn et al., 2014, 2008; Worthy et al., 2013; Worthy and Todd Maddox, 2014). Building on previous studies linking major depressive disorder (MDD) to abnormalities in the brain's reward circuitry, RL has been widely used to explain decision-making deficits in patients with MDD (Chen et al., 2015; Dombrovski et al., 2013; Yechiam et al., 2005).

There are various decision-making tasks, one of the simplest forms being the probabilistic two-armed bandit task. Owing to its simplicity, it has been neglected, as it is considered insufficient for studying complex cognitive processes, such as exploration (Cohen et al., 2007; Schulz and Gershman, 2019). However, even in this simple task, individuals display diverse behavioural patterns likely rooted in variations in neural states. Although behavioural models are intended to link these neural and behavioural patterns, most account for such behavioural diversity merely by randomness in action probabilities and exclude neural state diversity from consideration. This likely stems from the fact that most behavioural models have been developed to explain relatively low-dimensional behavioural patterns without explaining the high-dimensional diversity of neural states.

To address this limitation, we developed a new model inspired by the principles of the active inference framework (AIF) and determined whether the behaviourally relevant signals central to this model can account for variations in brain imaging data using a general linear model (GLM) in two ways. AIF is a unified theory of action and perception that attempts to elucidate decision-making processes by minimising a singular objective function, free energy (Parr et al., 2022). Free energy measures the discrepancy between an agent's internal model (its beliefs about the world) and the actual sensory data it receives, and, in the AIF, serves as a proxy for surprise (unexpectedness), which cannot be directly computed by agents.

The RL model assumes that people develop expected values (EV) for each choice option (i), which represent the reward (or punishment) they expect to receive following each choice. In this learning process, a Rescorla-Wagner model is generally implemented, which uses a fixed learning rate parameter (lr) for updating EVs based on the prediction error between the outcome ($r(t)$), and the current EV on trial (t) (Rescorla and Wagner, 1972).

$$EV_{i,t+1} = EV_{i,t} + lr \cdot \left(r(t) - EV_{i,t}\right) \tag{1}$$

Choices are made according to the probability calculated from EVs, which are updated and learned through observations, using the softmax action selection rule. The degree to which the option with the highest EV is chosen is determined by γ, referred to as the 'decision temperature'.

$$p\left(a_{i,t}\right) = \frac{\exp\left(\gamma \cdot EV_{i,t}\right)}{\sum_{i=1}^{K} \exp\left(\gamma \cdot EV_{i,t}\right)} \tag{2}$$

However, this method determines the probability of selecting each option solely based on its EVs, which cannot capture trial-by-trial variability in individual behaviour. In other words, it does not adequately account for within-subject fluctuations that unfold as the task progresses. The decision temperature parameter can partially compensate for this limitation by modulating the influence of EVs; namely, when choice inconsistency increases, the estimated temperature decreases. This results in individuals with more inconsistent behaviour being interpreted as placing less weight on value-based decisions, a feature often linked to impulsivity in decision-making (Findling et al., 2019). In this sense, this model is well-suited to explain the overall observed behavioural data. However, the model falls short when answering questions such as, 'Why did the person make that specific choice at that specific moment?'. This limitation is especially apparent when explaining moments of increased choice variability across individuals. We quantified choice variability in each trial with Shannon entropy (SE):

Fig. 1.

$$SE_t = -\sum_{i=1}^{2} P_{i,\,t} \log\left(P_{i,\,t}\right) \tag{3}$$

where $P_{i,\,t}$ is the frequency of selecting choice i on trial t. As shown in Fig. 2A, B, entropy is highest at the beginning of the task and decreases after approximately 10 trials, suggesting that most participants learned which bandit is better by that point. Interestingly, later in the task, there are certain trials where entropy increases again. The high entropy in the initial phase can be explained by the absence of knowledge, where EVs are close to 0.5, leading to action probabilities that approximate the chance level. However, the later increases in entropy tend to follow surprising events, such as consecutive negative outcomes from selecting the better bandit (the black dots in Fig. 2A, B).

The basic model in Eqs. (1–2) attributes such behaviour to an overall increase in randomness, as captured by the lower EV of better bandit consecutive negative outcomes. However, although entropy increases in these trials, it is not as high as in the early phase of the task, where total informational uncertainty dominates. This implies that the current increase in entropy is unlikely to stem from the same mechanism, and the model lacks an important explanatory term beyond EVs to account for action selection in these moments.

Such a term must capture how individuals respond to outcomes from the previous trial. A well-known strategy that directly reflects this is the win-stay-lose-shift (WSLS) strategy, which is entirely outcome-driven (Herrnstein et al., 1997; Otto et al., 2011; Steyvers et al., 2009). Prior research has suggested that both methods can be utilised in decision-making; some studies have even attempted to integrate both approaches (Ahn et al., 2014; Worthy et al., 2013). The WSLS-RL model, for example, combines both probabilities—those derived from EV and WSLS strategy—by weighing them with a parameter w.

$$P\left(a_{i,t}\right) = P(a_{i,t})_{WLS} \cdot w + P\left(a_{i,t}\right)_{RL} \cdot (1 - w) \tag{4}$$

In this model, parameter w and the probability of lose-shift (P_{WSLS}) correspond to the parameters that can explain the choice variability in specific trials. However, this method relies heavily on additional parameters and cannot explain the mechanism involved in the exchange between both strategies according to the state context.

Meanwhile, two aspects of AIF allow for a flexible model to address this limitation. First, rather than using EV directly, the expected free energy (EFE) is applied as the input to calculate choice probability. EFE represents the distance between the probability distribution of this EV and each agent's preference distribution for EV (Eq. (11)). This approach accounts for utility differences (Ahn et al., 2008) and captures the variations in EV precision, which can influence choice. As shown in Fig. 1A, even when EV values are identical, the computed EFE values vary depending on the precision of the EVs—higher precision leads to a greater EFE difference between the two choices. Similarly, even with the same EV and precision, the differences in utility for the rewards also affect the EFE gap between choices, with larger utility differences resulting in greater EFE separation. A higher EFE gap leads to more confident decision-making by increasing the probability of selecting one choice over the other. Second, the policy precision parameter (γ in Eq. (13)), which corresponds to the decision temperature in a softmax function, modulates the effect of EFE on action probability. This parameter is updated based on the differences between the EFE (G) predicted by the generative model and the free energy (F) calculated from the previous outcome (Eq. (14), Fig. 1B). As the discrepancy between the two increases and G_{error} grows, policy precision tends to decrease accordingly (Eq. (16)). That is, it modulates the weighting between model-based predictions and the influence of previous observation, as the parameter w does in the WSLS-RL model.

We propose a new behavioural model based on AIF for the probabilistic two-armed bandit task and demonstrate its advantages in not only the within-subject variability that emerges throughout the task, but

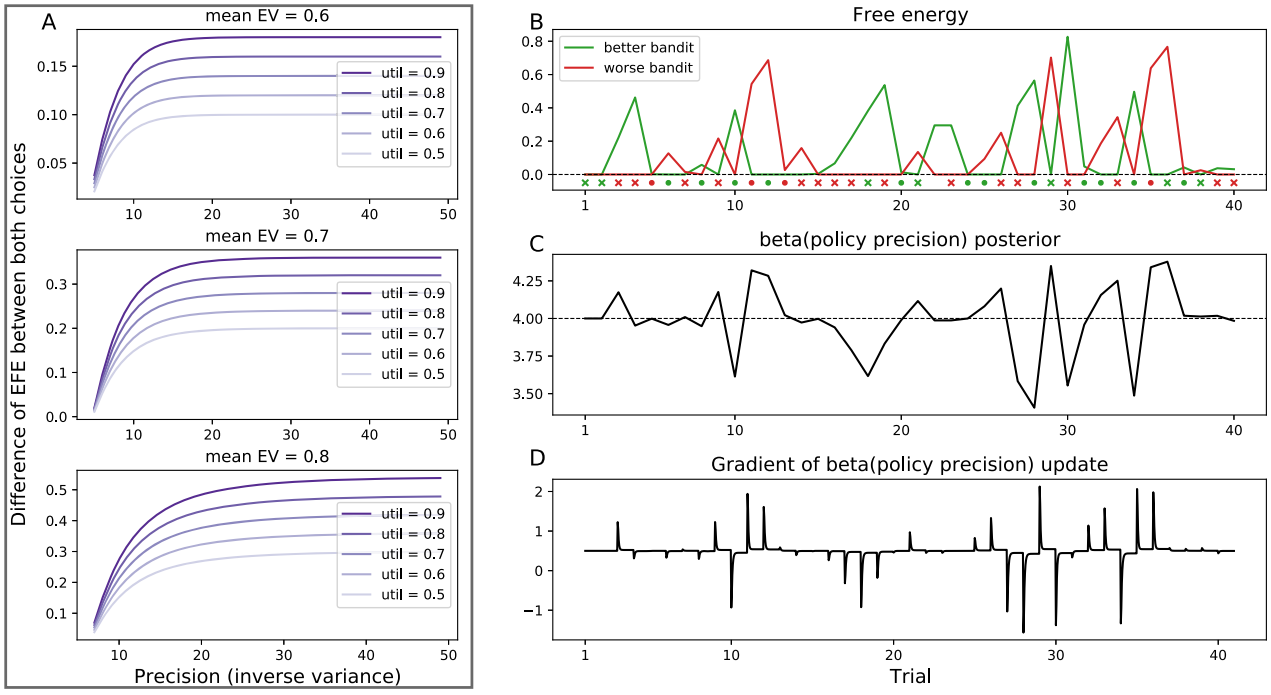


Fig. 1. Simulated expected free energy and example estimated results of AIF model. (A) Differences in expected free energy between the two actions increase as either the precision of the state increases or the utility increases, regardless of the state's expected value. (B–D) Estimated results for one participant. (B) Free energy values for both actions in each trial. These values are generally higher during the later phase of the trials because the variance of the state distribution is lower. In this phase, if the outcomes differ from the prediction, the distribution's distance to the predicted state is large. Conversely, free energy values before approximately the tenth trial are lower because the prediction precision is low; therefore, prediction error is also low. Each bandit's free energy tends to increase after it is selected and a losing outcome is observed. (C) As the policy precision is also updated according to the free energy, updated posterior values tend to follow the worse bandit's free energy while opposing the free energy of the better bandit. (D) While policy precision is updated, the rate of change is hypothesised to represent neural discharge.

