nature neuroscience

Article

https://doi.org/10.1038/s41593-023-01258-y

Neurocomputational mechanism of real-time distributed learning on social networks

Received: 23 February 2022

Accepted: 17 January 2023

Published online: 16 February 2023

Check for updates

Yaomin Jiang[1,2,3], Qingtian Mi[1,2,3] & Lusha Zhu[1,2,3]

Social networks shape our decisions by constraining what information we learn and from whom. Yet, the mechanisms by which network structures affect individual learning and decision-making remain unclear. Here, by combining a real-time distributed learning task with functional magnetic resonance imaging, computational modeling and social network analysis, we studied how humans learn from observing others' decisions on seven-node networks with varying topological structures. We show that learning on social networks can be approximated by a well-established error-driven process for observational learning, supported by an action prediction error encoded in the lateral prefrontal cortex. Importantly, learning is flexibly weighted toward well-connected neighbors, according to activity in the dorsal anterior cingulate cortex, but only insofar as social observations contain secondhand, potentially intertwining, information. These data suggest a neurocomputational mechanism of network-based filtering on the sources of information, which may give rise to biased learning and the spread of misinformation in an interconnected society.

Social networks channel communication and route information transmission in human society[1,2]. By constraining what information we receive and from whom, the structure of social networks has substantial impacts on how we form beliefs and make decisions, and how collective opinion and behavior are shaped and propagated[2,3]. Although many studies have demonstrated the influence of social networks on political, economic and social activities at the population level[2,4], the underlying neural and cognitive processes by which individuals learn from interconnected peers remain unclear. Answers to this question would shed light on the neural mechanisms of social learning in wider and more ecologically relevant contexts, and help understand collective maladaptation—such as herding and misinformation propagation[5,6]—in terms of the computational challenges faced by individuals trying to process entangled information in an interconnected society.

A window into the internal processes of learning on social networks is provided by observational learning—a well-established paradigm whereby individuals learn about an unknown environment from the decisions of one or multiple observees performing the same task as the observer[7–12]. Error-guided models of reinforcement learning (RL)[13] have robustly shown that unexpected actions of the observees (that is, action prediction error (aPE)) drive learning and parametrically explain neural activity in the observer's lateral prefrontal cortex (LPFC) and related brain regions in a variety of learning and choice imitation tasks[7–9,12]. Despite its success, however, this neurocomputational account provides an incomplete description of learning on social networks. Past research of observational learning typically assumes that decisions of different observees constitute isolated, independent signals that can be integrated unbiasedly by the observer[9,14]. Contrary to this assumption, however, substantial evidence in the social network literature suggests that choices selected by connected peers are often interrelated and vary in their informativeness[15,16]. Blindly relying on the conventional observational learning strategy without considering

[1]School of Psychological and Cognitive Sciences and Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing, China. [2]IDG/McGovern Institute for Brain Research, Peking University, Beijing, China. [3]Peking-Tsinghua Center for Life Sciences, Peking University, Beijing, China. e-mail: lushazhu@pku.edu.cn

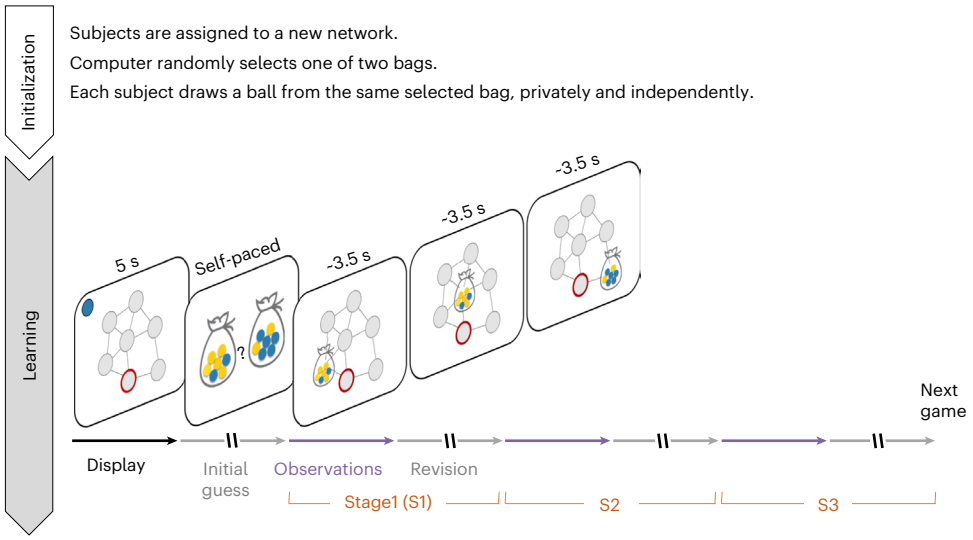


Fig. 1 | Task schematic. Each game is initialized by assigning subjects randomly to different nodes on a new network and selecting one of two bags that contain yellow and blue balls with opposing ratios (5:2 versus 2:5). Each subject, who does not know which bag was selected, privately draws a ball from the same selected bag with replacement, and needs to infer whether the underlying bag contains more blue or yellow balls in a series of decisions. All subjects are instructed that the chance of drawing a blue/yellow ball is either five-sevenths or two-sevenths, distributed independently and identically across all members embedded on the same network. After initialization, subjects are presented with the structure of the network (Display), one's own network location (red circle in Display), and the ball privately drawn from the bag (top left corner in Display). Subjects are then simultaneously asked to decide between two candidate bags (Initial guess), followed by three stages of observational learning (S1–S3). At each stage, a participant is presented with the most recent decisions by her immediate neighbors (Observations) and then provided with an opportunity to reassess her previous decision (Revision). During observations, neighbors' choices are revealed sequentially, in a clockwise order, starting from a randomly selected neighbor that varies across stages and between subjects.

the underlying connections that spread social influences is essentially ignoring the potential variations and repetitions in social signals[4,17]. Nevertheless, extant data suggest that social animals embedded in complex interacting webs demonstrate some level of sensitivity to the topological features of their immediate social environments[18,19], leaving open whether and how relevant structural information is incorporated into the learning processes.

To connect network geometries with the underlying learning mechanism in a quantitative and neurobiologically plausible manner, we draw insights from naïve social learning theories in the social network literature. Central to these theories is the idea that optimally using network structures to filter out uninformative or correlated information (for example, through Bayesian learning) is computationally demanding[4,17,20]. Heuristics that partially adjust for signal variations and repetitions save computation but sometimes lead to incorrect learning outcomes[4,17]. For instance, the canonical DeGroot learning model describes the dynamics of information aggregation as a simple Markov process that involves weighting and averaging others' information[21]. Under this heuristic, an individual exerts a constant level of influence on a particular social contact throughout the course of learning. This raises an intriguing question regarding the extent to which the between-individual differences in social influence and susceptibility to social influence can be linked to individuals' structural positions on their interacting network at the behavioral and neural levels[20,22].

Here, we propose to characterize observational learning in networked environments by incorporating the DeGroot learning heuristic into an RL-inspired account previously established for action imitation. We hypothesize that, similar to learning in simple social contexts, learning observationally from network neighbors involves the LPFC encoding of the aPE between an observed action and how expected that action was. Importantly, to integrate and reconcile observations from disparate neighbors, the brain needs to evaluate the relative importance of each observation and adjust the strength of learning (that is, learning rate) accordingly[23–25]. A key prediction derived from the DeGroot learning theory is that such social information evaluation is based on where and how the interacting individuals are embedded on the network that channels information transmission. As a result, the strength of learning from a particular observation may change according to the structural properties associated with the observer and observee on the network. Previous research into RL in nonsocial contexts has shown that the dorsal anterior cingulate cortex (dACC), frontoparietal areas and other regions are involved in tracking the dynamic changes in the RL learning rate[24,26]. On the bases of these previous data, we predict that responses in some of these brain regions to a social observation would vary parametrically with the network features of the specific observer and observee in a manner consistent with how those features modulate the behavioral learning rate.

We formulated a simple DeGroot learning model and evaluated its behavioral and neural implications by combining functional magnetic resonance imaging (fMRI) with a distributed learning task for action observation learning (Fig. 1 and Supplementary Fig. 1). The task was adapted from economic studies of information cascade[27] and housed in a variety of exogenously manipulated, seven-node, undirected and unweighted networks[17,28,29]. To focus on how network structures affect learning, the experiment used simulated networks where connections represent the routes of information transmission, abstracting away social affiliative associations that might be confounded with interpersonal influences[19].

Results

Stage-varying, degree-modulated learning on social networks

The proposed DeGroot learning model is built on the classic model of action observation learning, positing that learning from a neighbor's action is driven by an aPE between the observed and expected action, weighted by a learning rate. To delineate the rich network interactions, our model incorporates two additional assumptions. First, motivated by previous data on social influence that a well-connected individual has greater influence on her peers and is less susceptible to others' opinion[19,30,31], the proposed DeGroot learning model postulates that the extent to which one learns from an observed action scales with the

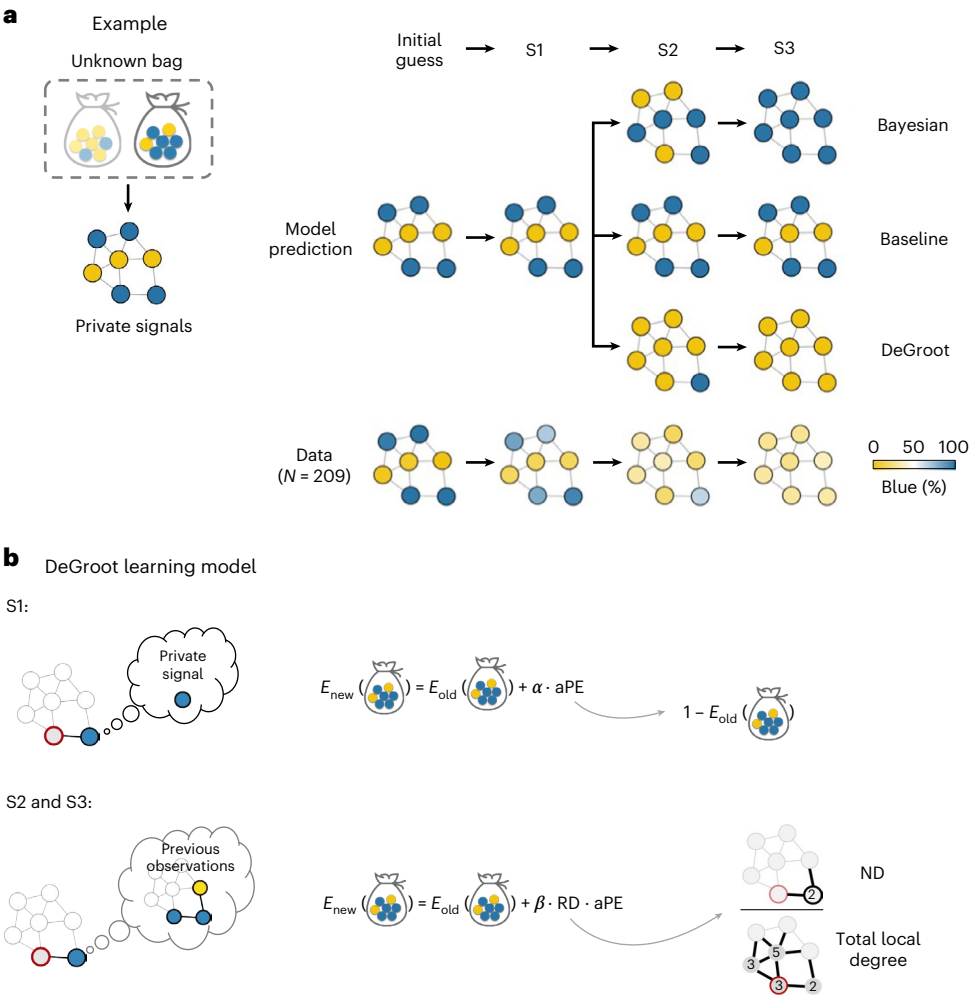


Fig. 2 | DeGroot learning model. a, Learning dynamics on an example network. The example also illustrates possible misinformation propagation under the DeGroot learning strategy. Left: the underlying bag selected by the computer and the private signal for each node given the selected bag. Right: simulated and actual choices on the network. Color in each node represents a simulated choice (model prediction) or the frequency of actual choices (data) over bags containing more blue or yellow balls. Unlike rational Bayesian learners, who gradually form a consensus on the correct underlying state, both the simulated DeGroot agents and actual participants converged toward the wrong estimation, biased by the inaccurate information from the central, most-connected individual on the network (see Supplementary Fig. 2 for model simulation on all networks employed by the study). **b**, DeGroot learning model illustration. Left: while an S1 observation reflects the neighbor's private signal, an observation in S2 or S3 additionally signals what the neighbor has learnt from her neighbors. Right: upon observing an action from a neighbor, the belief expectation about the unknown state (E_{old}) is updated through an aPE, defined as the discrepancy between the observed and expected action. Learning in S1 follows the typical setup for action observation learning, where the aPE signal is scaled by a baseline learning rate (α). In S2 and S3, aPE signals are weighted by the learning rate (β) and RD, with the latter being defined as the degree centrality of the observee relative to the total degree of the observer and all her direct neighbors on the network (Methods).

observee's network connectedness, relative to that of the observer. We quantify an individual's connectedness using her degree centrality, which captures the number of individuals to whom one is directly connected on a network; this is one of the most fundamental metrics for local prominence and immediate influence in social network analysis[1,32]. To reflect the mathematical requirement that a learning rate should be no greater than 1, our model adopts an assumption that the learning rate is weighted by a measure of relative degree (RD), defined as the observee's degree centrality with respect to the total degree centrality of the observer and all her direct neighbors on the network (Fig. 2b; see Methods for alternative specifications).

In addition, the DeGroot learning model posits that the degree-modulation effect on learning varies systematically over the course of information circulation, contingent on whether social observations differ in their informativeness. For example, when individuals learn from others' firsthand, isolated information, the DeGroot model may reduce to the standard observational learning algorithm, whereby an observer is similarly influenced by the received information regardless of the differences in the observees' network locations. Conversely, when learning from others' secondhand, possibly heterogeneous and intertwining information, the strength of learning will be modulated by the relative degree between the observee and observer on the network.

A real-time distributed learning game

A total of 217 unique subjects (31 fMRI participants) participated in the experiment in groups of seven (1 inside the fMRI scanner; 209 included in data analyses with 25 fMRI participants; see Methods for subject exclusion). The experiment consisted of 40 separate games on varying networks. The structures of the networks were preselected based on the separability of choice behavior simulated by different learning models (Extended Data Fig. 1 and Supplementary Fig. 2; Methods). In each game, a participant's goal was to infer an unknown state of the environment, which was common to all seven participants in the game. At the beginning of a game, seven subjects were assigned randomly to

different nodes on a new network, and a computer selected one of two underlying states at random. Each subject received a private signal that was independently and identically distributed conditional on the same selected state, and needed to make an initial guess about the underlying state (Fig. 1 and Supplementary Fig. 1; Methods). We hypothesized that a subject should rationally base her estimation on the private signal in this decision. The prediction was confirmed by our data, in which 98.34 ± 5.12% (mean ± intersubject s.d.) initial estimations matched subjects' private signals.

Critical for the purpose of this study, the participants were then allowed to revise their estimations in response to the choices previously selected by the neighbors to whom each was directly connected on the network. To allow for meaningful fMRI analyses, a subject was presented with her direct neighbors' previous decisions sequentially, one at a time, such that her neural responses could be linked directly to the action and network location of a particular observee. To allow for examining learning effects, the process of observing neighbors' actions and reassessing one's own estimation was carried out three times consecutively within each game (henceforth Stages 1–3, or S1, S2 and S3; Fig. 1).

Seven participants played the same game simultaneously, from their respective network locations via an intranet. That is, the participants were facing the same underlying state, the same network structure and display, and were making decisions at the same time in each game. Crucially, when a participant was witnessing her neighbors' choices, her neighbors were also presented with the choice information from their respective neighborhoods. Under this setup, information received by an observer is incorporated into this observer's subsequent decision and propagated gradually from the observer to her direct and indirect contacts along network connections in the later stages of the game. Throughout the experiment, all subjects were financially incentivized to guess as accurately as possible in all four decisions in each game (that is, initial guess and three reassessments) and had no incentive to mislead or collude with others. No feedback was provided on the accuracy of estimations during the experiment (Methods).

This task has two features important for evaluating the proposed DeGroot learning model at the neurocomputational level. First, the three learning stages (S1–S3) were set up identically within each game (Methods), allowing us to evaluate the hypothesis that learning is guided by aPE signals reflected by the LPFC activity in S1, S2 and S3 nonselectively. Second, despite their identical experimental setups, three learning stages differ in the type of information contained in neighbors' actions. Unlike S1 observations, which reflect observees' independent private signals, an S2 or S3 observation would additionally signal what the observee has learned from her neighbors (Fig. 2b, left), thereby becoming relatively more informative to the specific observer when the observee is better-connected and the observer is less-connected. Given this feature, an immediate implication of the proposed model is that learning from network neighbors should be modulated by the relative degree centralities between the observee and observer in S2 and S3, but not in S1 (Fig. 2b, right).

Degree modulates choices in S2 and S3, but not in S1

Behaviorally, participants adapted their decisions in response to neighbors' choices, such that the level of consensus within a network grew from 61.21 ± 1.29% (mean ± intergroup s.d.) in the initial guess to 88.10 ± 3.03% in the last (S3) decision (Extended Data Fig. 2). To characterize the overall learning effect, we first performed mixed-effects logistic regression for each separate learning stage, examining the extent to which the likelihood of participants altering their choices was influenced by social observations, and whether the strength of influence was modulated by the degree centralities of the observees and observers. We hypothesized that the likelihood that a participant aligned her estimation to an observation would be associated positively with the neighbor's degree centrality (ND) but negatively with the observer's own degree centrality (OD) in S2 and S3, but not in S1.

To evaluate the impact of ND, we summarized the sequence of observations revealed to a subject at each stage of each game, using two variables: (1) sum of observations across all direct neighbors (1, if an observed action differs from the observer's previous choice; otherwise, −1), and (2) weighted sum of observations across all direct neighbors, with the respective neighbor's degree serving as the weight. The inclusion of the unweighted and ND-weighted regressors helped to isolate the degree-modulation effect of interest from a baseline tendency of following the majority—a phenomenon widely reported in studies of group decision-making[5]. To remove any shared variances between the two regressors, we orthogonalized the ND-weighted regressor against the unweighted regressor, such that choices that were equally explainable by the two variables were attributed solely to the unweighted regressor. The regression coefficient for the ND-weighted sum of observations, therefore, served as a more stringent test on whether ND modulated learning.

Consistent with our hypothesis, the regression analyses showed that the likelihood of a participant modifying her decision was correlated positively with not only the unweighted sum of observations, but also the ND-weighted sum of observations, in both S2 and S3 (Fig. 3a and Supplementary Table 1a). The positive effects suggested that, in addition to following the majority, subjects were more likely to be swayed toward the decisions of highly connected neighbors, relative to those of poorly connected neighbors embedded on the same network. In stark contrast, in S1, the ND-weighted regressor showed no extra explanatory power above and beyond the unweighted regressor in predicting observers' subsequent choices.

