Attention networks engage the default mode network related to policy precision under uncertainty

Dayoung Yoon[1], Jaejoong Kim[2], Do Hyun Kim[3], Dong Woo Shin[4], Su Hyun Bong[1], Jaewon Kim[1], Hae-Jeong Park[5, 6, 7] Bumseok Jeong[1, 8]

1 Graduate School of Medical Science and Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Republic of Korea
2 Department of Psychiatry and Behavioral Sciences, University of Minnesota, MN 55454, USA
3 Department of Psychiatry, Dankook University College of Medicine, Cheonan, Korea
4 Department of Neurology, Ewha Womans University Mokdong Hospital, Seoul, Korea
5 Department of Cognitive Science, Yonsei University, Seoul, South Korea
6 Department of Nuclear Medicine, Department of Psychiatry, Graduate School of Medical Science, Brain Korea 21 Project, Yonsei University College of Medicine, Seoul, South Korea
7 Institute of Human Complexity and Systems Science, Center for Systems and Translational Brain Sciences, Yonsei University, Seoul, Republic of Korea
8 KAIST Institute for Health Science and Technology and Graduate School of Data Science, KAIST, Daejeon, Republic of Korea

Abstract

In decision-making, choices impact the present and cascade into future decisions, highlighting the importance of confidence when making a decision. Here, we investigated the meta-network level neural correlates of this confidence by estimating the instantaneous changes in precision in selecting actions and comparing them with brain state trajectories. To confirm the relationship between behavioral and neural signals, we leveraged inter-subject correlation to determine how similar the group-shared components of time-series of precision-triggered meta-network occurrences are across participants and how this similarity changes during decision-making. We found that policy precision is a proper behavioral signal to explain the meta-network level neural dynamics. It positively correlated with the default mode network (DMN)-dominant state, the occurrence of which is mutually exclusive with the dorsal attentional network-(DAN) and frontoparietal network (FPN)-dominant state, the activation of which is speculated to be associated with a highly uncertain state and arises from increased integration between the DAN, FPN, and DMN. Therefore, supporting the novel perspective that the DMN may reflect internal beliefs, these findings indicate that their integration promotes the DAN and FPN to exert attention to decrease uncertainty.

Introduction

In everyday life, we make various decisions, ranging from trivial to pivotal. In most cases, a decision does not conclude with itself but extends to the subsequent state that

engenders another decision. In these cases, both the decision made and the level of certainty regarding that decision play critical roles in knowledge formation. The confidence at the time of making a decision influences the agent's reaction to the outcome and their future decisions in the subsequent state. The **Bayesian brain** theory posits that the brain's primary role is "inference," asserting that perception occurs through Bayesian inference of belief about the environmental state (s in Fig. 1b) based on **prior belief** and sensory inputs (D and o in Fig. 1b). How much each information influences the inferred **posterior belief** is determined by the **precision** of each information.

The active inference framework, unifying the theory of action and perception, attempts to elucidate a series of decision-making by minimizing a unified objective function, **free energy**. This framework treats action selection as an inference problem as well. Therefore, not only the belief about states but also the belief about **policies** are inferred, and the precision of this belief was suggested [16]. With this framework (Fig. 1b), the agent infers both its current state (s_n) and optimal policies (π) based on the probability of both outcomes in each context (A) and transitions (B) depending on the policy. Given its prior preferences over outcomes (C), the expected free energy (G) of each policy is determined based on how closely the expected future outcomes align to preferred outcomes by following each policy, which is weighted by the precision over the policies (γ), indicating how confident the agent is that these policies will minimize G (for the details of the G calculation, please see SI Appendix, Supplementary Methods) [15].

Precision over the policies has been suggested to be encoded by dopamine [1, 18, 39]; however, the exact computational role of dopamine remains controversial [4, 19, 21, 38]. This controversy is attributed to two main reasons: first, dopaminergic neurons project to the whole brain, and the brain operates through highly complex and interconnected networks. Therefore, observing local and regional activation is insufficient, and understanding its pathways and subsequent effects through connectivity is crucial [2, 41, 42]. Second, any **computational model** of a biological signal remains hypothetical by nature [11, 14]. We argue that if a newly proposed signal can explain network-level dynamics in the context of a series of decision-making, it could at least validate the role of the signal in the neural process. This should precede further detailed studies on the hypothetical behavioral signals.

To achieve this, we first obtained network-level dynamics by estimating **dynamic functional connectivity** based on graph structures from functional magnetic resonance imaging (fMRI) data measured during decision-making tasks. Second, we selected a very simple **two-armed bandit task** for the behavioral experiment. The reason for selecting this task was to ensure that signals, such as the precision of prior belief (hereafter referred to as "prediction precision") and **prediction error**, which are not the focus of this study, could be updated following the same trajectory across participants, allowing us to isolate the precision of policies (hereafter referred to as "policy precision (γ)"). If there are three or more options available, the prediction precision of the selected option increases due to the selection, whereas the precision of the unselected options remains unchanged. Consequently, the trajectories of prediction precision across trials vary according to the agent's choices. However, when there are only two options, the outcome of one simultaneously reveals the outcome of the other, making the trajectory of prediction precision across trials identical regardless of the choice outcome (dbISC in Fig. 1f).

Allowing only the policy precision signal variation is to apply the ISC method (for the details of this approach, please see SI Appendix, Supplementary Methods) [20]. This method can avoid the requirement for a predetermined response model and increase the signal-to-noise ratio from the behavior signal. Therefore, it is adequate for validating the influence of unsettled signals [28]. In an earlier study, ISC was utilized to

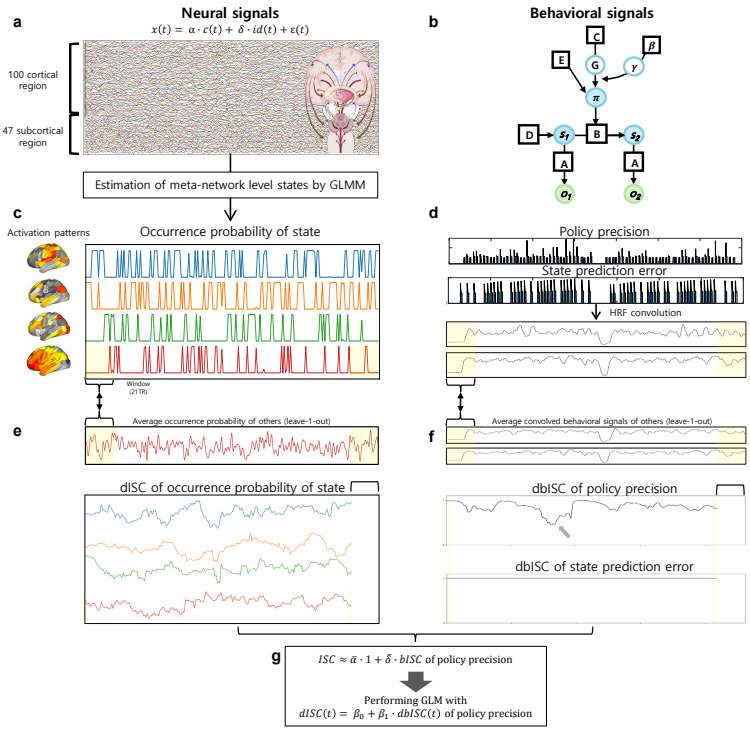


Figure 1. General flow of the experiment for investigating the association of dynamic changes of both brain state and policy precision. **a,** Using functional magnetic resonance imaging and behavioral datasets acquired during performing the two-armed bandit task, preprocessed time-series from 147 brain regions, including brainstem nuclei, were extracted. Additionally, behavioral parameters in the active inference framework (**b**: Graphical representations of state estimation and policy selection (adapted from Fig. 5 of Smith et al., 2022)) were estimated (s = states, o = observations, A = likelihood mapping between states and outcomes, D = initial state priors, B = transition between states, C = prior preferences of outcomes, G = expected free energy, E = prior over policies). **c,** Using the Graphical Laplacian Mixed Model (GLMM), meta-network level brain states and their occurrence probability time-series were estimated. We hypothesized that shared, ct(t), and idiosyncratic, id(t), components are included in the state-level time-series and a region-level time-series. **d,** Both policy precision and state prediction error in each trial were calculated, and both were convolved with the hemodynamic response function. **e, f,** By calculating the inter-subject correlation (ISC) of time-series over the sliding window, neuronal (e) and behavioral (f) dynamic inter-subject correlation (dISC) were determined, respectively. The dISC of state prediction error (SPE) is 1, indicating the shared component across subjects, whereas the dISC of policy precision signals varies in trials (suggesting its correspondence to the subject's idiosyncratic response during performing decision-making task (f). **g,** As the ISC of policy precision suddenly drops in a certain period (indicated by an arrow) and the correlation of idiosyncratic response is too high to be ignored (r = 0.45 - 0.96: Fig. 4a), the ISC of neural signals can be approximated by adding the amount of related idiosyncratic response scaled by its correlation. It can extend into the time-series equation with the dISC time-series to perform a generalized linear model (GLM) analysis. A GLM comparing the dISC of both neural and behavioral signals, policy precision-associated brain states were investigated.

pinpoint the neural correlates of common components in neural responses to naturalistic stimuli, such as movie clips. This synchronization of high-level brain functions is based on the premise that these neural responses are linear combinations of shared elements. These elements include common experimental stimuli, individual-specific idiosyncratic responses triggered by the stimuli, which align with behavioral signals, and inherent noise. This method assumed that the idiosyncratic response and noise are not correlated across subjects. They were averaged out by calculating correlations so that the ISC value reflects the amount of neural activity related to the shared responses. However, when the inter-correlation of idiosyncratic responses across subjects is considerable, the ISC value should consider that the amount of idiosyncratic response is related as much as the ISC (equation in Fig. 1g).

The only interesting signal (policy precision) is included in idiosyncratic responses by making the non-interesting behavioral signal the same across participants. The behavioral signals are estimated from the behavioral data by fitting the active inference model. Subsequently, they are convolved with hemodynamic response function (HRF) for comparison with neural signals (Fig. 1d). The dynamics of ISC (dISC) can also be calculated using a brief sliding window with the time-series of both neural and behavioral signals [23] so that the generalized linear model (GLM) comparing the ISC of the neural and behavioral signals can be applied. In other words, we took an alternative approach, assuming that the changes in the ISC of neural activity recruiting high-level brain function would align with changes in the ISC of the behavioral model signals, such as policy precision that is supposed to represent the brain function during a decision-making task (Fig. 1g).

After determining that the policy precision was a significant signal modulating the dynamics of **brain states**, we investigated the correlations between policy precision and occurrence probability of each state with a typical GLM. Next, to determine the underlying cause of those correlations of the policy precision signal, we calculated the edge connectivity at each time point through **phase coherence** [6]. Using this, we calculated the graph metrics of integration and recruitment [3] to determine the correlation of connectivity patterns at levels lower than those of the meta-network.

Results

Computational model based on active inference framework: precision, positive prediction error (PPE) and negative prediction error (NPE) trajectories

The two-armed bandit task involved two parts: whether participants gained or avoided losing upon winning and whether they avoided gaining or losing the reward upon making an incorrect choice (Fig. 2a). Furthermore, there were two choices, with one having a 70% likelihood of winning and the other having 30%. Each part consisted of 40 trials, totaling 80 trials per participant. Participants had to infer the choice with a better chance of winning over time to maximize the accumulated reward. Therefore, the probability of a better choice was the hidden state value inferred by the agents during the task, and prediction precision corresponded to the predictive state value. A value far from 0.5 reflected more precise predictive precision, and vice versa. The trajectories in Fig. 2b and 2c shows the average state value, with the shaded areas indicating the standard deviation across participants. Although there are some variances, the general trajectories closely follow the average line.