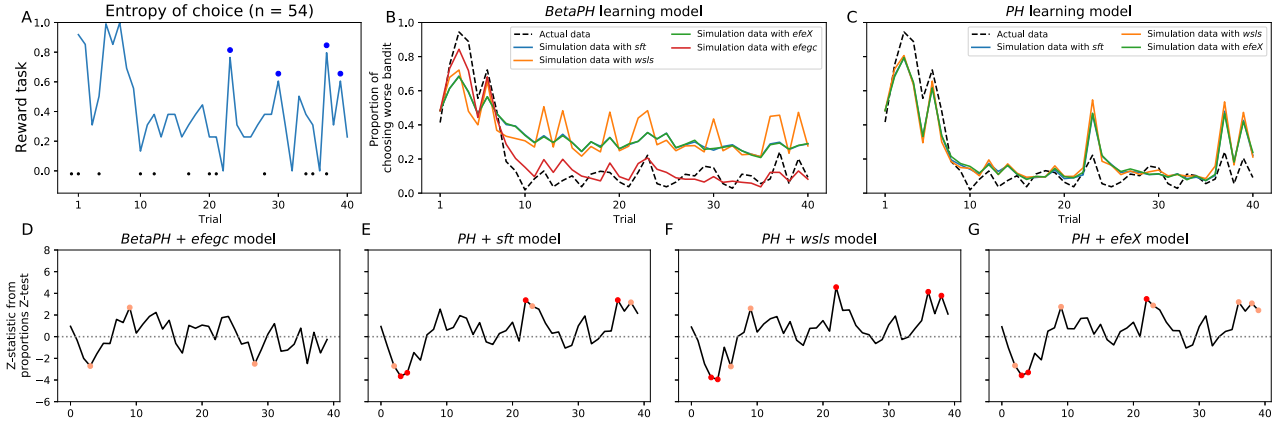


Fig. 2. Superiority in explaining the inter-subject variability of choices. (A, B) The entropy of choices in reward task trial out of 54 participants. The black dots at the bottom of the plots mark the trial where choosing the worse bandit leads to a loss outcome (12 trials in total, corresponding to 30 %. The blue dots at the top of the plots mark late increase entropy trials. (B, C) Proportion of participants choosing a worse bandit in each trial. The cosine similarity of the *efegc* model with actual data is 0.98 in (B), and those of others are 0.94, 0.94, and 0.93 in (C), respectively. The dashed lines indicate the actual proportion and solid lines are generated from simulation data of each choice model. (D-G) Z-statistics from proportions Z-tests assessing the difference in choice distributions between simulated and empirical data at each trial. Larger absolute values indicate greater discrepancies. Significant trials are marked with dots and the red dots indicate trials with $|Z| > 3.2$.

also the inter-subject variability in choice behaviour. We argue that these advantages stem from the model's ability to capture the notion that confidence in decision making, as reflected in decision variability, is influenced by disconfirmatory evidence over and above the change in pragmatic action value that such evidence induces. To evaluate how well the new model captures and reproduces the variability of the choices observed at the group level, we generated simulated choice data using the parameter sets estimated from each model. We then computed the cosine similarity between the actual and simulated choice proportions to quantify their correspondence. Additionally, we identified trials where the distribution of simulated choices significantly diverged from that of the actual data using the proportions Z-tests, and found that these trials tend to exhibit high choice uncertainty, as indexed by SE. Notably, unlike other models, the proposed model was able to reproduce participants' choices in trials with later entropy increase, indicating its superior ability to account for choice variability under uncertainty

(Fig. 2D). In addition, the validity of the policy precision signal (rate of change in policy precision, Fig. 1D) has been confirmed through two experiments: one involving the analysis of neural signal data and another examining whether behaviour parameters provide better explanatory power in describing traits of depressive patients.

In Experiment 1, using neural signal data obtained via task-functional magnetic resonance imaging (fMRI), we used two methods to confirm whether the policy precision signal is related to neural activity (Fig. 3). First, we conducted a conventional GLM approach, where models are generated by including the stimuli of behavioural tasks as regressors. The second method addresses the limitations of the conventional GLM by using inter-subject correlation (ISC) (Nastase et al., 2019). We determined whether the variance in brain imaging data across individuals could be explained by the variance in behavioural signals and, similarly, whether adding policy precision signal improves model performance measured by Akaike Information Criterion (AIC) and Bayes Information Criterion (BIC).

In Experiment 2, we conducted the decision-making task involving a population that included patients diagnosed with MDD and healthy controls (HC). In patients with MDD, the risk of suicide is nearly 20 times higher than that in the general population (Turecki et al., 2019). According to the Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5), 'impairment in decision-making' is an important criterion for diagnosing MDD as 'indecisiveness'. Previous studies have analysed patients with MDD with serious suicidal behaviour and found that their ability to learn and choose actions leading to rewarding outcomes is impaired (Brown et al., 2020; Clark et al., 2011; Dombrovski et al., 2019). Although most studies on suicide have shown impaired decision-making and its associated neural correlates within the reward domain, impairments in the loss domain have rarely been investigated, especially in young patients with MDD and suicidal risk. Furthermore, the review by (Chen et al., 2015) highlighted inconsistent findings

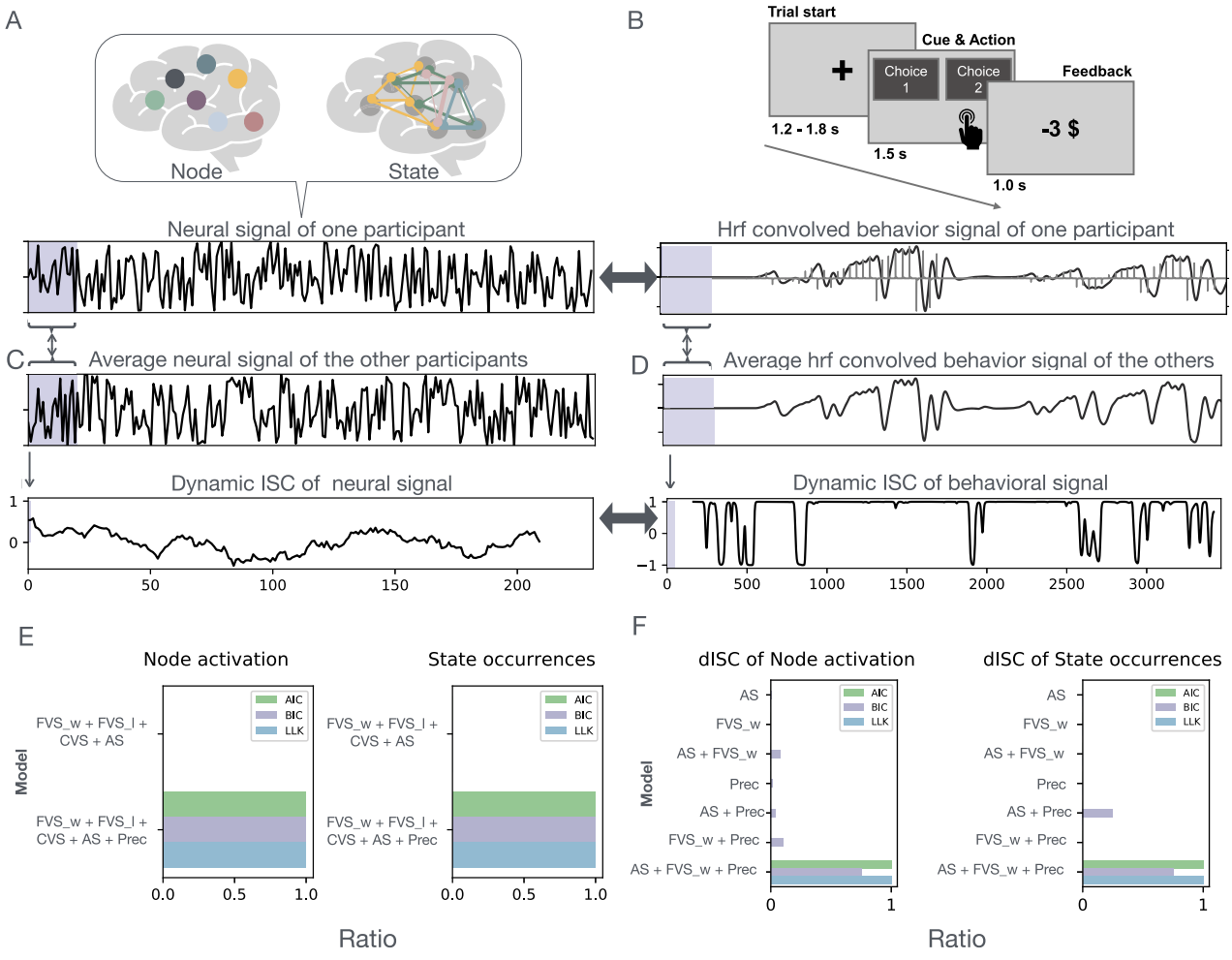


Fig. 3. Schematic illustration of the analysis for explaining neural signals by behavioural signals. (A) Neural signals are extracted at two levels from functional magnetic resonance imaging (fMRI) data measured during decision-making tasks. (B) The task begins with gaze fixation on a central cross for a random duration ranging from 1.2 to 1.8 s; subsequently, the meaningful experimental stimulus is presented. The 'Cue' refers to the presentation of two option figures and lasts for 1.5 s during which the participant had to choose one. After 1.5 s, the choice results are presented as the monetary reward or loss. We defined this presentation result as 'Feedback' and lasted for 1.0 s. The images of each choice are shown as meaningless black-and-white pictures, and different sets of images are shown during each part of the task. Selection should be made in 1.5 s, during which the Cue images are presented on the screen. The policy precision signal is generated by the calculation process as in Fig. 1D, then aligned to the time-point of Feedback presentation because the gradient update is assumed to occur by the free energy induced by Feedback. (C, D) By calculating the inter-subject correlation (ISC) of time-series over the sliding window, neural and behavioural dynamic inter-subject correlation (dISC) were obtained, respectively. (E) Performance comparison of conventional GLM. (F) Performance of dISC GLM. The x-axis represents the ratio of the number of regions where each model was measured as best (smallest values of each criterion. Note that, as higher log-likelihood (LLK) values indicate better fit, we used –LLK so that lower values indicate better fit), and the y-axis shows the independent variables included in each GLM model. CVS: cue visual stimulus; FVS w: feedback visual stimulus of win; FVS l: feedback visual stimulus of lose; AS: action selection; Prec: policy precision signal.

across the studies employing reinforcement learning tasks in the reward and punishment domains. In this study, we determine whether estimated parameters from a decision-making task involving reward and loss domains can distinguish MDD subgroups categorised by suicidal ideation, and whether the newly proposed AIF model offers improved adaptability for this purpose.