To assess how learning was biased by an observer's OD, we further compared the influence of the ND-weighted sum of observations on the observer's subsequent decision when the observer was endowed with high versus low degree centrality. As predicted by our model, we found that the probability of an observer modifying her decision was associated negatively with the interaction between OD- and ND-weighted sum of observations, in both S2 and S3 (Fig. 3b and Supplementary Table 1b,c for regression estimates and robustness check). The negative interaction suggested a decreased susceptibility to social observations when a participant was highly relative to poorly connected. In S1, by comparison, no systematic variation was observed in participants' susceptibility to neighbors' actions with their OD.

DeGroot learning model explains learning behavior

To formally test the proposed model and derive latent variables that might reflect the underlying neurocognitive operations, we fit the DeGroot learning model with each participant's choice behavior (Supplementary Fig. 3 and Supplementary Tables 2 and 3; Methods). We compared the proposed model against two benchmark models: a baseline action imitation model (Baseline model), which assumes that the network-related information is completely ignored throughout learning, and a Bayesian learning model, which assumes that the information regarding the network structure is used optimally by all participants (Supplementary Note 2; Methods).

Using both the Bayesian information criterion (BIC) based on the in-sample model fit and the out-of-sample prediction accuracies based on a fivefold cross-validation procedure (Methods), we found that the proposed DeGroot model outperformed the alternative models (Fig. 3c, Extended Data Fig. 3 and Supplementary Fig. 4). Across subjects, estimates from the DeGroot learning model were consistent with results of logistic regression, such that subjects whose behavior was better characterized by DeGroot learning showed a more pronounced logistic behavioral sensitivity to both ND and OD in S2 and S3 (Fig. 3d). In addition to the Bayesian and Baseline models, the proposed model outperformed a range of alternative models that could explain choice behavior based on assumptions differing in either which network

https://doi.org/10.1038/s41593-023-01258-y

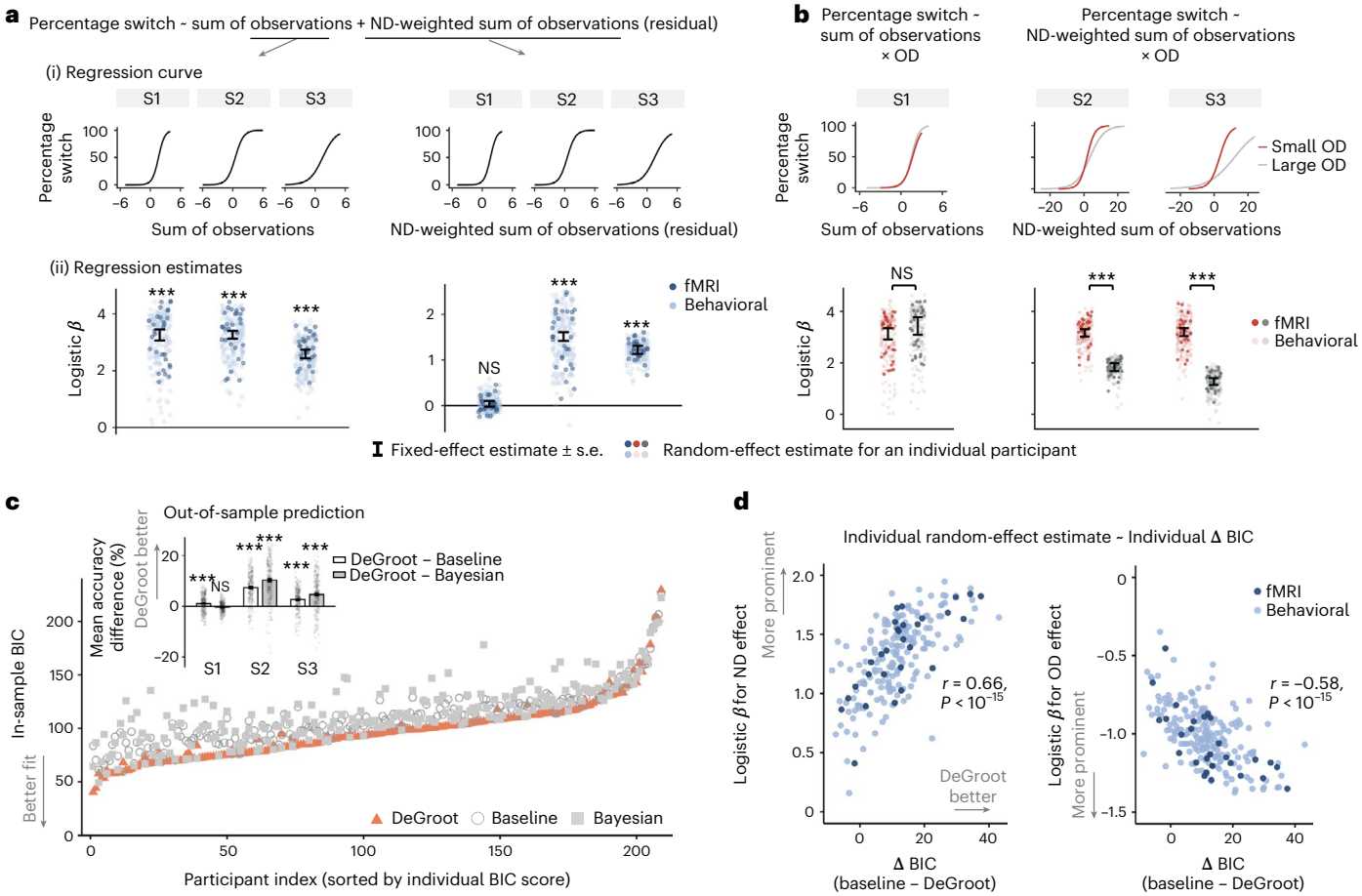


Fig. 3 | Behavioral evidence. a, Mixed-effects logistic regression analyses show that ND modulates learning in S2 and S3, but not in S1. (i) The probability of an observer changing her previous decision as predicted by the estimated logistic regression model at each separate stage. (ii) Fixed-effects regression coefficients and the corresponding s.e., overlaid by the random effects coefficient associated with each observer (see Supplementary Table 1 for statistics). **b**, Observer's OD is associated negatively with the susceptibility to observations in S2 and S3, but not in S1. For illustration purposes, large (gray) and small (orange) ODs were defined by median splits on OD across networks for each participant. **c**, Comparisons of in-sample model fits using the BIC of each participant show that the DeGroot model explains behavioral choices better than the Bayesian and Baseline models (aggregated BIC, DeGroot = 20,407.41; Baseline = 22,827.97; Bayesian = 23,755.10). Inset: out-of-sample prediction accuracy is superior for the DeGroot model compared with Bayesian and Baseline models in S2 and S3. **d**, Consistent results between logistic regression and model estimation. Individual BIC differences between the DeGroot and Baseline models are plotted against the individual logistic regression estimates for ND (left) and OD (right) effects, respectively, exploiting the fact that the DeGroot model differs from the Baseline model only in the assumption regarding the degree-modulation effect in S2 and S3 (Methods). The model-free ND (or OD) effect is captured by the coefficient of individual random effects with respect to the ND-weighted sum of observations (or OD × ND-weighted sum of observations), averaged across S2 and S3 within each subject. Each dark/light dot represents an fMRI/behavior participant. Error bars in **a** and **b** represent the s.e. of fixed-effect estimates in the logistic regression, in **c** represent intersubject s.e.m., both based on $N = 209$ behavioral participants. ***$P < 0.001$; NS, not significant; z tests in **a** and **b**; t tests in **c**, all two-sided, Bonferroni-corrected when appropriate.

metric might modulate learning (for example, eigenvector centrality, betweenness, and so on) or how parameters were specified when formulating error-driven learning (Supplementary Table 4; Methods). To address the potential concern regarding the between-network variations in learning behavior, we estimated models for each separate network (pooling over subjects), and observed similar results at the across-network level (Extended Data Fig. 3; Methods).

Right LPFC tracks aPE estimates in S1, S2 and S3

Having established the DeGroot learning model at the behavioral level, we then tested whether fMRI activity reflected the model-derived aPE estimates, and did so consistently across three learning stages. We performed event-related fMRI analyses of participants' neural responses, on an observation-by-observation basis, at each time in each learning stage when participants were presented with a neighbor's action (GLM1; Methods). Averaging the regression coefficients of aPE estimates across three learning stages (Methods), we observed significant aPE-related signals in the right LPFC (rLPFC), which has been implicated in representing notions of prediction error signals in action observation learning[7–9] and as part of the 'mirror' system encoding the executed and observed actions in a range of interpersonal scenarios[33] (Fig. 4a, Extended Data Fig. 4 and Supplementary Table 5). The rLPFC encoding reflects the observation-by-observation changes in aPE estimates and could not be attributed to the correlation between the rLPFC activity and the aPE estimate associated with a single observee (for example, one's most-connected (MC) neighbor) across learning stages (Extended Data Figs. 5 and 6). The rLPFC encoding remained significant after controlling for a range of decision variables (GLM2; cluster-wise family-wise-error (FWE)-corrected $P < 0.05$, with cluster-forming threshold uncorrected P ($P_{unc.}$) < 0.001; Methods).

Importantly, and consistent with our model prediction, the rLPFC cluster tracked the aPE estimates stably in S1, S2 and S3, such that

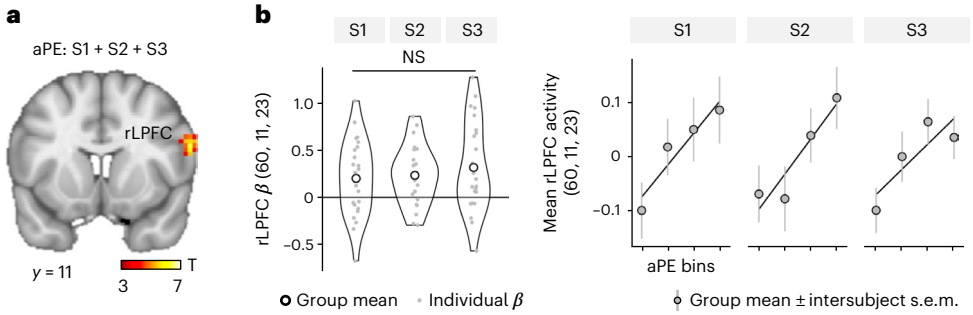


Fig. 4 | rLPFC tracks the value estimate of aPE in S1, S2 and S3. a, Statistical parametric map with respect to aPE estimates at observation onsets, computed by averaging the regression coefficients for aPE estimates across S1, S2 and S3 within each subject (peak voxel MNI coordinates: x, y and z = 60, 11 and 23; cluster-wise FWE-corrected $P < 0.05$, with cluster-forming threshold $P_{unc.} < 0.001$). **b**, rLPFC cluster demonstrates similar effect sizes for individual aPE estimates across stages. Left: violin plots for the distribution of individual rLPFC β values of aPE estimates, extracted separately for each stage from the significant rLPFC cluster as identified in Fig. 4a (one-way repeated-measures ANOVA, $F(2,48) = 0.55$, $P = 0.582$). Right: mean fMRI activity extracted from the same rLPFC cluster and binned by aPE estimates in each separate stage. Each dot represents an fMRI subject. Each circle represents a group mean. Error bars represent intersubject s.e.m. in the fMRI sample ($N = 25$).

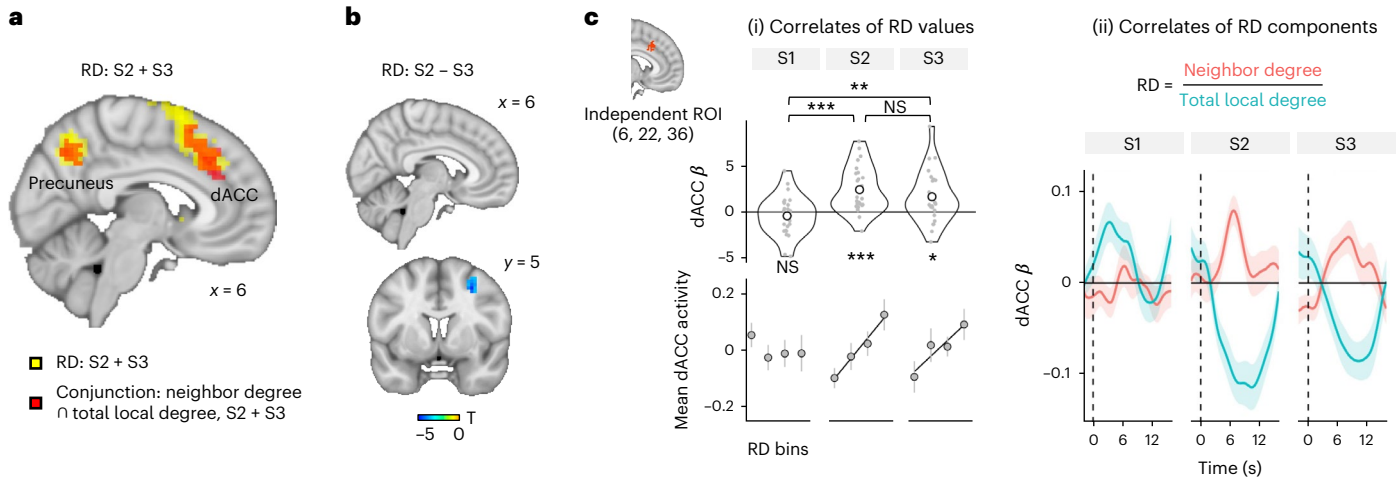


Fig. 5 | Activity in the dACC correlates with RD values in S2 and S3. a, dACC shows RD-related signals on an observation-by-observation basis in S2 and S3. Regions shaded in yellow indicate clusters where activity significantly correlates with RD values at observation onsets in S2 and S3, calculated by averaging the regression coefficients for RD across S2 and S3 within each participant (cluster-wise FWE-corrected $P < 0.05$, with cluster-forming threshold $P_{unc.} < 0.001$). Regions shaded in red indicate clusters scaling both positively with neighbor's degree and negatively with total local degree at observation onsets in S2 and S3, as revealed by a whole-brain conjunction analysis (cluster-wise FWE-corrected $P < 0.05$, with cluster-forming threshold $P_{unc.} < 0.001$). **b**, Paired comparison of RD correlates between S2 and S3. The only significant cluster locates in the middle frontal gyrus (MNI, x, y and z = 30, 5 and 47; cluster-wise FWE-corrected $P < 0.05$, with cluster-forming threshold $P_{unc.} < 0.001$). **c**, dACC ROI, independently defined by Neurosynth[39] using the topic term 'cognitive control'. (i) Top: β values with respect to RD extracted at observation onsets of separate learning stages from the same independent ROI (S1: $\beta = -0.43 \pm 0.44$, two-sided t test, $t_{24} = -0.99$, $P = 0.332$; S2, $\beta = 2.41 \pm 0.50$, $t_{24} = 4.85$, $P = 6.05 \times 10^{-5}$; S3, $\beta = 1.64 \pm 0.56$, $t_{24} = 2.91$, $P = 0.008$; across-stage, one-way repeated-measures ANOVA, $F(2,48) = 10.65$, $P = 1.48 \times 10^{-4}$). Bottom: mean dACC activity binned by RD values in each stage. (ii) Timecourse analyses with respect to the neighbor's degree and total local degree at each stage in the same dACC ROI. Vertical dashed lines indicate the observation onset. Error bars in i and ii indicate intersubject s.e.m. in the fMRI sample ($N = 25$). $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, two-sided t tests, Bonferroni-corrected.

within-subject comparisons found no systematic difference across learning stages in the neural responses to aPE estimates, both within the identified rLPFC cluster (Fig. 4b) and at the whole-brain level (cluster-wise FWE-corrected $P < 0.05$, with cluster-forming threshold $P_{unc.} < 0.001$).

In addition to aPE signals at observation onsets, we also observed—at the choice time—classic neural correlates of computational signals implicated by the temporal difference form of learning[34]. This included the value estimates of reward expectation associated with the observer's decision in the orbitofrontal cortex (OFC) and the model-derived probability of modifying previous decisions in the anterior cingulate cortex (Extended Data Fig. 7). Findings at the choice time, together with the aPE signals at the observation time, indicated that action observation learning on complex social networks may be supported by an error-driven process reminiscent of those seen in non-networked environments.

dACC represents RD-related signals in S2 and S3

Using similar analyses, we looked for brain regions tracking the changes in relative degree centralities between the observee and the observer (that is, RD), and did so consistently in S2 and S3 (GLM1; Methods). Averaging the regression coefficients of RD values across S2 and S3, we found a strong correlation in a network of brain regions, including the dACC extending to the adjacent presupplementary motor area (preSMA), precuneus, bilateral anterior insula, visual cortex and other areas (Fig. 5a and Supplementary Table 5). The loci of activation in

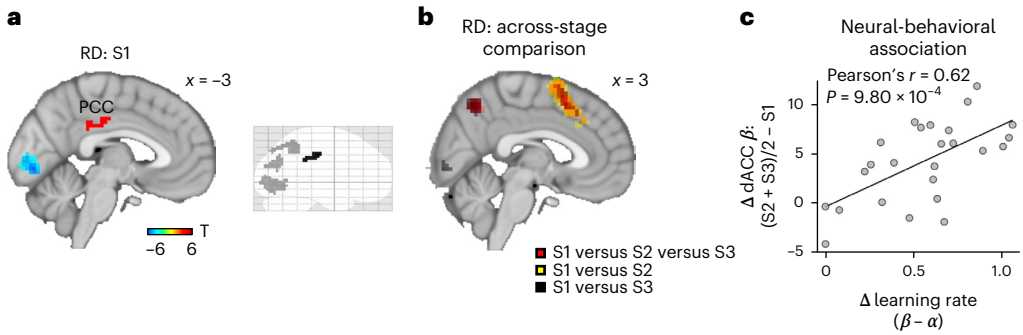


Fig. 6 | Activity in the dACC does not correlate with RD values in S1. a, Neural correlates of RD values at observation onsets in S1 (cluster-wise FWE-corrected $P < 0.05$, with cluster-forming threshold $P_{unc.} < 0.001$). **b**, Paired comparisons with respect to RD correlates across learning stages (all cluster-wise FWE-corrected $P < 0.05$, with cluster-forming threshold $P_{unc.} < 0.001$). Whole-brain within-subject ANOVA comparing RD correlates across three stages (red) is overlaid with post hoc analyses comparing RD correlates in S2 versus S1 (yellow) and S3 versus S1 (black; two-sided). **c**, Across fMRI subjects, the dACC beta values with respect to RD in S2 and S3, relative to that in S1, are correlated positively with the individually estimated learning rates in S2 and S3, relative to that in S1. Each dot represents an fMRI subject ($N = 25$).

the dACC/preSMA were similar to those seen in the past experiments where subjects adjusted behavioral strategies, such as learning rate, in response to environmental changes[24,26,35–38].

In support of the observation-by-observation representation of RD, the correlation between the dACC activity and RD values was seen robustly in the first, second, odd-numbered, even-numbered, early and late observations in S2 and S3, with no systematic difference in the correlation effect sizes across observations (Extended Data Fig. 6). The dACC encoding remained significant after controlling for a wide variety of decision-related and unrelated variables, including the visual properties of the network layout, general cognitive functions of learning and decision-making previously associated with the dACC, and other nuisance effects (GLM3; Extended Data Fig. 8; Methods). Moreover, activity in the dACC was correlated simultaneously with the neighbor's degree (numerator in RD) and total local degree (denominator in RD), with opposing signs, at observation onsets in both S2 and S3. The effect was seen not only in a region of interest (ROI) in the dACC, defined independently using an automated online meta-analysis[39] (Fig. 5c, right), but also in a conjunction analysis at the whole-brain level (GLM4; Fig. 5a; Methods).