State prediction error (SPE) signals can be positive or negative based on trial outcomes. With the better choice having a 70% likelihood of winning, the NPE signals are expected in 12 trials (40×0.3) of each task part, whereas the PPE signals occur in

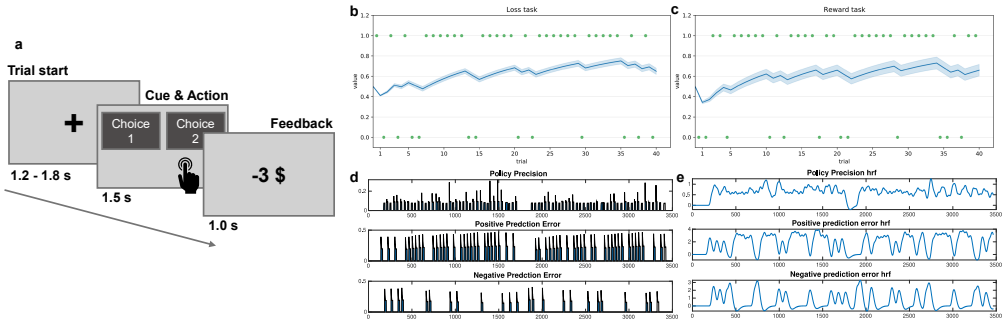


Figure 2. Sequence of the experimental task and state value trajectory and the behavioral signals simulated as neural firing and the time-series of each signal convolved with the hemodynamic response function. **a**, The task begins with gazing at the centered cross for a random duration ranging from 1.2 to 1.8 s; subsequently, the meaningful experimental stimulus is presented. The "cue" indicates the presentation of two options of figures, and it lasts for 1.5 s during which the participant must choose one. After 1.5 s, the results of the choice are presented as monetary reward or loss. We defined this presentation result as "feedback", which lasts 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 done in 1.5 s during which the cue images are presented on the screen. **b-c,** The trajectory of the average inferred state value indicates that the probability of one choice option is better than the other in each trial. The initial values start from 0.5 as there is no information about which option is better. The green dots indicate the observation of each trial after selection (choosing the better option will give win = 1, lose = 0). The differences between the green dot and the blue line of each trial indicate the state prediction error of that trial. d, Example of one participant's simulated neural firing signals generated from each behavioral signal calculated using the active inference model. **e,** The hemodynamic response function convolved time-series in **d**.

the remaining 28 trials. This can be misconstrued as an NPE signal only arising from losses. However, in this task, the prediction error was the difference between the observed probability (always 0 or 1) of one choice being better and the inferred probability updated through the trials. It is not about winning or losing but rather about selecting the better option, which may lead to either outcome. When selecting the better option and losing, or vice versa, the outcome is 0. Conversely, when selecting the better option and winning, or vice versa, the outcome is 1 (Fig. 2b). Therefore, the prediction error signal's sign is determined by task settings rather than the individual choice outcomes. Hence, the feedback of each trial can be categorized based on the outcomes (which vary across participants) and the sign of the prediction errors (which is consistent across participants). Naturally, trials generating NPE signals often coincide with wins, whereas those generating PPE signals are common. However, a large number of participants experienced losses in trials associated with NPE (SI Appendix, Figures S1 and Tables S1). Although winning or losing and the sign of the prediction error signals did not perfectly align, this discrepancy was more noticeable during the first 10 trials (Fig. 5a)

In contrast, the policy precision signals varied across participants based on their trial outcome and performance. The detailed precision updating process is explained in the Methods section. Briefly, the agents calculated the expected free energy (EFE, G) of each policy on the trial t, and the precision (γ) corresponded to the reliability of the agent in a preferred state if they follow the calculated G. The initial value of γ was

calculated from the inverse of fixed prior value of β ($\frac{1}{\beta_0}$) [40]. Thus, the probability that the agent will choose each policy π_0 will be

$$\pi_0 \leftarrow \sigma(-\gamma G), \tag{1}$$

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

$$\pi \leftarrow \sigma(-F - \gamma G), \tag{2}$$

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

$$G_{error} \leftarrow (\pi - \pi_0) \cdot (-G). \tag{3}$$

The above equation reflects the level of (dis)agreement between EFE (G and π_0) and the variational free energy of the observation (F and π) [40]. Thus, the β whose inverse means γ is updated by G_{error}.

$$\begin{aligned} G_{error} &\leftarrow (\pi - \pi_0) \cdot (-G). \\ \beta_{\text{update}} &\leftarrow \beta - \beta_0 + G_{error} \\ \beta &\leftarrow \beta - \beta_{\text{update}}/\psi \\ \gamma &\leftarrow \frac{1}{\beta} \end{aligned} \tag{4}$$

Therefore, policy precision reflects the reliability of the EFE based on new observations representing winning or losing results. Fig. 2d illustrates an individual participant's time course for SPE and policy precision.

Four representative brain states and their dynamics

The graph Laplacian mixture model (GLMM) identifies meta-networks representing intra- and inter-network interactions. The GLMM learns the graph structure of the states by estimating the Laplacian matrices and captures the dynamics of brain states that are consistent with the timing of various task paradigms in the HCP dataset [24, 34]. Four key states were estimated and they are depicted in Fig. 3. Using the mean activation patterns, we grouped the regions as a priori resting-state networks (RSNs) [43] for each state (Fig. 3a-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 network (FPN). 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, whereas it was negatively correlated with the DMN and SOM network. State 4 shows generally low activations of all networks except for the limbic (LIM) network and an anterior-to-posterior gradient.

To examine subcortical activation patterns, we conducted an analysis of variance (ANOVA) comparing the average activation of each subcortical region across states (SI Appendix, Tables S2). States 2 and 3 showed opposite patterns: state 2 showed higher hippocampus and hypothalamus activation than did state 3 (Fig. 3i-l). State 3 showed higher activations in the anterior thalamus and caudate nucleus than did other states. Regarding dopamine regions, only the regions in the left hemisphere, including the SNc, VTA, and habenula, showed significant differences across states, with higher activation in states 2 and 3 than in states 1 and 4. Conversely, the right part of the internal

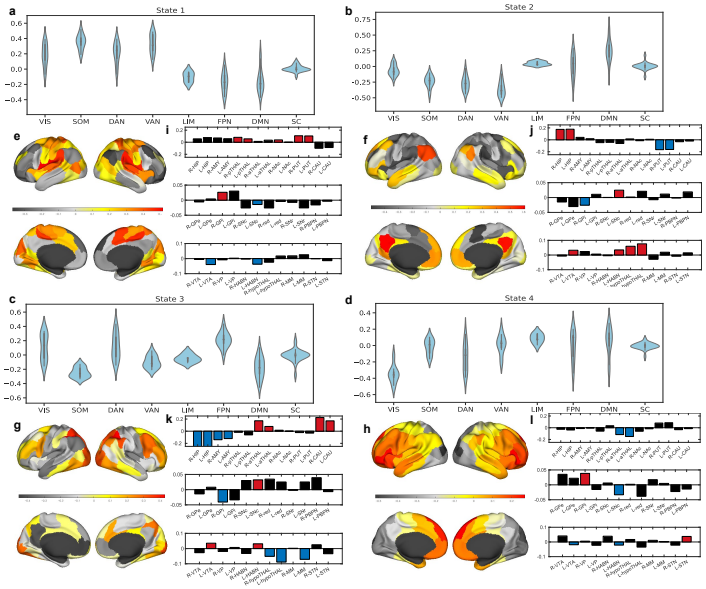


Figure 3. Results of the graph Laplacian mixture model (GLMM) analysis. **a-h,** Mean activation pattern across participants for each brain state estimated from the GLMM. **a-d**, shows the activation values of regions included in each a priori resting state network (RSN). **i-l,** Mean activations in the subcortical region (SC) of each state. Full names and numbers of each region are described in Tables S2. The regions with significant differences compared with the other states in an analysis of variance (ANOVA) ($p < 0.01$) are colored red if their activation is higher than that of others and blue if lower.

segment of the globus pallidus showed significant differences, with higher activation in states 1 and 4 than in states 2 and 3. Fig. 4c displays the average time courses of each state, which we refer to as the original time-series to differentiate it from the dISC. State 1 accounted for the highest proportion (SI Appendix, Figures S2a), whereas the other states appeared fairly evenly throughout the task. The occurrence probability of state 3 increased during the early stages and in the middle of the task. Furthermore, according to the dISC time-series, the ISC of state 3 changed similarly to that of state 2, indicating that the increase in state 3 is accompanied by a decrease in state 2.

Brain state dynamics explained by the policy precision signal

By comparing the dISC time-series of policy precision and occurrence probability of each brain state, we determined whether the inter-subject variability in the policy precision signal can explain the variability in the dynamics of the brain state. Higher dISC values indicate that the occurrence probabilities of the brain state are similar to that of others during that window, and both similarly low and high activation patterns are possible. Therefore, if the dISC of the policy precision signal has a significant positive correlation to the dISC of state occurrences, then we can consider the policy precision as an essential signal that modulates the changes in the neural states. The direction of this modulation can be confirmed by the result of the original time-series. Therefore, we conducted a GLM analysis using dISC and original time-series with Equations [5] and [6] respectively. Analysis of the original time-series was conducted with the HRF-convolved signals [17] of the following task stimulus variables as regressors: cue visual stimulus (CVS), feedback visual stimulus of win (FVS_w), feedback visual stimulus of lose (FVS_l), action selection (AS). Given the nature of the occurrence of

the state, logistic regression was chosen as the appropriate modeling technique.

$$\text{dISC of state k} = \beta_0 + \text{dISC of Precision} \cdot \beta_1 \tag{5}$$

$$\text{Occurrence of state k} = \sigma(\beta_0 + \text{CVS} \cdot \beta_1 + \text{FVS_w} \cdot \beta_2 + \text{FVS_l} \cdot \beta_3 + \text{AS} \cdot \beta_4 + \text{Precision} \cdot \beta_5) \tag{6}$$

Subsequently, a secondary analysis was employed, through a non-parametric one-sample test comparing beta coefficients with those of a null distribution generated from 1,000 randomly shifted time-series. All states dynamics were related to the policy precision signal changes (Fig. 4d) and its effects were positive on states 1 and 2 (p-value $= 0.002$, < 0.0001, respectively) and negative on state 3 (p-value < 0.0001) (Fig. 4e), suggesting the involvement of state 3 when policy precision is low (SI Appendix, Tables S3).

State 3 is inversely associated with policy precision with uncertain feedback

The occurrence probability of state 3 was high when policy precision was low. More specifically, we investigated the dynamics of state 3 at the feedback presentation, the outcome of which, win or lose, will determine the ensuing policy precision signal. This inverse relation between state 3 and policy precision is attributable to the increase of state 3 with uncertain feedback, irrespective of simply the outcome or sign of the prediction error.

Although we used the FVS separated by the win or lose outcome in Equation [6] and found a positive correlation between state 3 and both types of FVS (SI Appendix, Tables S3), FVS can also be categorized by the sign of the SPE signal, as the prediction error signal is assumed to align with the timing of the feedback presentations. Therefore, to determine whether the effect of the feedback on state occurrences depends on the outcome or prediction error, we categorized trials into four feedback types based on the prediction error's sign [P or N in Equation [7]] and outcome [W or L in Equation [7]] and conducted a GLM analysis. Next, a two-way ANOVA, including two factors of outcome and sign of the prediction error using beta coefficients of each type of feedback, was conducted.

$$\text{Occurrence of state k} = \sigma(\beta_0 + \text{CVS} \cdot \beta_1 + \text{P_W} \cdot \beta_2 + \text{P_L} \cdot \beta_3 + \text{N_W} \cdot \beta_4 + \text{N_L} \cdot \beta_5 + \text{AS} \cdot \beta_6 + \text{Precision} \cdot \beta_7) \tag{7}$$

The interaction between the prediction error and outcome significantly affected the occurrences of all states (SI Appendix, Tables S5). In post-hoc analysis, the occurrence of state 3 was greater in P_L and N_W than in P_W and N_L (SI Appendix, Tables S5). Based on these results, we speculated that state 3 is more likely to occur in cases where the signs of the outcome and prediction error are opposite (P_L and N_W) rather than simply being due to the sign of an outcome or prediction error. Therefore, we categorized the feedback type as "congruent" when a winning outcome was achieved in the PPE trial or a losing outcome in the NPE trial and as "incongruent" when a losing outcome was achieved in the PPE trial or a winning outcome in the NPE trial. As mentioned, incongruent outcomes are less frequent than congruent (SI Appendix, Tables S1). Then, what makes the incongruency correlated with the occurrence of state 3? As shown in Figure 5a, most of the incongruent outcomes occur during the early phase of the tasks when the agent is not confident regarding which choice is better. In the middle of the task, such as trials 16 and 24 of the loss task and trials 23, 37, and 39 of the reward task, approximately 24% of the participants were wrong as they changed the choice based on the results from the few trials before, including consecutive NPE results (SI Appendix, Figures S1). Even for participants who won in those trials, consecutive