2. Results

2.1. Model construction and the results of parameter recovery and model comparison

We constructed 20 models, comprising five learning models combined with four choice models. Two of the learning models utilised scalar state values; one is described in Eq. (1) (the Rescorla-Wagner model, see Supplementary Table. 1) and the other is the Pearce-Hall model (Pearce and Hall, 1980). The others utilised either beta or normal distributions for state value. The latter corresponds to the Kalman-filter model (Kalman, 1960). With beta distribution, to prevent beliefs from becoming quickly inflexible, two different options are applied. One applies the 'learning rate' using a similar mechanism to that of the Pearce-Hall model (Eq. (8)), and the other applies the 'decay' to prior belief during belief updating (Eq. (6)). Both the learning rate and decay terms are changed based on the Bayes-factor surprise, a ratio between the subjective probability of an observation under the current beliefs and prior beliefs (Eq. (7)) (Liakoni et al., 2021).

The four choice models included two previous models of the softmax rule (*sft* in Supplementary Table. 1) and the WSLS-RL (*wsls* in Supplementary Table. 1) model. The other two models used EFE, referred to as the *efeX* and *efegc* models. The *efeX* model excluded the process of policy precision update, making it similar to the softmax model. By contrast, the *efegc* model incorporated dynamic changes in policy precision in Eq. (16) (where 'gc' stands for gamma change), resembling the WSLS-RL model, as this update modulates the weighting between the influences of prediction and previous outcome. The parameter compositions and recovery results of each model are described in Supplementary Table. 1 with descriptions of each parameter. The model construction is detailed in the Materials and methods section. Parameter estimation was accomplished using the variational Bayes method (Friston et al., 2007), and the approximations of log model evidence were subjected to random-effect Bayesian model selection (Rigoux et al., 2014). First, we grouped 20 models into four 'families' by action models (Penny et al., 2010). This method yields family-specific exceedance probabilities, representing the probability that the learning model has a higher frequency than the other included models on the group level. The beta distribution with learning rate model outperformed in *efegc* model, while the Pearce-Hall model was best with the others (Supplementary Figure 5A). Using the best combinations of each action model, we compared the four models and the model with *efegc* was best for loss and reward tasks (Supplementary Figure 5B).

2.2. Evaluating the model's capacity to capture individual differences in choice

Fig. 2A, B illustrates the entropy of binary choices across individuals. On the first trial, selection variability was highest because there was no difference in value between bandits, suggesting that this early variance was driven by random selection. By contrast, the increase in variance observed following the tenth trial was attributed to a distinct mechanism related to the direct influence of the previous trial's result, encoded as the lose-shift probability (*ls*) in the WSLS-RL model or free energy (Fig. 1B) in the *efegc* model. The softmax and *efeX* model account for this by the increased selection randomness due to a temporary reduction in the value difference between two options, similar to the early phase. Therefore, in the late increase entropy trials (the blue dots in Fig. 2A), it becomes difficult for models to accurately distinguish between participants who made different choices. As a result, when simulations are conducted using such models, the proportion of each choice tends to be close to 0.5 due to the reliance on randomness (Fig. 2B, C). This phenomenon was observed even in the Pearce-Hall model, which served as the best learning model for those action models combination—despite showing improved cosine similarity with the actual data. Notably, the WSLS-RL model also failed to explain this pattern. This may be due to the fact that, whereas the degree of switching in response to previous outcomes is modulated by belief precision across the trials in the *efegc* model, it remains fixed in the WSLS-RL.

2.3. Neuroimaging analysis

The ability to modulate the relative influence of prior belief and responses to previous outcomes by belief precision across trials in the *efegc* model may underlie its advantage in capturing the behavioural patterns observed in actual data. A key component in this process is an additional signal, represented by the rate of change in policy precision. To evaluate the plausibility of this signal as a candidate mechanism for brain function, we compared behavioural and neural signals at both the regional and network levels. This approach allowed us to examine localised brain activity, as well as alterations in large-scale brain networks that are critical for cognitive processing (Fig. 3A). We estimated four key states using the graph Laplacian mixture model (GLMM), as described by (Ricchi et al., 2022). A detailed description of the neural signal extraction process is provided in the Materials and methods section.

Using the mean activation patterns, we grouped the regions into a priori resting-state networks (RSNs) based on (Thomas Yeo et al., 2011) (Supplementary Figure. 1A-D). State 1 was positively correlated with activation of the visual (VIS) and somatomotor (SOM) networks, dorsal attention network (DAN), and ventral attention network (VAN) and negatively correlated with the default mode network (DMN) and frontoparietal control network (FPN) and limbic network (LIM). State 2 was positively correlated with the DMN and negatively correlated with State 1′s network. State 3 was positively correlated with the DAN and FPN and negatively correlated with the DMN and SOM network. State4 shows generally low activations for most networks, while those of the LIM and VAN are relatively high.

In a conventional GLM, two models were included. The regressors for these models were based on two types of stimuli: cue-visual stimuli (CVS) and feedback-visual stimuli (FVS), as well as regressors for the subjects' behavioral actions in making a choice (AS). Given that the feedback stimulus is different based on the outcome, the FVS regressor was further divided into two separate regressors for win outcomes (FVS_w) and loss outcomes (FVS_l). We discovered that the model incorporating policy precision signals (Prec) yielded lower AIC and BIC values, which corresponds to better performance, in explaining both regional and network-level neural activity (Fig. 3E).

Furthermore, in a separate GLM that used dynamic inter-subject correlation (dISC) of behavioral signals to explain the time-varying ISC of neural signals, the model with policy precision signals also demonstrated the best performance (Fig. 3F). In this analysis, the cue stimulus presentation time was identical across all subjects, so it was not included as a regressor. Since the dISC for both FVS_l and FVS_w yielded identical values, only one was included in the model.

Using beta coefficients from the best model, we performed one-sample *t*-test to identify which regional activations or interactions are related to the policy precision signals (Supplementary Table. 3). Across both tasks, similar regions included in the DMN were positively correlated with the policy precision signal (Fig. 4), whereas regions included in the FPN and DAN showed negative correlations. These results are consistent with the results of the state level in that occurrences of State 2 showed a positive correlation with policy precision signals, whereas State 3 exhibited a negative correlation (Supplementary Table. 4). Using the occurrences of each state, we computed three of the state dynamics

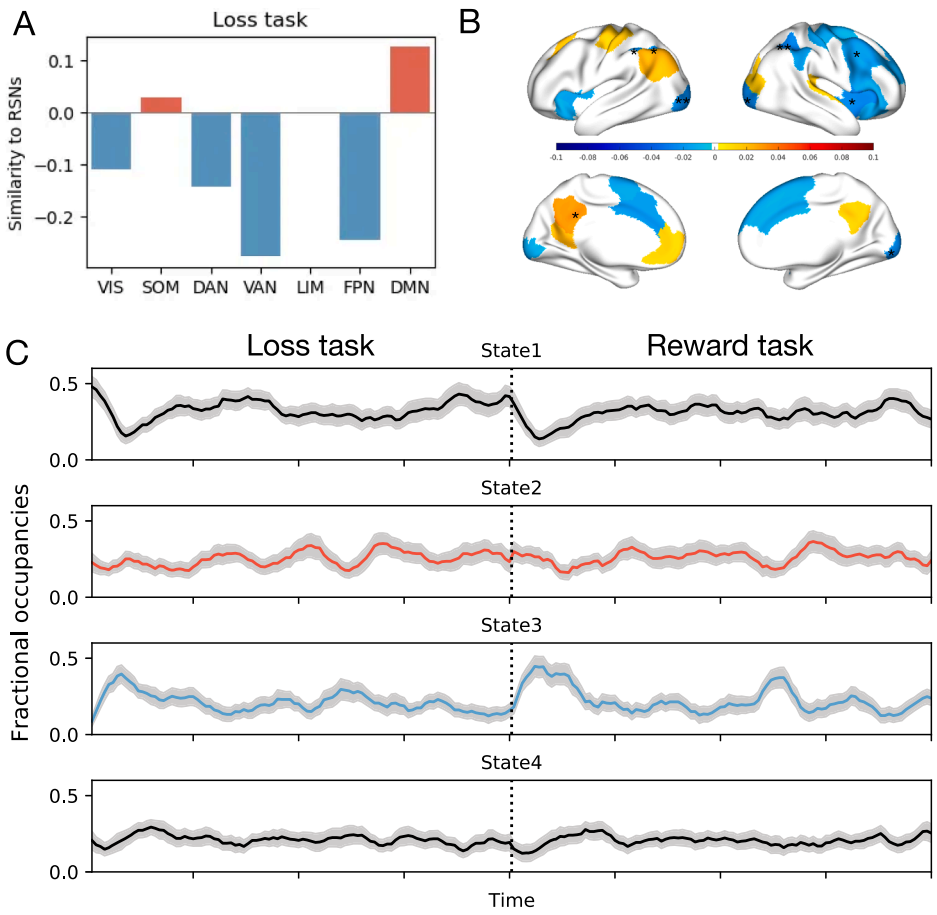


Fig. 4. Results of a one-sample *t*-test using beta coefficient values result from GLM to investigate the regions related to the policy precision signal. (A) Cosine similarity between beta coefficient values of significantly related regions ($p < 0.05$) and the seven networks of intrinsic functional connectivity from [32]. Regions included within the default mode network (DMN) show positive correlations, while those included in the dorsal attention network (DAN) and frontoparietal network (FPN) show negative correlations. (B) Coloured regions have significant correlations with the policy precision signal ($p < 0.05$). *, ** indicate significance after controlling for the false discovery rate (FDR) (* for $q < 0.05$, ** for $q < 0.01$) using the Benjamini–Hochberg procedure. (C) Dynamic fractional occupancies of each state estimated using a sliding window with a length of seven repetition times (TRs). The left-hand side of the central dashed line corresponds to data collected during the loss task, and the right-hand side corresponds to data from the reward task. Although there was a 10-second resting period between the two tasks, it is omitted from the figure for clarity.

metrics for each participant: fractional occupancy (FO), dwell time (DT), and transition probability (TP) (Vidaurre et al., 2018). FO refers to the proportion of time a participant spends in each state across the entire task. DT indicates the average temporal duration of consecutive visits to a given state, reflecting how long a state tends to persist once entered. TP quantifies the likelihood of switching from one state to another, capturing the dynamics of state-to-state transitions. We then examined whether these metrics were associated with behavioural parameters using Pearson correlation analyses. Among the brain state metrics, those that showed significant correlations with behavioural parameters included the TP from State 3 to 2 and FO of State 2 (positive correlation with '*util*' and '*lr*'), as well as the TP of remaining in State 4 (State 4 to 4) and the DT of State 4 (both negatively correlated with '*util*' and '*lr*') (Supplementary Table. 5).