Consistent with the DeGroot model prediction, activity in the dACC demonstrated similar response patterns to RD values in S2 and S3. Neural β values extracted separately for S2 and S3 from the same independent dACC ROI were both highly significant and showed no systematic difference in their effect sizes (Fig. 5c). A whole-brain within-subject comparison of RD correlates further confirmed that, except for a cluster confined to the middle frontal gyrus (Fig. 5b), no other cluster responded differently to RD values in S2 versus S3 at cluster-wise FWE-corrected $P < 0.05$.

dACC does not respond to RD values in S1

By contrast, in S1, the same analysis revealed no significant correlation with RD values at observation onsets in the dACC or other frontal regions, in either positive or negative direction (Fig. 6a; GLM1; Methods). Instead, we observed positive correlations with RD values in a circumscribed cluster in the posterior cingulate cortex (PCC), and negative correlations restricted to the precuneus and visual cortex (Fig. 6a and Extended Data Fig. 9). The identified regions in the PCC and precuneus have been implicated recently in encoding features of real-world social networks, even when such network features were task-irrelevant[40–42].

To more formally examine the spatial expression of the RD correlates and test the stage-varying involvement of the dACC, we searched the whole brain for voxels that responded similarly (conjunction analyses) or differently (analysis of variance (ANOVA)) to RD values across stages. Conjunction analyses showed substantial overlapping responses to RD values between S2 and S3 in regions including the dACC, but no significant overlap among S1, S2 and S3 for either positive or negative activation to RD values (Extended Data Fig. 8). Moreover, using a whole-brain ANOVA, we compared the RD correlates directly across stages within subjects, and identified a significant stage effect in several brain regions including the dACC (Fig. 6b). According to the post hoc paired comparison, this stage difference was attributable to the increased correlation between dACC signals and RD values in S2 versus S1, and S3 versus S1 (Fig. 6b), but not between S2 and S3, in either positive or negative direction (Fig. 5b; see also Fig. 5c for ROI analyses). Together, these data provide consistent evidence suggesting that the neural correlates of RD in S1 were segregated spatially from those in S2 and S3 in a manner consistent with the DeGroot prediction.

dACC sensitivity to RD predicts choice sensitivity to RD

To relate the dACC encoding of RD values to choice behavior, we tested whether, across subjects, the extent to which dACC activity reflected RD values was predictive of the behavioral effects of RD on learning. We used the individual value estimate of learning rate in S2 and S3 (that is, β as in Fig. 2b) as a measure of how strongly RD affected learning at these stages (zero effects on learning when $\beta = 0$). We plotted the individual behavioral estimate of β against the dACC beta of RD that was averaged from S2 and S3 in each subject, controlling for the respective baseline effects in S1 (Fig. 6c). The data showed that subjects with higher learning rates in S2 and S3 than in S1 exhibited greater dACC sensitivity to RD values at observation onsets in S2 and S3 than in S1. The between-subject association was also highly significant and denoted similar effect sizes when tested separately in S2 and S3 (S2: Pearson's $r = 0.52$, $P = 0.008$; S3: $r = 0.56$, $P = 0.004$; S2 versus S3: $\beta = -0.13 \pm 0.25$, $P = 0.593$). Of note, this neural-behavioral association was not a spurious effect arising from the double dipping of data[43]. This is because the dACC beta with respect to RD reflects the neural responses to the exogenously manipulated network features, which are independent of participants' choice behavior, model specification or data estimation.

VMPFC encodes updated belief expectation in S1, S2 and S3

The above results thus raised the question of how social observations from disparate neighbors were integrated in the brain to inform the subsequent decision. Unlike previous learning experiments, where subjects typically make a choice immediately after an observation, our experiment required participants to cache a sequence of social information until they were asked to make a decision. Thus, a sensible

strategy based on our model would be to maintain an expectation about the unknown state and sequentially update the expectation using either the unweighted (in S1) or RD-weighted (in S2 and S3) prediction error signals each time an observation is witnessed.

This hypothesis immediately led to two neural predictions. First, signals reflecting the value estimate for updated belief expectation (E_{new}; Fig. 2b) might be represented in brain regions previously implicated in tracking expectations, like the OFC or ventromedial prefrontal cortex (VMPFC)[34]. That is, in addition to the classic learning signals for belief expectations of the currently chosen option at the choice time (Extended Data Fig. 7a), we would also expect—at observation times—the neural representation of the value estimate for the updated E_{new} associated with the option previously selected by the observer. Similar to aPE signals, we hypothesized that signals related to E_{new} estimates would be seen on an observation-by-observation basis and across three stages nonselectively. The second prediction was motivated by the DeGroot model that, compared with S1, incorporating an aPE signal into belief expectation in S2 and S3 would involve additional modulatory inputs. Thus, regions representing E_{new} estimates might demonstrate increased functional connectivity in S2 and S3 compared with S1, with regions related to tracking, representing or implementing modulatory signals in service of learning, such as the dACC.

In line with the first prediction, we found a strong positive correlation between E_{new} estimates and activity in a number of brain regions including the VMPFC, even after controlling for the potentially confounding variables such as aPE estimates and RD (Fig. 7a and Supplementary Table 5; GLM5; Methods) and other nuisance effects in the same regression model (cluster-wise FWE-corrected $P < 0.05$, with cluster-forming threshold $P_{unc.} < 0.001$; GLM6; Methods). As predicted, we observed stable neural representation of E_{new} estimates across S1, S2 and S3, such that within-subject comparisons identified no significant difference in neural responses to E_{new} estimates across stages either within the VMPFC cluster (Fig. 7b) or at the whole-brain level (cluster-wise FWE-corrected $P < 0.05$, with cluster-forming threshold $P_{unc.} < 0.001$).

To test the second prediction, we performed an ROI-based psychophysiological interaction (PPI) analysis to investigate whether the dACC would show differential functional coupling with the VMPFC when RD was versus was not needed for scaling aPE signals. The PPI analysis was performed between the identified VMPFC region and the independently defined ROI of dACC as used in the RD analyses (Methods). By comparing the connectivity strength averaged over S2 and S3 against that in S1, we observed increased coupling between the VMPFC and dACC, with no significant difference in the coupling effect sizes between S2 and S3 (Fig. 7c). These findings were further supported by an exploratory whole-brain PPI analysis (Extended Data Fig. 10).

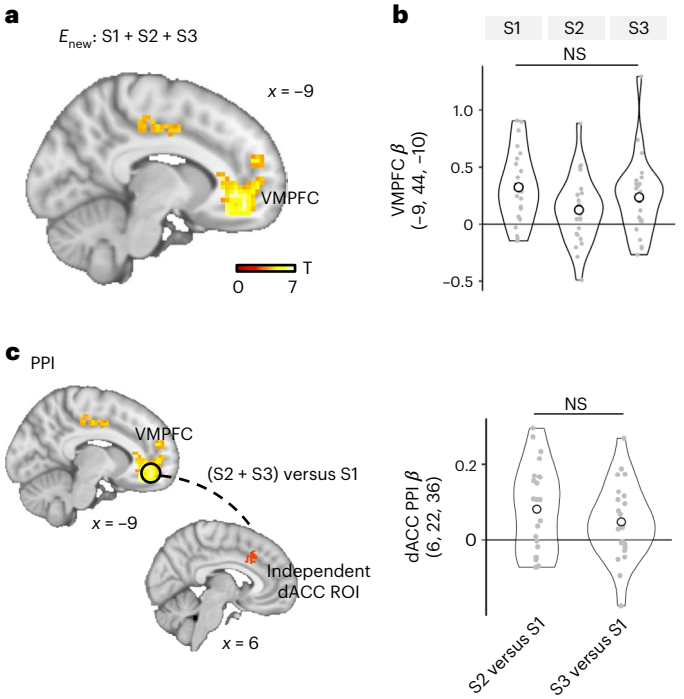


Fig. 7 | VMPFC tracks the value estimate of updated E_{new} at the time of observation in S1, S2 and S3. a, Statistical parametric map with respect to E_{new} estimates at observation onsets, computed by averaging the regression coefficients for E_{new} across S1, S2 and S3 for each subject (MNI: x, y and z, = −9, 44 and −10; cluster-wise FWE-corrected $P < 0.05$, with cluster-forming threshold $P_{unc.} < 0.001$). **b**, Similar effect sizes in the VMPFC encoding of E_{new} estimates across stages, as demonstrated by violin plots for the distribution of neural β values for E_{new} estimates (one-way repeated-measures ANOVA, $F(2,48) = 2.33$, $P = 0.108$). The neural β values were extracted for each separate stage from the significant VMPFC cluster as identified in **a**. **c**, Increased functional connectivity between the independent dACC ROI and the VMPFC ROI at observation onsets in S2 and S3, relative to that in S1 ($\beta = 0.10 \pm 0.02$, two-sided t test, $t_{24} = 4.41$, $P = 1.88 \times 10^{-4}$; Methods). Left: the dACC cluster is the same independently defined ROI. The VMPFC region is defined as a 6-mm sphere around the peak activation as identified in **a**. Right: no systematic difference in the effect sizes of functional coupling between S2 and S3 ($\beta = 0.03 \pm 0.02$, two-sided paired t test, $t_{24} = 1.35$, $P = 0.188$). Each dot represents an fMRI subject ($N = 25$).

Discussion

Information flowing in a large-scale, interconnected society is often entangled, conflated and sometimes superfluous[6,16]. This poses a computational challenge for social learning, during which agents need to reconcile disparate sources of signals based on their informativeness. Previous RL research on individual learning in nonsocial contexts has shown that humans can accurately estimate how relevant a learning signal is in predicting the future and use this estimate to adjust RL learning rates[23–25]. On social networks, however, optimally evaluating the predictive value of each observation is cognitively demanding, sometimes even prohibitive[4,17]. Indeed, failure to effectively aggregate information from connected peers has long been hypothesized to underlie herding, social influence biases, misinformation propagation and other forms of collective maladaptation.

Combining fMRI, formal theories of social and observational learning and social network analysis, we explored the possibility that, to balance computational costs, the brain approximates the relative informativeness of a social signal based on the structural properties of the network that routes information transmission. Grounding the DeGroot learning heuristic into an RL-inspired account used previously for action observational learning, the study provided behavioral and neural evidence that learning in complex, interconnected environments can be realized by an error-driven process reminiscent of those seen in non-networked environments. Importantly, the learning rate fluctuated according to a signal related to network degree centrality, indexed by the dACC activity at the time of witnessing others' actions, but only insofar as the social observations varied in their informativeness.

The observed dACC response to the degree centrality of the observee relative to that of the observer in S2 and S3 is consistent with past evidence showing a key role of this region in facilitating behavioral flexibility and adjusting RL learning rates[24,26,35–38]. Importantly, our data emphasized the absence of a dACC response to the same RD signal in S1, when the network structure was irrelevant to learning. This absence argues against the possibility that the dACC engagement identified in S2 and S3 was due to some low-level visual processing of the network displays, or due to other more general dACC functions—such as detecting errors or monitoring social conflicts[35,44]—that would be involved across all learning stages nonselectively. Alternatively, the selective

dACC encoding is consistent with a broader proposal regarding this region, suggested by past neurophysiological and neuroimaging evidence, as representing task-relevant (but not irrelevant) information for supporting behavioral changes and guiding appropriate action selection[35]. Our data thus implicate the involvement of a high-level, controlled process in evaluating the sources of social information in service of learning, and argue against a model of blind, automatic discrimination among social contacts in explaining social information aggregation. More broadly, the observed adjustment between the initial and latter stages of learning echoes previous studies of news circulation on social media, which have emphasized the importance of separating the spread of first- versus secondhand information in understanding phenomena such as rumor dissemination and amplification[16,45].

While our findings highlighted a role of the dACC specific to S2 and S3, we also observed that, in S1, activity in the PCC, precuneus and visual cortex correlated with measures of degree centralities. One possibility is that the observed S1 activation is associated with the recognition or representation of network features, which facilitates the flexible usage of those features in the latter stages. Indeed, the loci of S1 activation were similar to those implicated in encoding the centralities and other characters of one's acquaintances on the real-world network, when the individual was required to view pictures or videos of those acquaintances[40–42]. Alternatively, it is also possible that the S1 activation reflects some low-level processing of the network stimuli (for example, visual processing), which typically shows a more pronounced activation when the stimulus is new (as in S1) than when the stimulus has been recently processed (as in S2 and S3)[46]. Future studies need to clarify these possibilities by using more naturalistic experimental approaches that eliminate the visual display of network geometries from the behavioral learning task.

Degree centrality has long been hypothesized to have a close relationship with social influence in small group interaction[1,32]. Our finding of degree-modulated learning is consistent with two broad accounts previously proposed for how network centrality affects learning. The first has its basis in human and nonhuman studies that emphasize the role of the structural position in social behavior, suggesting that structurally comparable individuals are facing similar interacting environments, therefore exhibiting similar behavior toward one another[15,18]. In the context of learning, network positions constrain the opportunities to obtain information from peers and the degree centrality quantifies such network-based individual differences. A second, but not mutually exclusive, possibility has its basis in the dynamic nature of network topology: knowledgeable or successful individuals tend to become highly connected, thus one's degree centrality may serve to signal an individual's capability or social status to other individuals[18,47]. Under this possibility, social animals may have evolved to preferentially follow the more 'connected' or 'prestigious' cospecies, even in controlled experiments where the network structure is fixed and locations are randomized. Compatible with these proposals, our data additionally highlighted a dual effect of centrality on learning: a higher degree centrality not only amplified one's social influence, but also reduced one's susceptibility to others' influence. This duality corroborates evidence from popular social media, demonstrating that more influential individuals are usually less susceptible to peers' influence, compared with their less influential counterparts[48]. The results further point to a possibility that, while social influence and susceptibility to social influence are often considered as distinct personal attributes[44], they may be jointly affected by an internal learning system, which approximates the predictive value of others' information relative to one's own, to cope with the complexity of the social environment.

Previous research on the neurocomputation of social learning has typically focused on highly simplified social settings, leaving open whether and how putative mechanisms identified in simplistic setups can support behavior in more complex, ecologically relevant environments. As an initial step toward addressing this issue, our study raises key questions for future research. First, owing to the fundamental role of degree centrality in network analyses and its close connections with a range of network characteristics, we cannot rule out the possibility that alternative network features may also affect learning. In addition to the degree centrality, which parameterizes the immediate effect of social influence, learning may be affected by, for example, measures proposed to capture the long-term (eigenvector centrality), sequential (closeness), circular (clustering coefficient), global or local mediation (betweenness or constraint coefficient) effects seen in information transmission[1]. Our focus on degree centrality reflects the assumption that the brain may be more sensitive to simple, straightforward topological features, especially in complex decisions. Indeed, across analyses, there was no evidence that alternative metrics outperformed RD in explaining behavioral data (Supplementary Table 4). Future investigation is needed to more firmly isolate and compare the potential influences of various network features at the behavioral and neural levels.

Second, across network structures, we found no systematic differences in how well the DeGroot learning model explained data (Supplementary Fig. 5). Yet, the network geometries included in the study were but a sample of immense possibilities in the real world, leaving open the generality and scalability of the proposed model. It is possible that the brain may follow other learning algorithms when facing a different set of networks—for example, deploying Bayesian strategies when making decisions in a line, one of the simplest forms of directed network[27]. Hybrid learning is also possible according to a recent behavioral study suggesting a mixture of Bayesian and DeGroot learning in a relatively more educated (but not less educated) sample[29]. Also, our experiment focuses on relatively small, seven-node networks, and does not directly speak to larger, more naturalistic settings. We speculate that, by relying merely on local information, the proposed model may be particularly suitable for scaling up, as individuals in large social groups typically have access only to local information, but not global knowledge about the entire network. A deeper investigation of the tradeoff between learning complexity and effectiveness[49] will be valuable for understanding whether, and under what circumstances, the current findings can be extended to greater varieties of network structures.

Another possible issue arises concerning the explicit display of sociograms in the behavioral task. Social animals normally do not have the graph visualization of social relationships during decision-making and usually acquire such information through a slow learning process. We examined this more naturalistic setup in a follow-up behavioral study ($N = 91$), where subjects first learned network structures and then played the same distributed learning game but with no ostensive network display throughout the experiment. The choice data were largely consistent with those seen in the original experiment and were also best explained by the DeGroot learning model (Supplementary Note 3). It remains to be explored how the brain learns and represents the structure of social connections and how such representations support the dACC signaling in service of information integration.

Finally, our experiment employed simulated networks, where connections represent routes of information transmission. This stylized setup allows for isolating observability—a basic determinant of social information transmission—from other factors that are implied by interpersonal connections and might also affect learning. For example, in real-life situations, social connections reflect not just observability, but also social bonds, affiliation and other interpersonal ties[19]. Also, individuals' network locations are not completely random, but tend to signal certain characteristics of the individuals (for example, extraversion, social economic status, and so on)[18,47]. These differences raise questions for future research regarding whether and how the current findings can be extended to capture perceivers' intuitions about the functional consequences and correlates of centrality in the contexts where social connections imply more than just observability.

Social networks have been widely hypothesized to play a key role in many large-scale social phenomena, including vaccine hesitancy,

voting behavior and fake news proliferation, yet the exact mechanism by which social connections contribute to these phenomena remains unclear. The current study sheds light on this topic from a unique neurocognitive perspective, by elucidating how individuals experience and interact with a networked environment. Our data provide neural evidence for a bounded rational, network-dependent filtering of social information, which may result in the spread of misinformation and biased consensus among connected peers. More broadly, this work demonstrates the possibility of developing computationally tractable and neurobiologically plausible tools for investigating the complex interplay between social behavior and social embedding at the neural level, which may have the potential to translate upward for tackling phenomena in wider society.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41593-023-01258-y.