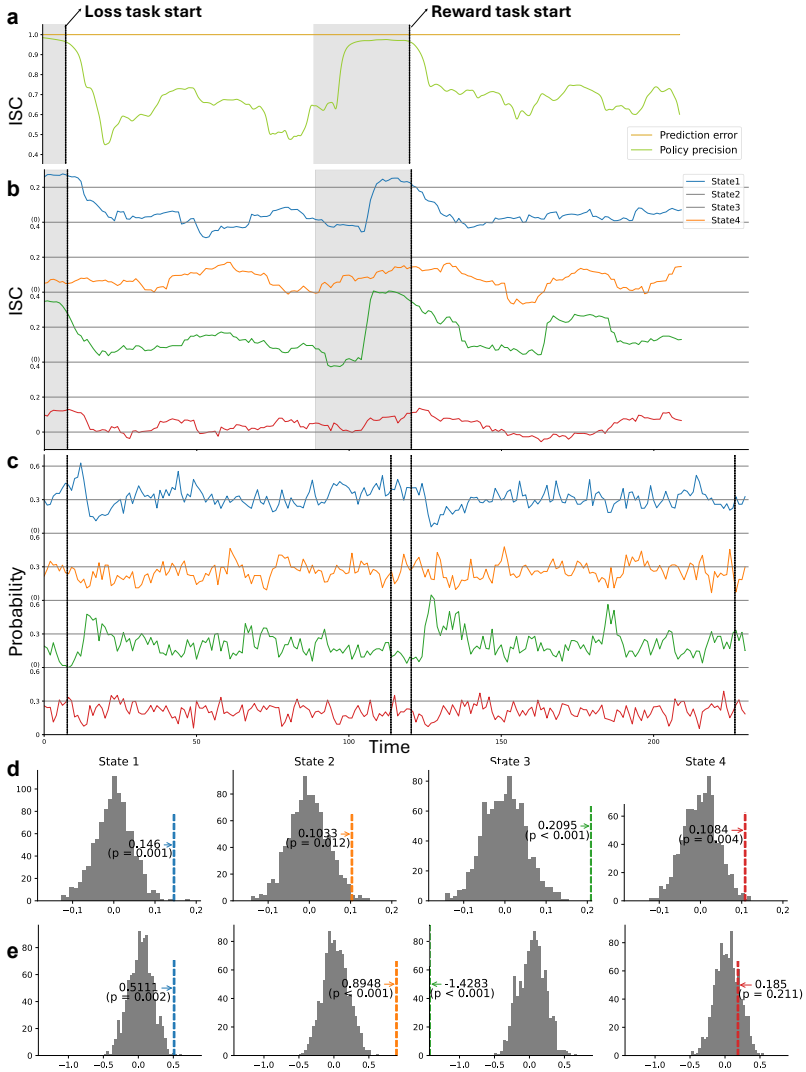


Figure 4. Average time-series of dynamic inter-subject correlation (dISC) of the behavioral signals, state occurrence probability, and the results of permutation test of policy precision signals. **a,** dISC time-series of the behavioral signal, policy precision, and prediction error. The inter-subject correlation (ISC) of policy precision varied across trials, whereas the prediction error was fixed at a value of 1 as it was consistent across participants. The gray color in the background indicates the length of the window between the before and the start of each task (21 volumes, 31.5 s). The start point of each task is indicated using dotted lines. **b,** dISC time-series of occurrence probability of each brain state. **c,** Original time-series of occurrence probability of each brain state estimated using the graph Laplacian mixture model (GLMM). **d, e** The results of one-sample tests with the beta coefficient of the precision signal from a general linear model (GLM) comparing the dISC and original time-series, respectively, are shown.

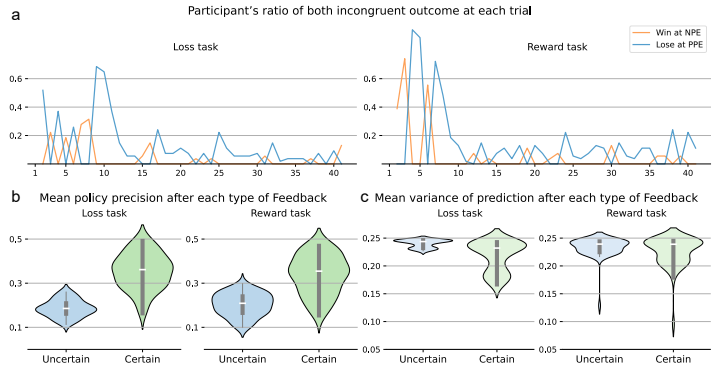


Figure 5. Average number of each type of feedback and the association with precision. **a,** the ratio of participants who win in the NPE trial or lose in the PPE trial. The y-axis indicates the ratio over the 54 participants. **b,** The average policy precision values are calculated by extracting the peak value of policy precision during the next trial of each uncertain and certain feedback as the policy precision values reflect the result from the outcome of the trial before. **c,** The average variance of prediction, which is the inverse of the prediction precision of each type of feedback. Both average policy and prediction precisions are lower after uncertain feedback (blue) than after certain feedback (green).

results that were against their belief regarding the better choice must have caused uncertainty; however, the degree was lesser than for those who changed their choice. Therefore, we defined "uncertain" feedback for the previous trial of trials that caused more than 20% of participants to change their choice (incongruent outcome) and "certain" feedback for the remaining trials. After the first 10 trials, trials 15 and 23 were included in the uncertain feedback for the loss task and trials 22, 36, and 38 for the reward task. We conducted a GLM analysis with Equation [8] and the same non-parametric one-sample test, and the result showed that both certain and uncertain feedback were positively correlated with state 3 (p-value = 0.003, < 0.0001, respectively) (SI Appendix, Tables S6).

$$\text{Occurrence of state k} = \sigma(\beta_0 + \text{CVS}\cdot\beta_1 + \text{Certain_F}\cdot\beta_2 + \text{Uncertain_F}\cdot\beta_3 + \text{AS}\cdot\beta_4 + \text{Precision}\cdot\beta_5) \tag{8}$$

We conducted a paired t-test comparing beta coefficient values of certain and uncertain feedback, and the beta coefficient of uncertain feedback was significantly greater than that of certain feedback (p-value = 0.0002).

We then calculated the average policy precision value across trials after both types of feedback, and found that it was significantly higher after certain feedback than after uncertain feedback (Fig. 5b). We also estimated the variance of state prediction using the Hierarchical Gaussian Filter model [25], where the first level variance represents the inverse of prediction precision. Subsequently, we calculated the average variance across trials of each feedback type and found that it was significantly lower in certain feedback than in uncertain feedback (Fig. 5c).

Integration and recruitment among functional networks with uncertain feedback

As a result of the state-level analysis, we found that state 3 was related to uncertainty and involved processing uncertain states when the precision of belief was low or observations that ran counter to current beliefs were encountered. State 3 subsequently leads to low policy precision, making it negatively correlated with state 3 and positively correlated with state 2. To determine how state 3 is activated by uncertainty, we

examined the connectivity changes between regions modulated by uncertain feedback that caused the incongruent results. First, we examined the connectivity characteristics by calculating integration and recruitment among seven RSNs using adjacency matrices estimated by GLMM (Fig. 6a). Integration corresponds to the interconnectivity between different network systems, whereas recruitment reflects the intraconnectivity of each network [3]. In state 3, integrations, including either the FPN or LIM network, were the highest (upward arrows in Fig. 6a), whereas the recruitments of all networks were the lowest (downward arrows in Fig. 6a). As the adjacency matrices estimated from GLMM were not continuous values with time, the connectivity values were calculated using the leading eigenvector of the BOLD phase coherence matrices [6] for by performing GLM analysis. The integration and recruitment of each RSNs were calculated using connectivity matrices to reduce the dimensions. We performed a GLM analysis with Equation [9] to investigate the effect of both types of feedback.

$$\text{Integration and recruitment} = \beta_0 + \text{CVS}\cdot\beta_1 + \text{Certain_F}\cdot\beta_2 + \text{Uncertain_F}\cdot\beta_3 + \text{AS}\cdot\beta_4 + \text{Precision}\cdot\beta_5 \tag{9}$$

Discussion

The findings of this study can be summarized as follows: First, policy precision is an important signal to explain the meta-network level changes in the brain. Second, the uncertainty of prediction (prediction precision) increases the integration between the DMN and the other networks, including the DAN and FPN, which may relate to occurrences of state 3, where the DAN and FPN are dominant. Lastly, when participants change their action selections by observations that are against their belief states, integration between the DMN and VAN increases. Subsequently, integrations between the VAN and LIM network as well as the DAN and SOM network increase and the recruitment of the FPN decreases until the changed action is taken.

The dISC time-series of the occurrence probability of state 3 was higher when the occurrence probability of state 3 increased. The original time-series analysis showed that, when there was insufficient evidence (at the beginning of each task) or feedback that did not meet expectations (causing incongruent outcome in the subsequent trial by changing one's choice), that is, uncertain feedback, state 3 was primarily active. In state 3, the DAN and FPN are predominantly activated, and these networks are involved in executive function and goal-oriented, cognitively demanding tasks [26, 46]. Based on our results, we speculate that the core functions of the DAN and FPN are related to exerting attentional efforts in an uncertain state where precision cannot be sufficiently high. To determine how state 3 is activated according to uncertainty, we investigated the connectivity between brain regions that is supposed to lead to the activation of meta-network states. Most of the integration, including the DMN, was higher with uncertain than with certain feedback.

In addition, state 2, where the DMN was dominantly activated, and integrations, including the FPN, especially with the DMN, were the lowest, was significantly positively correlated with the policy precision signal. Although the DMN is generally regarded as a resting network activated during rest, mind-wandering, and self-referential tasks [9, 10, 32], enhanced suggestions explaining the function of the DMN focusing on intrinsic information rather than just resting or doing nothing, have been raised [5, 30, 33]. It has been suggested that the DMN detects the associative relevance of internal and external stimuli and enables value coding [35]. From this viewpoint, an alternative proposal argues that understanding the self-referential function of the DMN entails linking it to the choice of the best actions within the environmental context, necessitating recursive exploration. [22]. Following this suggestion, there is a concept

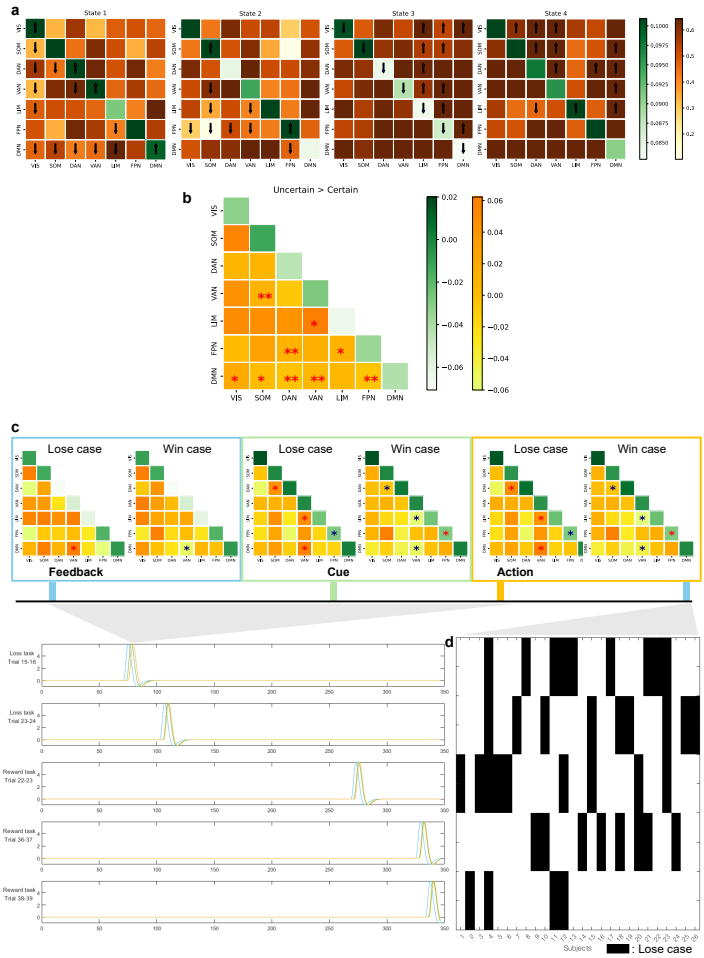


Figure 6. Integration and recruitment of networks with each type of feedback. **a,** shows the integration and recruitment of each state. Upward arrows marked in the upper triangles of matrices indicate that those integration or recruitment were highest among the states, whereas downward arrows marked in the lower triangles of matrices indicate that those were the lowest values among the states. **b, c,** show the average beta coefficient values of each feedback for both integration and recruitment across participants. The diagonal values (green) indicate the recruitment of each network, reflecting the intraconnections within the network, and the other values (yellow) indicate the integration between two networks, reflecting interconnections of networks. * indicates significantly higher (red) or lower (navy) values than the other feedback (p-value $<$ 0.01 and ** indicates significance after Bonferroni correction). **d,** shows the results of repeated-measures ANOVA investigating the differences between the case when changing one's selection so that one lost on the subsequent trial (black) and when keeping the selection and winning on the subsequent trial (white). Out of 130 cases (5 times per 26 participants), there were 39 lose and 91 win cases.