2.4. Model validation in population, including the patients with MDD

Fig. 4.

Finally, we determined whether the parameters estimated from the new model could better discriminate the clinical characteristics. To this end, we conducted the behavioural task with a population that included patients diagnosed with MDD. The task composition was slightly modified by setting the probability of winning on the better bandit to 0.75 and mixing loss and reward tasks within the same session. Two task sets, 1 and 2, were used ($n = 74$ and $n = 43$, respectively). Model comparison among action model families indicated that the beta distribution with learning rate parameter provided the best fit for *efegc* model, while the Rescorla–Wagner model was best fit for the other models, except for the loss task in the task set 2 group (Supplementary Figure 4C-F).

We conducted three analyses using the estimated behaviour parameter values from the best combination models of each action model: (1) logistic regression predicting group membership between healthy controls (HC; $n = 58$) and patients with major depressive disorder (MDD; $n = 58$); (2) logistic regression distinguishing suicidal risk (SR; $n = 37$) group from the non-suicidal risk group (NSR; $n = 21$); (3) analysis of covariance (ANCOVA) predicting Hamilton Depression Rating Scale (HAM-D) scores (Yi et al., 2005) from the behavioural parameters. All analyses controlled for age, sex, education level, and task set. Behavioural parameters estimated from the loss and reward tasks of each model were initially included as independent variables.

All parameters estimated from the reward task, except *lr* from the *sft* and *efeX* models, were significant predictors distinguishing between the control and MDD groups. By contrast, when differentiating between NSR and SR groups, significant predictors emerged from the loss task, and this pattern was observed only in the *efegc* model. Regarding the prediction of HAM-D scores, the results are similar to those of the HC vs. MDD logistic regression.

3. Discussion

To explain psychiatric symptoms or real human cognitive actions, it is essential to utilise biologically plausible models rather than relying solely on those that demonstrate the best performance. This approach ensures that the model is accurate and reflective of the complexities of neural processes. Our model included distributional representation during the learning process, calculated expected energy during action selection, and optimised policy precision by observing each trial. Although this may seem overly complex, including biologically plausible parameters and signals significantly enhances explanatory power in neural signal data and validates this advantage in the clinical data.

The main advantage of this model lies in its ability to account for the context-dependent interplay between model-based predictions and the influence of previous observation using a single free energy function, thereby avoiding additional weighting parameters. By optimising policy precision based on outcomes from previous trials, the model not only improved task performance but also enhanced the explanatory power of the free parameter in capturing clinical characteristics (Fig. 5). Furthermore, the rate of change in policy precision emerged as a meaningful signal in explaining neural dynamics (Fig. 3E), particularly when inter-subject correlations (ISC) were considered, leading to improved explanatory power regarding variability in neural responses (Fig. 3F).

In the neural signal analysis, the regions with a positive correlation with the policy precision signal were predominantly affiliated with the DMN, whereas those with a negative correlation were associated with the DAN and FPN. These regional findings align with network-level results, which revealed a positive correlation with the occurrence of State 2 (dominated by the DMN) and a negative correlation with State 3 (dominated by the DAN and FPN). The anti-correlation between these states has been consistently observed in prior studies, where the DMN is typically associated with the resting (i.e., task-negative) state, and the DAN and FPN with task-positive processing. Although the DMN is generally activated during rest, mind-wandering, and self-referential thought (Damoiseaux et al., 2006; Dastjerdi et al., 2011; Power et al., 2011), recent evidence suggests its role extends to the processing of intrinsic information during task performance (Buckner and Carroll, 2007; Power et al., 2011; Raichle and Raichle, 2001; Weber et al., 2022). In line with this, State 2, dominated by DMN activity, was the second most frequently observed state during task execution in our study (Supplementary Figure. 3B), supporting the functional involvement of the DMN during goal-directed behaviour.

The DMN has been suggested to facilitate the detection of associative relevance between internal and external stimuli and to contribute to value coding (Roy et al., 2012). From this perspective, an alternative proposal argues that the DMN supports recursive exploration by selecting optimal actions within environmental contexts, potentially operating at higher levels of the inferential hierarchy to minimise free energy at lower levels (Carhart-Harris and Friston, 2010; Lehmann et al.,

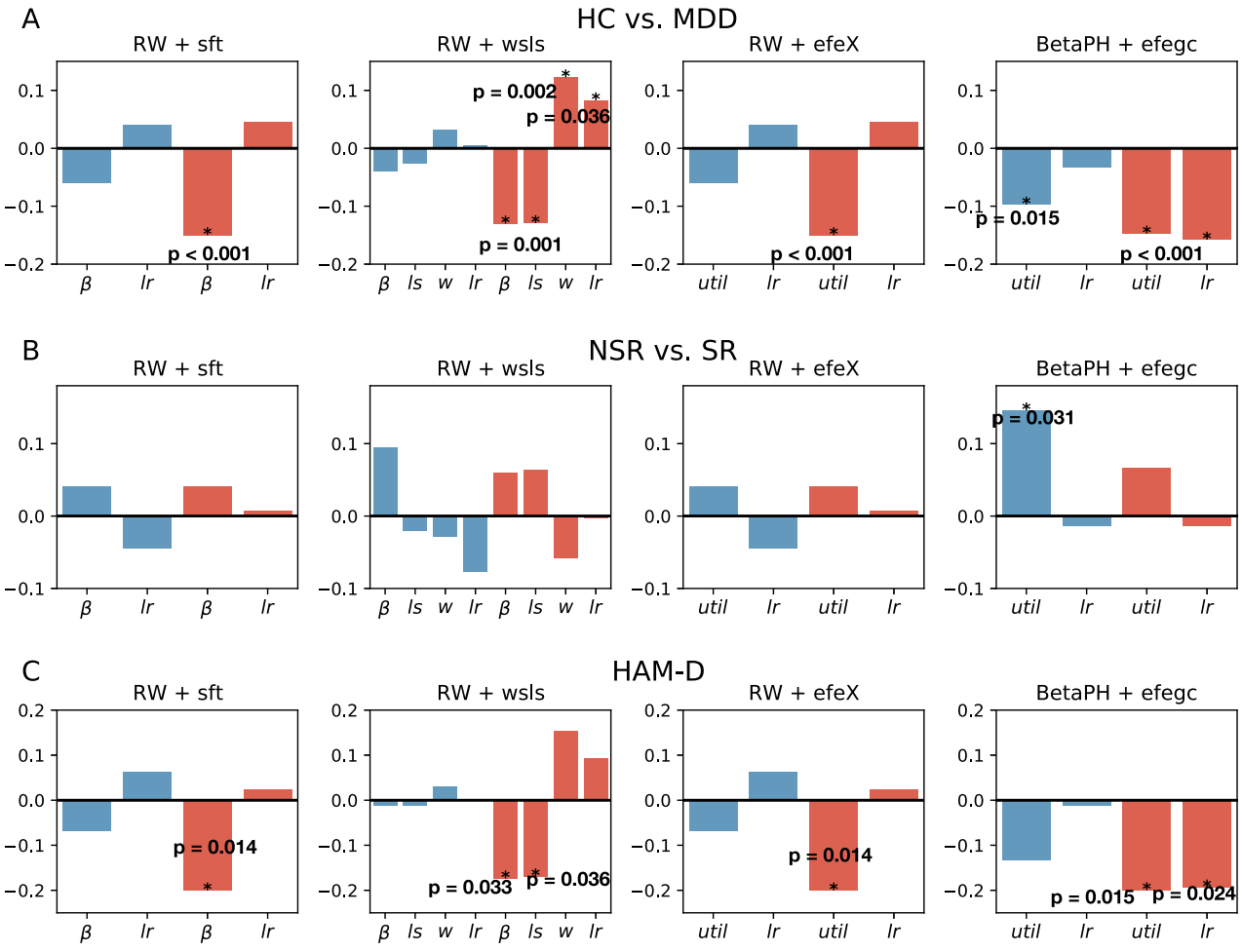


Fig. 5. Results of three types of analyses using parameter values estimated from four behaviour models. (A, B) Results of logistic regression analysis predicting group using parameters. (C) Results of analysis of covariance (ANCOVA) predicting HAM-D scores and each individual parameter value, respectively. Asterisks (*) indicate statistically significant predictors ($p < 0.05$). Blue and red bars indicate parameters from the loss task and the reward task, respectively.

2023). Additionally, the hippocampus, which was active in State 2 (Supplementary Figure. 1 J), is implicated in internal model construction and planning (Miller et al., 2017; Turner et al., 2022), further linking DMN functionality with internal modelling processes. Conversely, the activation of State 3, characterised by dominant DAN and FPN activity, was associated with decreased policy precision—likely reflecting increased uncertainty and the need for externally directed attention.

Correlation analyses between behavioural parameters inferred from the model and state dynamic metrics revealed that longer DT in State 4 was associated with increased choice inconsistency. By contrast, the metrics related to State 2 showed the opposite pattern. Further examination revealed a mutually inhibitory relationship between States 2 and 4: an increase in State 2 DT was associated with reduced transitions involving State 4, and vice versa (Supplementary Figure. 3D). This suggests that, while anti-correlation between States 3 and 2 related to the policy precision signal may reflect a cooperative dynamic supporting task performance via policy precision modulation, State 4 may interrupt the functional role of State 2. State 4, despite its generally lower nodal activation, showed relatively higher activation in the LIM and VAN and was characterised by high connectivity (Supplementary Figure. 2).

Previous studies suggest that VAN is activated by surprising external stimuli (Asplund et al., 2010; Corbetta and Shulman, 2002) and plays a role in regulating the balance between the DMN and DAN (Corbetta et al., 2008; Goulden et al., 2014). Conversely, DAN activity has been shown to suppress VAN, supporting sustained goal-directed attention. The stronger this suppression is, the better is the task performance (Vossel et al., 2014). Based on this, we speculate that, when surprise-induced VAN activation occurs due to unexpected outcomes, a rapid suppression by DAN may be necessary to restore attentional focus and maintain learning based on the internal model. This dynamic may reduce the DT in State 4 and facilitate appropriate engagement of States 2 and 3, thereby improving task performance and behavioural consistency. Moreover, these results align with the view that reactive strategy to previous observation are not a distinct category opposed to model-based (MB) control, but rather reflect a relative state over the MB continuum, where MB influence is diminished. From this perspective, transitions between States 2 and 3 correspond to a shift in action strategy characterised by reduced MB control and increased reliance on alternative mechanisms. The engagement of task-positive regions observed in State 3 may serve a compensatory role, supporting decision-making in the face of reduced MB influence.