References

1. Wasserman, S. & Faust, K. *Social Network Analysis* (Cambridge Univ. Press, 1994).
2. Borgatti, S. P., Mehra, A., Brass, D. J. & Labianca, G. Network analysis in the social sciences. *Science* **323**, 892–895 (2009).
3. Momennejad, I. Collective minds: social network topology shapes collective cognition. *Phil. Trans. R. Soc. B* **377**, 20200315 (2022).
4. Jackson, M. O. *Social and Economic Networks* (Princeton Univ. Press, 2010).
5. Toelch, U. & Dolan, R. J. Informational and normative influences in conformity from a neurocomputational perspective. *Trends Cogn. Sci.* **19**, 579–589 (2015).
6. Lazer, D. M. J. et al. The science of fake news. *Science* **359**, 1094–1096 (2018).
7. Burke, C. J., Tobler, P. N., Baddeley, M. & Schultz, W. Neural mechanisms of observational learning. *Proc. Natl Acad. Sci. USA* **107**, 14431–14436 (2010).
8. Suzuki, S. et al. Learning to simulate others' decisions. *Neuron* **74**, 1125–1137 (2012).
9. Dunne, S. & O'Doherty, J. P. Insights from the application of computational neuroimaging to social neuroscience. *Curr. Opin. Neurobiol.* **23**, 387–392 (2013).
10. Behrens, T. E. J., Hunt, L. T., Woolrich, M. W. & Rushworth, M. F. S. Associative learning of social value. *Nature* **456**, 245–249 (2008).
11. Zhu, L., Mathewson, K. E. & Hsu, M. Dissociable neural representations of reinforcement and belief prediction errors underlie strategic learning. *Proc. Natl Acad. Sci. USA* **109**, 1419–1424 (2012).
12. Jiang, Y., Wu, H.-T., Mi, Q. & Zhu, L. Neurocomputations of strategic behavior: from iterated to novel interactions. *Wiley Interdiscip. Rev. Cogn. Sci.* **13**, e1598 (2022).
13. Sutton, R. S. & Barto, A. G. *Reinforcement Learning* (MIT, 1998).
14. Zhang, L. & Gläscher, J. A brain network supporting social influences in human decision-making. *Sci. Adv.* **6**, eabb4159 (2020).
15. McPherson, M., Smith-Lovin, L. & Cook, J. M. Birds of a feather: homophily in social networks. *Annu. Rev. Sociol.* **27**, 415–444 (2001).
16. Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W. & Starnini, M. The echo chamber effect on social media. *Proc. Natl Acad. Sci. USA* **118**, e2023301118 (2021).
17. Mobius, M. & Rosenblat, T. Social learning in economics. *Annu. Rev. Econ.* **6**, 827–847 (2014).
18. Kulahci, I. G. & Quinn, J. L. Dynamic relationships between information transmission and social connections. *Trends Ecol. Evol.* **34**, 545–554 (2019).
19. Basyouni, R. & Parkinson, C. Mapping the social landscape: tracking patterns of interpersonal relationships. *Trends Cogn. Sci.* **26**, 204–221 (2022).
20. Molavi, P., Tahbaz-Salehi, A. & Jadbabaie, A. A theory of non-Bayesian social learning. *Econometrica* **86**, 445–490 (2018).
21. DeGroot, M. H. Reaching a consensus. *J. Am. Stat. Assoc.* **69**, 118–121 (1974).
22. Golub, B. & Jackson, M. O. Naïve learning in social networks and the wisdom of crowds. *Am. Econ. J. Microecon.* **2**, 112–149 (2010).
23. Dayan, P., Kakade, S. & Montague, P. R. Learning and selective attention. *Nat. Neurosci.* **3**, 1218–1223 (2000).
24. Behrens, T. E. J., Woolrich, M. W., Walton, M. E. & Rushworth, M. F. S. Learning the value of information in an uncertain world. *Nat. Neurosci.* **10**, 1214–1221 (2007).
25. Radulescu, A., Shin, Y. S. & Niv, Y. Human representation learning. *Annu. Rev. Neurosci.* **44**, 253–273 (2021).
26. Leong, Y. C., Radulescu, A., Daniel, R., DeWoskin, V. & Niv, Y. Dynamic interaction between reinforcement learning and attention in multidimensional environments. *Neuron* **93**, 451–463 (2017).
27. Anderson, L. R. & Holt, C. A. Information cascades in the laboratory. *Am. Econ. Rev.* **87**, 847–862 (1997).
28. Grimm, V. & Mengel, F. Experiments on belief formation in networks. *J. Eur. Econ. Assoc.* **18**, 49–82 (2018).
29. Chandrasekhar, A. G., Larreguy, H. & Xandri, J. P. Testing models of social learning on networks: evidence from two experiments. *Econometrica* **88**, 1–32 (2020).
30. Paluck, E. L., Shepherd, H. & Aronow, P. M. Changing climates of conflict: a social network experiment in 56 schools. *Proc. Natl Acad. Sci. USA* **113**, 566–571 (2016).
31. Paluck, E. L. & Shepherd, H. The salience of social referents: a field experiment on collective norms and harassment behavior in a school social network. *J. Pers. Soc. Psychol.* **103**, 899–915 (2012).
32. Friedkin, N. E. *A Structural Theory of Social Influence* (Cambridge Univ. Press, 1998).
33. Ramsey, R., Kaplan, D. M. & Cross, E. S. Watch and learn: the cognitive neuroscience of learning from others' actions. *Trends Neurosci.* **44**, 478–491 (2021).
34. Niv, Y. Reinforcement learning in the brain. *J. Math. Psychol.* **53**, 139–154 (2009).
35. Heilbronner, S. R. & Hayden, B. Y. Dorsal anterior cingulate cortex: a bottom-up view. *Annu. Rev. Neurosci.* **39**, 149–170 (2016).
36. Sheth, S. A. et al. Human dorsal anterior cingulate cortex neurons mediate ongoing behavioural adaptation. *Nature* **488**, 218–221 (2012).
37. Jocham, G., Neumann, J., Klein, T. A., Danielmeier, C. & Ullsperger, M. Adaptive coding of action values in the human rostral cingulate zone. *J. Neurosci.* **29**, 7489–7496 (2009).
38. O'Reilly, J. X. et al. Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proc. Natl Acad. Sci. USA* **110**, E3660–E3669 (2013).
39. Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C. & Wager, T. D. Large-scale automated synthesis of human functional neuroimaging data. *Nat. Methods* **8**, 665–670 (2011).
40. Parkinson, C., Kleinbaum, A. M. & Wheatley, T. Spontaneous neural encoding of social network position. *Nat. Hum. Behav.* **1**, 0072 (2017).
41. Zerubavel, N., Bearman, P. S., Weber, J. & Ochsner, K. N. Neural mechanisms tracking popularity in real-world social networks. *Proc. Natl Acad. Sci. USA* **112**, 15072–15077 (2015).

42. Morelli, S. A., Leong, Y. C., Carlson, R. W., Kullar, M. & Zaki, J. Neural detection of socially valued community members. *Proc. Natl Acad. Sci. USA* **15**, 201712811–201712816 (2018).
43. Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S. F. & Baker, C. I. Circular analysis in systems neuroscience: the dangers of double dipping. *Nat. Neurosci.* **12**, 535–540 (2009).
44. Cascio, C. N., Scholz, C. & Falk, E. B. Social influence and the brain: persuasion, susceptibility to influence and retransmission. *Curr. Opin. Behav. Sci.* **3**, 51–57 (2015).
45. Sutton, J. et al. A cross-hazard analysis of terse message retransmission on Twitter. *Proc. Natl Acad. Sci. USA* **112**, 14793–14798 (2015).
46. Grill-Spector, K., Henson, R. & Martin, A. Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn. Sci.* **10**, 14–23 (2006).
47. Almaatouq, A., Noriega-Campero, A., Alotaibi, A., Krafft, P. M. & Pentland, A. Adaptive social networks promote the wisdom of crowds. *Proc. Natl Acad. Sci. USA* **117**, 11379–11386 (2020).
48. Aral, S. & Walker, D. Identifying influential and susceptible members of social networks. *Science* **337**, 337–341 (2012).
49. Bossaerts, P. & Murawski, C. Computational complexity and human decision-making. *Trends Cogn. Sci.* **21**, 917–929 (2017).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2023

Methods

Participants

A total of 217 healthy volunteers (127 females; mean age ± s.d. = 20.25 ± 2.68 years) were recruited from Peking University, China. Of these participants, 31 completed the distributed learning experiment inside an fMRI scanner (14 females; age = 20.16 ± 1.97 years), while the remaining 186 subjects performed the same experiment outside the scanner. All participants reported having normal or corrected-to-normal eye vision, no color blindness, and no history of neurological or psychiatric illnesses. All fMRI participants were right-handed. All subjects provided informed consent, and the study was approved by the Committee for Protection of Human Subjects at Peking University, China.

A total of 209 subjects were included in the behavioral analyses, and 25 of these participants were also used for the fMRI data analyses. Specifically, 6 out of 186 behavioral subjects were excluded from analyses due to insufficient variations in choice behavior ($N = 4$; repeating the same decision based on their private signals in 95% of decisions), abnormal reaction time ($N = 1$; 5 s.d. above the group mean), or reporting strong disbelief in the experimental setting in the postexperiment survey ($N = 1$). Of 31 fMRI participants, 4 were excluded from neuroimaging analyses but included in behavioral analyses due to excessive motion, and 2 were excluded from both neural and behavioral analyses due to scanning issues that did not interfere with their group members' experiment (flickering display, $N = 1$; unable to finish scanning but able to complete the experiment outside the scanner, $N = 1$).

Procedure

Participants were assigned randomly into groups of seven for the experiment. Upon arrival, seven participants received instructions together to ensure that the instructions were identical and known to be identical by all group members. The instructions explained the task and the graph display of social networks, including what a node and a link represent on a graph (Supplementary Note 1). Subjects were truthfully instructed that (1) an individual on a network can see the decisions made by all her immediate connections, (2) an individual's decision will be seen by all her immediate connections and (3) an individual's decision is not directly observable beyond her immediate connections. To ensure comprehension, subjects were required to complete a quiz and three practice games on networks whose structures were different from those used in the actual experiment.

At the beginning of the distributed learning experiment, one subject within a group was moved into the scanning room, while the remaining six subjects were seated in an adjacent computer room with dividers to prevent communication and seeing each other's computer screen. All seven participants were connected via an intranet. Subjects played 40 separate games, divided into three (scanning) sessions with short breaks in between. The order of the games was randomized across subject groups. As a game can proceed to the next stage only after all seven group members have submitted their decisions, the participants were encouraged (but not mandated) to select their choices within 5 s, so that the total scanning time could be reasonably constrained. To incentivize subjects to guess as accurately as possible in all decisions during the experiment, subjects were informed truthfully and identically that each participant would be rewarded based on the accuracy of 30 randomly selected decisions from all her decisions (that is, four decisions in each of 40 games), plus a show-up fee (60 CNY for behavioral participants and 150 CNY for fMRI participants). To prevent intergame learning, no feedback on choice accuracy was provided to subjects during the experiment. A schematic representation of the distributed learning game and the timeline of the experiment are shown in Supplementary Fig. 1 and dynamically illustrated in Supplementary Video 1.

Networks and private signals

Each distributed learning game entailed a predetermined combination of a network structure and private signal for every node on the network. Extended Data Fig. 1a shows all 40 sets of experimental stimuli used in the experiment. These stimuli were selected based on an exhaustive simulation over all possible seven-node, connected, unweighted and undirected networks (853 nonisomorphic networks in total) and all possible combinations of private signals ($2^7 = 128$ possibilities) on each of 853 networks. The stimuli were chosen with the goal of (1) ensuring the dissociation power in simulated choices across three learning models (DeGroot, Bayesian, and Baseline learning models), while (2) maintaining reasonable variations across selected networks in the topological characteristics relevant for learning and information propagation.

To quantify these objectives, we computed the following nine indices for each of 128 × 853 candidate stimuli. The first three were discriminability indices for simulated choices between any two basic learning models (Bayesian versus DeGroot, Bayesian versus Baseline and DeGroot versus Baseline). These metrics were defined as the proportion of choices simulated by one model, given a network structure and private signals on the network, that disagreed with the simulation from the other model. The metrics were averaged over decisions in S2 and S3 and across seven decision-makers simulated from their respective network locations. The initial guess and S1 decision were excluded from the discriminability indices because three models make same predictions in those choices. We ensured that, for the selected stimuli, the average discriminability between any two models was as high as the top 3% according to the empirical distribution of each discriminability index over all candidate stimuli (Extended Data Fig. 1b).

The second category included six widely used network-level parameters relevant for information transmission[50–53]. They are (1) average degree centrality across all nodes on a network, (2) standard deviation of degree centrality across all nodes on a network, (3) Freeman centralization, (4) diameter, (5) average path length and (6) the overall clustering coefficient. For each of these indices, we ensured there was no systematic mismatch in the distribution range between the selected networks and the pool of all possible candidate networks (Extended Data Fig. 1c).

Network locations for fMRI/behavioral participants were determined pseudorandomly on each network structure. A post hoc check on the locations of fMRI subjects was performed to ensure sufficient variations in the RD values of the fMRI participants' direct neighbors. Notably, we observed no systematic difference in the choice behavior between the fMRI and behavioral participants in the experiment (Supplementary Table 3).

To allow for evaluating how node properties affect learning while controlling for the potential influence of the overall network topology, five network structures were used twice in the experiment (last ten stimuli in Extended Data Fig. 1a). Each structure was associated with a reallocation of participants and private signals on the network (see Supplementary Fig. 6 for within-network comparisons). Unless otherwise specified, all analyses in the current study were carried out on all 40 games. Behavioral and neural results remained unchanged in the analyses of 35 games, excluding the data from five games containing the network structures with which the participants interacted for the second time during the experiment.

DeGroot learning model

In each game, a DeGroot learner is assumed to form a private belief about the underlying state based on the private signal she received at the beginning of the game, and then update this belief continuously in an error-driven manner, every time she witnesses an observation from the social network. The model assumes differential updating rules in S2 (S3) versus S1. In S1, the update follows:

$$E_{i,1}^{k}(\mathrm{B}) = E_{i,1}^{k-1}(\mathrm{B}) + \alpha\left[\mathrm{B}_{i,1}^{k} - E_{i,1}^{k-1}(\mathrm{B})\right],$$

$$E_{i,1}^{k}(\mathrm{Y}) = E_{i,1}^{k-1}(\mathrm{Y}) + \alpha\left[\mathrm{Y}_{i,1}^{k} - E_{i,1}^{k-1}(\mathrm{Y})\right],$$

where $E^k_{i,1}(\mathrm{B})$ ($E^k_{i,1}(\mathrm{Y})$) is agent i's belief that the underlying state of the game is blue (yellow) after witnessing the kth observation in S1, α is the learning rate in S1, and $\mathrm{B}^k_{i,1}$ ($\mathrm{Y}^k_{i,1}$) is the kth observation of agent i at S1, which, without loss of generality, is set to be 1, if the observee selects blue (yellow), and −1 otherwise. Under the experimental setting, $E^k_{i,s}(\mathrm{B}) = -E^k_{i,s}(\mathrm{Y})$. Agent i's belief at the beginning of S1 is determined by a free parameter (γ), and her belief at the end of S1 is denoted as $E^{k\max}_{i,1}(\mathrm{B})$ and $E^{k\max}_{i,1}(\mathrm{Y})$, where $k$max refers to the last observation in the given stage.

In S2 and S3, the update is assumed to follow:

$$E^k_{i,s}(\mathrm{B}) = E^{k-1}_{i,s}(\mathrm{B}) + \beta \times \mathrm{RD}^k_{i,s} \times \left[\mathrm{B}^k_{i,s} - E^{k-1}_{i,s}(\mathrm{B})\right],$$

$$E^k_{i,s}(\mathrm{Y}) = E^{k-1}_{i,s}(\mathrm{Y}) + \beta \times \mathrm{RD}^k_{i,s} \times \left[\mathrm{Y}^k_{i,s} - E^{k-1}_{i,s}(\mathrm{Y})\right],$$

where $s = 2$ or 3 (that is, S2 or S3), β is the learning rate shared by S2 and S3 (see Alternative models in Methods and Supplementary Table 4 for evidence supporting this specification), and $\mathrm{RD}^k_{i,s}$ represents the relative degree, defined as the degree centrality of the observee k relative to the total local degree of agent i (see Fig. 2b for illustration). Learning in S2 and S3 runs continuously from that of the previous stage, such that $E^0_{i,s}(\mathrm{B}) = E^{k\max}_{i,s-1}(\mathrm{B})$ and $E^0_{i,s}(\mathrm{Y}) = E^{k\max}_{i,s-1}(\mathrm{Y})$ for each stage s.

To convert agent i's belief into choice probabilities, the model assumes that the agent's choices follow a standard softmax function used widely in learning and decision-making research:

$$P(A_{i,s} = \mathrm{B}) = \frac{1}{1 + \exp\left(-\lambda E^{k\max}_{i,s}(\mathrm{B})\right)},$$

where $A_{i,s}$ is agent i's choice at stage s, $s = 1,2,3$, and λ is a free parameter for the inverse temperature capturing choice randomness. Altogether the model contains four free parameters: learning rates α and β, initial belief parameter γ and inverse temperature λ.

For fMRI analyses, we focused on the aPE estimate associated with the observer's previous decision in a game. For example, if a participant selected blue in the previous stage of a game, then, upon witnessing the kth observation in current stage s, we looked for the neural correlates of aPE defined as $\mathrm{B}^k_{i,s} - E^{k-1}_{i,s}(\mathrm{B})$ (Fig. 2b).

Model estimation

We calibrated the DeGroot learning model with participants' behavior by maximizing the log-likelihood of the model prediction at both individual and network levels. For individual estimation, we assumed that a subject's choices across games were generated by the same set of DeGroot learning parameters, and applied the maximum likelihood estimation with a grid search for each parameter. Specifically, we fit the model for each subject by maximizing the log of choice probability of observed data, $\sum_g \left[\log\left(P(A_{i,\mathrm{ini}}|g)\right) + \sum_{s=1,2,3} \log\left(P(A_{i,s}|g)\right)\right]$, pooled over game g given a subject i.

To account for the potential variations in learning parameters across networks, we calibrated the model by pooling data over all subject groups given each game. That is, we assumed that each game was associated with a set of DeGroot learning parameters shared by all subjects regardless of their groups and network locations. We fit the model by maximizing the log of choice probability of the observed data, $\sum_i \left[\log\left(P(A_{i,\mathrm{ini}}|g)\right) + \sum_{s=1,2,3} \log\left(P(A_{i,s}|g)\right)\right]$, pooled over subject i given a game g.

Unless otherwise stated, behavioral and neuroimaging analyses presented in the current study were based on results derived from individual-level estimations. Best-fitting parameters from both individual- and network-level estimations are summarized in Supplementary Table 2.

Model evaluation

For in-sample model comparison, we computed the standard BIC based on the in-sample model fits, estimated at either the individual (Extended Data Fig. 3a, left) or network level (Extended Data Fig. 3b, left). For out-of-sample comparison, we evaluated the between-subject out-of-sample prediction accuracy for candidate models (Extended Data Fig. 3a, middle). This involved dividing participants randomly into five groups, and predicting the choice behavior of one group of participants using the median of the parameters computed from the participants in the other four groups. We computed the mean prediction accuracy based on 1,000 random groupings (to guard against the potential biases in assigning subjects into groups) and, within each grouping, five times of leaving one group out (to guard against the potential biases in assigning hold-out samples). Finally, we performed Bayesian model selection among the proposed and competing models, assuming that there is a fixed but unknown distribution of different underlying models across individuals. We first inverted models using a variational Bayes approach implemented in the VBA toolbox[54]. The model log-evidence estimated for each participant and each model was then submitted to a group-level random effects analysis, assuming the occurrence of the model for each participant as a multinomial random variable with a Dirichlet conjugate prior (Extended Data Fig. 3a, right). This analysis generated exceedance probability that measures the plausibility that a given model is more frequently implemented by participants than any other model in the comparison set. Similar procedures were also used for network-level estimation results (Extended Data Fig. 3b, middle and right).

Alternative models

The proposed DeGroot learning model was compared against two sets of competing models. First, to evaluate how sophisticated participants are in integrating networked information, we compared the DeGroot learning model against the Baseline and Bayesian learning models, which either completely ignores or makes rational use of the network structure during learning. Second, to validate the specific algorithmic configuration under the DeGroot learning hypotheses, we compared the proposed model against 18 competing models, all maintaining the main DeGroot learning assumptions but differing in the specific implementations (see model comparison in Fig. 3, Extended Data Fig. 3 and Supplementary Table 4).