outlining how the DMN functions at elevated levels of inferential hierarchy, which diminishes the free energy associated with lower levels [7]. In addition to the DMN, the hippocampus was activated in state 2 (Fig. 3j). Recent studies examining the function of the hippocampus focusing on the internal model [27, 45] support the notion that the role of DMN is linked to the internal model. Presentation of cue also showed positive correlations with this state, although the significance was lower than that for policy precision. This result suggests that when a cue is provided, activating state 2 corresponds to making a suitable choice based on one's accumulated internal beliefs. In other words, the correlation between policy precision and state 2 is related to how much the participant's actions in making choices are based on their internal beliefs. As uncertainty rose from the outcome of the previous trial, leading to the activation of state 3, this hindered the activation of state 2. Additionally, it resulted in reduced policy precision, which could create a stronger positive correlation between policy precision and the probability of state 2, while also establishing a negative correlation with state 3. Previous studies have focused on the anti-correlated relationship between the FPN and DMN; their interplay has recently received attention [8]. Our findings support both aspects: the former by the anti-correlations of states 2 and 3, alternating according to the uncertainty, and the latter by the determining that the increasing integration between the DMN, FPN, and DAN in response to uncertain feedback and the integration of DMN with other networks in state 3 was highest, despite the low activation level. Based on this interaction, our results support the hypothesis that information transmission from the DMN, where internal information is processed, is required for the proper functioning of the DAN and FPN.

The roles of the VAN and DAN in attention processes have been well investigated. Goal-directed, sustained attention is associated with the DAN, whereas transient attention to salient or surprising stimuli is linked to the VAN [36]. Additionally, inhibitory connections from the DAN to the VAN have been studied in relation to suppressing responses to meaningless or low-information stimuli, which is crucial for maintaining sustained attention [49]. The DAN is thus associated with more internally directed cognitive processes, whereas the VAN is linked to externally directed cognitive functions [50]. The higher integration between the VAN and both the DMN and LIM network in the case of changing choice by the external stimuli contradicting one's beliefs is consistent with these previous findings. This suggests that heightened immediate attention to external stimuli results in choices that oppose internal goals. Although this might cause the wrong choice, as seen herein, these integrations may facilitate flexible decision-making in unstable situations, thereby reducing the rigidity of learning processes and required integration that normally increases in uncertain feedback.

Although the simplicity of the experimental task is a limitation of the current study, it was an essential element for using the ISC method to identify policy precision as a suitable behavioral signal, which is the core objective of this study. However, this study confirmed that the policy precision signal is a validated signal, and based on this, future studies investigating policy precision can be developed to include more complex behavioral tasks, such as those with three or more choices or added volatility. In addition, we used GLMM to obtain representative brain states; various methods are available to estimate brain state, such as the hidden Markov model [47]. In the current study, we selected the GLMM because it is suitable for extracting graph structures from the task-fMRI data. However, different methods can be applied to estimate brain state dynamics depending on the research objectives in further studies.

Materials and Methods

Participants

Fifty-four volunteers participated in this study (15 female and 39 male; mean [standard deviation] age was 22.4 [3.2] years; all individuals were right-handed and had no history of psychiatric or neurological disorders). All participants provided written informed consent for participating in all the procedures and data usage before the start of the study. The experimental procedures were approved by the Institutional Review Board at KAIST and conducted in accordance with the Declaration of Helsinki. Informed consent was obtained from all study participants.

fMRI Data Acquisition and Analysis

All participants underwent MRI using a 3.0-T MRI scanner (Siemens Verio Syngo). 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 iso-voxel) and T2-weighted (T2w; TE = 330 ms, TR = 2200 ms, FOV = 224 × 224 mm, and 0.7 mm iso-voxel) structural images were acquired. 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 iso-voxel with no gap, 64 slices, phase encoding = anterior to posterior) were performed using a 32-channel 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 and pressed an MRI-compatible button to indicate their decision. In the pretraining session, prior to the first scan, the participants were acquainted with the apparatus and tasks were performed during scanning. The sequences executed during this pretraining session were not encountered later in the experiment. For details about the fMRI data preprocessing, please see SI Appendix, Supplementary Methods.

After preprocessing, we acquired the averaged region of interest (ROI) time-series extracted from the regions from the Schaefer Atlas [37], including 100 regions covering the cortex, and the Melbourne Subcortex Atlas [44] covering the subcortex. In addition, to include other subcortical regions related to important neurotransmitters, including dopamine and norepinephrine, we added some regions from the CIT168 atlas [29] and Harvard Ascending Arousal Network Atlas [13] to obtain signals from a total of 147 ROIs (SI Appendix, Tables S9).

To investigate the dynamic changes in different levels of neural interactions, we obtained neural signals at the network and edge levels. At the edge level, we used phase coherence to calculate the connectivity between two regions. This method is adept at capturing temporally delayed relationships and exhibits high sensitivity in detecting ultraslow oscillatory processes that underlie the formation of RSNs, a phenomenon corroborated by recent experimental and computational research [12, 31]. In our study, we used a recently developed data-driven approach named Leading Eigenvector Dynamics Analysis that can reduce the dimensionality by considering only the relative phase of BOLD signals and capturing instantaneous phase-locking (PL) patterns [6, 48]. However, in this study, we did not employ k-means clustering to identify representative states based on the leading eigenvector. Instead, we used the outer product of the leading eigenvectors of the PL patterns to reconstruct the connectivity patterns at each time point. K-means clustering was not used as it does not consider the graph-based connectivity of brain structures that generate signals when distinguishing clusters.

Therefore, to extract the representation of the overall brain state corresponding to network-level interactions, we employed the GLMM, an approach that involves directly estimating the graph structure that generates the signals, which serves as an alternative

to k-means clustering for capturing the representation of the entire brain state [34]. This method extracts the representative network states described in three aspects, two of which represent the whole population and one of which could be interpreted as subject-specific. The former is the mean activation pattern of each node and the graph Laplacians of the connectivity patterns (Fig. 3e-h, SI Appendix, Figures S3), whereas the latter is the probability of each state being assigned to each time point. Using the probabilities of each state, we calculated dynamic functional connectivity metrics, including the fractional occupancy, dwell time, and transition probability of each participant (SI Appendix, Figures S2).

1 acknowledgments

This study was supported by the Medical Scientist Training Program from the Ministry of Science & ICT of Korea, and the Brain Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science & ICT (NRF-2022M3E5E8081200 and NRF-2016M3C7A1914448).

2 Data Availability

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

3 Supporting Information

Supplementary Methods, Figures S1–S5, and Tables S1–S10 are presented in Supplementary file.

Figures S1

Figures S2

Figures S3

Figures S4

Figures S5

Tables S1

Tables S2

Tables S3

Tables S4

Tables S5

Tables S6

Tables S7

Tables S8

Tables S9

Tables S10

4 Competing Interests

The authors declare no competing interests.

5 Technical Terms

Bayesian brain Theory that considers the brain as a statistical organ of inference predicting current and future events based on experience.

prior belief The probability distribution or density of the causes of data that encode beliefs about those causes prior to observing the data.

posterior belief The conditional probability results from updating the prior belief with information.

Precision In statistics, the inverse variance or precision matrix quantifies certainty about variables and is also called the concentration matrix.

Free energy An information theory measure that provides an upper bound on the surprise when sampling data from a generative model.

Policy A sequence of actions, guiding its decision-making process to achieve specific objectives.

computational model An explanation of a complex system's operation using an algorithmic or mechanistic approach.

dynamic functional connectivity Changes of functional connectivity and statistical dependencies between different brain areas over a short time.

two-armed bandit task A classic reinforcement learning task that models decision-making by choosing the better option between two choices.

prediction error The difference between the observed outcomes and predicted value.

Brain state The evolving pattern of self-organized neural activity specific to a condition.

Phase coherence Cosine difference of phases between each pair of brain areas.

References

1. R. A. Adams, M. Moutoussis, M. M. Nour, T. Dahoun, D. Lewis, B. Illingworth, M. Veronese, C. Mathys, L. De Boer, M. Guitart-Masip, K. J. Friston, O. D. Howes, and J. P. Roiser. Variability in Action Selection Relates to Striatal Dopamine 2/3 Receptor Availability in Humans: A PET Neuroimaging Study Using Reinforcement Learning and Active Inference Models. *Cerebral Cortex*, 30(6), 2020.

2. D. S. Bassett and O. Sporns. Network neuroscience, 2 2017.

3. D. S. Bassett, M. Yang, N. F. Wymbs, and S. T. Grafton. Learning-induced autonomy of sensorimotor systems. *Nature Neuroscience*, 18(5):744–751, 2015.

4. J. D. Berke. What does dopamine mean?, 2018.

5. R. L. Buckner and D. C. Carroll. Self-projection and the brain. *Trends in Cognitive Sciences*, 11(2), 2007.

6. J. Cabral, D. Vidaurre, P. Marques, R. Magalhães, P. Silva Moreira, J. Miguel Soares, G. Deco, N. Sousa, and M. L. Kringelbach. Cognitive performance in healthy older adults relates to spontaneous switching between states of functional connectivity during rest. *Scientific Reports*, 7(1), 12 2017.

7. R. L. Carhart-Harris and K. J. Friston. The default-mode, ego-functions and free-energy: A neurobiological account of Freudian ideas, 2010.

8. M. Chen, Y. He, L. Hao, J. Xu, T. Tian, S. Peng, G. Zhao, J. Lu, Y. Zhao, H. Zhao, M. Jiang, J. H. Gao, S. Tan, Y. He, C. Liu, S. Tao, L. Q. Uddin, Q. Dong, and S. Qin. Default mode network scaffolds immature frontoparietal network in cognitive development. *Cerebral Cortex*, 33(9), 2023.

9. J. S. Damoiseaux, S. A. Rombouts, F. Barkhof, P. Scheltens, C. J. Stam, S. M. Smith, and C. F. Beckmann. Consistent resting-state networks across healthy subjects. *Proceedings of the National Academy of Sciences of the United States of America*, 103(37), 2006.

10. M. Dastjerdi, B. L. Foster, S. Nasrullah, A. M. Rauschecker, R. F. Dougherty, J. D. Townsend, C. Chang, M. D. Greicius, V. Menon, D. P. Kennedy, and J. Parvizi. Differential electrophysiological response during rest, self-referential, and non-self-referential tasks in human posteromedial cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 108(7), 2011.

11. P. Dayan and L. F. Abbott. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. 2001.

12. G. Deco and M. Kringelbach. Correction to Metastability and Coherence: Extending the Communication through Coherence Hypothesis Using a Whole-Brain Computational Perspective [Trends in Neurosciences, 39, 3, (2016), 125-135], doi: 10.1016/j.tins.2016.01.001, 2016.

13. B. L. Edlow, E. Takahashi, O. Wu, T. Benner, G. Dai, L. Bu, P. E. Grant, D. M. Greer, S. M. Greenberg, H. C. Kinney, and R. D. Folkerth. Neuroanatomic connectivity of the human ascending arousal system critical to consciousness and its disorders. *Journal of Neuropathology and Experimental Neurology*, 71(6), 2012.

14. K. Friston. The free-energy principle: A unified brain theory?, 2010.

15. K. Friston, T. FitzGerald, F. Rigoli, P. Schwartenbeck, and G. Pezzulo. Active inference: A process theory. *Neural Computation*, 29(1):1–49, 1 2017.

16. K. Friston, P. Schwartenbeck, T. FitzGerald, M. Moutoussis, T. Behrens, and R. J. Dolan. The anatomy of choice: Active inference and agency. *Frontiers in Human Neuroscience*, (SEP), 2013.