In a related study by (Moutoussis et al., 2021), a single latent dimension termed ‘decision acuity’ was identified, capturing shared variance across multiple decision-making tasks. This construct aligns closely with policy precision (γ) or its principal component in the active inference framework. Their findings showed that intra- and inter-connectivity within FPN, medial prefrontal cortex (MPC), orbitofrontal cortex, medial and lateral (OFC), opercular cortex (OPC), posterior cingulate cortex (PCC), right dorsolateral prefrontal cortex (RDC), somatosensory and motor areas (SMT), and visual regions modules significantly predicted decision acuity. As shown in Supplementary Figure 2, State 4 exhibited high integration within RSNs overlapping these modules. Although these results are from resting-state fMRI, such baseline connectivity patterns likely influence task-state neural dynamics, suggesting a degree of relevance to our findings.

Moreover, while task-fMRI was not conducted in Experiment 2, the observed behavioural parameter differences between healthy controls and individuals with MDD or high suicidal risk may be attributable to differences in brain state dynamics. Specifically, modulating State 4 DT to maintain levels similar to those observed in healthy individuals during surprise events may aid in symptom alleviation or serve as an objective marker for assessing MDD risk.

Patients with MDD exhibited significantly lower total scores in the reward task, whereas no such group difference was observed in the loss task (Supplementary Figure. 7A). Lower performance in the reward task could stem either from impaired learning or from inconsistent choice behaviour despite learning. The former would be reflected in the parameters of the learning model, while the latter would manifest in the action model parameters. Indeed, both β and *util*, which govern decision consistency and reward sensitivity, were significantly reduced in patients with MDD across all action models. Additionally, parameters such as *w* and *ls*, which also contribute to decision variability, were elevated in the MDD group. Among learning parameters, the learning rate *lr* showed a significant reduction in MDD, although this effect was not observed in the *sft* and *efeX* models. These findings align with the literature reporting impairments in reward valuation processes among individuals with MDD (Huys et al., 2021; Rupprechter et al., 2018). By contrast, loss-related parameters appeared relatively preserved compared to healthy controls. Notably, scores from the reward and loss tasks were highly correlated (Supplementary Figure. 7B). Although no significant differences emerged between the NSR and SR groups in overall task scores, the SR group tended to show lower performance in the reward task but paradoxically higher performance in the loss task (Supplementary Figure. 7E). Reflecting this trend, the *util* parameter in the loss domain—estimated using the *efegc* model—was increased in the SR group, suggesting that altered sensitivity to negative outcomes could be a key feature in identifying suicide risk among individuals with MDD. Previous studies have noted that MDD is often associated with diminished reward sensitivity alongside heightened punishment sensitivity, which may result in an exaggerated gain–loss asymmetry during value-based decision-making (Chen et al., 2015; Dombrovski et al., 2013). The current findings extend this perspective by suggesting that such asymmetries may contribute to the development of suicidal risk. However, as this subgroup analysis included a small number of people for each NSR and SR group ($n = 21$ and 37, respectively), future studies with larger samples remain warranted to confirm our findings.

A behaviour model that can reflect the various mental states of people is warranted to better explain mental states through the model. In addition, there are limitations to understanding individual characteristics through parameters obtained from behavioural modelling. Therefore, high-dimensional brain signal data that reflects individual diversity must be integrated. This makes behavioural modelling even more critical. Rather than focusing solely on better explaining behavioural outcomes, we must prioritise the development of biologically plausible models, identify their neural correlates through brain signals, and understand mental symptoms in terms of the differences or abnormalities that arise from these connections. This will explain the differences in various brain states through models and serve as a foundation for personalised neuromodulatory treatment through accurate correlation between behaviour and individual brain activity.

4. Materials and methods

4.1. Participants

4.1.1. Experiment 1

A total of 54 volunteers participated in Experiment 1 (15 females and 39 males; mean [standard deviation] age: 22.4 [3.2] years; all participants were right-handed and had no history of psychiatric or neurological disorders). All participants provided written informed consent for all the procedures and data usage before the start of the study. The experimental procedures were approved by the Ethics Committee of Korea Advanced Institute of Science and Technology.

4.1.2. Experiment 2

A total of 117 participants aged 18 to 34 years were recruited through the outpatient clinic of the Samsung Medical Center in Seoul, South Korea, between July 2018 and October 2020. MDD was diagnosed by two psychiatrists (H. Kim and H.J. Jeon) based on clinical interviews. Those whose diagnosis was categorised as MDD per DSM-5 were included in the MDD group. The exclusion criteria were: (1) MDD with

psychotic features; (2) comorbidity of any other psychiatric illnesses including bipolar disorder, schizophrenia, delusional disorder, delirium, neurocognitive disorder, intellectual disability, and other mental disorders due to another medical condition; (3) history of substance-related disorders except for tobacco-related disorders within 12 months; (4) primary neurologic illness or history of brain damage; and (5) a history of major physical illness. HCs were recruited from the community using an advertisement for the Clinical Trial Center of Samsung Medical Center. A total of 117 individuals were initially enrolled, of whom 112 returned for a follow-up visit one month later. Among them, 105 completed a second follow-up visit another month later, during which they completed the questionnaire assessment. The control group included participants without any current or history of psychiatric illness and who consistently scored $\leq$7 on the HAM-D scale at all three visits (Yi et al., 2005). Suicidal risk was assessed using the suicidality module of the Mini International Neuropsychiatric Interview (Kim et al., 2017). Participants who scored $\geq$6 on the suicidality scale at any of the three visits were classified into the suicidal risk group (SR). This standard includes individuals who experienced suicidal ideation within the past month or those who, despite the absence of current suicidal thoughts, had a lifetime history of suicide attempts and reported a desire to harm or injure themselves during the past month. Among them, eight individuals showed an increase in suicidality from level 0 at the first visit to level 1 or 2 at the second and/or third visits. Of these eight participants, seven were diagnosed with MDD.

Among the 59 participants who met the criteria for the control group at the first visit, four individuals scored >7 on the HAM-D scale at one or more of the second and third visits. Of these, three had a score of 8 on only one occasion, with low overall symptom severity and no suicidal risk (score = 0 at all visits); thus, they were retained in the control group. The remaining participant scored >8 at both follow-up visits, with reported suicidality scores of 9 and 2 at the second and third visits, respectively, and was therefore excluded from the control group. This participant was the only one among those whose suicidal risk increased after an initial score of 0 at the first visit and who was not diagnosed with MDD. Therefore, a total of 58 participants were classified into the control group, 37 into the suicidal risk (SR) group, and 21 into the non-suicidal risk (NSR) group.

The study design was approved by the Institutional Review Board of Samsung Medical Center (IRB No 2018–04–137). All participants provided written informed consent before participating in this study in accordance with the Declaration of Helsinki.

4.2. fMRI data acquisition

In Experiment 1, all participants underwent MRI using 3.0-T MRI (Siemens Verio Syngo Scanner). Both T1-weighted (T1w; echo time [TE] = 2.02 ms, repetition time

[TR] = 2400 ms, field of view (FOV) = 224 × 224 mm, and 0.7 mm isovoxel) and T2-weighted (T2w; TE = 330 ms, TR = 2200 ms, FOV = 224 × 224 mm, and 0.7 mm isovoxel) structural images were scanned. A total of 240 echo-planar imaging scans of BOLD responses (multiband factor = 4, TE = 32 ms, TR = 1.5 s, flip angle = 50°, FOV = 225 × 221 mm, matrix size = 110 × 107, 2.0 mm isovoxel with no gap, 64 slices, phase encoding = anterior to posterior) were performed using a 32channel head coil. Field map images (TR = 731 ms, TE: 4.92 ms and 7.38 ms, flip angle = 90°) were also acquired to correct distorted images. In every trial, the participants performed the experimental task with both hands to press an MRI-compatible button to make their decision. In the pretraining session, prior to the first scan, the participants were acquainted with the apparatus and tasks performed during scanning. Sequences executed during this pretraining session were not encountered later in the experiment. Preprocessing and processing of the neuroimaging data were performed using AFNI (v23.1.00) (Cox, 1996; Cox and Hyde, 1997), FSL (v6.0.5.2), FreeSurfer (v7.1.1), Workbench (v1.5.0), and Advanced Normalization Tools (ANTs). Following the Human Connectome Project, a minimal processing pipeline for structural images, structural T1w and T2w images (Glasser et al., 2013), were processed, involving automatic segmentation of surfaces and reconstruction using FreeSurfer 7.1.1 (https://surfer.nmr.mgh.harvard.edu/) (Dale et al., 1999; Fischl et al., 2002). The generated volumes and surface files were converted to the CIFTI file format using ciftify (https://edickie.github.io/ciftify/) (Dickie et al., 2019) including MSMSulc surface realignment (Robinson et al., 2014).

Raw time-series of task fMRI were performed with slice-timing correction, realignment, and skull stripping using AFNI commands (3dTshift, 3dvolreg, and 3dAutomask) (Taylor et al., 2018). Realigned data were unwarped using a fieldmap in FSL. A bandpass filter (highpass = 0.01) was applied to the unwarped functional data and motion artifact time-series before regression. Using algorithms in ANTs, we registered the unwarped functional data into a 2-mm MNI152 T1w template space via skull-stripped T1w images preprocessed with AFNI. The eroded white matter (WM) and ventricular cerebrospinal fluid masks segmented using FreeSurfer were moved to the functional image space using the transformation matrix generated via ANTs in the previous step. PCA (3dpca in AFNI) was applied to the eroded ventricular mask in bandpass-filtered, scaled functional data, and ventricular PCA time-series was acquired. Using the fast ANATICOR algorithm in AFNI (Jo et al., 2010) on an eroded WM mask, a voxel-wise local WM regressor was generated.

3dTproject in AFNI, all regressors generated were projected into a regression matrix. Finally, the bandpass-filtered, scaled functional data were denoised, including pre-whitening, to remove the autocorrelation property using 3dREMLfit (residual maximum likelihood) in AFNI. Using transformation matrices and warping images generated by ANT algorithms, the denoised functional data were directly moved to a 2-mm MNI152 T1w template space and projected to the CIFTI format using the ciftify package (Dickie et al., 2019).