Baseline learning model. The Baseline learning model entails an error-driven process with no degree-modulation effect:

$$E^k_{i,s}(\mathrm{B}) = E^{k-1}_{i,s}(\mathrm{B}) + \alpha\left[\mathrm{B}^k_{i,s} - E^{k-1}_{i,s}(\mathrm{B})\right],$$

$$E^k_{i,s}(\mathrm{Y}) = E^{k-1}_{i,s}(\mathrm{Y}) + \alpha\left[\mathrm{Y}^k_{i,s} - E^{k-1}_{i,s}(\mathrm{Y})\right],$$

where notations are the same as those in the DeGroot learning model. The model contains three free parameters, a learning rate α that is constant across stages, an initial belief parameter γ and inverse temperature λ.

Bayesian learning. A key difference between the Bayesian and DeGroot (Baseline) learning models is that Bayesian learning is not a Markovian process, and a Bayesian learner updates her belief based on all the information available to her at the decision time, which includes her private signal, all her past actions and observations in a game. The study examines two variants of Bayesian learning models: the baseline and noisy Bayesian learning. The baseline Bayesian model demonstrates theoretically appealing properties (Supplementary Fig. 2), whereas the noisy Bayesian model provides better fit to the experimental data (Extended Data Fig. 3). These models are built on a recursive algorithm[28,55] under the assumption that all agents on the network are

Bayesian learners, and know that others are Bayesian (and know that others know about this, and so on)[56,57].

Intuitively, a Bayesian learner needs to infer the private signal received by each of seven agents on the network, or to infer $\boldsymbol{d} = (d_1, \ldots, d_7)$, where d_i represents agent i's private signal (blue or yellow ball). The agent starts with a set of all possible $\boldsymbol{d}$ vectors, denoted as D, which contains 2^7 possible combinations of private signals on a specific seven-node network, and gradually eliminates the $\boldsymbol{d}$ vectors that are inconsistent with the information history of the agent in the game. At the initial guess of a game ($t = 0$), an agent i's information history contains only her private signal, $I_{i,0} = \{\omega_i\}$, where $I_{i,0}$ denotes i's information history at $t = 0$, and ω_i denotes i's actual private signal. Based on this information, agent i will reduce D to $D_{i,0}$ by deleting $\boldsymbol{d}$ vectors with $d_i \neq \omega_i$ (that is, removing half of the space in D that contains the opposite color of ω_i). Given $D_{i,0}$, agent i computes her posterior belief about the underlying state θ using the Bayes' rule: $Q_{i,0}(I_{i,0}, \theta) = \frac{\sum_{\boldsymbol{d} \in D_{i,0}} P(\boldsymbol{d}|\theta)P(\theta)}{\sum_{\theta' \in \{B,Y\}} \sum_{\boldsymbol{d} \in D_{i,0}} P(\boldsymbol{d}|\theta')P(\theta')}$,

where $P(\theta) = 0.5$ is the prior belief in support of a possible state θ. A standard softmax function is used for mapping an agent's posterior belief into decisions.

In S1 ($t = 1$), agent i's information history becomes $I_{i,1} = \{\omega_i, A_{i,0}, A_{j,0}\}$, where $A_{i,0}$ and $A_{j,0}$ are the initial guesses selected by i and i's direct neighbor j, respectively. Without loss of generality, here we assume that i has only one direct neighbor on the network, but the same algorithm will apply when i has multiple network neighbors. Starting from $D_{i,0}$, agent i eliminates the $\boldsymbol{d}$ vectors that are incompatible with her observation $A_{j,0}$, by simulating what j would have selected in the initial guess given each possible $\boldsymbol{d}$ in $D_{i,0}$. Algorithmically, this can be achieved by carrying out the same computational procedure as that in $t = 0$, but from j's perspective, conditional on each $\boldsymbol{d}$ in $D_{i,0}$. To map others' posterior belief to others' possible decisions, the baseline Bayesian model assumes that an agent believes that others will always select the option with the highest posterior probability, a deterministic assumption that is relaxed in the noisy Bayesian model.

In S2 ($t = 2$) and S3 ($t = 3$), the computational procedure is essentially the same as that in $t = 1$, but with more iterations. For example, in S2, agent i needs to simulate neighbor j's choice in $t = 1$ conditional on each possible $\boldsymbol{d}$ in $D_{i,1}$, which further involves the recursive computation of actions that would have been selected by neighbor j's neighbors in $t = 0$. By gradually taking into account the topological structure of the network, the algorithm allows an agent to infer the choices of others locating further and further away from her on the network. Importantly, by conditioning on each possible initial state, $\boldsymbol{d} = (d_1, \ldots, d_7)$, the algorithm allows for predicting the decision by each network agent in each stage using the Bayes' rule, circumventing complex computations of correlated choices on networks when $\boldsymbol{d}$ is unknown (see Supplementary Note 2 for formal modeling).

The noisy Bayesian model differs from the baseline Bayesian model in assuming that the probability of a neighbor choosing a particular action is proportional to this neighbor's posterior probability in support of that action (that is, $P(A_{j,t} = \theta) \propto Q_{j,t}(I_{j,t}, \theta)$). The baseline Bayesian model was used for designing and selecting the experimental stimuli (Fig. 2a and Extended Data Fig. 1). Both models were used in behavioral data estimation. Due to the superior explanatory power of the noisy Bayesian model relative to the baseline variant, we presented the results from the noisy Bayesian model in behavioral data analyses in main texts (Fig. 3c, Supplementary Fig. 4 and 5 and Supplementary Tables 2 and 3), supplemented by the comparison of estimation results between the baseline Bayesian and other models (Extended Data Fig. 3).

DeGroot learning with alternative node parameters. To evaluate the assumption that DeGroot learning is modulated by the degree centralities of the observee relative to the observer, we replaced the value of RD in the proposed model with the following alternative node properties used widely in social network analysis[1,4,58]: (1.1) eigenvector centrality, (1.2) betweenness centrality, (1.3) closeness centrality, (1.4) constraint (an inverse measure of network brokerage, 1 − constraint was used in the model) and (1.5) node-level clustering coefficient (Supplementary Table 4). Similar to RD, we calculated the relative value for each of these measures, by normalizing each measure associated with the observee with respect to that of the observer and the observer's direct neighbors on a network. The resulting ratio was then used in a new learning model as the modulator for the learning rate in S2 and S3 (but not in S1), in the same way as in our proposed DeGroot learning model.

DeGroot learning with alternative degree normalization. To evaluate whether and how the observee's degree centrality was normalized when being incorporated in the learning process, we replaced RD in the proposed model with (2.1) the non-normalized degree centrality of the observee, (2.2) the ratio of degree centrality between the observee and the observer, (2.3) the ratio of degree centrality between the observee and the observer's direct neighbors (excluding the observer's own degree from the denominator of RD) and (2.4) the ratio of degree between the observee and the total number of unique connections within the observer's local neighborhood (which avoids double-counting connections shared between two individuals within the neighborhood in the normalization term).

DeGroot learning with alternative stage dependency. To evaluate the assumption concerning the stage-varying modulation effects, we constructed alternative models positing that the learning rate was scaled by the value of RD in either (3.1) all three stages, (3.2) S2 only or (3.3) S3 only.

DeGroot learning with alternative learning rate specification. To evaluate whether learning across S1, S2 and S3 was associated with the same or different learning rates, we examined alternative specifications in which learning was either associated with a common learning rate across all stages (4.1), or a unique learning rate for each learning stage (4.2). These alternative models maintained the assumption that the learning rate in S2 and S3 (but not S1) was additionally scaled by the respective value of RD, as in the proposed model.

DeGroot learning with alternative observational learning strategy. The last set of models test an alternative possibility that an individual learns only from her MC neighbor within her neighborhood, but not from the other direct neighbors (non-MC neighbors). We formulated models assuming that an individual learns from the MC neighbor only when the MC neighbor has more connections than the observer (5.1), or regardless of how connected the MC neighbor is (5.2). We assume learning is weighted by either the ratio between the MC observee's degree centrality and the total local degree (in 5.1 and 5.2) or the ratio between the MC observee's degree centrality and the observer's own degree centrality (5.3). Finally, rather than assuming zero influence by non-MC neighbors, we tested whether neighbors affect learning with differential strength, by incorporating separate learning rate parameters for the MC and non-MC neighbors (5.4).

fMRI data acquisition and preprocessing

We collected fMRI images using a 3 T Siemens Prisma scanner equipped with a 32-channel head coil at the Center for MRI Research of Peking University. Before the task, high-resolution structural T1-weighted scans were acquired using a magnetization-prepared rapid gradient echo sequence with the following parameters: repetition time = 2,530 ms, echo time = 2.98 ms, flip angle = 7°, field of view (FoV) = 224 × 256 mm^2, inter-slice gap = 0.5 mm, voxel size = 0.5 × 0.5 × 1 mm^3 and number of slices = 192. During the task, functional images were acquired using echo-planar T2* images with BOLD contrast, and angled 30° with respect to the anterior commissure–posterior commissure line to minimize susceptibility artifacts in the orbitofrontal area.

The scanning parameters were as follows: repetition time = 2,000 ms, echo time = 30 ms, flip angle = 90°, field of view = 224 × 224 mm^2, voxel size = 3.5 × 3.5 × 3.5 mm^3, inter-slice gap = 0.7 mm and number of slices = 33. Visual stimuli were presented using a mirror mounted on the MRI head coil, and responses were acquired via an MRI-safe button response pad.

Image preprocessing and analyses were performed in SPM12 (Wellcome Trust Center for Neuroimaging, Institute of Neurology, UCL; www.fil.ion.ucl.ac.uk/spm/software/spm12/). For each fMRI session, the preprocessing of neuroimaging data included, in order: slice time correction, motion correction (aligning to the mean image), coregistration, spatial normalization to the Montreal Neurological Institute (MNI) template (final image resolution of 3 × 3 × 3 mm^3) and spatial smoothing using a Gaussian kernel of 6-mm full width at half maximum. All images were high-pass filtered in the temporal domain (width 128 s), and autocorrelation of the hemodynamic responses was modeled as an AR(1) process.

fMRI data analysis

We implemented six GLMs for model-based fMRI analysis[59,60] used widely in neuroimaging studies of learning and decision-making. In all GLMs, a distributed learning game was modeled as a series of discrete events using stick functions. These included fixation onset, network display, initial guess onset, initial guess submission and—for each learning stage—the display onset of one's own choice, the onset of each and every observation, the onset of the reassessment decision and the submission of the reassessment. Parametric modulators such as the value estimates for aPE and belief expectation were derived using each fMRI participant's best-fitting parameters from the individual-level estimation of the DeGroot learning model. The parametric modulator of RD was calculated based on the locations of the observee and observer on the given network, independent of participants' behavior and model estimation.

We made the following assumptions in all GLM models. First, rather than modeling each observation onset as a separate GLM regressor, we entered the onsets of all observations of an fMRI participant within a learning stage into one GLM condition (regressor), with the across-observation variations in aPE, RD or other decision-related variables being captured by the respective parametric modulators. This allowed us to flexibly account for the varying numbers of observations of a subject across networks (that is, the inevitable variations in an observer's own degree centrality in the experiment) and boost the statistical power for detecting the observation-by-observation changes in fMRI signals within each stage. Second, to facilitate the between-stage comparison, we assumed three separate GLM conditions (regressors) for S1, S2 and S3, each including the onsets of all observations in the respective learning stage. Moreover, as the DeGroot learning model assumes all decisions (initial guess and three reassessments in a game) are made in the same way based on belief expectation, we evaluated the neural correlates of belief expectation estimates at the choice time assuming a shared condition (regressor) for choice submission within each game.

Specifically, GLM1 served to identify the neural correlate of the aPE estimates and RD values. Each game was modeled as having the following eight conditions (regressors): (1) fixation onset, (2) network display, (3) S1 observation onsets (including all events of observation onsets in S1, same below), (4) S2 observation onsets, (5) S3 observation onsets, (6) choice onsets (including all events of decision onsets, same below), (7) choice submission and (8) onsets of one's own choice display. We entered the aPE estimate associated with the observed action and the value of RD between the observee and the observer's neighborhood as the parametric modulators for the onset of the specific observation at the corresponding learning stage. For each decision submission, we entered the value estimate of belief expectation for the chosen option as well as the button pressed (left/right) as the parametric modulators for the respective decision. In all GLMs in the current study, parametric modulators associated with the same event were orthogonalized against one another to remove any shared variances, such that the regression coefficient reflected the unique contribution of each regressor in explaining the variances in neural signals. These regressors were convolved with the canonical hemodynamic response function and entered into a regression analysis against each fMRI subject's BOLD signals. The six vectors of head motion parameters derived from preprocessing were also included as nuisance regressors in all GLM analyses. Regression maps for aPE (RD) of each fMRI participant were first computed between learning stages based on the specific hypothesis derived from the DeGroot learning model (for example, aPE maps were averaged over S1, S2 and S3 within participants), and then taken into the standard random effects group-level analyses. All whole-brain analyses were thresholded and displayed at the cluster-wise FWE-corrected $P < 0.05$, with cluster-forming threshold $P_{unc.} < 0.001$, as reported by SPM12. Unless otherwise specified, all GLMs were structured in the same way (but with varying parametric modulators) and corrected for multiple comparisons using the same statistical procedure.

GLM2 served to examine the robustness of aPE encoding by controlling for nuisance effects related to action observation. This model was identical to GLM1 except that it included the following variables as additional parametric modulators for each observation onset in each learning stage: the order of observation display, the color selected by the observee (yellow/blue) and the consensus level among choices by all network members at the beginning of the particular learning stage (the proportion of the dominant choice on the network).

GLM3 served to test for brain activity whose variance was explained uniquely by RD values, while controlling simultaneously for potentially confounding variables as regressors of no interest. The regression model was identical to GLM1 except that it contained additional parametric modulators for all observation onsets at all learning stages, reflecting either the visual properties of the network layouts, nuisance effects related to action observation or variables previously associated with dACC functioning. These included the Euclidean distance between the observee's network position and the centroid of the network, the Euclidean distance between the observer's and observee's locations on the network display, order of observation display, color selected by the observee (yellow/blue), aPE magnitude, variance in attained observations within the learning stage, level of conflict between social observations and observer's own belief (proportion of attained observations within the learning stage that were different from the observer's previous decision), updated belief expectation estimate with respect to the observer's previous decision and choice difficulty reflected by the distance in the value estimates of belief expectation between two choice options.

GLM4 served to evaluate whether and how fMRI signals reflected the numerator and denominator in RD, by replacing the parametric modulator of RD in GLM1 with two variables—the observee's degree and the total degree centrality within the observer's local neighborhood—for all observation onsets at all learning stages in GLM1.

GLM5 served to identify brain regions responding to the belief expectation estimates, on an observation-by-observation basis, while controlling for potentially confounding variables. We entered the model-derived value estimates of belief expectation with respect to the observer's previous decision as the parametric modulator for observation onsets in each learning stage. To control for the correlation of belief expectation with the aPE and RD signals, as well as the potential effect of belief accumulation resulting from the sequential update by subjects, we also included the following variables of no interest as parametric modulators for observation onsets in all learning stages: aPE estimate for the observed action, RD value between the observee and observer and the order of observation display in the particular learning stage.

GLM6 served to examine the robustness of the belief expectation signals at observation onsets. We added to GLM5 the following decision-related variables as parametric modulators for all observation onsets in all learning stages: the product of RD and aPE estimates, the

color selected by the neighbor, the Euclidean distance between the observee's network position and the centroid of the network display, the Euclidean distance between the observer's and observee's locations on the network display, and the consensus level among choices by all network members at the beginning of the particular learning stage.

PPI analysis

We performed PPI analyses at both the ROI and whole-brain levels using SPM12. The ROI-based PPI analysis tested functional coupling between the VMPFC and the dACC at observation onsets in S1 versus S2/S3. We used the same ROI of the dACC defined independently by an online meta-analysis of 598 studies using the term 'cognitive control' in Neurosynth[39] (peak voxel MNI coordinates: x, y and z = 6, 22 and 36; Fig. 5c). The VMPFC cluster was defined by a sphere with a 6-mm radius centered at the peak VMPFC activation for belief expectation estimates as in Fig. 7a (MNI coordinates: x, y and z = −9, 44 and −10). The PPI model contained the following regressors for observation onsets: (1) BOLD timeseries extracted from the VMPFC cluster, (2) the dummy variable indicating S2 and S3 against S1 and (3) the product of the above two regressors. We also performed an exploratory PPI analysis at the whole-brain level, seeded in the same VMPFC cluster with the identical sets of regressors (Extended Data Fig. 10).

Conjunction analyses

Conjunction analyses were performed based on the minimum T-statistics in SPM12 testing for a logical AND (against the conjunction null) between group-level T-maps[61]. For all analyses, conjunction maps were thresholded at a conjunction P value corresponding to a whole-brain cluster-level $P_{FWE} < 0.05$, with cluster-forming threshold $P_{unc.} < 0.001$.

To evaluate whether there existed overlapping encoding of neighbors' degree (numerator of RD) and total local degree (denominator of RD), with opposing signs, at the observation onsets in S2 and S3, we first averaged each regression fit of interest from GLM4 for each participant across S2 and S3 (that is, the average of neighbor's degree across S2 and S3 and that of total local degree), and separately took them into the standard random effects analyses. The resulting second-level T-maps were then used for two separate conjunction analyses testing the overlap between (1) the positive encoding of neighbors' degree and negative encoding of total local degree (Fig. 5a) and (2) the negative encoding of neighbors' degree and positive encoding of total local degree (no significant overlap at cluster-level $P_{FWE} < 0.05$, with cluster-forming threshold $P_{unc.} < 0.001$).

To evaluate whether overlapping activation to RD existed at observation onsets across S1, S2 and S3, we first computed, for each learning stage, the second-level T-map for RD based on GLM1 regression fits. The resulting maps were then used in two conjunction analyses that tested the overlap in (1) the positive and (2) negative encoding of RD (both were not significant at cluster-level $P_{FWE} < 0.05$, with cluster-forming threshold $P_{unc.} < 0.001$). To evaluate whether the lack of overlap among S1, S2 and S3 was due to effects in S2 and S3, we performed similar analyses by combining the second-level T-maps of RD in S2 and S3, and evaluating the overlap in (1) the positive RD encoding between S2 and S3 (Extended Data Fig. 8) and (2) negative RD encoding between S2 and S3 (no significant overlap at cluster-level $P_{FWE} < 0.05$, with cluster-forming threshold $P_{unc.} < 0.001$).

fMRI timecourse analyses

We extracted the preprocessed BOLD timeseries as the average of voxels within the independent dACC ROI (as shown in Fig. 5c). The extracted BOLD series were further regressed out the head motion to control for potential motion-related artifacts, applied a high-pass filter (cutoff, 128 s) to remove low-frequency drifts and oversampled by a factor of 20 to obtain a time resolution of 0.1 s. For each observation, an 18-s window (180 timepoints) time locked to observation onsets (2 s before the onset) was applied. To obtain the parameter estimate timecourse for each learning stage, we performed linear regression for each time point, separately for S1, S2 and S3, to estimate the effects of variables of interest on the extracted brain activity and then concatenated the regression β values across timepoints. Standardized regression coefficients are shown in Fig. 5c.