17. K. J. Friston, P. Fletcher, O. Josephs, A. Holmes, M. D. Rugg, and R. Turner. Event-related fMRI: Characterizing differential responses. *NeuroImage*, 7(1), 1998.

18. J. M. Galea, S. Bestmann, M. Beigi, M. Jahanshahi, and J. C. Rothwell. Action reprogramming in Parkinson's disease: Response to prediction error is modulated by levels of dopamine. *Journal of Neuroscience*, 32(2), 2012.

19. A. A. Hamid, J. R. Pettibone, O. S. Mabrouk, V. L. Hetrick, R. Schmidt, C. M. Vander Weele, R. T. Kennedy, B. J. Aragona, and J. D. Berke. Mesolimbic dopamine signals the value of work. *Nature Neuroscience*, 19(1), 2015.

20. U. Hasson, Y. Nir, I. Levy, G. Fuhrmann, and R. Malach. Intersubject Synchronization of Cortical Activity during Natural Vision. *Science*, 303(5664), 2004.

21. S. Lammel, B. K. Lim, and R. C. Malenka. Reward and aversion in a heterogeneous midbrain dopamine system, 2014.

22. K. Lehmann, D. Bolis, K. J. Friston, L. Schilbach, M. J. Ramstead, and P. Kanske. An Active-Inference Approach to Second-Person Neuroscience. *Perspectives on Psychological Science*, 2023.

23. G. Majumdar, F. Yazin, A. Banerjee, and D. Roy. Emotion dynamics as hierarchical Bayesian inference in time. *Cerebral Cortex*, 33(7), 2023.

24. H. P. Maretic and P. Frossard. Graph Laplacian Mixture Model. *IEEE Transactions on Signal and Information Processing over Networks*, 6:261–270, 2020.

25. C. D. Mathys, E. I. Lomakina, J. Daunizeau, S. Iglesias, K. H. Brodersen, K. J. Friston, and K. E. Stephan. Uncertainty in perception and the Hierarchical Gaussian filter. *Frontiers in Human Neuroscience*, 8, 2014.

26. V. Menon. Large-scale brain networks and psychopathology: A unifying triple network model, 2011.

27. K. J. Miller, M. M. Botvinick, and C. D. Brody. Dorsal hippocampus contributes to model-based planning. *Nature Neuroscience*, 20(9):1269–1276, 9 2017.

28. S. A. Nastase, V. Gazzola, U. Hasson, and C. Keysers. Measuring shared responses across subjects using intersubject correlation. *Social Cognitive and Affective Neuroscience*, 14(6), 2019.

29. W. M. Pauli, A. N. Nili, and J. Michael Tyszka. Data Descriptor: A high-resolution probabilistic in vivo atlas of human subcortical brain nuclei. *Scientific Data*, 5, 2018.

30. J. M. Pearson, S. R. Heilbronner, D. L. Barack, B. Y. Hayden, and M. L. Platt. Posterior cingulate cortex: Adapting behavior to a changing world, 2011.

31. A. Ponce-Alvarez, G. Deco, P. Hagmann, G. L. Romani, D. Mantini, and M. Corbetta. Resting-State Temporal Synchronization Networks Emerge from Connectivity Topology and Heterogeneity. *PLoS Computational Biology*, 11(2), 2015.

32. J. D. Power, A. L. Cohen, S. M. Nelson, G. S. Wig, K. A. Barnes, J. A. Church, A. C. Vogel, T. O. Laumann, F. M. Miezin, B. L. Schlaggar, and S. E. Petersen. Functional Network Organization of the Human Brain. *Neuron*, 72(4), 2011.

33. M. E. Raichle and M. E. Raichle. Searching for a baseline: Functional imaging and the resting human brain. *Nature Reviews Neuroscience*, 2(10), 2001.

34. I. Ricchi, A. Tarun, H. P. Maretic, P. Frossard, and D. Van De Ville. Dynamics of functional network organization through graph mixture learning. *NeuroImage*, 252, 2022.

35. M. Roy, D. Shohamy, and T. D. Wager. Ventromedial prefrontal-subcortical systems and the generation of affective meaning, 2012.

36. S. Sadaghiani and M. D'Esposito. Functional characterization of the cingulo-opercular network in the maintenance of tonic alertness. *Cerebral Cortex*, 25(9), 2015.

37. A. Schaefer, R. Kong, E. M. Gordon, T. O. Laumann, X.-N. Zuo, A. J. Holmes, S. B. Eickhoff, and B. T. T. Yeo. Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cerebral Cortex*, 28(9):3095–3114, 9 2018.

38. W. Schultz. Dopamine reward prediction error coding. *Dialogues in Clinical Neuroscience*, 18(1), 2016.

39. P. Schwartenbeck, T. H. FitzGerald, C. Mathys, R. Dolan, and K. Friston. The dopaminergic midbrain encodes the expected certainty about desired outcomes. *Cerebral Cortex*, 25(10), 2015.

40. R. Smith, K. J. Friston, and C. J. Whyte. A step-by-step tutorial on active inference and its application to empirical data. *Journal of Mathematical Psychology*, 107, 2022.

41. S. M. Smith, K. L. Miller, G. Salimi-Khorshidi, M. Webster, C. F. Beckmann, T. E. Nichols, J. D. Ramsey, and M. W. Woolrich. Network modelling methods for FMRI. *NeuroImage*, 54(2):875–891, 1 2011.

42. O. Sporns, J. Faskowitz, A. S. Teixeira, S. A. Cutts, and R. F. Betzel. Dynamic expression of brain functional systems disclosed by fine-scale analysis of edge time series. *Network Neuroscience*, 5(2):405–433, 5 2021.

43. B. T. Thomas Yeo, F. M. Krienen, J. Sepulcre, M. R. Sabuncu, D. Lashkari, M. Hollinshead, J. L. Roffman, J. W. Smoller, L. Zöllei, J. R. Polimeni, B. Fisch, H. Liu, and R. L. Buckner. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, 106(3):1125–1165, 9 2011.

44. Y. Tian, D. S. Margulies, M. Breakspear, and A. Zalesky. Topographic organization of the human subcortex unveiled with functional connectivity gradients. *Nature Neuroscience*, 23(11), 2020.

45. V. S. Turner, R. O. O'Sullivan, and M. A. Kheirbek. Linking external stimuli with internal drives: A role for the ventral hippocampus, 10 2022.

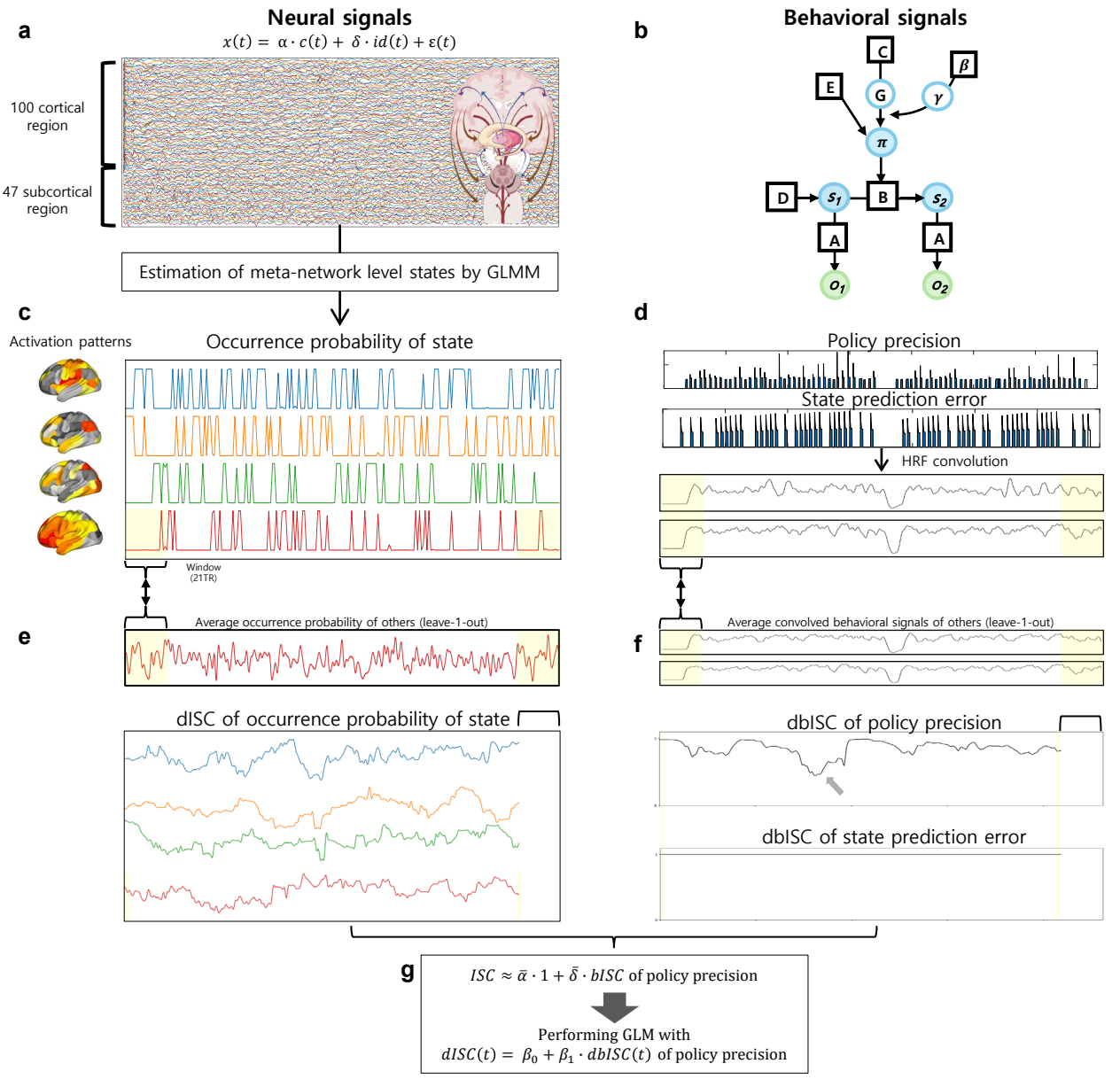
46. L. Q. Uddin, B. T. Yeo, and R. N. Spreng. Towards a Universal Taxonomy of Macro-scale Functional Human Brain Networks, 2019.

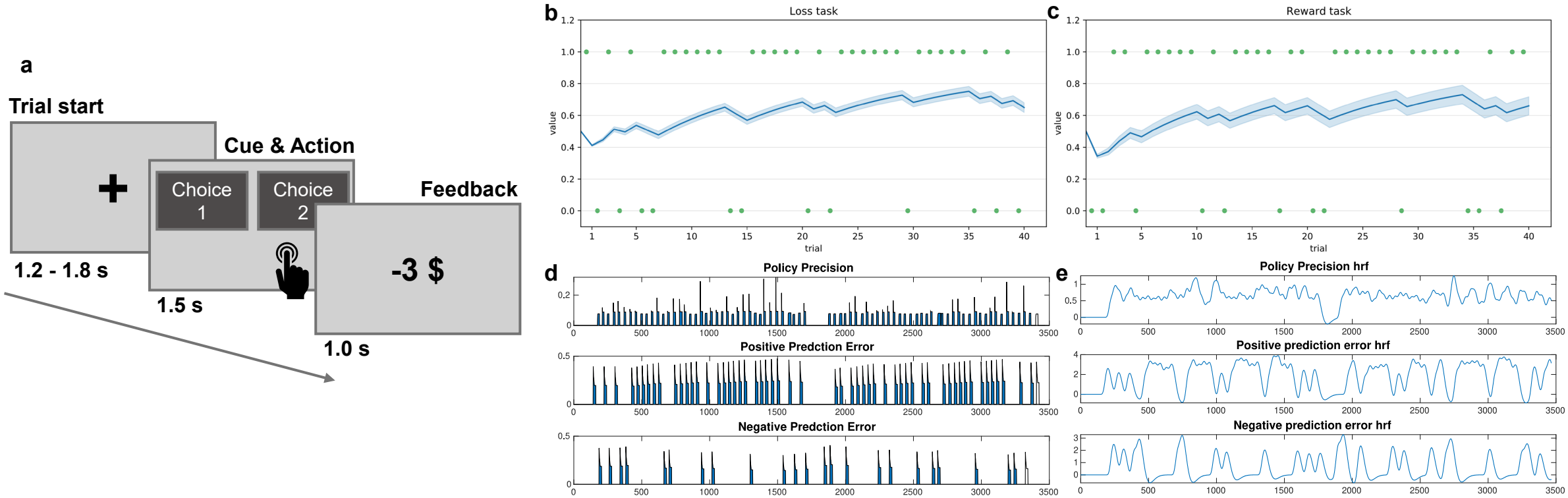
47. D. Vidaurre, R. Abeysuriya, R. Becker, A. J. Quinn, F. Alfaro-Almagro, S. M. Smith, and M. W. Woolrich. Discovering dynamic brain networks from big data in rest and task, 2018.