4.3. Neural signal extraction

After preprocessing, we acquired the averaged ROI time-series extracted from the Schaefer Atlas (Schaefer et al., 2018), including 100 regions covering the cortex, and the Melbourne Subcortex Atlas (Tian et al., 2020) covering the subcortex. Additionally, to include other subcortical regions related to the release of important neurotransmitters, including dopamine and norepinephrine, we added some regions to the CIT168 atlas (Pauli et al., 2018) and Harvard Ascending Arousal Network Atlas (Edlow et al., 2012) to obtain signals from a total of 147 ROIs. To extract the representation of overall brain states associated with network-level interactions, we employed the Graph Laplacian mixture model (GLMM). This method identifies meta-networks representing both intra- and inter-network interactions by estimating the Laplacian matrices that characterise the graph structure and dynamics of brain states. This method showed state classifications consistent with the timing of the various task paradigms in the HCP dataset (Maretic and Frossard, 2020; Ricchi et al., 2022). GLMM provides information on brain state activation patterns (estimating means), functional connectivity structure (Laplacians), and their temporal dynamics (probability of occurrences over time). Four major states were identified and depicted in Supplementary Figure. 1.

4.4. Behaviour task and modelling

4.4.1. Experiment 1

Participants faced two choices: one with a 70 % likelihood of winning and the other with 30 %. The number of trials was 40 per part; thus, a total of 80 trials were performed by each participant. Participants were explicitly instructed that one of the two bandits had a higher probability of winning and were asked to infer which one was better to win as the trial progressed to maximise the accumulated reward. Participants were expected to learn about the reward value of each option by trial-and-

error and utilise computed reward values to make their decision. Therefore, the probability of one choice being better represents the hidden state value inferred by the agents during the task. This state could be modelled using either a beta or a normal distribution. The inferred value over one choice could be updated in every trial, even if that choice was not selected, because winning or losing by selecting one choice simultaneously means losing or winning by selecting the other choice. The expected value can be represented by either scalar probability or a distribution form. Using a scalar value corresponds to the Rescorla-Wagner model and Pearce-Hall models. In the Pearce-Hall model, the learning rate in is changed across the trial based on the absolute value of prediction error.

$$EV_{i,\ t+1} = EV_{i,\ t} + lr \cdot \left| r(t-1) - EV_{i,\ t-1} \right| \cdot \left(r(t) \ - \ EV_{i,\ t} \right) \tag{5}$$

Meanwhile, using a distribution form, we can utilise the variance of state when calculating behavioural measures in the model, such as the learning rate or free energy. Each trial corresponds to the Bernoulli trial with outcomes of win or lose, and state value $q(\theta)$ can be represented by a beta distribution, where θ is the probability of one choice being better bandit. The beta distribution is parameterised as:

$$\theta \sim Beta(\alpha, \beta) \tag{6}$$

where α and β are equal to one plus the number of win or lose outcomes, respectively, for a given choice. In this case, as the task progress, the learning process can be represented simply by adding one according to the outcome of each trial (o_t), as the likelihood of outcome at each trial is a binomial distribution, which is conjugate to the beta distribution:

$$\begin{aligned} \alpha_{t+1} &= \rho_t \cdot \alpha_t + (1 \ - \ \rho_t) \cdot \alpha_0 + \delta_{a_t, o_t} \ \beta_{t+1} \\ &= \rho_t \cdot \beta_t + (1 \ - \ \rho_t) \cdot \beta_0 + \left(1 \ - \ \delta_{a_t,\ o_t}\right) \end{aligned} \tag{7}$$

δ indicates the Kronecker delta and takes 1 if both variables a_t and o_t are equal (when selecting choice *1* results in a win or selecting choice *0* results in a lose), and 0 otherwise. ρ is the decay parameter ($0 < \rho < 1$). Smaller values lead the concentration parameters to converge toward their initial values, effectively corresponding to a forgetting process in which learning from observations fades over time. This parameter is dynamically updated across trials as follows:

$$\rho_t = \frac{m \cdot PS}{1 + m \cdot PS} \ m = \frac{v}{1 - v}$$

$$PS \ = \ - \ln q_t(o_t; \theta)$$

$$q_t(o_t; \theta) \ = \ \frac{\alpha_t \delta_{a_t, o_t} + \beta_t \left(1 \ - \ \delta_{a_t, o_t}\right)}{\alpha_t + \beta_t}$$

The likelihood of observing o_t given the current belief is denoted as $q_t(o_t;\theta)$, which can be efficiently computed under a beta distribution. When an outcome o_t that was unlikely under the current belief is observed, the level of surprise increases. This mismatch between prior beliefs and observed outcomes is quantified as predictive surprise (PS). An increased PS reduces the relative influence of the prior belief during belief updating, thereby allowing the new observation to exert a stronger effect on the posterior belief. Meanwhile, parameter m, which depends on volatility parameter ($0 < v < 1$), modulates how strongly the surprise influences learning. Volatility refers to the degree to which the underlying generative probability has drifted, making it less compatible with the current belief. Higher volatility implies greater environmental instability and thus, a stronger need to adapt by giving more weight to recent observations.

In addition, another model modulating learning rate was constructed as follows:

$$\alpha_{t+1} = \alpha_t + \delta_{a_t, o_t} \cdot lr \cdot \rho_{t-1} \ \beta_{t+1} = \beta_t + \left(1 \ - \ \delta_{a_t,\ o_t}\right) \cdot lr \cdot \rho_{t-1} \tag{8}$$

In this model, ρ_t is always positive as per Eq. (7) and serves as a substitute for the absolute value of the prediction error in the Pearce-Hall model.

State value $q(\theta)$ can be represented by a normal distribution as $\theta \sim Norm(\mu,\ \sigma^2)$ with a prior mean of 0.5, and typically, a large standard deviation of 10. In this case, corresponding to the Kalman-filter model (Kalman, 1960), the likelihood of outcome at each trial can be represented as a normal distribution, and the standard deviation of the observation (σ_o in Eq. (9)) can either be fixed or defined as a free parameter. That modulates how much prior belief will be updated:

$$\mu_{t+1} = \ \mu_t \ + \ \frac{\sigma_t^2 \sigma_o^2}{\sigma_t^2 + \sigma_o^2} \ (r(t) \ - \ \mu_t) \tag{9}$$

As an action model, we used a conventional softmax function (Eq. (2)), WSLS-RL (Eq. (3)), and two models based on AIF scheme utilising EFE. Formally, we express the EFE of a choice a on trial t as:

$$G_t(a) = D_{KL}[q(o_t|a_t = a) \| \ p(o_t)] + E_{q(\theta)}[H][o_t|\theta, a_t = a]] \tag{10}$$

The preference distribution of outcomes, $p(o_t)$, is encoded using *util* parameter for each participant as ($0 < util < 1$):

$$P(o_t)[win, \ lose] \ = \ \left[\frac{exp^{util}}{exp^{util} + 1}, \frac{1}{exp^{util} + 1}\right] \tag{11}$$

The probability of selecting choice a is then calculated as:

$$P_t(a) = \sigma(-\gamma \cdot G_t(a)) \tag{12}$$

In addition, the precision of beliefs about actions γ, can be inferred instead of a fixed value (Friston et al., 2014). In the *efeX* model, the following adjustment is excluded. In the *efegc* model, the initial value of γ is calculated from the fixed prior value of β_0 (Smith et al., 2022). Thus, the probability that the agent will choose each policy π_0 will be

$$\pi_0 \leftarrow \sigma(-\gamma G) \tag{13}$$

As precision weights the EFE to determine the probability of choosing each policy, the results of trial t are given as either a win or a loss, and the variational free energy (VFE) of each policy F over that observation is calculated. Subsequently, the probability of choosing each policy after a subsequent observation is:

$$\pi \leftarrow \sigma(-F - \gamma G) \tag{14}$$

and the error of the agent's EFE (G_{error}) is given as:

$$G_{error} \leftarrow (\pi - \pi_0) \cdot (-G) \tag{15}$$

The above equation reflects the level of (dis)agreement between the EFE (G and π_0) and the VFE of the observation (F and π) (Smith et al., 2022). Thus, the β whose inverse means γ is updated based on G_{error}.

$$\gamma_{update} \leftarrow 1 \Big/ \gamma - 1 \Big/ \beta_0 + G_{error} \tag{16}$$

$$1 \Big/ \gamma \leftarrow 1 \Big/ \gamma - \gamma_{update} \Big/ \psi$$

Therefore, policy precision reflects the reliability of the EFE based on new observations that represent either winning or losing outcomes. If the result of a chosen action betrays expectations, precision decreases, causing the probability of action selection to rely less on the expected free energy. Simultaneously, this increases the effect of free energy that is calculated from the previous outcome.

4.4.2. Experiment 2

The overall paradigm of the decision-making task is similar to that of Experiment 1. However, participants performed 80 trials at once, and each trial's decision was either a loss or reward specific. In reward decision-making task trials, choice options are presented with a blue boundary, while red boundaries are used in loss trials (Supplementary Figure. 4). One option was associated with a high probability of reward (75 % of reward) while another option had a high probability of non-

reward. Reward and loss schedules were semi-randomised and counterbalanced with two schedules (74 participants for set 1 and 43 for set 2). The same behaviour models were used as in Experiment 1.

4.5. GLM analysis

To identify the shared dynamics of neural engagement across participants who performed the same experimental task, we calculated the dISC for each level of the time-series (Majumdar et al., 2023; Nastase et al., 2019). A sliding window with a length of three repetition times (TRs) corresponded to approximately the length of one trial (3.7–4.3 s) and a 1 TR overlap. This was calculated by correlating the neural signal of one participant for each region within each window with the mean signal within the same window for the same region from the remaining participants (leave-one-out method). The ISC method increases the signal-to-noise ratio (SNR) to detect stimulus-induced inter-regional correlations. The improvement in the SNR arises from filtering out intrinsic neural dynamics (for example, responses arising from intrinsic cognitive processes unrelated to ongoing stimulus processing) and non-neural artifacts (for example, respiratory rate and motion) that can influence network correlation patterns within the brain but are not correlated across participants. The time-series of neural signal from node i of participant j is $x_j^i(t)$; this signal consists of a linear combination of the shared component c(t), idiosyncratic responses to the stimulus for participant j $id_j^i(t)$, and spontaneous activity unrelated to the stimulus $\varepsilon_j^i(t)$. Therefore, the time series of participants a and b can be expressed as Eq. (17), with shared c(t).

$$x_a^i(t) = \alpha_a^i \cdot c(t)^i + \delta_a^i \cdot id_a^i(t) + \varepsilon_a^i(t)\ x_b^i(t) = \alpha_b^i \cdot c(t)^i + \delta_b^i \cdot id_b^i(t) + \varepsilon_b^i(t) \quad (17)$$

$$ISC_{a,b}^i \approx \sqrt{\alpha_a^i \cdot \alpha_b^i}$$

In the previous ISC study, the correlation between participants a and b was approximated to the $\sqrt{\alpha_a^i \cdot \alpha_b^i}$ because id(t) and $\varepsilon(t)$ are not systematically correlated across participants, whereas c(t) is perfectly correlated (Nastase et al., 2019). However, in the current study, the idiosyncratic response conforms to the policy precision signal, in that id (t) is not perfectly uncorrelated across participants. The ISC of the policy precision signals varied from 0 to 1, with a considerably high value. In the original logic of ISC, this idiosyncratic response should be averaged to a small value close to zero due to its inconsistent characteristic, such that only c(t) can be isolated; however, correlations across idiosyncratic responses of the participants exist. Therefore, the neural ISC values corresponded to the values including not only the effect of the shared component but also the idiosyncratic response component in the extent of ISC of the policy precision signal for participant j. Thus, the neural ISC value of each node i can be approximated as:

$$ISC^i \approx \beta_0 + \beta_1 \cdot bISC \quad (18)$$

where we can assume that the intercept term β_0 corresponds to the effect of α^i in Eq. (17) and the β_1 corresponds to the effect of ISC of behavioural signal (bISC), such as policy precision. If the bISC is 0, the result is equivalent to the third line of Eq. (17). This can be applied to the time-series data of each participant (j) using dynamic ISC (dISC) of neural and dISC of behavioural signals (dbISC in Eq. (19)) for performing GLM analysis.

$$dISC_j^i(t) = \beta_{0j} + \beta_{1j} \cdot dbISC_j \quad (19)$$

Since action timings (AS) and feedback visual stimulus (FVS) are also different across participants, we performed a GLM analysis that included dISC of these signals, both with and without dISC of policy precision signal. We then compared the model performance using AIC, BIC scores.

