Competing Decision-Making Systems Are Adaptively Chosen Based on Individual Differences in Brain Connectivity

Y. Yang[1]*, C. Sibert[2], A. Stocco[1]

[1]Department of Psychology, Physics, University of Washington; Seattle, WA, USA 98105
[2]Science and Engineering, University of Groningen, Groningen, Netherlands

*Corresponding author: Y. Yang. Email: chery@uw.edu; A. Stocco. Email: stocco@uw.edu

Abstract: Experiential decision-making could be explained as a result of either memory-based or reinforcement-based processes. Here, for the first time, we show that individual preferences between a memory-based and a reinforcement-based strategy, even when the two are functionally equivalent in terms of expected payoff, are adaptively shaped by individual differences in resting-state brain connectivity between the corresponding brain regions. Using computational cognitive models to identify which mechanism was most likely used by each participant, we found that individuals with comparatively stronger connectivity between memory regions prefer a memory-based strategy, while individuals with comparatively stronger connectivity between sensorimotor and habit-formation regions preferentially rely on a reinforcement-based strategy. These results suggest that human decision-making is adaptive and sensitive to the neural costs associated with different strategies.

One-Sentence Summary: Individual preferences between two decision-making strategies can be predicted by resting-state brain connectivity.

Two different frameworks have been proposed to explain how humans change their behavior from experience. According to one class of theories, previous decisions and their outcomes are recorded as long-term memory traces, and decision-making is guided by their sequential retrieval and examination from long-term memory *(1, 2)*. The decision for which the best outcomes can be remembered is then chosen. We will refer to this as the "Declarative" framework. The other class of theories posits that decision-making from experience is guided by basic mechanisms of reinforcement learning, whereby the value of the outcome of each previous decision incrementally modifies an internal cached value associated with each option *(3, 4)*. When making a decision, the option with the highest expected value is chosen. We will refer to this as the "Procedural" framework. Despite making similar predictions about the outcome of decision-making processes (*5*), these two mechanisms depend on different neural resources. In the Declarative framework, decisions from experience depend on the brain circuits involved in the storage and retrieval of episodes in long-term memory, such as the medial temporal lobe *(6)* and the lateral prefrontal cortex *(7)*. In the Procedural framework, decisions from experience rely on brain circuits for implicit reward learning and habit formation, such as the basal ganglia *(8)* and the supplementary motor area *(9)*.

Both systems are concurrently active at any given time *(10, 12)*. So, why should individuals choose to rely on one system over the other? One likely explanation is that humans have inherently limited processing capacity *(13)*, and therefore adopt different procedures to maximize the outcome of a decision while consuming the minimum amount of cognitive resources *(14, 15)*. And because the cost of a decision also depends on the neural resources it consumes, individuals with different characteristics would make decisions in different ways. This paper puts forward the hypothesis that individuals rely on Declarative or Procedural processes based on their relative efficiency of their corresponding neural circuits. A common index of neural efficiency is represented by the resting-state connectivity between regions within a circuit, i.e., the degree of correlation in their spontaneous activity between pairs of regions *(16)*. Higher correlations at rest reflect tighter coupling of neural dynamics and greater exchange and integration of communication between regions *(17)*. Thus, relatively higher functional connectivity within the Declarative or the Procedural circuit should predict greater reliance on the corresponding system.

An important obstacle in testing our hypothesis is that the choice between the two systems does not depend solely on their mental and neural costs but also on their relative effectiveness in a given task. For example, it has been argued that the Procedural system would be preferred when decisions are probabilistic and the stimuli are difficult to verbalize *(12, 18, 19)*. If, for some reason, one system is better suited for a given task than the other, an individual preference based on neural efficiency would be overridden by the effectiveness of the alternate system. To investigate the relationship between brain connectivity and preferred decision-making processes requires a task where the Procedural system is as effective as the Declarative system strategy. One such task is the Incentive Processing task (IPT: *(20)*, which guarantees that either system yields the same expected payoff. In this task, participants repeatedly guess whether a hidden number is greater or smaller than five by pressing one of two buttons, and receive monetary feedback for correct guesses. Once the choice is made, the number is revealed and feedback is provided (Figure 1A). Unbeknownst to participants, the number is chosen *after* their choice and follows a predefined feedback schedule. Under these conditions, preferences for either the Declarative or Procedural decision-making systems should depend only on the neural costs for each individual participant. Because the number of wins and losses is fixed and predefined, behavioral data from the IPT cannot be analyzed using accuracy measures. Instead, individual behavioral differences can be measured by computing the probability of choosing a different option (i.e., from "win" to "lose" or vice versa) after receiving feedback from the previous trial, referred to as the *shift probability*.