Statistical analyses

No statistical methods were used for determining sample sizes. The size of the fMRI sample is comparable with those reported in previous research of RL and social learning[26,62,63]. Blinding is irrelevant in this study since the experiment did not involve multiple conditions. Statistical comparisons were performed using with appropriate inferential methods, as indicated in the figure captions. For parametric statistical tests that require normality assumptions, data distribution was assumed to be normal, but this was not formally tested. Behavioral data were collected using a custom-written Python (v.2.7) program. Statistical analyses were conducted using R (v.3.3) and Matlab (v.R2017a). SPM12 was used for analyzing neuroimaging data, and the VBA toolbox was used for Bayesian model selection.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Data underlying the findings of this study are available at Open Science Framework: https://osf.io/8rbs4/.

Code availability

Code supporting the findings of this study is available at Open Science Framework: https://osf.io/8rbs4/.

References

50. Freeman, L. C. Centrality in social networks conceptual clarification. *Soc. Netw.* **1**, 215–239 (1978).
51. Watts, D. J. Collaborative learning in networks. *Proc. Natl Acad. Sci. USA* **109**, 764–769 (2012).
52. Watts, D. J. *Small Worlds* (Princeton Univ. Press, 2018).
53. Centola, D. The spread of behavior in an online social network experiment. *Science* **329**, 1194–1197 (2010).
54. Daunizeau, J., Adam, V. & Rigoux, L. VBA: a probabilistic treatment of nonlinear models for neurobiological and behavioural data. *PLoS Comput. Biol.* **10**, e1003441 (2014).
55. Chandrasekhar, A. G., Larreguy, H. & Xandri, J. P. *Testing Models of Social Learning on Networks: Evidence from a Lab Experiment in the Field* Working Paper Series No. 21468 (National Bureau of Economic Research, 2015).
56. Acemoglu, D., Dahleh, M. A., Lobel, I. & Ozdaglar, A. Bayesian learning in social networks. *Rev. Econ. Stud.* **78**, 1201–1236 (2011).
57. Gale, D. & Kariv, S. Bayesian learning in social networks. *Games Econ. Behav.* **45**, 329–346 (2003).
58. Burt, R. S., Kilduff, M. & Tasselli, S. Social network analysis: foundations and frontiers on advantage. *Annu. Rev. Psychol.* **64**, 527–547 (2013).
59. O'Doherty, J. P., Hampton, A. & Kim, H. Model-based fMRI and its application to reward learning and decision making. *Ann. N. Y. Acad. Sci.* **1104**, 35–53 (2007).
60. Gläscher, J. P. & O'Doherty, J. P. Model-based approaches to neuroimaging: combining reinforcement learning theory with fMRI data. *Wiley Interdiscip. Rev. Cogn. Sci.* **1**, 501–510 (2010).
61. Friston, K. J., Penny, W. D. & Glaser, D. E. Conjunction revisited. *NeuroImage* **25**, 661–667 (2005).
62. Cockburn, J., Man, V., Cunningham, W. A. & O'Doherty, J. P. Novelty and uncertainty regulate the balance between exploration and exploitation through distinct mechanisms in the human brain. *Neuron* **110**, 2691–2702.e8 (2022).

63. Suzuki, S., Jensen, E. L. S., Bossaerts, P. & O'Doherty, J. P. Behavioral contagion during learning about another agent's risk-preferences acts on the neural representation of decision-risk. *Proc. Natl Acad. Sci. USA* **113**, 3755–3760 (2016).

Acknowledgements

We thank Y. Yin, Y. Wang and Y. Dong for assistance with intranet setup, data collection and code validation. We also thank the National Center for Protein Sciences and the high-performance computing platform at the Center for Life Sciences at Peking University for facilitating data acquisition and computation. This work is supported by NSFC (32071095 to L.Z.), STI2030-Major Projects (2022ZD0205104 to L.Z.), China Postdoctoral Science Foundation (2022TQ0013 to Q.M.) and Center for Life Sciences at Peking University.

Author contributions

Y.J. and L.Z. designed the study. Y.J. conducted the experiments. Y.J., Q.M. and L.Z. analyzed the data. Y.J. and L.Z. wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41593-023-01258-y.

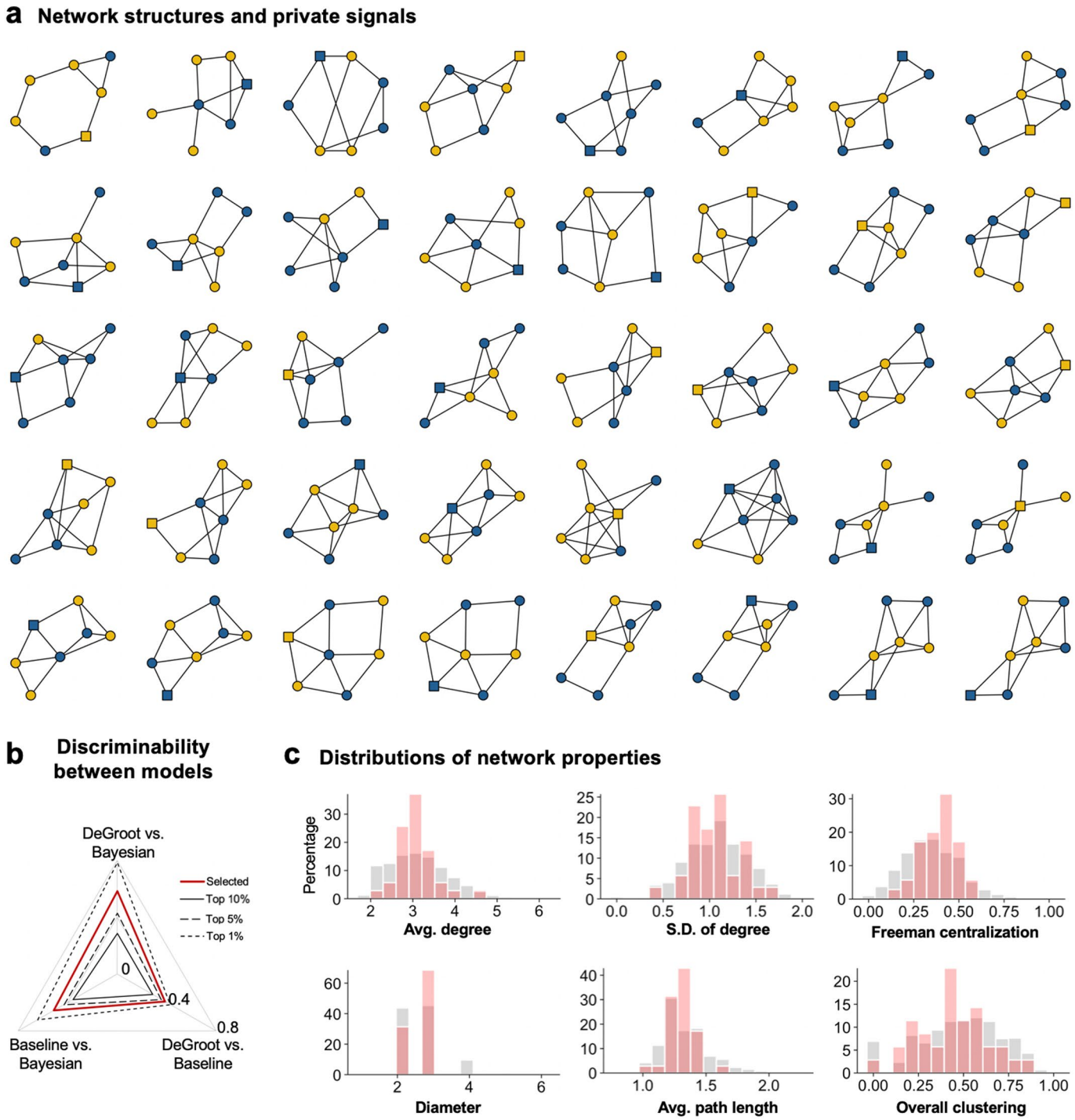
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41593-023-01258-y.

Correspondence and requests for materials should be addressed to Lusha Zhu.

Peer review information *Nature Neuroscience* thanks A. Bhandari, S. Sul and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

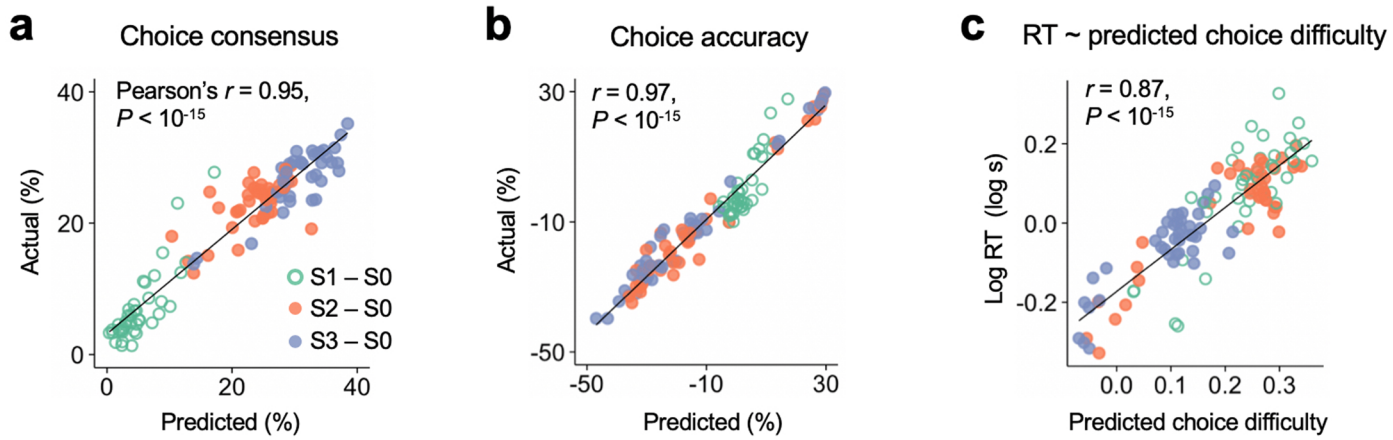
Reprints and permissions information is available at www.nature.com/reprints.

https://doi.org/10.1038/s41593-023-01258-y



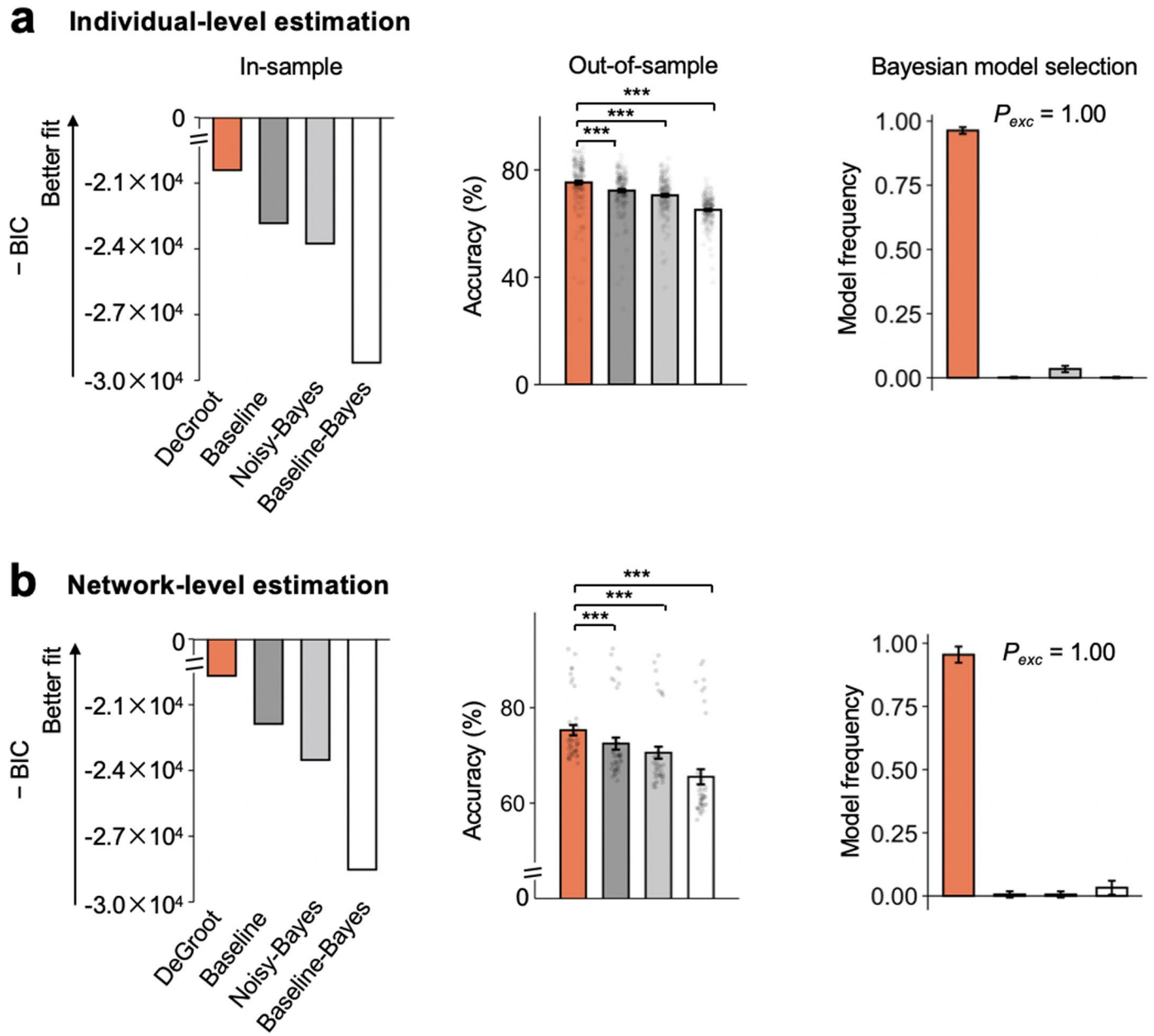
Extended Data Fig. 1 | Stimuli. (a) Network structures and private signals. Blue/yellow represents the color of the private signal at a specific node. Circle/square represents the location of a behavioral/fMRI participant, which was assigned pseudo-randomly (Methods). Network structures were displayed using the 'force-directed' algorithm implemented in MATLAB (R2017a), with minor adjustments to node coordinates for avoiding overlapping edges. During the experiment, the network structure was presented to all participants, but one's private signal was only known to the particular subject. **(b)** The empirical distribution of discriminability indices calculated from the pre-experiment simulation based on the DeGroot, Baseline, and Bayesian learning models (Methods). The radar plot illustrates the level of discriminability between any two candidate models. Each axis represents a model pair. For any point on an axis, the distance from the center of the plot corresponds to the discriminability of the given model pair (ranging from 0 to 1). The dotted, dashed, and solid lines in black represent, respectively, the 99th, 95th, and 90th percentiles of discriminability indices computed from all 128 × 853 candidate stimuli. The red line represents the discriminability averaged over 40 sets of stimuli employed in the experiment. The discriminability index is defined as the proportion of choices simulated by one model, given a network structure and private signals on the network, that disagree with the simulation from the other model, and is averaged over decisions in S2 and S3 and across 7 agents simulated from different network locations (Methods). **(c)** Selected networks demonstrate reasonable distribution ranges in the topological features relevant for information transmission (Methods). Each gray/red histogram depicts the distribution of the network parameter computed by pooling candidate/selected stimuli.

https://doi.org/10.1038/s41593-023-01258-y



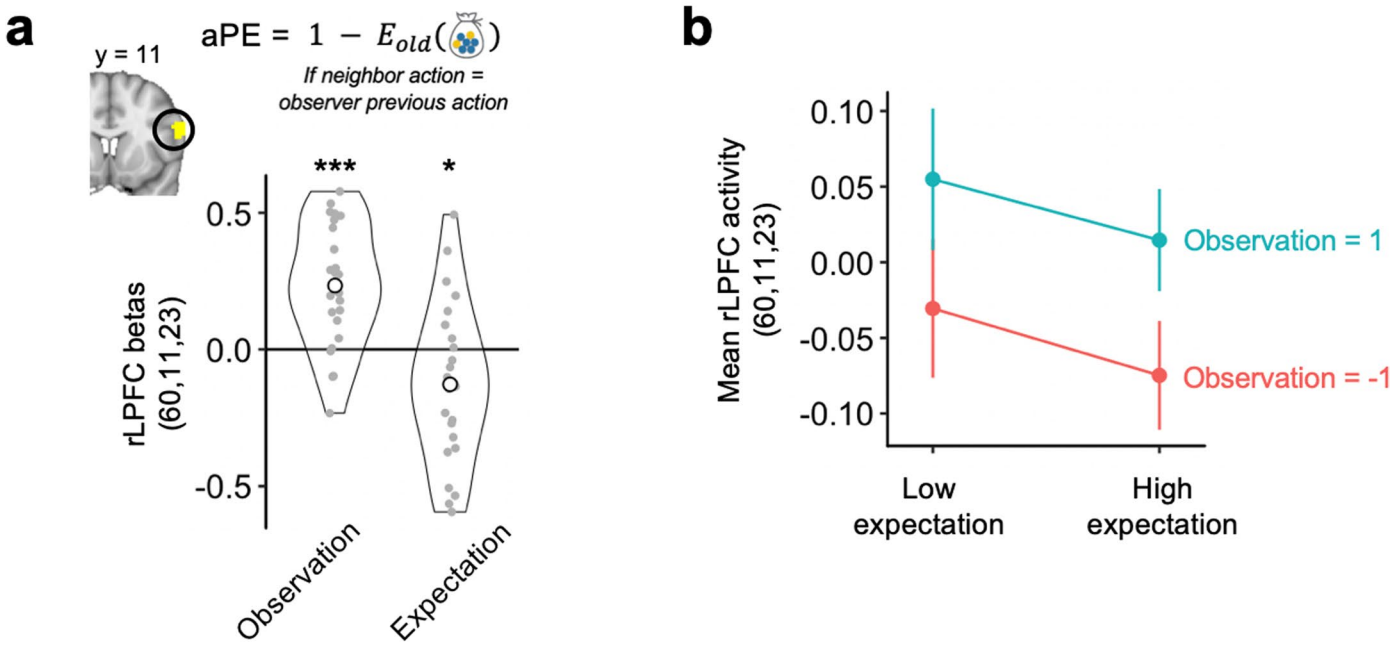
Extended Data Fig. 2 | The DeGroot learning model explains multiple aspects of behavior across stages. The actual vs. model-predicted choice consensus **(a)** and choice accuracy **(b)**. Choice consensus is defined as the proportion of participants embedded on the same network who choose the same (dominant) option at a given stage in a game. Choice accuracy is defined as the proportion of participants on the same network whose choices are consistent with the most likely underlying state given the distribution of private signals in the game. **(c)** Model-predicted choice difficulty vs. reaction time (RT), a widely used empirical measure for choice difficulty. Model-derived choice difficulty is defined as the entropy of the softmax action probability calculated from the DeGroot learning model (that is, more difficult when the model-derived action probability is closer to 0.5). To control for the influences of the initial guesses (which are affected by the distribution of private signals), effects of initial guesses (S0) are subtracted from all the above measures. Each dot represents the average value across all subject groups given a stage and a game, colored by the stage (N = 40 games).