To validate the relevance of policy precision signals in explaining neural signals, we performed a conventional GLM to identify which regions or networks are related to this signal. Using beta coefficient values estimated from the GLM, we conducted a one-sample t-test to examine if those values are significantly positive or negative.

5. Conclusion

Our findings demonstrate that the biologically plausible AIF model provides a better reflection of individual states than conventional RL models. This highlights the potential of developing behavioural models in this manner to advance precision medicine in psychiatry.

Data and code availability

The data underlying this article will be shared upon reasonable request to the corresponding author. The code is available at https://github.com/ydy-dyd-y/AIF_dFC.

CRediT authorship contribution statement

Dayoung Yoon: Writing – original draft, Formal analysis, Conceptualization. **Jaejoong Kim:** Conceptualization. **Do Hyun Kim:** Data curation. **Dong Woo Shin:** Data curation. **Su Hyun Bong:** Project administration. **Jaewon Kim:** Project administration. **Hae-Jeong Park:** Writing – review & editing. **Hong Jin Jeon:** Data curation. **Bumseok Jeong:** Writing – review & editing, Writing – original draft, Conceptualization.

Declaration of competing interest

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

Acknowledgments

This study was supported by the Medical Scientist Training Program from the Ministry of Science & ICT of Korea [NRF-2016M3C7A1914448], grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI) [RS-2024-00440131] and the Brain Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science & ICT [RS-2025-00562004] and [NRF-2022M3E5E8081200]. We would like to thank Editage (www.editage.co.kr) for English language editing.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2025.121479.

References

Ahn, W.Y., Busemeyer, J.R., Wagenmakers, E.J., Stout, J.C., 2008. Comparison of decision learning models using the generalization criterion method. Cogn. Sci. https://doi.org/10.1080/03640210802352992.

Ahn, W.Y., Vasilev, G., Lee, S.H., Busemeyer, J.R., Kruschke, J.K., Bechara, A., Vassileva, J., 2014. Decision-making in stimulant and opiate addicts in protracted abstinence: evidence from computational modeling with pure users. Front. Psychol. 5. https://doi.org/10.3389/fpsyg.2014.00849.

Asplund, C.L., Todd, J.J., Snyder, A.P., Marois, R., 2010. A central role for the lateral prefrontal cortex in goal-directed and stimulus-driven attention. Nat. Neurosci. 13, 507–512. https://doi.org/10.1038/nn.2509.

Brown, V.M., Wilson, J., Hallquist, M.N., Szanto, K., Dombrovski, A.Y., 2020. Ventromedial prefrontal value signals and functional connectivity during decision-making in suicidal behavior and impulsivity. Neuropsychopharmacology 45. https://doi.org/10.1038/s41386-020-0632-0.

Buckner, R.L., Carroll, D.C., 2007. Self-projection and the brain. Trends. Cogn. Sci. 11. https://doi.org/10.1016/j.tics.2006.11.004.