Shift probabilities were computed for the different types of feedback ("win", "lose", or "neutral" when the number is exactly five) and the two different types of experimental blocks that were used in the task ("Mostly Lose", in which 6 loss trials pseudo-randomly interleaved with either 1 neutral and 1 reward trial, 2 neutral trials, or 2 reward trials, and "Mostly Win", in which 6 reward trials pseudo-randomly interleaved with either 1 neutral and 1 loss trial, 2 neutral trials, or 2 loss trials). Under these conditions, individual preferences for Declarative vs. Procedural-based decision-making should maximally reflect the underlying efficiency of the corresponding circuit.

To test this hypothesis, we analyzed a subset of 200 participants from the Human Connectome Project *(21)* for whom neural and behavioral measures from the IPT as well as resting-state fMRI data were available. Figure 1B illustrates the methodology of this study. To determine whether a participant relied on the Procedural or Declarative mechanisms, each participant's behavior was fitted to two parametrized computational models, one implementing Declarative decision-making and one using the Procedural system.

In the Declarative model, each decision/outcome pair is stored as an episodic trace. Because of forgetting, the availability $A(t)$ of each episodic trace at time t decays according to a power law *(22)*, i.e. $A(\mathrm{t}) = \sum_n \log\,[t - t(n)]^{-d}$. In this equation, $t(n)$ is the time at which the same decision/outcome pair has been experienced for the n-th time and d is an individual-specific forgetting rate *(23)*. A decision is made by retrieving the most available memory of a "win" outcome and executing the corresponding decision (i.e., "more" or "less"). Memories are selected probabilistically by adding transient noise to each memory's availability, drawn from a logistic distribution with zero mean and scale parameter μ.

In the Procedural model, the two possible decisions ("more" or "less") are represented as two actions. Each action has an associated value V that is updated over time. After the n-th decision, the value $V(n)$ of an action is updated using a reinforcement-learning-like equation *(24)*: $V(\mathrm{n}) = V(n-1) + \alpha[R - V(n-1)]$. In this case, R is the reward associated with an outcome, (1 for "win", -1 for "lose"), and α is the learning rate. Actions are selected probabilistically by adding transient noise to each action's value, drawn from a logistic distribution with zero mean and scale parameter v. Both models were implemented in the same architecture, ACT-R *(25)*, so that all of the other cognitive, perceptual, and motor components could be kept equal between the two. Thus, each model depends on one learning parameter (d or α) and one noise parameter (μ or v). Each combination of parameters was explored using a grid search and the distribution of their shift probabilities for each block was recorded. Each participant was then assigned to either the Declarative or Procedural group based on which version of the model had the greatest log-likelihood of producing their observed behavioral data.