https://doi.org/10.1038/s41593-023-01258-y



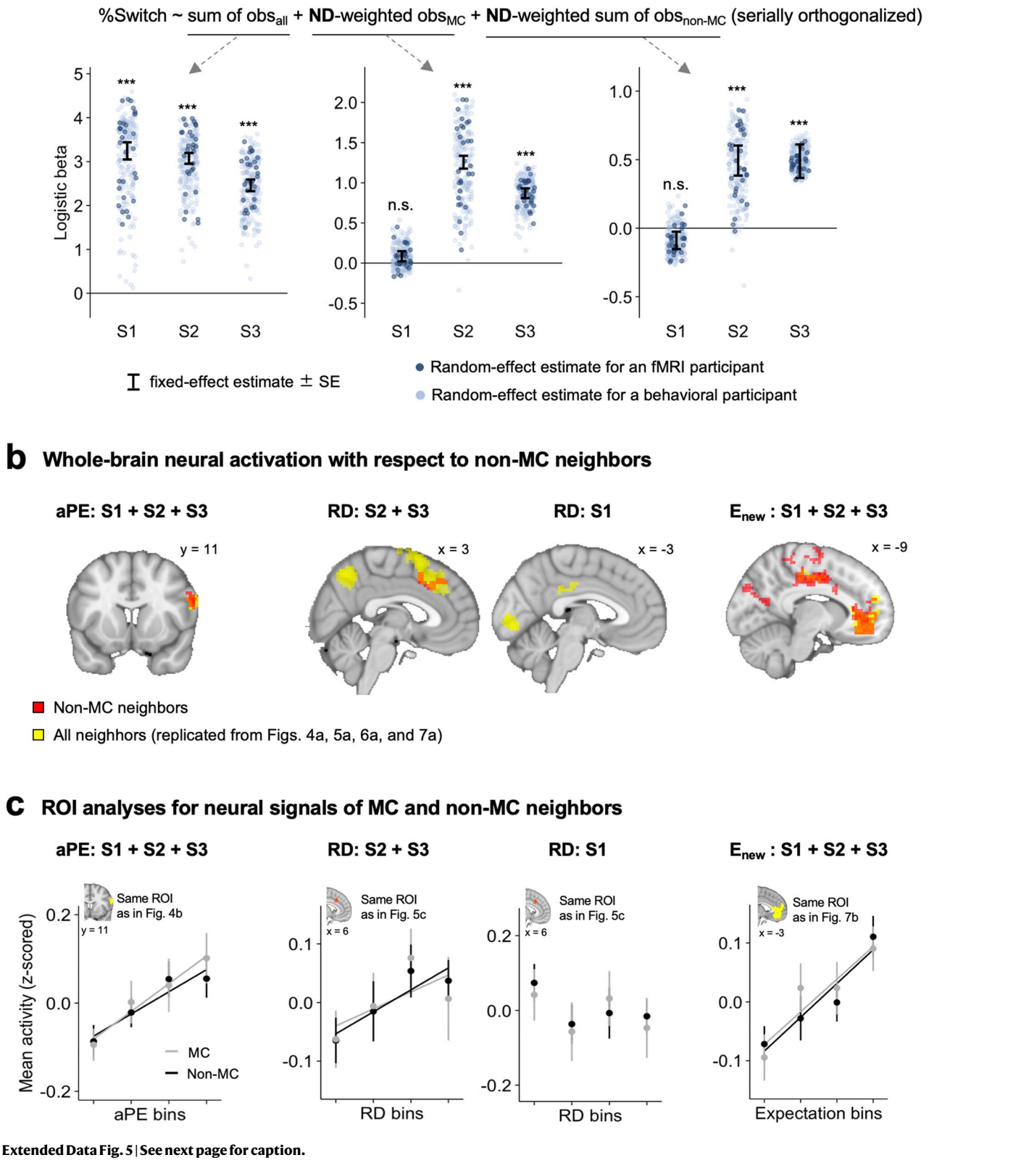
Extended Data Fig. 3 | Model comparisons based on (a) individual-level and (b) network-level estimations. Goodness-of-fit comparisons are based on the in-sample BIC scores (left), out-of-sample prediction power based on a five-fold cross-validation procedure (middle), and model frequencies and exceedance probability (P_{exc}) calculated by Bayesian model selection (right; Methods). Two variants of the Bayesian learning models (Noisy-Bayes and Baseline-Bayes) are considered (see Methods and Supplementary Note 2). Error bars in the middle panels of (**a**) represent mean ± intersubject SEM, of (**b**) represents mean ± intergame SEM, and in the right panels of (**a**, **b**) represent estimated model frequencies ± SD of the Dirichlet distribution. Individual-level analysis are based on N = 209 behavioral subjects, and network-level analysis are based on N = 40 games. $^{***}P < 0.001$, two-sided paired *t*-tests, all Bonferroni-corrected.

https://doi.org/10.1038/s41593-023-01258-y



Extended Data Fig. 4 | Activity in the right LPFC is correlated with the observed action and belief expectation estimate, with opposing signs. **(a)** Neural betas with respect to two components in an action prediction error (aPE) signal, the observed actions (1 if observation matches the observer's prior choice; otherwise, −1) and belief expectation estimates (E_{old}) associated with the observer's prior decision. The beta values were separately extracted from the rLPFC cluster as identified by aPE estimates in Fig. 4a. **(b)** Visualization of rLPFC responses to observed actions and belief expectation estimates. Consistent with the aPE assumption, the mean rLPFC activity is higher when an observation is consistent with the observer's prior decision (observation = 1) than when the observation differs from the observer's prior decision (observation = −1). Also, the mean rLPFC activity demonstrates a negative main effect for high- vs. low-value estimates of belief expectation (E_{old}), based on median splits on belief expectation estimates for each fMRI participant. Each dot in violin plots represents a subject. Error bars represent intersubject SEM in the fMRI sample (N = 25). $^{*}P < 0.05$, $^{***}P < 0.001$, two-sided t-tests.

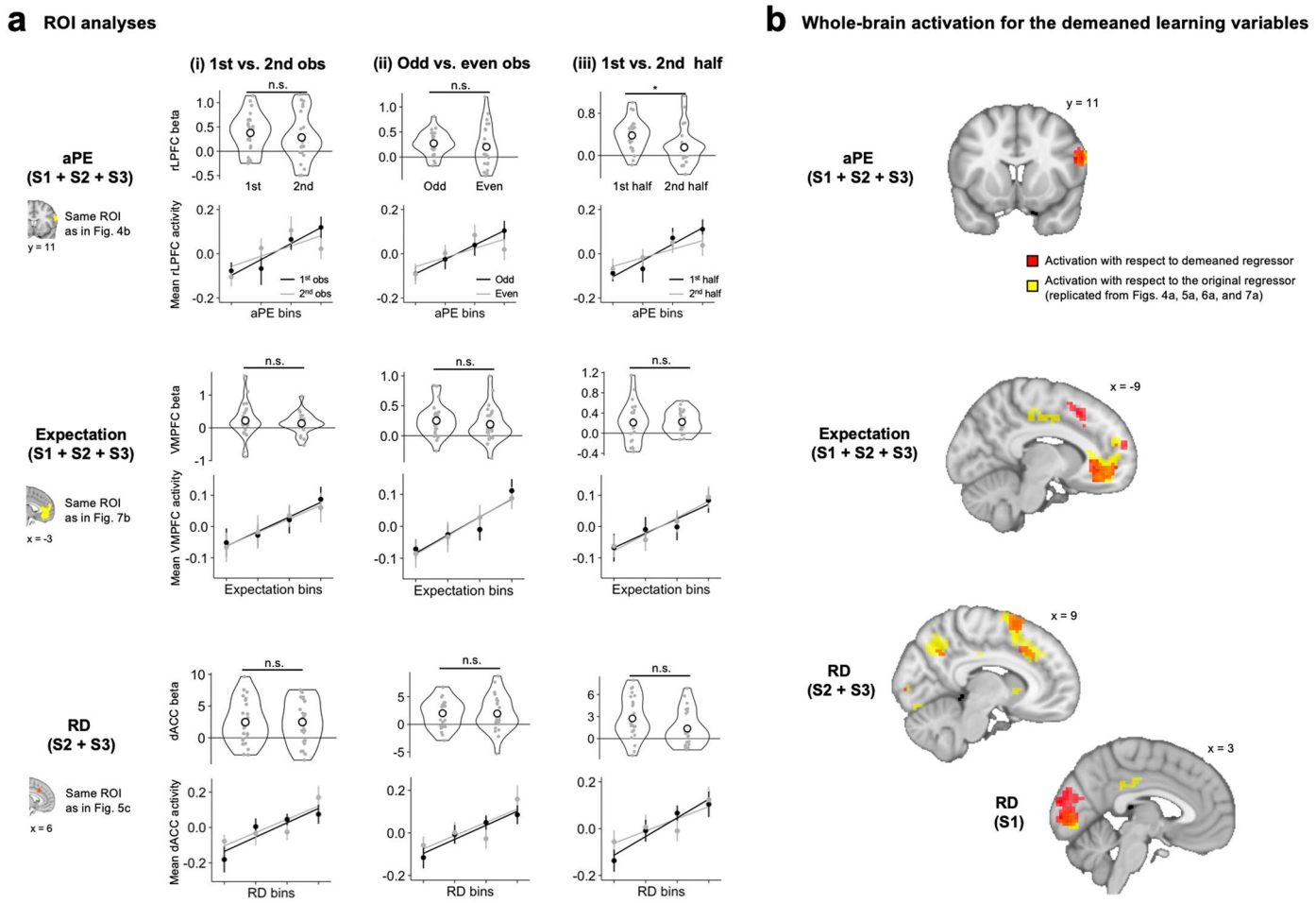
https://doi.org/10.1038/s41593-023-01258-y



Extended Data Fig. 5 | See next page for caption.

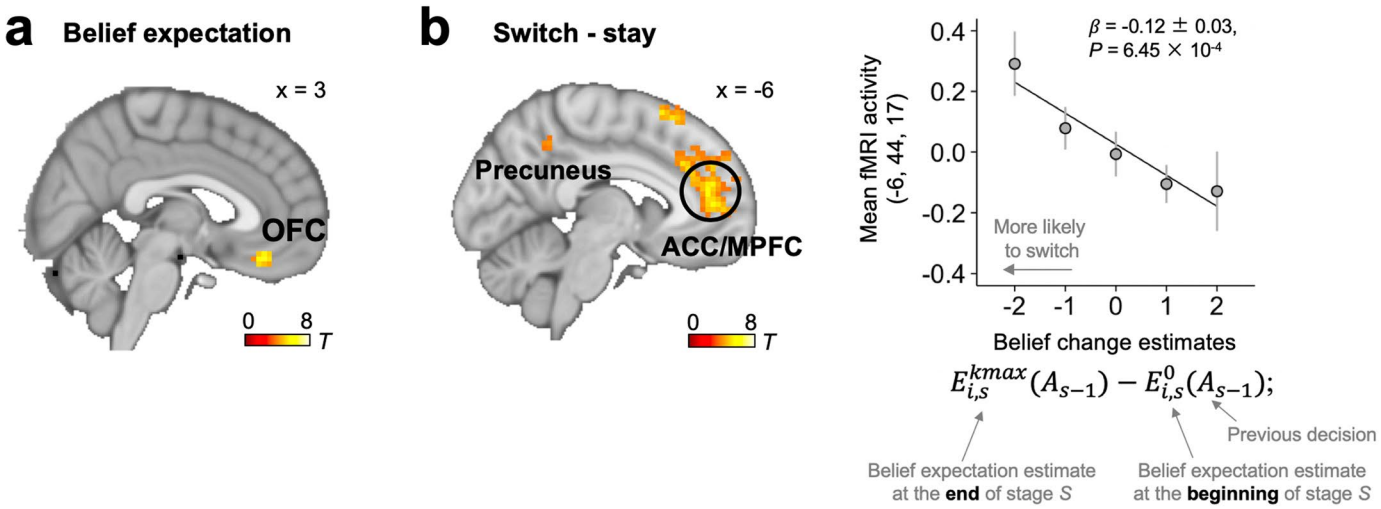
Extended Data Fig. 5 | Behavioral and neural evidence that learning is not limited to the most-connected (MC) neighbor but also to the other (non-MC) neighbors. (a) Mixed-effects logistic regression for each separate learning stage. The regression analyses are similar to those in Fig. 3a but with the following 3 regressors, serially orthogonalized: (i) the unweighted sum of observations of all neighbors, (ii) action by the MC neighbor scaled by her degree centrality, and (iii) the ND-weighted sum of observations for all non-MC neighbors. Serial orthogonalization ensures the regression coefficient for the third regressor reflects only the variances in choice behavior that can be uniquely explained by this last regressor (see also Methods and Supplementary Table 4 for corresponding model-based analyses). **(b)** Whole-brain analyses show significant neural responses at the onsets of observations from non-MC neighbors (all thresholded at cluster-wise FWE-corrected $P < 0.05$, with cluster-forming threshold $P_{unc.} < 0.001$). **(c)** ROI analyses comparing effects between the MC and non-MC neighbors. Mean fMRI activity are separately extracted for the MC and non-MC neighbors from respective ROIs and binned by the corresponding estimate values. $^{***}P < 0.001$, two-sided z-tests. Error bars in (a) represent the SE of fixed-effect estimates in the logistic regression of the behavioral sample (N = 209), error bars in (c) represent the intersubject SEM of the fMRI sample (N = 25).

https://doi.org/10.1038/s41593-023-01258-y



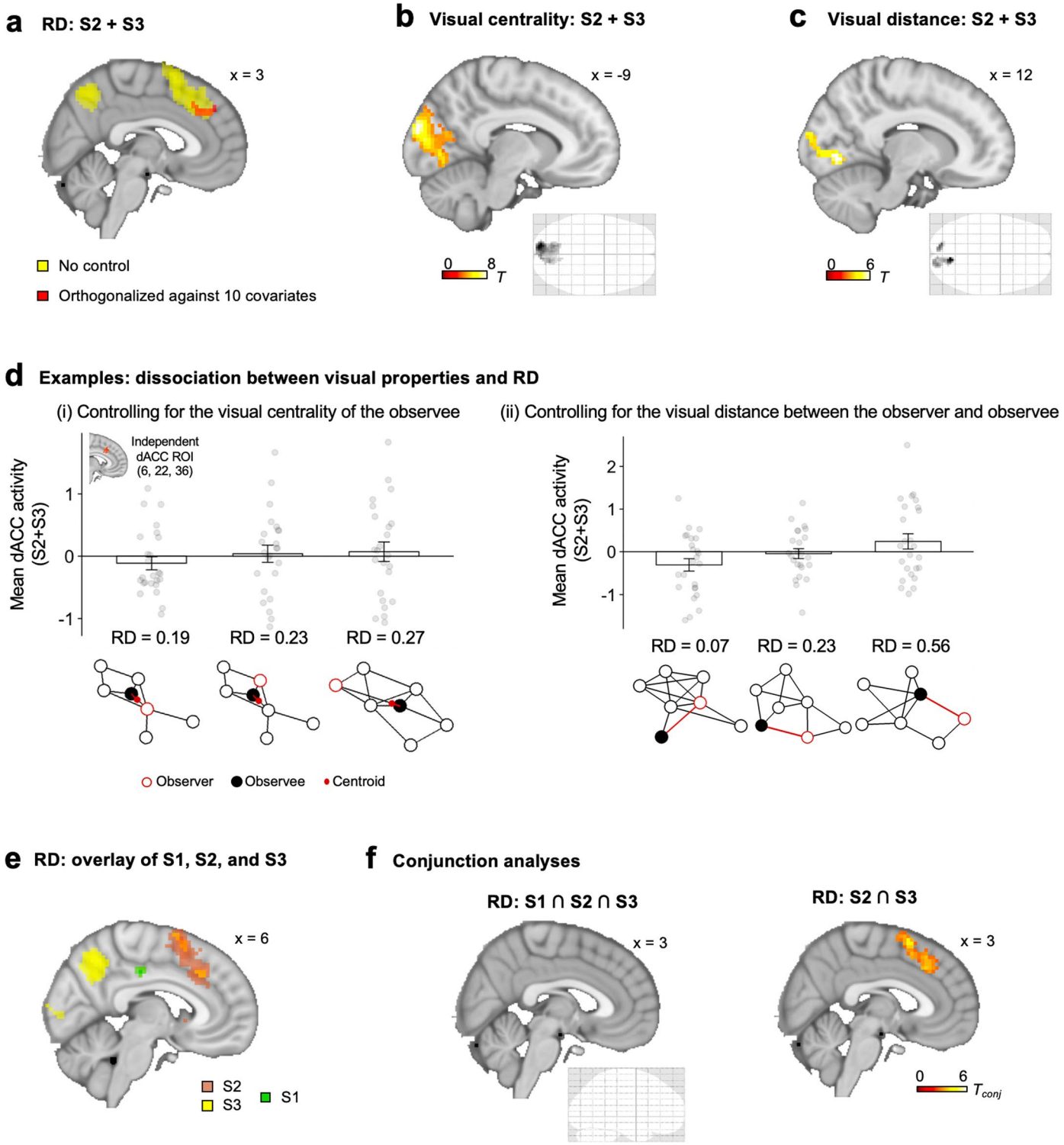
Extended Data Fig. 6 | Evidence for neighbor-by-neighbor neural representations. (a) ROI analyses comparing neural responses between the 1st vs. 2nd, odd- vs. even-numbered, 1st half vs. 2nd half observations within a learning stage. The violin plots in the upper panels show the effect sizes within each corresponding ROI (two-sided *t*-tests, all $P < 0.05$, uncorrected). The lower panels visualize the effects by plotting the mean BOLD activity extracted from each ROI against bins of ascending values. **(b)** Whole-brain analyses demonstrating that the observed neural activation could not be entirely attributed to the game-to-game variations. Statistical parametric maps show neural responses at observation onsets to learning variables demeaned within each game (clusters in red), overlaid with the activation with respect to the original values (clusters in yellow). All thresholded at cluster-wise FWE-corrected $P < 0.05$, with cluster-forming threshold $P_{unc.} < 0.001$, except for RD-related activation in S2 and S3, which is thresholded at $P_{unc.} < 0.001$ with cluster size $K > 20$. The reduced RD-related activation following the removing of between-game variations is consistent with the definition of RD, as it contains both neighbor-by-neighbor (neighbor degree) and game-by-game (total local degree) variances. Error bars represent intersubject SEM in the fMRI sample (N = 25). n.s., not significant; * $P < 0.05$; all two-sided paired *t*-tests, uncorrected.