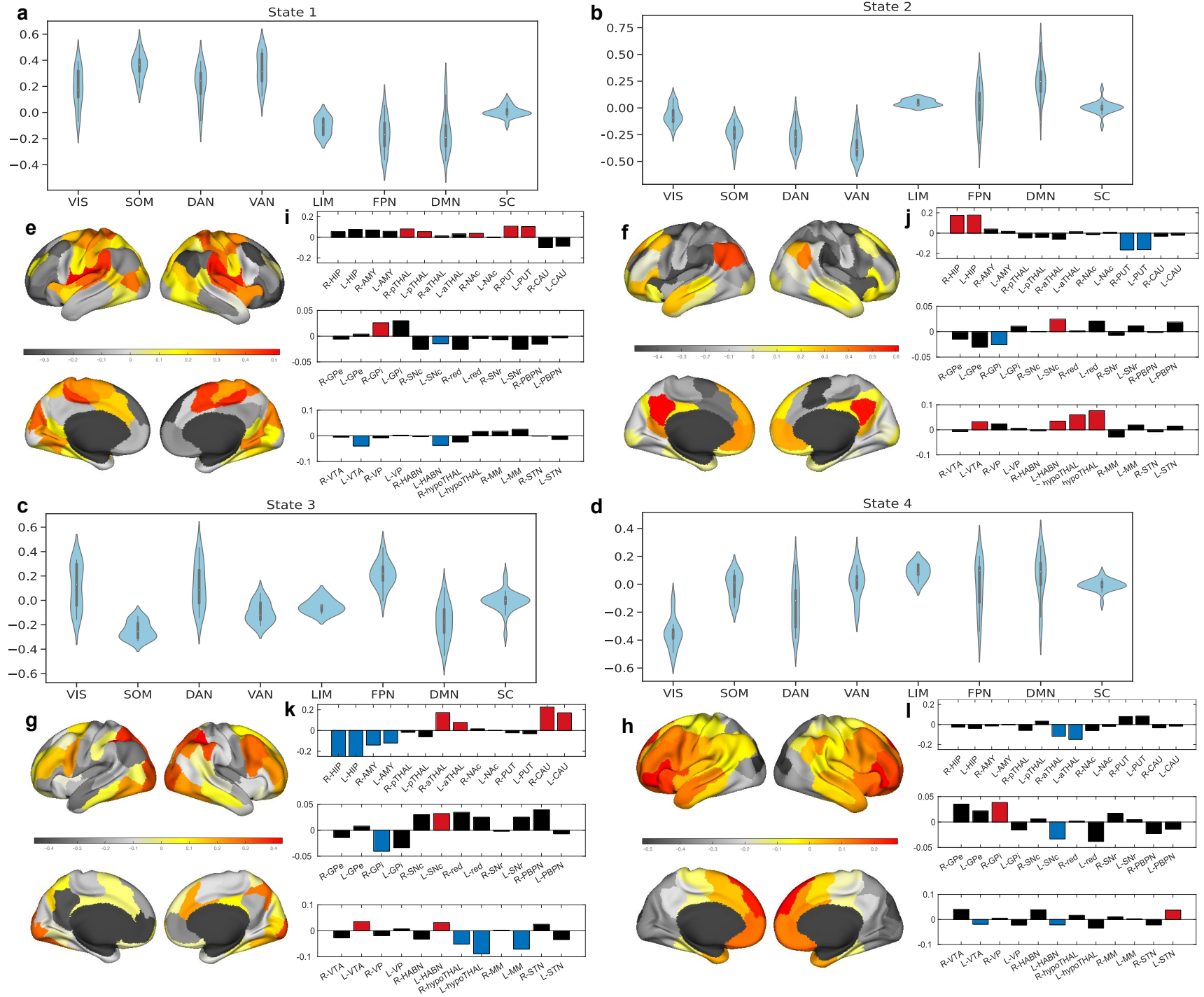
48. J. Vohryzek, G. Deco, B. Cessac, M. L. Kringelbach, and J. Cabral. Ghost Attractors in Spontaneous Brain Activity: Recurrent Excursions Into Functionally-Relevant BOLD Phase-Locking States. *Frontiers in Systems Neuroscience*, 14, 4 2020.

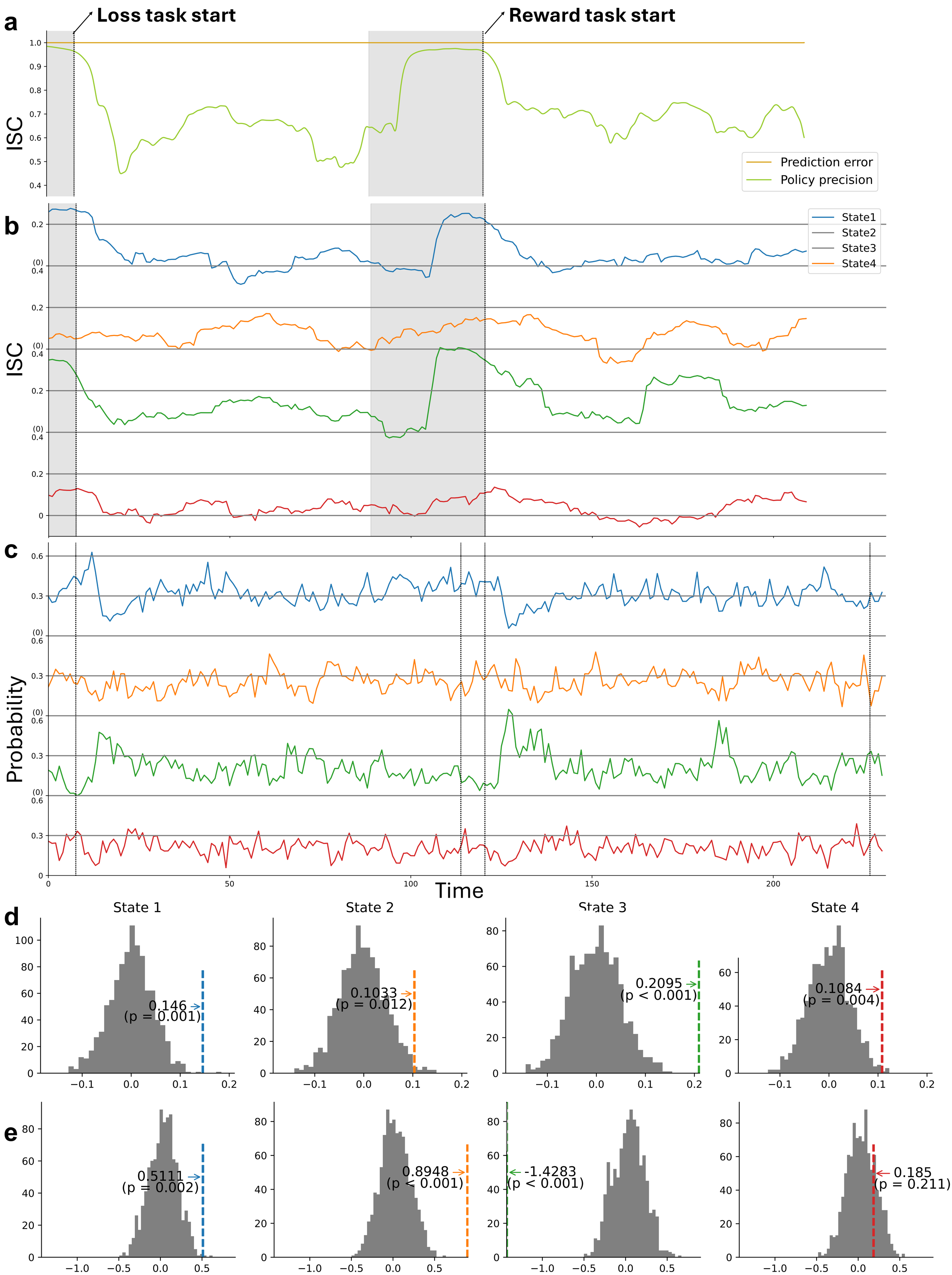
49. S. Vossel, J. J. Geng, and G. R. Fink. Dorsal and ventral attention systems: Distinct neural circuits but collaborative roles. *Neuroscientist*, 20(2), 2014.

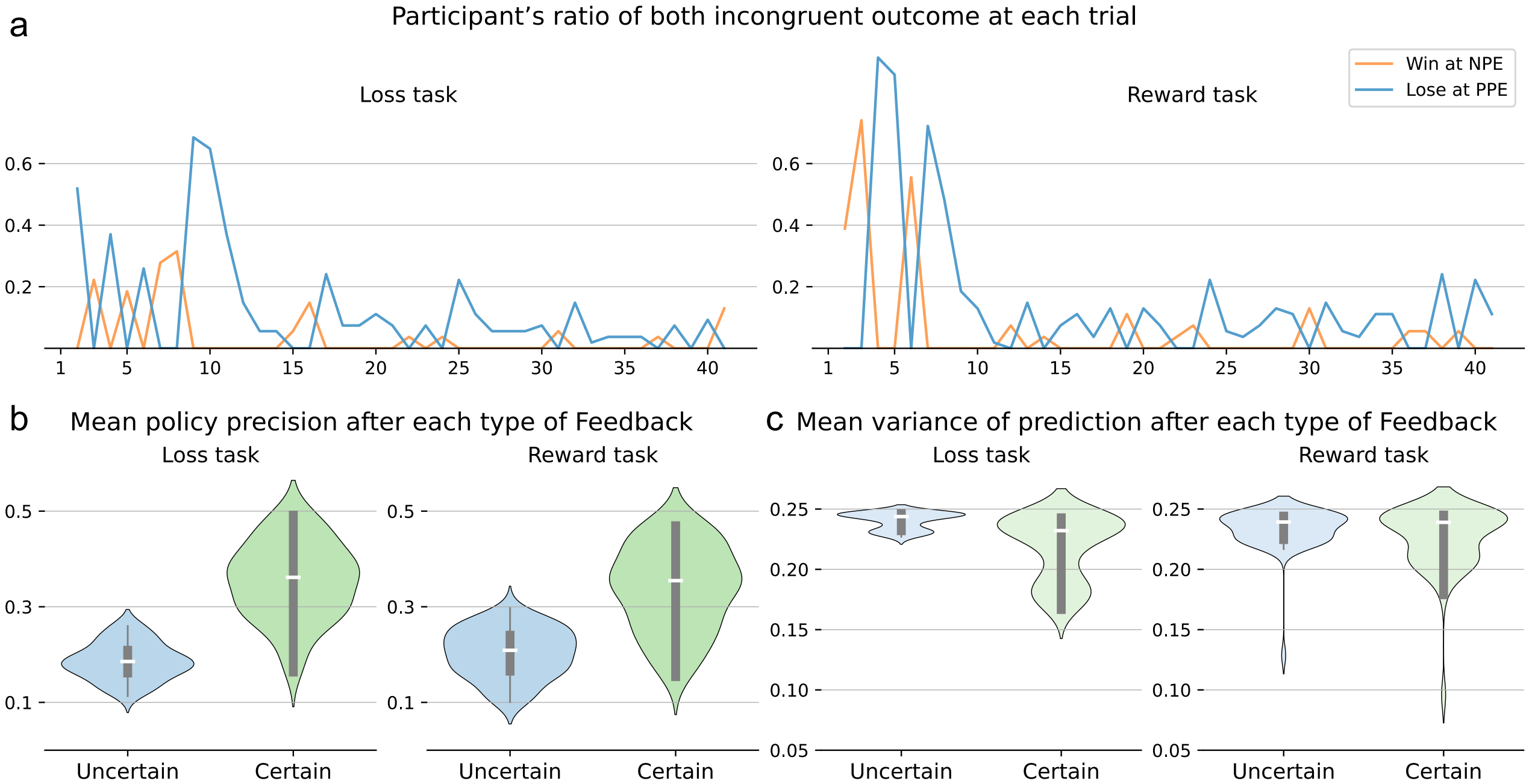
50. S. Vossel, R. Weidner, J. Driver, K. J. Friston, and G. R. Fink. Deconstructing the architecture of dorsal and ventral attention systems with dynamic causal modeling. *Journal of Neuroscience*, 32(31), 2012.