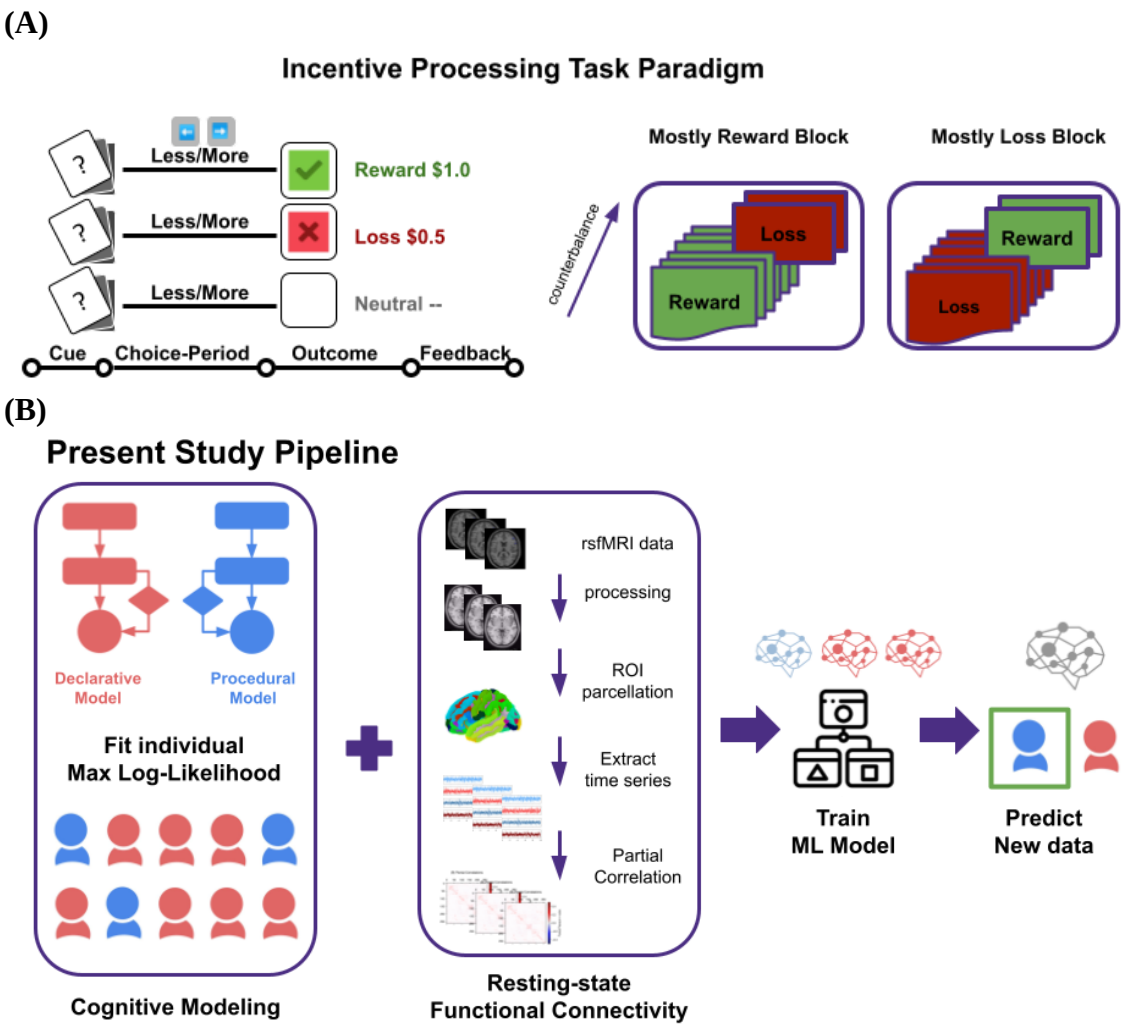


Fig. 1. (A) Incentive processing task paradigm. (B) The pipeline of the present study. First, each subject's behavioral data are fit by two distinct cognitive models, Declarative vs. Procedural model through maximum log-likelihood. Second, each subject's neuroimaging data are processed to construct resting-state functional connectivity matrices. Lastly, machine learning models are trained to predict unseen subject's decision-making process from the neuroimaging data.

Although indistinguishable in most cases, the two models behave differently in the IPT. Because, in the Declarative model, the availability of a memory grows with the number of times it has been experienced as the task continues, feedback from the most recent decision has less influence on future choices, which are instead affected by the longer history of previous decisions. The Procedural model, on the other hand, does not maintain a memory of previous choices and adjusts the value of an action after every feedback; thus, it is more sensitive to the most recent outcome. Because of this, we expect that individuals best fitted by the Declarative model would be less prone to changing their choice preferences immediately after feedback, and thus would exhibit smaller shift probabilities. And because the two mechanisms rely on different neural substrates, we expect that regional differences in BOLD activation would reflect the neural cost associated with each. Therefore, we expect individuals best fitted by Declarative models to show greater activity in regions associated with episodic encoding and retrieval, such as the

hippocampus and the ventral frontal cortex. Correspondingly, we expect individuals best fit by the Procedural model to display greater activity in the circuits associated with habit formation, such as the medial frontal cortex and the basal ganglia. Finally, we expect the type of decision-making processes used to depend on the patterns of functional connectivity at rest. To test this latter prediction, a classifier was trained to predict whether a participant would be best fit by a Declarative or Procedural model from the same individual's functional connectivity data. We expect that functional connections that successfully predict the reliance on the Declarative mechanism for decision-making would be found between memory encoding and retrieval regions. Conversely, we expected that functional connections predictive of reliance on the Procedural mechanism reward-based learning would be found in sensorimotor cortices and the basal ganglia circuit.