https://doi.org/10.1038/s41593-023-01258-y



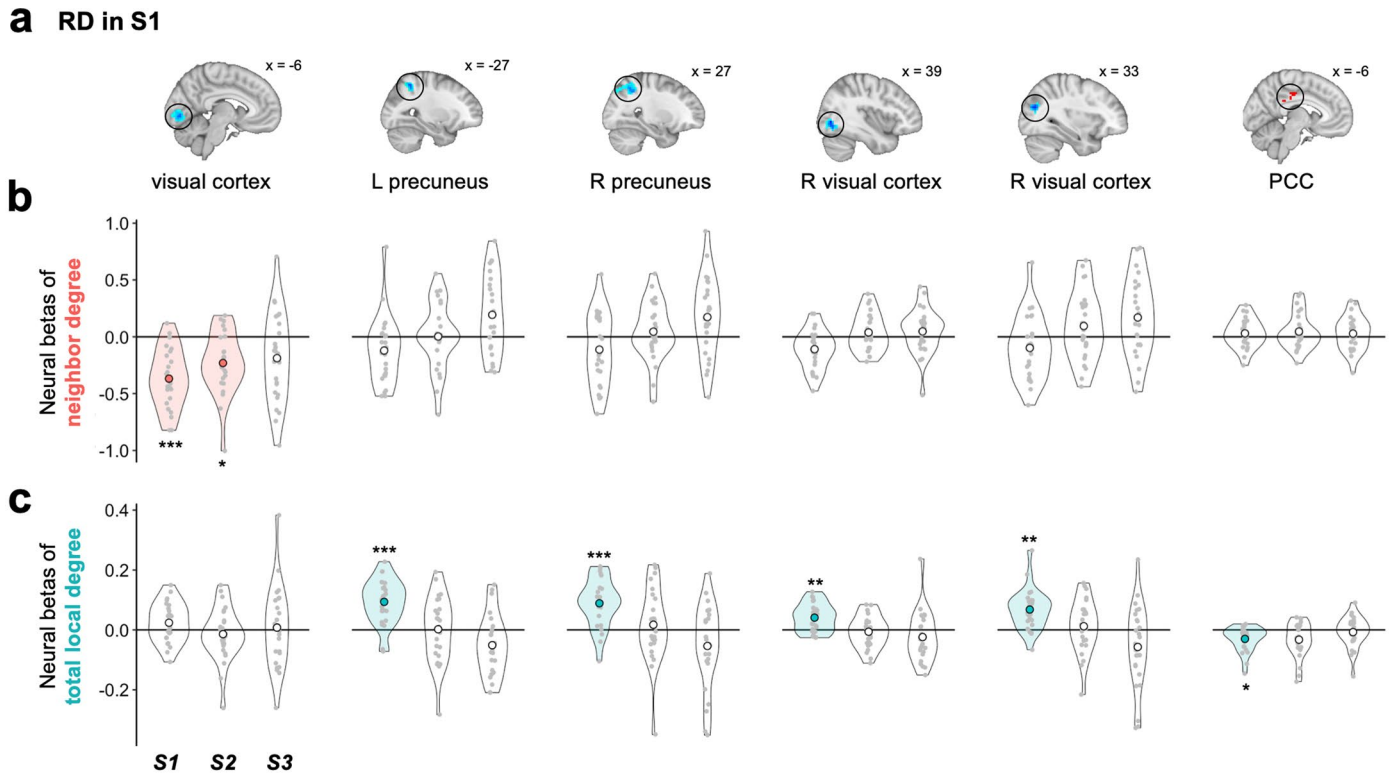
Extended Data Fig. 7 | Choice-related neural activity at decision time. (a) BOLD activity in the orbitofrontal cortex (OFC) is positively correlated with the value estimate of belief expectation for the chosen option at the time of decision submission (cluster-wise FWE-corrected $P < 0.05$, with cluster-forming threshold $P_{unc.} < 0.001$). **(b)** BOLD activity in the anterior cingulate cortex (ACC) and neighboring medial prefrontal cortex (MPFC) at decision submission reflects the model-derived tendency of modifying one's prior estimation. Left: BOLD activity at decision time in the ACC/MPFC, precuneus, and inferior parietal lobule (not shown) is higher when a subject revises her previous decision than when the subject sticks to the same decision (cluster-wise FWE-corrected $P < 0.05$, with cluster-forming threshold $P_{unc.} < 0.001$). No region shows a decreased response to switch vs. stay at choice time at the same statistical threshold. Right: Results of mixed-effects linear regression show that the mean fMRI signal extracted from the identified ACC/MPFC cluster (peak voxel MNI coordinates: $x, y, z = -6, 44, 17$; as shown in the left panel) at the decision time is negatively correlated with the amount of change in the model-derived belief expectation from the beginning to the end of the corresponding learning stage (that is, belief change estimates). The x-axis represents the model-derived belief change within a learning stage, rounded to the nearest integer for illustration. Error bars represent intersubject SEM in the fMRI sample (N = 25).

https://doi.org/10.1038/s41593-023-01258-y



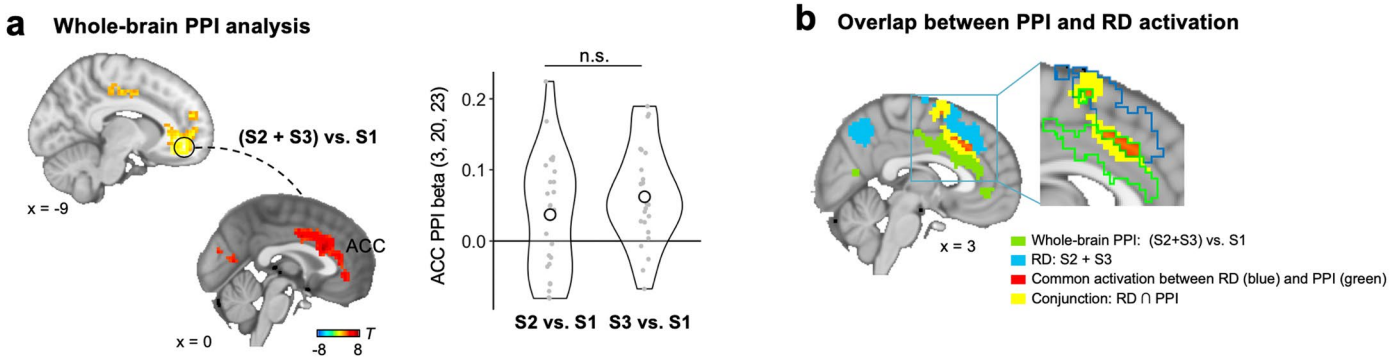
Extended Data Fig. 8 | Robustness and specificity of the encoding of RD values in the dACC in S2 and S3. (a) Extent of the dACC responses to RD values in S2 and S3. Yellow clusters reflect the maximal extent of activation to RD, *without* controlling for *any* decision variables. Red clusters reflect the activation to the residuals of RD, after being orthogonalized against 10 variables-of-no-interest as parametric modulators (GLM3, Methods). All thresholded at cluster-wise FWE-corrected $P < 0.05$, with cluster-forming threshold $P_{unc.} < 0.001$. **(b–c)** The visual cortex, but not the dACC, tracks the observees' visual centralities in the network display **(b)** and the visual distance between the observer and observee's locations **(c)**. Visual centrality is defined as the Euclidean distance between the observee's network position and the centroid of the network display. All thresholded at cluster-wise FWE-corrected $P < 0.05$, with cluster-forming threshold $P_{unc.} < 0.001$. **(d)** Example networks where the observees are associated with the same visual centralities (left) or the same visual distance between the observer and observee (right), yet the dACC activity varies with the RD values. **(e)** Overlay of RD-related activation in S1, S2, and S3 in a single map. For illustration purpose, all maps are shown at $P_{unc.} < 0.001$ with $K > 10$. **(f)** Whole-brain conjunction analyses for RD correlates. Left: A three-way conjunction on RD correlates among S1, S2, and S3 identified no significant overlap in either the positive or negative responses to RD (Methods). Right: A two-way conjunction analysis between S2 and S3 on RD correlates identified a significant overlap for the positive correlations with RD in the dACC between S2 and S3. No overlap was identified in the whole-brain conjunction analysis for the negative correlation with RD between S2 and S3 at the same threshold. All thresholded at cluster-wise FWE-corrected $P < 0.05$, with cluster-forming threshold $P_{unc.} < 0.001$. Error bars represent intersubject SEM in the fMRI sample (N = 25).

https://doi.org/10.1038/s41593-023-01258-y



Extended Data Fig. 9 | Whole-brain and ROI analyses for RD correlates at observation onsets in S1. Brain regions where BOLD activity correlates with RD values at observation onsets in S1 (cluster-wise FWE-corrected $P < 0.05$, with cluster-forming threshold $P_{unc.} < 0.001$; same as Fig. 6a but shown in different cuts; GLM1, Methods). **(b–c)** ROI analyses of each region in (a), with respect to neighbor's degree (numerator of RD) and total local degree (TLD; denominator of RD), respectively. Unlike the dACC, where fMRI signals correlated with both neighbor degree and TLD in S2 and S3, the neural encoding of RD identified in S1 is driven by one of the two components in an RD signal. That is, a cluster in the visual cortex tracks only neighbor degree, whereas clusters in the precuneus, posterior cingulate cortex (PCC), and visual cortex track only TLD. Effect sizes in these regions in S2 and S3 are also included for completeness. Violin plot with color represents significant effects. Each dot represents a subject (N = 25). $^{*}P < 0.05$, $^{**}P < 0.05$, $^{***}P < 0.001$, two-sided t-tests, Bonferroni-corrected.

https://doi.org/10.1038/s41593-023-01258-y



Extended Data Fig. 10 | Whole-brain psychophysiological interaction (PPI) analysis testing differential functional connectivity with the VMPFC in S2 and S3 vs. S1. (a) Left: Increased functional connectivity between the seed region in the VMPFC (6-mm sphere around the peak activation as identified in Fig. 7a) and a cluster in the anterior cingulate cortex (ACC) at observation onsets in S2 and S3, relative to S1 (cluster-wise FWE-corrected $P < 0.05$, with cluster-forming threshold $P_{unc.} < 0.001$; Methods). Right: No systematic difference in the effect sizes of functional coupling between S2 and S3 (two-sided paired *t*-test, $t_{24} = -1.20$, $P = 0.242$), as revealed by the PPI betas extracted from the significant cluster in the ACC as identified in the left panel. Each dot represents a subject (N = 25). **(b)** Overlay of the whole-brain PPI activation (green; as in (a)), RD correlates in S2 and S3 (blue; as in Fig. 5a), the overlap between the first two activation maps (red), and the result of a formal whole-brain conjunction analysis between regions correlating with RD values in S2 and S3 and regions demonstrating differential functional connectivity with the VMPFC seed region in S2 and S3 vs. S1 (yellow; cluster-wise FWE-corrected $P_{conj} < 0.05$, with cluster-forming threshold $P_{unc.} < 0.001$).

nature portfolio

Corresponding author(s): Lusha Zhu

Last updated by author(s): Dec 21, 2022

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Python (2.7)
Data analysis	R (3.3), Matlab (R2017a), SPM (12), VBA toolbox

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The data that support the findings of this study are available at Open Science Framework: https://osf.io/8rbs4/.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Two independent samples were recruited. The first one consisted of 186 behavioral participants and 31 fMRI participants. These subjects participated in the experiment in groups of 7, with one fMRI subject being randomly paired with six behavioral participants in each group. No statistical methods were used for determining this sample size. The size of the fMRI sample is comparable to those reported in previous publications on RL or social learning (ref 26,62,63). The second one consisted of 91 behavioral participants. They also participated in the experiment in groups of 7. The sample size was predetermined based on a power analysis, assuming the same effect size as in our first sample, the alpha level 0.05, and the power 85%.
Data exclusions	For the first sample, six out of 186 behavioral subjects were excluded from analyses due to insufficient variations in choice behavior (N=4; repeating the same decision based on their private signals in 95% of decisions), abnormal reaction time (N=1; five standard deviations above group mean), or demonstrating strong disbelief of the experimental setting in the post-experiment survey (N=1). Of 31 fMRI participants, four were excluded from neuroimaging analyses but included in behavioral analyses due to excessive motion, and two were excluded from both neural and behavioral analyses due to scanning issues that did not interfere with their group members' experiment (unstable stimuli display N=1; unable to finish scanning but completed the task outside the scanner N=1). For the second sample, no participants were excluded from analyses.
Replication	We conducted two independent experiments in total. Main behavioral results in our first experiment were replicated in our second experiment, using a different experimental design (Supplementary Note 3). We also provided all information needed to conduct replication experiment in the manuscript.
Randomization	Participants were randomly assigned to groups of seven. The order of games were randomized across groups. Choice positions (left/right) were also randomized across subjects.
Blinding	Blinding is irrelevant in this study since the experiment did not involve multiple conditions. All participants experienced all games in random orders predetermined by the computer program with no interaction with the experimenter.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☐	☒ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	We recruited two independent samples of healthy volunteers from Peking University, China. The first one consisted of 217 participants including 127 females; mean age ± S.D. = 20.25 ± 2.68 years. Of these participants, 31 completed the experiment inside a fMRI scanner including 14 females; age = 20.16 ± 1.97 years. The second one consisted of 91 participants including 49 females; mean age ± S.D. = 22.29 ± 2.38 years. All participants reported having normal or corrected-to-normal eye vision, no color blindness, and no history of neurological or psychiatric illnesses. All fMRI participants were right-handed.

Recruitment	All participants were recruited from Peking University, China, via an online recruitment system of the Neuroeconomics Lab. There was no potential self-selection bias, and we made every reasonable effort to minimize the potential for other biases in subjects recruitments that could impact results.
Ethics oversight	The study was approved by the Committee for Protection of Human Subjects at Peking University, China.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Magnetic resonance imaging

Experimental design

Design type	Task fMRI; event-related design
Design specifications	Each fMRI participant played 40 separate distributed learning games, divided into three scanning sessions (13 or 14 games per session) with short breaks in between.
Behavioral performance measures	We recorded participants' decisions in each game (including one initial guess and three reassessments) as well as their response time.

Acquisition

Imaging type(s)	Functional and structural fMRI images
Field strength	3 Tesla
Sequence & imaging parameters	Prior to the task, high-resolution structural T1-weighted scans were acquired using a magnetization-prepared rapid gradient echo sequence with the following parameters: repetition time (TR) = 2530 ms, echo time (TE) = 2.98 ms, flip angle = 7°, field of view (FoV) = 224 mm × 256 mm, inter-slice gap = 0.5 mm, voxel size = 0.5 mm × 0.5 mm × 1 mm, number of slices = 192. During the task, functional images were acquired using echo-planar T2* images with blood-oxygenation-level-dependent (BOLD) contrast, and angled 30 degrees with respect to the anterior commissure-posterior commissure (AC-PC) line to minimize susceptibility artifacts in the orbitofrontal area. The scanning parameters were as follows: TR = 2000 ms, TE = 30 ms, flip angle = 90°, FoV = 224 mm × 224 mm, voxel size = 3.5 mm × 3.5 mm × 3.5 mm, inter-slice gap = 0.7 mm, number of slices = 33.
Area of acquisition	Whole-brain

Diffusion MRI ☐ Used ☒ Not used

Preprocessing

Preprocessing software	Image preprocessing and analyses were performed in SPM12 (Wellcome Trust Centre for Neuroimaging Institute of Neurology, UCL). For each fMRI session, the preprocessing of neuroimaging data included, in order: slice time correction, motion correction (aligning to the mean image), coregistration, spatial normalization to the Montreal Neurological Institute (MNI) template (final image resolution of 3 mm × 3 mm × 3 mm), and spatial smoothing using a Gaussian kernel of 6-mm full width at half maximum. All images were high-pass filtered in the temporal domain (width 128 s) and autocorrelation of the hemodynamic responses was modeled as an AR(1) process.
Normalization	Functional and structural images were spatially normalized by warping subject-specific images to the reference brain in MNI coordinate space (final image resolution of 3 mm × 3 mm × 3 mm).
Normalization template	MNI template
Noise and artifact removal	The head motion was corrected in each session by using the alignment procedure in SPM12, and the resulting 6 vectors of head motion parameters were included as nuisance regressors in all GLM analyses.
Volume censoring	No volume censoring was performed.

Statistical modeling & inference

Model type and settings	We implemented standard general linear models (GLMs) for model-based univariate fMRI analysis. First-level analyses were conducted using fixed-effect models. In all GLMs, a distributed learning game was modeled as a series of discrete events using stick functions. This included fixation onset, network display, initial guess onset, initial guess submission, and—for each learning stage—the display of own guess, the onset of each and every observation, the onset of revision, and the submission of revision. These events were then grouped into regressors, which were modulated by variables of interest(see below). Second-level analyses were conducted using random-effect models.
Effect(s) tested	Our main GLM served to identify the neural correlate of action prediction error(aPE) and relative degree (RD) estimates, on an observation-by-observation basis, separately for each learning stage. The model included the following regressors: fixation onset, network display, S1 observation onsets (including separate events of all observation onsets in S1), S2 observation onsets, and S3 observation onsets, decision onsets (including separate events of all decision onsets of a game), decision submission (all decision submission), and display of own choices. We entered aPE estimate associated with the

observed action and RD between the observee and the observer's neighborhood a
onset at each learning stage, and the value estimate of belief expectation for the chosen option and button pressed (left/ right) as parametric modulators for each respective choice submission. aPE and belief expectation were derived using each fMRI participant's best fitting parameters from the DeGroot-RL model. RD was calculated based on the structure of the network in each respective game, independent of participants' behavior and model estimation.

For all GLMs in the current study, parametric modulators associated with the same event were orthogonalized against one another to remove any shared variances, such that the regression coefficient reflected the unique contribution of each regressor in explaining the variances in neural signals. These regressors were convolved with the canonical hemodynamic response function and entered into a regression analysis against each fMRI subject's BOLD signals. The six vectors of head motion parameters derived from preprocessing were also included as nuisance regressors in all GLM analyses. Regression map for aPE (RD) of each fMRI participant were first computed across learning stages (e.g., averaged over S2 and S3 within participant) based on the specific hypothesis derived from the DeGroot-RL model, and then taken into random-effects group-level analyses. The group-level effects were tested against zero with one-sample t-tests.

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☒ Both

Anatomical location(s): The ROI in the dorsal anterior cingulate cortex was independently defined from Neurosynth (http://neurosynth.org), with peak voxel MNI coordinates: x, y, z = 6, 22, 36 (as shown in Fig. 5c). Other functional ROIs were defined by activated clusters in whole-brain GLM analyses. The ROI in the right lateral prefrontal cortex (rLPFC, Fig. 4b) was defined as the significant rLPFC cluster identified in Fig. 4a (peak x, y, z = 60, 11, 23). The ROI in the ventromedial prefrontal cortex (VMPFC, Fig. 7b) was defined as the significant VMPFC cluster identified in Fig. 7a (peak x, y, z = -9, 44, -10).

Statistic type for inference (See Eklund et al. 2016): All reported effects are whole brain family-wise error (FWE) cluster level $P < 0.05$ corrected, after thresholding at $P < .001$ uncorrected.

Correction: $P < 0.05$ cluster-wise FWE corrected, with cluster-forming threshold $P < 0.001$ uncorrected.

Models & analysis

n/a	Involved in the study
☐	☒ Functional and/or effective connectivity
☒	☐ Graph analysis
☒	☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity: We performed the psychophysiological interaction (PPI) analyses at both the ROI and whole-brain levels using SPM 12. In the ROI-based PPI analysis, we analyzed functional connectivity between the VMPFC and the dACC ROI that was independently defined by an online meta-analysis of 598 studies using the term "cognitive control" in Neurosynth (peak voxel MNI coordinates: x, y, z = 6, 22, 36; as in Fig. 5c). The VMPFC cluster was defined by a sphere with a 6-mm radius centered at the peak VMPFC activation for belief expectation estimates as in Fig. 7a (MNI coordinates: x, y, z = -9, 44, -10). The PPI model contained the following regressors for observation onsets: (i) BOLD time-series extracted from the VMPFC cluster, (ii) the dummy variable indicating S2 and S3 against S1, and (iii) the product of the above two regressors. We also performed an exploratory PPI analysis at the whole-brain level, seeded in the same VMPFC cluster with the identical sets of regressors (Extended Data Fig. 10).