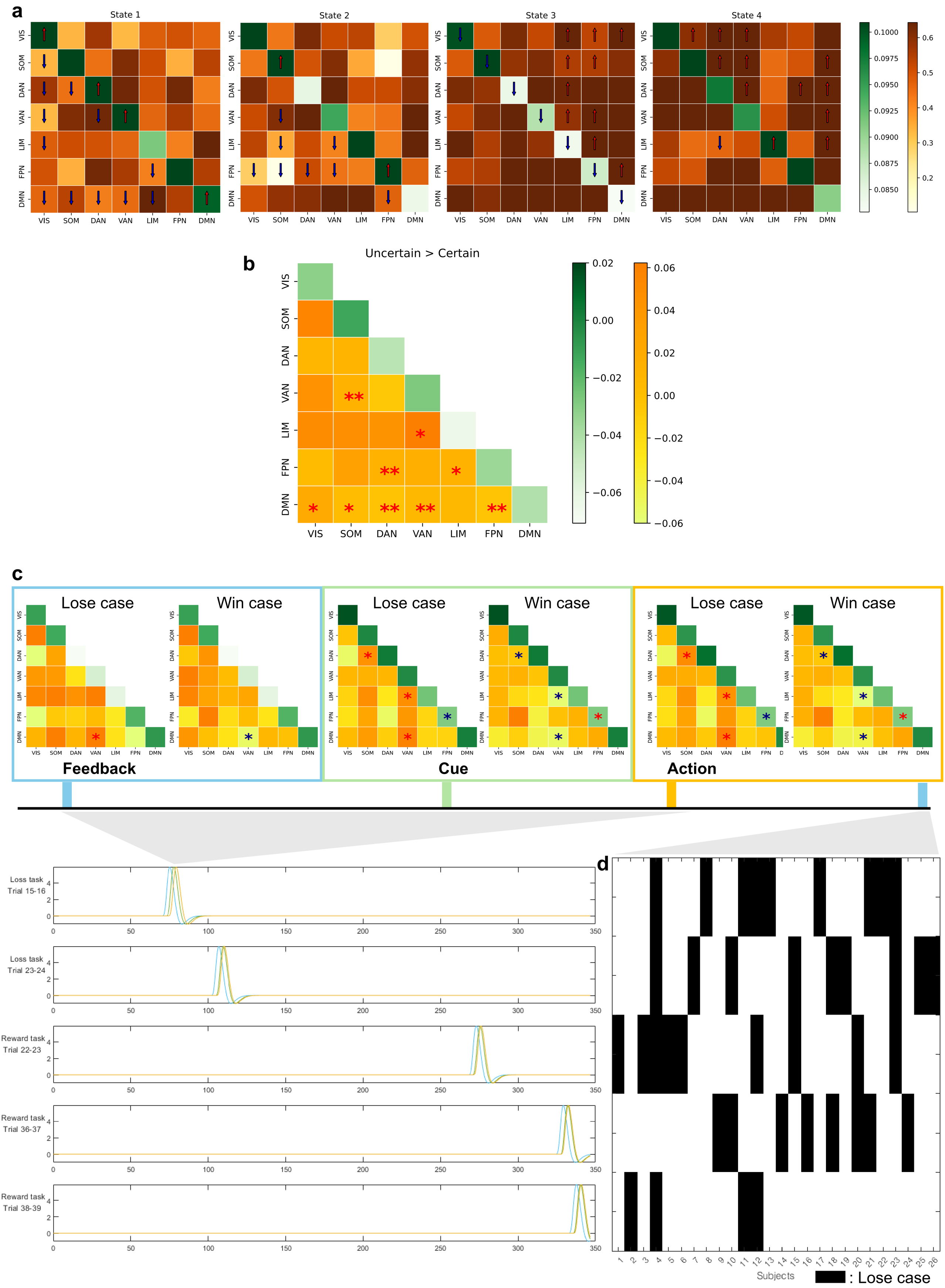

a
Neural signals
$x(t) = \alpha \cdot c(t) + \delta \cdot id(t) + \varepsilon(t)$
100 cortical region
47 subcortical region
Estimation of meta-network level states by GLMM
b
Behavioral signals
C
E
G
β
γ
π
D
s_1
B
s_2
A
A
o_1
o_2
c
Activation patterns
Occurrence probability of state
Window (21TR)
d
Policy precision
State prediction error
HRF convolution
e
Average occurrence probability of others (leave-1-out)
f
Average convolved behavioral signals of others (leave-1-out)
dISC of occurrence probability of state
dbISC of policy precision
dbISC of state prediction error
g
$ISC \approx \bar{\alpha} \cdot 1 + \bar{\delta} \cdot bISC$ of policy precision
Performing GLM with
$dISC(t) = \beta_0 + \beta_1 \cdot dbISC(t)$ of policy precision

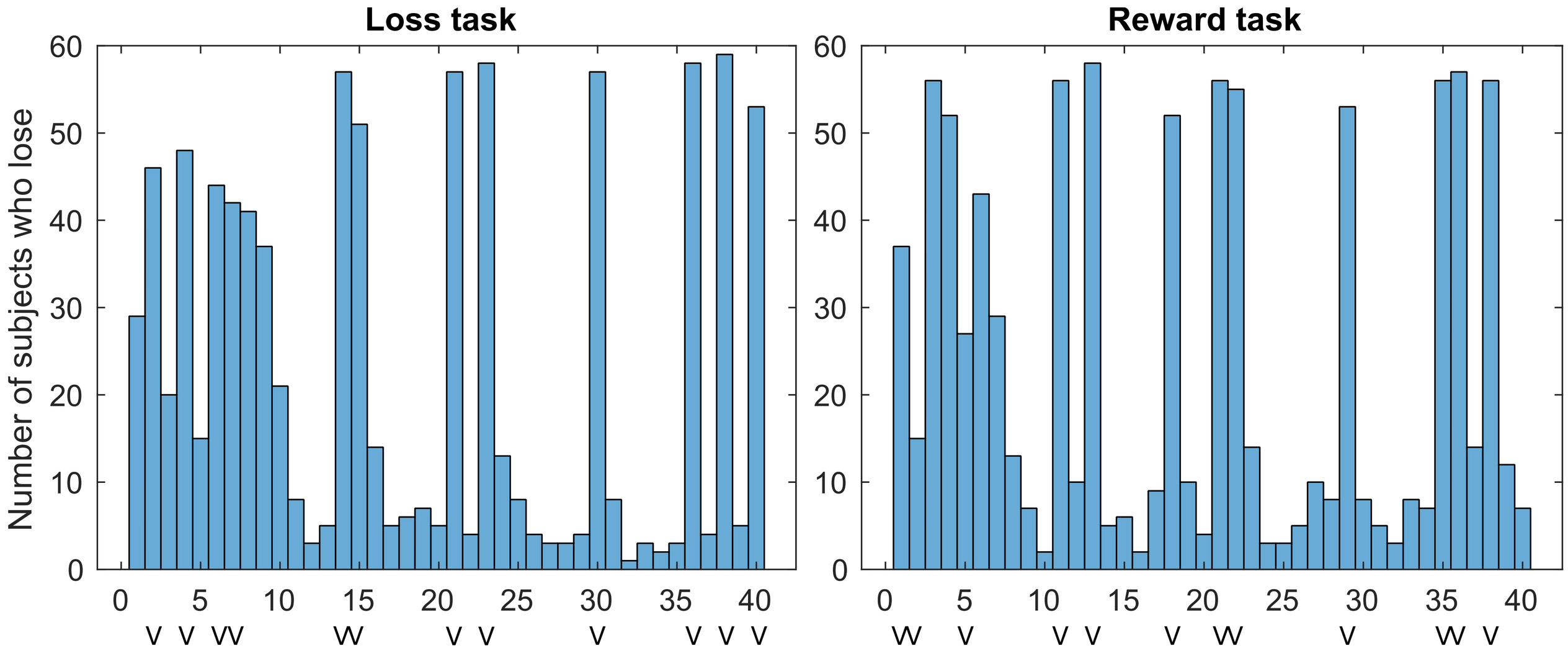

a
Trial start
Cue & Action
Feedback
Choice 1
Choice 2
-3 $
1.2 - 1.8 s
1.5 s
1.0 s
b
Loss task
value
trial
c
Reward task
value
trial
d
Policy Precision
Positive Predction Error
Negative Predction Error
e
Policy Precision hrf
Positive prediction error hrf
Negative prediction error hrf

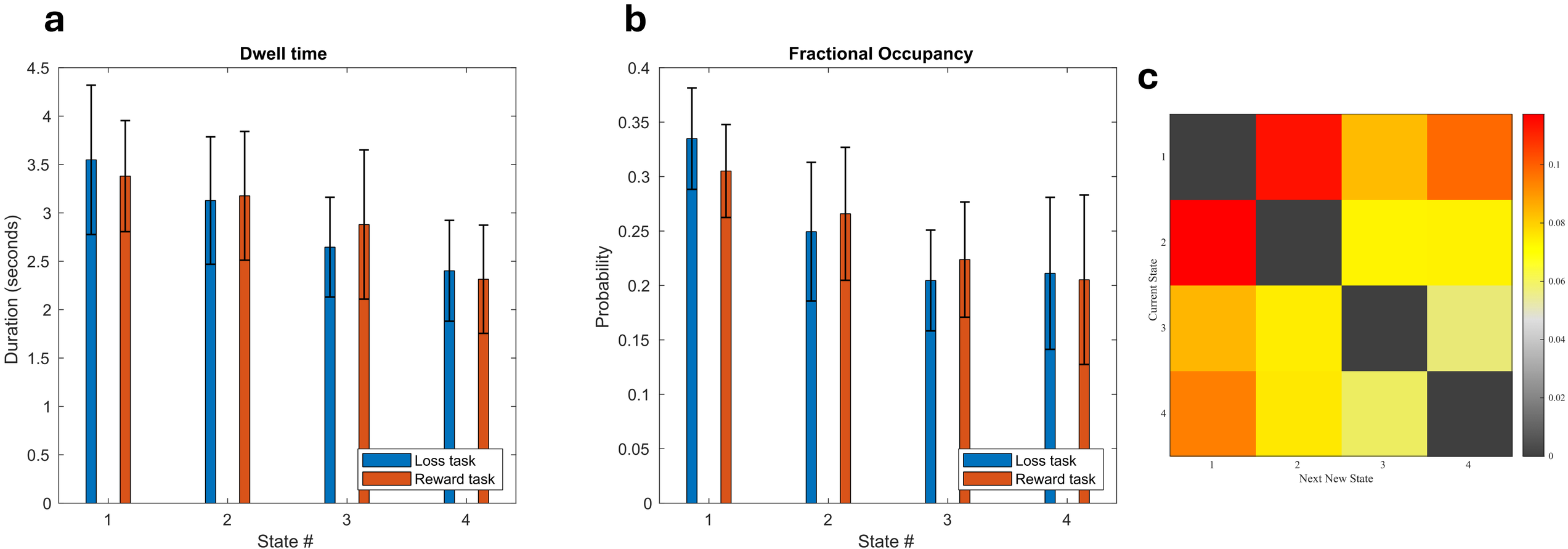

a
State 1
b
State 2
c
State 3
d
State 4
VIS
SOM
DAN
VAN
LIM
FPN
DMN
SC
e
f
g
h
i
j
k
l
R-HIP
L-HIP
R-AMY
L-AMY
R-pTHAL
L-pTHAL
R-aTHAL
L-aTHAL
R-NAc
L-NAc
R-PUT
L-PUT
R-CAU
L-CAU
R-GPe
L-GPe
R-GPi
L-GPi
R-SNc
L-SNc
R-red
L-red
R-SNr
L-SNr
R-PBPN
L-PBPN
R-VTA
L-VTA
R-VP
L-VP
R-HABN
L-HABN
R-hypoTHAL
L-hypoTHAL
R-MM
L-MM
R-STN
L-STN