Results

Decision-Making Process Identification

By excluding participants who did not complete the IPT and two sessions of resting-state fMRI scanning, a total of 199 participants were fit by the two ACT-R models. Of these, 127 were best fit by the Declarative model and thus were included in the Declarative group. The remaining 72 individuals were best fit by the Procedural model and included in the Procedural group (Figure 2). A logistic mixed-effects model was conducted using orthogonal contrast coding as implemented in the lme4 package in R. Group (Declarative vs. Procedural), Block Type (Mostly Win vs. Mostly Loss), and Feedback (Win vs. Loss; Figure 1A) were treated as fixed effects, and individual subjects were treated as random effects. (Full statistical results are shown in Table S1).

As expected, the switch probability was found to be statistically different between the two groups ($z = -6.11$, $p < 0.001$), and consistent with the predicted model differences. Specifically, participants in the Procedural group showed a significantly greater probability of shifting their behavior after a "Loss" feedback, while those in the Declarative group rarely did so. This finding was consistent with our understanding of the memory retrieval process, with the Declarative system becoming less sensitive over time to individual items of feedback in relation to the accumulation of historical decision/outcome pairs.

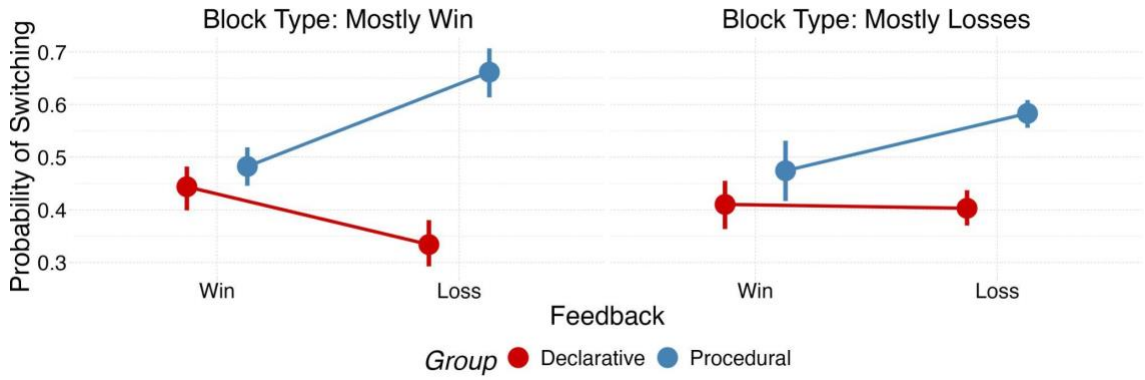


Fig. 2. The probability of response switching by two groups of individuals identified as either preferring Declarative or Procedural strategy in this decision-making task. Of 199 participants, 127 subjects were best fit by the Declarative model, and 72 were best fit by the Procedural model.

Group Differences in BOLD Activations During Task

In addition to finding significant behavioral differences between two groups of participants, we also found relevant and consistent differences in their task-based fMRI data. In line with previous studies using the same dataset (*26*), our investigation focused on the difference in brain activity between Mostly Win and Mostly Lose blocks. A *T*-test was performed to identify brain regions more active in one group over the other. The test was weighted using the relative log-likelihood of each participant's being fitted by either of the two models, so that participants who had a stronger preference for one process were weighted more (See Materials and Methods). The test identified several brain regions that show significant BOLD signal differences between the Declarative and Procedural groups (Figure 3; Table S2). Specifically, participants in the Declarative group showed greater brain activity in regions involved in memory encoding (e.g., parahippocampal gyri) and retrieval (ventral frontal cortex), while the Procedural group showed greater brain activity in areas involved in habit formation (caudate nucleus) and error prediction and correction (dorsal anterior cingulate cortex, insula).