Carhart-Harris, R.L., Friston, K.J., 2010. The default-mode, ego-functions and free-energy: a neurobiological account of Freudian ideas. Brain. https://doi.org/10.1093/brain/awq010.
Chen, C., Takahashi, T., Nakagawa, S., Inoue, T., Kusumi, I., 2015. Reinforcement learning in depression: a review of computational research. Neurosci. Biobehav. Rev. https://doi.org/10.1016/j.neubiorev.2015.05.005.
Clark, L., Dombrovski, A.Y., Siegle, G.J., Butters, M.A., Shollenberger, C.L., Sahakian, B. J., Szanto, K., 2011. Impairment in risk-sensitive decision-making in older suicide attempters with depression. Psychol. Aging 26. https://doi.org/10.1037/a0021646.
Cohen, J.D., McClure, S.M., Yu, A.J., 2007. Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. Philosophical Transactions of the Royal Society B: Biological Sciences. https://doi.org/10.1098/rstb.2007.2098.
Corbetta, M., Patel, G., Shulman, G.L., 2008. The reorienting system of the Human brain: from environment to theory of mind. Neuron 58, 306–324. https://doi.org/10.1016/j.neuron.2008.04.017.
Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. Nat. Rev. Neurosci. 3. https://doi.org/10.1038/nrn755.
Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput. Biomed. Res. 29. https://doi.org/10.1006/cbmr.1996.0014.
Cox, R.W., Hyde, J.S., 1997. Software tools for analysis and visualization of fMRI data. NMR Biomed. 10. https://doi.org/10.1002/(SICI)1099-1492(199706/08)10:4/5<171::AID-NBM453>3.0.CO;2-L.
Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis: I. Segmentation and surface reconstruction. Neuroimage 9. https://doi.org/10.1006/nimg.1998.0395.
Damoiseaux, J.S., Rombouts, S.A.R.B., Barkhof, F., Scheltens, P., Stam, C.J., Smith, S.M., Beckmann, C.F., 2006. Consistent resting-state networks across healthy subjects. Proc. Natl. Acad. Sci. u S. a 103. https://doi.org/10.1073/pnas.0601417103.
Dastjerdi, M., Foster, B.L., Nasrullah, S., Rauschecker, A.M., Dougherty, R.F., Townsend, J.D., Chang, C., Greicius, M.D., Menon, V., Kennedy, D.P., Parvizi, J., 2011. Differential electrophysiological response during rest, self-referential, and non-self-referential tasks in human posteromedial cortex. Proc. Natl. Acad. Sci. u S. a 108. https://doi.org/10.1073/pnas.1017098108.
Dickie, E.W., Anticevic, A., Smith, D.E., Coalson, T.S., Manogaran, M., Calarco, N., Viviano, J.D., Glasser, M.F., Van Essen, D.C., Voineskos, A.N., 2019. Ciftify: a framework for surface-based analysis of legacy MR acquisitions. Neuroimage 197. https://doi.org/10.1016/j.neuroimage.2019.04.078.
Dombrovski, A.Y., Hallquist, M.N., Brown, V.M., Wilson, J., Szanto, K., 2019. Value-based choice, contingency learning, and suicidal behavior in mid- and late-life depression. Biol. Psychiatry 85. https://doi.org/10.1016/j.biopsych.2018.10.006.
Dombrovski, A.Y., Szanto, K., Clark, L., Reynolds, C.F., Siegle, G.J., 2013. Reward signals, attempted suicide, and impulsivity in late-life depression. JAMa Psychiatry 70. https://doi.org/10.1001/jamapsychiatry.2013.75.
Edlow, B.L., Takahashi, E., Wu, O., Benner, T., Dai, G., Bu, L., Grant, P.E., Greer, D.M., Greenberg, S.M., Kinney, H.C., Folkerth, R.D., 2012. Neuroanatomic connectivity of the human ascending arousal system critical to consciousness and its disorders. J. Neuropathol. Exp. Neurol. 71. https://doi.org/10.1097/NEN.0b013e3182588293.
Findling, C., Skvortsova, V., Dromnelle, R., Palminteri, S., Wyart, V., 2019. Computational noise in reward-guided learning drives behavioral variability in volatile environments. Nat. Neurosci. 22. https://doi.org/10.1038/s41593-019-0518-9.
Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., Van Der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A.M., 2002. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 33. https://doi.org/10.1016/S0896-6273(02)00569-X.
Friston, K., Mattout, J., Trujillo-Barreto, N., Ashburner, J., Penny, W., 2007. Variational free energy and the Laplace approximation. Neuroimage 34. https://doi.org/10.1016/j.neuroimage.2006.08.035.
Friston, K., Schwartenbeck, P., FitzGerald, T., Moutoussis, M., Behrens, T., Dolan, R.J., 2014. The anatomy of choice: dopamine and decision-making. Philos. Trans. R. Soc. B: Biol. Sci. 369. https://doi.org/10.1098/rstb.2013.0481.
Glasser, M.F., Sotiropoulos, S.N., Wilson, J.A., Coalson, T.S., Fischl, B., Andersson, J.L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J.R., Van Essen, D.C., Jenkinson, M., 2013. The minimal preprocessing pipelines for the Human Connectome Project. Neuroimage 80, 105–124. https://doi.org/10.1016/j.neuroimage.2013.04.127.
Goulden, N., Khusnulina, A., Davis, N.J., Bracewell, R.M., Bokde, A.L., McNulty, J.P., Mullins, P.G., 2014. The salience network is responsible for switching between the default mode network and the central executive network: replication from DCM. Neuroimage 99. https://doi.org/10.1016/j.neuroimage.2014.05.052.
Herrnstein, R.J., Rachlin, H., Laibson, D.I., 1997. The Matching Law: Papers in Psychology and Economics. Cambridge, MA. Kacelnik, New York, p. 53. A.
Huys, Q.J.M., Browning, M., Paulus, M.P., Frank, M.J., 2021. Advances in the computational understanding of mental illness. Neuropsychopharmacology. https://doi.org/10.1038/s41386-020-0746-4.
Huys, Q.J.M., Maia, T.V., Frank, M.J., 2016. Computational psychiatry as a bridge from neuroscience to clinical applications. Nat. Neurosci. https://doi.org/10.1038/nn.4238.
Jo, H.J., Saad, Z.S., Simmons, W.K., Milbury, L.A., Cox, R.W., 2010. Mapping sources of correlation in resting state FMRI, with artifact detection and removal. Neuroimage 52. https://doi.org/10.1016/j.neuroimage.2010.04.246.
Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. J. Fluids Eng., Trans. ASME 82. https://doi.org/10.1115/1.3662552.
Kim, K., Kim, S.W., Myung, W., Han, C.E., Fava, M., Mischoulon, D., Papakostas, G.I., Seo, S.W., Cho, H., Seong, J.K., Jeon, H.J., 2017. Reduced orbitofrontal-thalamic functional connectivity related to suicidal ideation in patients with major depressive disorder. Sci. Rep. 7. https://doi.org/10.1038/s41598-017-15926-0.
Lehmann, K., Bolis, D., Friston, K.J., Schilbach, L., Ramstead, M.J.D., Kanske, P., 2023. An active-inference approach to second-person neuroscience. Perspect. Psychol. Sci. https://doi.org/10.1177/17456916231188000.
Liakoni, V., Modirshanechi, A., Gerstner, W., Brea, J., 2021. Learning in volatile environments with the bayes factor surprise. Neural Comput. 33. https://doi.org/10.1162/neco_a_01352.
Maia, T.V., Frank, M.J., 2011. From reinforcement learning models to psychiatric and neurological disorders. Nat. Neurosci. https://doi.org/10.1038/nn.2723.
Majumdar, G., Yazin, F., Banerjee, A., Roy, D., 2023. Emotion dynamics as hierarchical bayesian inference in time. Cereb. Cortex 33. https://doi.org/10.1093/cercor/bhac305.
Maretic, H.P., Frossard, P., 2020. Graph laplacian mixture model. IEEe Trans. Signal. Inf. Process. Netw. 6, 261–270. https://doi.org/10.1109/TSIPN.2020.2983139.
Miller, K.J., Botvinick, M.M., Brody, C.D., 2017. Dorsal hippocampus contributes to model-based planning. Nat. Neurosci. 20, 1269–1276. https://doi.org/10.1038/nn.4613.
Montague, P.R., Dolan, R.J., Friston, K.J., Dayan, P., 2012. Computational psychiatry. Trends. Cogn. Sci. https://doi.org/10.1016/j.tics.2011.11.018.
Moutoussis, M., Garzón, B., Neufeld, S., Bach, D.R., Rigoli, F., Goodyer, I., Bullmore, E., Fonagy, P., Jones, P., Hauser, T., Romero-Garcia, R., St Clair, M., Vértes, P., Whitaker, K., Inkster, B., Prabhu, G., Ooi, C., Toseeb, U., Widmer, B., Bhatti, J., Villis, L., Alrumaithi, A., Birt, S., Bowler, A., Cleridou, K., Dadabhoy, H., Davies, E., Firkins, A., Granville, S., Harding, E., Hopkins, A., Isaacs, D., King, J., Kokorikou, D., Maurice, C., McIntosh, C., Memarzia, J., Mills, H., O'Donnell, C., Pantaleone, S., Scott, J., Fearon, P., Suckling, J., van Harmelen, A.L., Kievit, R., Guitart-Masip, M., Dolan, R.J., 2021. Decision-making ability, psychopathology, and brain connectivity. Neuron 109. https://doi.org/10.1016/j.neuron.2021.04.019.
Nastase, S.A., Gazzola, V., Hasson, U., Keysers, C., 2019. Measuring shared responses across subjects using intersubject correlation. Soc. Cogn. Affect. Neurosci. 14. https://doi.org/10.1093/scan/nsz037.
Otto, A.R., Taylor, E.G., Markman, A.B., 2011. There are at least two kinds of probability matching: evidence from a secondary task. Cognition 118. https://doi.org/10.1016/j.cognition.2010.11.009.
Parr, T., Pezzulo, G., Friston, K.J., 2022. Active Inference: the Free Energy Principle in Mind, Brain, and Behavior. MIT Press.
Pauli, W.M., Nili, A.N., Michael Tyszka, J., 2018. Data descriptor: a high-resolution probabilistic in vivo atlas of human subcortical brain nuclei. Sci. Data 5. https://doi.org/10.1038/sdata.2018.63.
Pearce, J.M., Hall, G., 1980. A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. Psychol. Rev. 87. https://doi.org/10.1037/0033-295X.87.6.532.
Penny, W.D., Stephan, K.E., Daunizeau, J., Rosa, M.J., Friston, K.J., Schofield, T.M., Leff, A.P., 2010. Comparing families of dynamic causal models. PLoS. Comput. Biol. 6. https://doi.org/10.1371/journal.pcbi.1000709.
Power, J.D., Cohen, A.L., Nelson, S.M., Wig, G.S., Barnes, K.A., Church, J.A., Vogel, A.C., Laumann, T.O., Miezin, F.M., Schlaggar, B.L., Petersen, S.E., 2011. Functional network organization of the human brain. Neuron 72. https://doi.org/10.1016/j.neuron.2011.09.006.
Raichle, M.E., Raichle, M.E., 2001. Searching for a baseline: functional imaging and the resting human brain. Nat. Rev. Neurosci. 2. https://doi.org/10.1038/35094500.
Rescorla, R.A., Wagner, A.R., 1972. A theory of pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. Clasical Conditioning II: Current Research and Theory.
Ricchi, I., Tarun, A., Maretic, H.P., Frossard, P., Van De Ville, D., 2022. Dynamics of functional network organization through graph mixture learning. Neuroimage 252. https://doi.org/10.1016/j.neuroimage.2022.119037.
Rigoux, L., Stephan, K.E., Friston, K.J., Daunizeau, J., 2014. Bayesian model selection for group studies - revisited. Neuroimage 84. https://doi.org/10.1016/j.neuroimage.2013.08.065.
Robinson, E.C., Jbabdi, S., Glasser, M.F., Andersson, J., Burgess, G.C., Harms, M.P., Smith, S.M., Van Essen, D.C., Jenkinson, M., 2014. MSM: a new flexible framework for multimodal surface matching. Neuroimage 100. https://doi.org/10.1016/j.neuroimage.2014.05.069.
Roy, M., Shohamy, D., Wager, T.D., 2012. Ventromedial prefrontal-subcortical systems and the generation of affective meaning. Trends. Cogn. Sci. https://doi.org/10.1016/j.tics.2012.01.005.
Rupprechter, S., Stankevicius, A., Huys, Q.J.M., Steele, J.D., Seriès, P., 2018. Major depression impairs the use of reward values for decision-making. Sci. Rep. 8. https://doi.org/10.1038/s41598-018-31730-w.
Schaefer, A., Kong, R., Gordon, E.M., Laumann, T.O., Zuo, X.-N., Holmes, A.J., Eickhoff, S.B., Yeo, B.T.T., 2018. Local-global parcellation of the Human cerebral cortex from intrinsic functional connectivity MRI. Cereb. Cortex 28, 3095–3114. https://doi.org/10.1093/cercor/bhx179.
Schulz, E., Gershman, S.J., 2019. The algorithmic architecture of exploration in the human brain. Curr. Opin. Neurobiol. https://doi.org/10.1016/j.conb.2018.11.003.
Smith, R., Friston, K.J., Whyte, C.J., 2022. A step-by-step tutorial on active inference and its application to empirical data. J. Math. Psychol. 107. https://doi.org/10.1016/j.jmp.2021.102632.
Steyvers, M., Lee, M.D., Wagenmakers, E.J., 2009. A Bayesian analysis of human decision-making on bandit problems. J. Math. Psychol. 53. https://doi.org/10.1016/j.jmp.2008.11.002.

Taylor, P.A., Chen, G., Glen, D.R., Rajendra, J.K., Reynolds, R.C., Cox, R.W., 2018. FMRI processing with AFNI: some comments and corrections on "exploring the impact of analysis software on task fMRI results.". bioRxiv.

Thomas Yeo, B.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zöllei, L., Polimeni, J.R., Fisch, B., Liu, H., Buckner, R.L., 2011. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J. Neurophysiol. 106, 1125–1165. https://doi.org/10.1152/jn.00338.2011.

Tian, Y., Margulies, D.S., Breakspear, M., Zalesky, A., 2020. Topographic organization of the human subcortex unveiled with functional connectivity gradients. Nat. Neurosci. 23. https://doi.org/10.1038/s41593-020-00711-6.

Turecki, G., Brent, D.A., Gunnell, D., O'Connor, R.C., Oquendo, M.A., Pirkis, J., Stanley, B.H., 2019. Suicide and suicide risk. Nat. Rev. Dis. Primers. https://doi.org/10.1038/s41572-019-0121-0.

Turner, V.S., O'Sullivan, R.O., Kheirbek, M.A., 2022. Linking external stimuli with internal drives: a role for the ventral hippocampus. Curr. Opin. Neurobiol. https://doi.org/10.1016/j.conb.2022.102590.

Vidaurre, D., Abeysuriya, R., Becker, R., Quinn, A.J., Alfaro-Almagro, F., Smith, S.M., Woolrich, M.W., 2018. Discovering dynamic brain networks from big data in rest and task. Neuroimage. https://doi.org/10.1016/j.neuroimage.2017.06.077.

Vossel, S., Geng, J.J., Fink, G.R., 2014. Dorsal and ventral attention systems: distinct neural circuits but collaborative roles. Neuroscientist. 20. https://doi.org/10.1177/1073858413494269.

Weber, S., Aleman, A., Hugdahl, K., 2022. Involvement of the default mode network under varying levels of cognitive effort. Sci. Rep. 12. https://doi.org/10.1038/s41598-022-10289-7.

Worthy, D.A., Hawthorne, M.J., Otto, A.R., 2013. Heterogeneity of strategy use in the Iowa gambling task: a comparison of win-stay/lose-shift and reinforcement learning models. Psychon. Bull. Rev. 20. https://doi.org/10.3758/s13423-012-0324-9.

Worthy, D.A., Todd Maddox, W., 2014. A comparison model of reinforcement-learning and win-stay-lose-shift decision-making processes: a tribute to W.K. Estes. J. Math. Psychol. 59. https://doi.org/10.1016/j.jmp.2013.10.001.

Yechiam, E., Busemeyer, J.R., Stout, J.C., Bechara, A., 2005. Using cognitive models to map relations between neuropsychological disorders and human decision-making deficits. Psychol. Sci. 16. https://doi.org/10.1111/j.1467-9280.2005.01646.x.

Yi, J.S., Bae, S.O., Ahn, Y.M., Park, D., 2005. Validity and reliability of the korean version of the Hamilton Depression rating Scale (K-HDRS). J. Korean Neuropsychiatr. Assoc. 44.