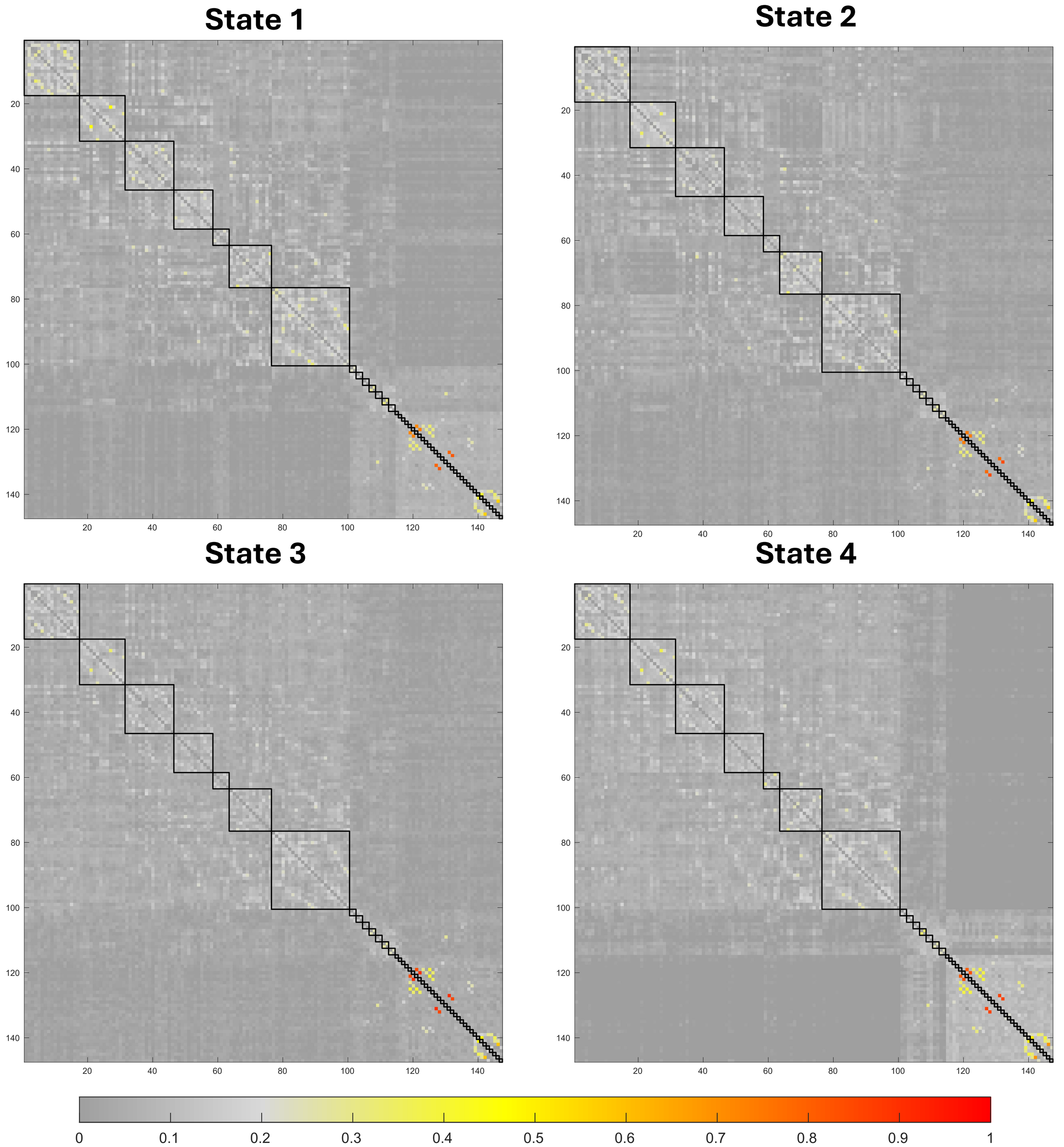

a
Loss task start
Reward task start
ISC
Prediction error
Policy precision
b
State1
State2
State3
State4
ISC
c
Probability
Time
d
State 1
State 2
State 3
State 4
0.146
(p = 0.001)
0.1033
(p = 0.012)
0.2095
(p < 0.001)
0.1084
(p = 0.004)
e
0.5111
(p = 0.002)
0.8948
(p < 0.001)
-1.4283
(p < 0.001)
0.185
(p = 0.211)

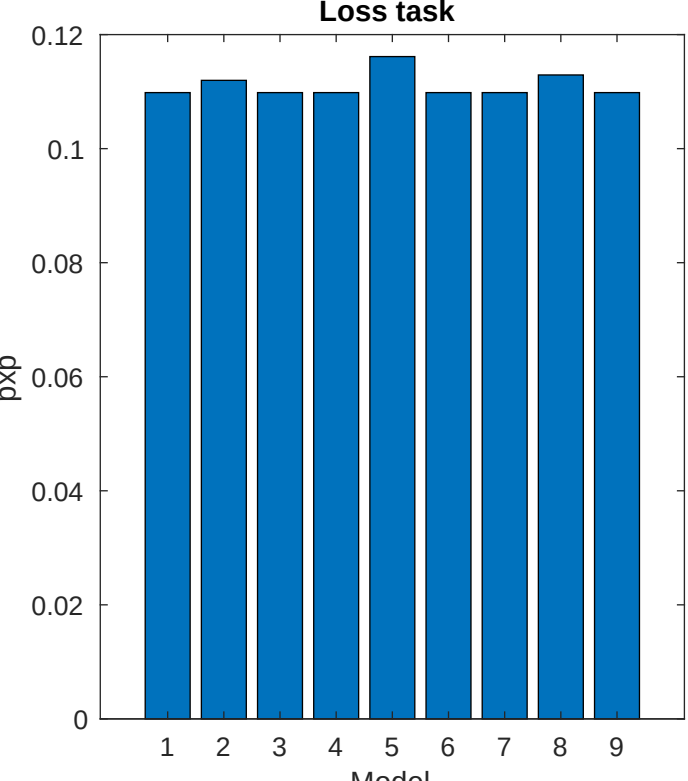

a
Participant's ratio of both incongruent outcome at each trial
Loss task
Reward task
Win at NPE
Lose at PPE
0.6
0.4
0.2
1
5
10
15
20
25
30
35
40
b
Mean policy precision after each type of Feedback
Loss task
Reward task
0.5
0.3
0.1
Uncertain
Certain
c
Mean variance of prediction after each type of Feedback
Loss task
Reward task
0.25
0.20
0.15
0.10
0.05
Uncertain
Certain

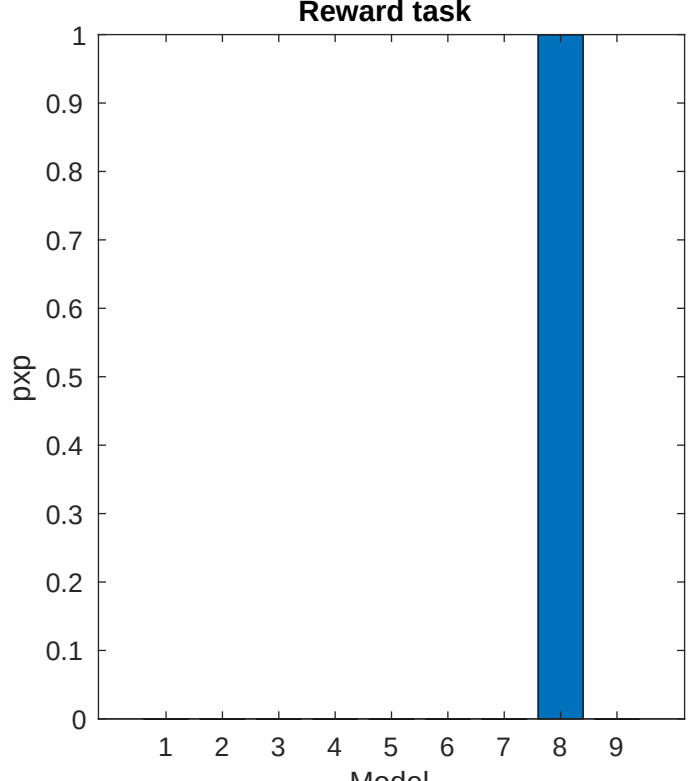

a
State 1
State 2
State 3
State 4
VIS
SOM
DAN
VAN
LIM
FPN
DMN
0.1000
0.0975
0.0950
0.0925
0.0900
0.0875
0.0850
0.6
0.5
0.4
0.3
0.2
b
Uncertain > Certain
0.02
0.00
−0.02
−0.04
−0.06
0.06
0.04
0.02
0.00
−0.02
−0.04
−0.06
c
Lose case
Win case
Feedback
Cue
Action
Loss task Trial 15-16
Loss task Trial 23-24
Reward task Trial 22-23
Reward task Trial 36-37
Reward task Trial 38-39
d
Subjects
: Lose case

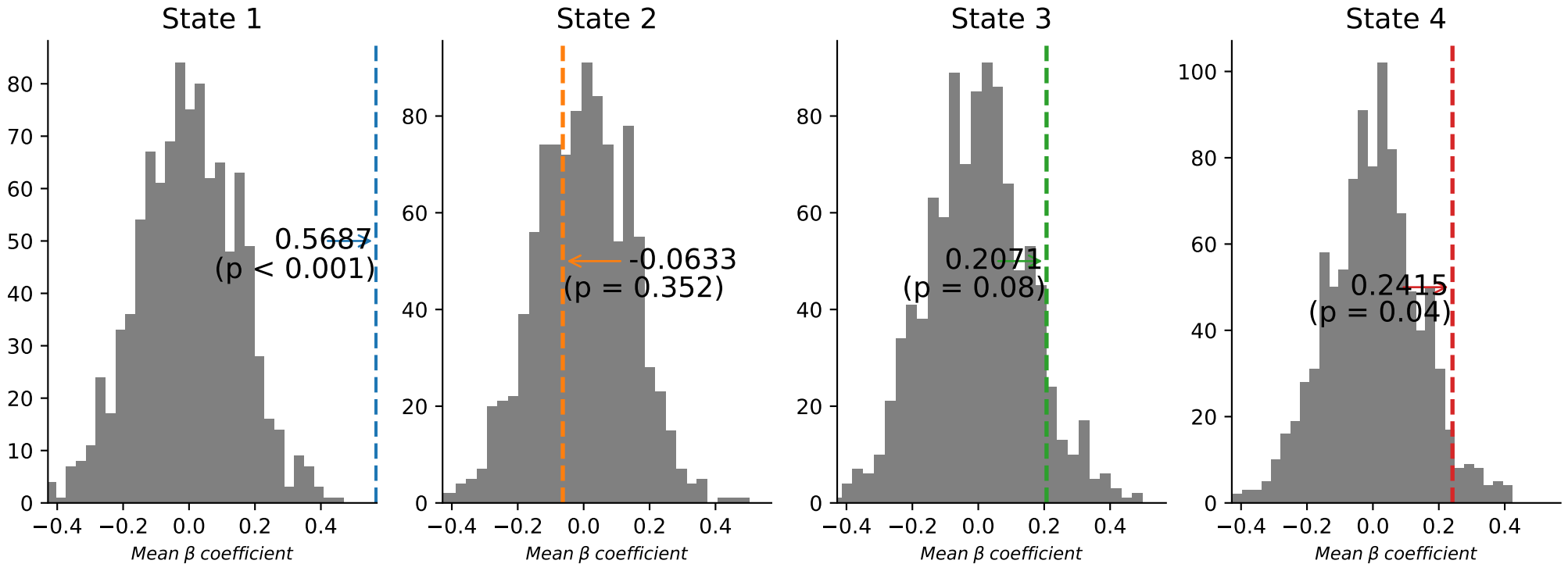

Loss task
Reward task
Number of subjects who lose
0
10
20
30
40
50
60
0
5
10
15
20
25
30
35
40
V V VV W V V V V V V
W V V V V W V W V


a
Dwell time
Duration (seconds)
State #
Loss task
Reward task
b
Fractional Occupancy
Probability
State #
Loss task
Reward task
c
Current State
Next New State


State 1
State 2
State 3
State 4
20
40
60
80
100
120
140
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1


Loss task
pxp
0.12
0.1
0.08
0.06
0.04
0.02
0
1 2 3 4 5 6 7 8 9
Model


Reward task
pxp
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1 2 3 4 5 6 7 8 9
Model


State 1
0.5687
(p < 0.001)
State 2
-0.0633
(p = 0.352)
State 3
0.2071
(p = 0.08)
State 4
0.2415
(p = 0.04)
Mean β coefficient
Mean β coefficient
Mean β coefficient
Mean β coefficient