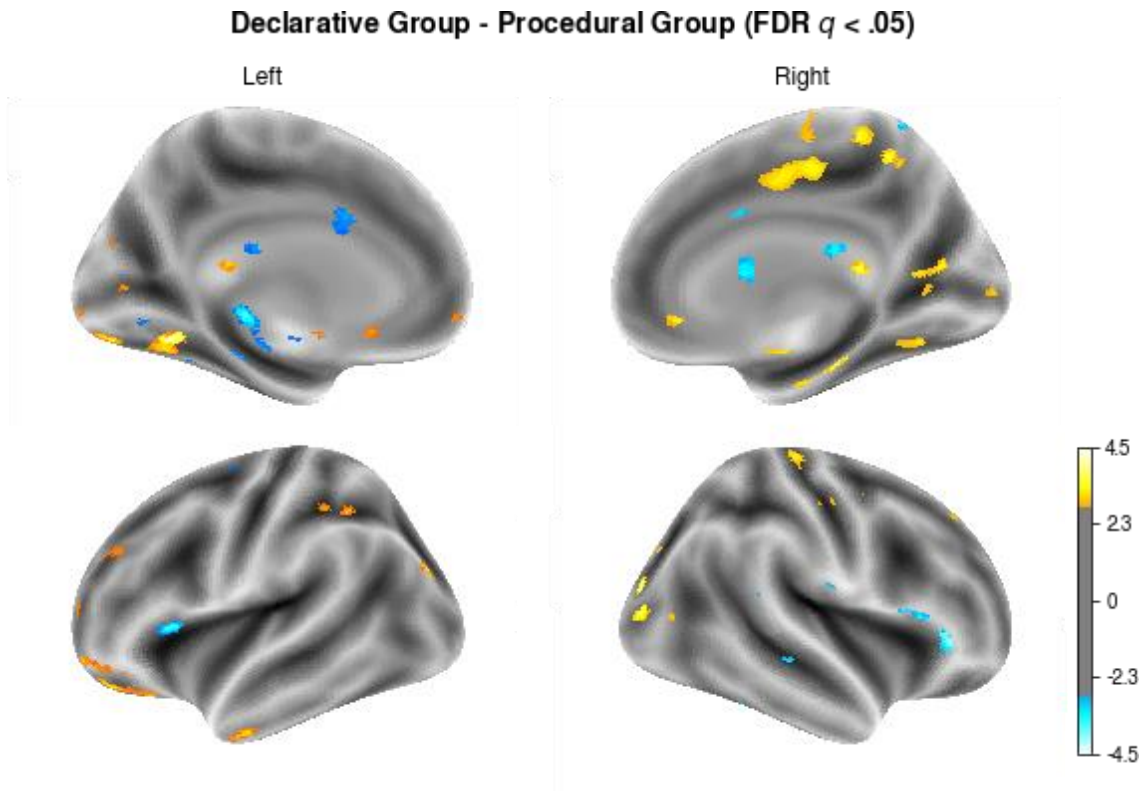


Fig. 3. Differences between Declarative and Procedural participants during the incentive processing task. Regions shown are significant at FDR $q < 0.05$, corresponding to a voxel-wise threshold of $T >= 2.83$.

Predicting Individual Differences in Decision-Making Processes Through Functional Connectivity

Functional connectivity data consisted of pairwise partial correlation matrices between each pair of the 264 regions in the Power parcellation scheme (*27*). To identify which functional connectivity features predict whether a participant would belong to the Declarative or Procedural group, we used a logistic Least Absolute Shrinkage and Selection Operator (LASSO) (*28*) regression model, which reduces the large number of potential predictors while retaining the

most predictive features. The LASSO model was implemented using the glmnet package in R (2010). To ensure generalization, we chose the optimal LASSO hyperparameter λ (0.0239) using a nested cross validation method (See Supplementary Materials). Having the optimal λ, we refit the LASSO model using the Leave-One-Out (LOO) cross-validation approach and obtained the correlation coefficient betas (β) for subsequent brain connectivity analysis.

The prediction performance was evaluated by regular LOO cross-validation and nested cross validation approaches. After refitting the model with optimal λ, the cross-validation accuracy was 99.58%, and ROC-AUC was 0.9996. Nested CV model performance was averaged, resulting in a mean training accuracy score of 100%, and a mean testing accuracy score of 78%. These results indicate that the LASSO model was successful at predicting an individual's preferred decision-making process (Declarative or Procedural) from resting-state brain functional connectivity.

Finally, we examined which functional connectivity features were predictive of individual preferences in decision making. After LASSO regularization, 60 functional connections (approximately 0.01% of the total) had non-zero β parameters, suggesting a very sparse neurofunctional connectivity. Since the ultimate effect of a β parameter on the predicted group assignment depends on the polarity of the underlying functional connectivity, a positive β value has different implications if applied to a positive or negative partial correlation between two regions. Consider, for example, a connectivity value associated with a positive β value: if the underlying connectivity is positive ($r > 0$), then the degree of functional connectivity can be taken as a vote in favor of the Declarative system. If, on the other hand, the underlying connectivity value is negative ($r < 0$), then it should be counted as a vote for the Procedural system.

To make the interpretation of the values unambiguous, we multiplied the β matrix with the average functional connectivity matrix, obtaining a new group-level weighted averaged correlation matrix W. The matrix W can now be interpreted unambiguously, since $W > 0$ predicts reliance on a Declarative process and $W < 0$ predicts reliance on a Procedural process. Figures 4(A) and 4(C) show these different functional connections.

Connections predictive of Declarative processes involved clearly different networks than connections predictive of Procedural processes, as shown in Fig 4(B) and Fig 4(D). As expected, reliance on Declarative processes in decision-making was predicted by greater connectivity in the networks of regions associated with task control (frontoparietal networks and attention networks) and episodic memory (memory retrieval network). Reliance on Procedural learning processes, instead, was predicted by greater connectivity in sensorimotor salience, and the subcortical networks including the basal ganglia.

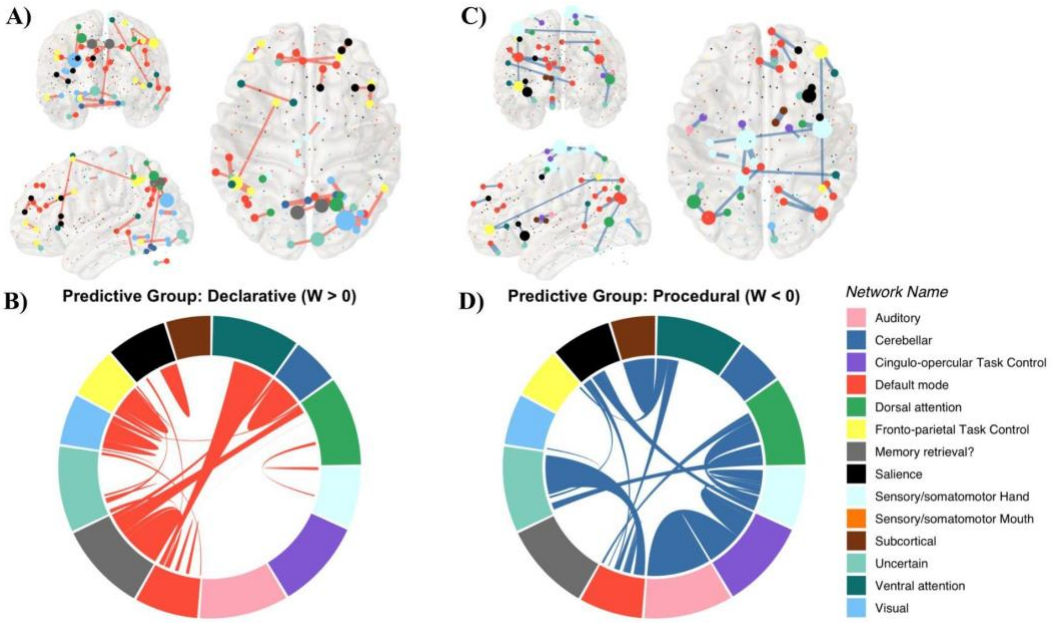


Fig. 4. The distribution of predictive connections in Declarative and Procedural functional networks**. (A)** Anatomical location of the function connection predictive of Declarative group assignment; nodes represent regions, node colors indicate the network they belong to; node size indicated the importance of a region (sum of connectivity values). (**B)** Distribution of the Declarative functional connections within the Power parcellation networks. (**C)** Anatomical location of the functional connections predictive of Procedural group assignment). (**D)** Distribution of the Procedural connections across the Power parcellation networks.

Discussion

This study shows that individuals rely on different mechanisms when deciding from experience, and that this preference is adaptive and reflects individual differences in the functional connectivity between each process' corresponding circuitry. Specifically, individuals exhibiting stronger connectivity between and within frontoparietal and memory retrieval regions tend to use Declarative strategies that are more reliant on episodic encoding and retrieval, while individuals with stronger connectivity in cingulate, sensory, and basal ganglia regions tend to rely on habitual actions and reinforcement learning. An individual's preference can be, in fact, successfully predicted from their underlying functional connectivity.

Although our results shed new light on the neural bases of experiential decision-making, a number of limitations must be acknowledged. First, participants were assigned to the Declarative or Procedural group based on the log-likelihood of a corresponding cognitive model. Because no other ground-truth labels were available, these classifications should be interpreted with caution. Another limitation is that the incentive processing task differs from most decision-making paradigms, as there is no winning rule for participants to learn from the feedback. Thus, although this task was particularly well suited for the current study, more work is needed to determine whether these findings would translate to more realistic situations.

In addition, our study assumes that greater functional connectivity reflects greater efficiency. While much experimental evidence points to this, the mechanisms by which connectivity translates to computational efficiency are not clear. One proposed solution is that connectivity reflects communication efficiency across regions. If so, connectivity should correspond to decision-making noise in our models. In fact, a follow-up analysis of individual model parameters shows that individual preferences for one process over the other largely follow

differences in the estimated noise parameter. That is, individuals whose estimated Procedural noise was larger than the estimated Declarative noise were significantly more likely to be fit by the Declarative model, and vice-versa (see Fig S3A-B and Table S4).

Although we have found evidence that individuals prefer to implement decision-making processes using the circuits that show the greatest connectivity, our analysis does not directly support a causal direction. To conclusively show a direct causal effect, it would be necessary to experimentally alter functional connectivity within the Declarative and Procedural circuits and measure subsequent changes in behavior in the same task. Such an experimental intervention is difficult to carry out in humans, although they are in principle conceivable using pharmacological interventions that directly target one circuit or intracranial direct stimulation of these circuits in patients with implanted electrodes or ECoG grids.

Finally, our study still does not directly address the source of differences in functional connectivity. Two hypotheses are possible. In principle, these differences could reflect underlying genetic or structural differences in brain function. Alternatively, it is possible that differences in functional connectivity simply reflect one's past history of relying on one particular process (and its underlying circuit) over the other. In this case, individuals might have a preference (let's say, for the Declarative system) and repeated use leads to corresponding increases in functional connectivity (let's say, in the memory retrieval network). The two explanations might be intertwined, i.e., small initial differences in the efficiency of these circuits might lead to greater use of one process over the other, and practice over time might result in additional gains in efficiency.

Further investigation is needed to fully untangle these questions, but these results nevertheless represent a promising step forward in directly connecting the mechanical underpinnings of brain circuitry with observed human behavior in decision-making.

References

1. C. Gonzalez, V. Dutt, Instance-based learning: integrating sampling and repeated decisions from experience. *Psychol. Rev.* **118**, 523–551 (2011).
2. N. Stewart, N. Chater, G. D. A. Brown, Decision by sampling. *Cogn. Psychol*. **53**, 1–26 (2006).
3. Y. Niv, Reinforcement learning in the brain. *J. Math. Psychol*. **53**, 139–154 (2009).
4. D. Lee, H. Seo, M. W. Jung, Neural basis of reinforcement learning and decision making. *Annu. Rev. Neurosci*. **35**, 287–308 (2012).
5. S. E. Chelian, J. Paik, P. Pirolli, C. Lebiere, R. Bhattacharyya, "Reinforcement learning and instance-based learning approaches to modeling human decision making in a prognostic foraging task" *in 2015 Joint IEEE International Conference on Development and Learning and Epigenetic Robotics* (ICDL-EpiRob 2015), pp. 116–122.
6. L. R. Squire, S. Zola-Morgan, The medial temporal lobe memory system. *Science*. **253**, 1380–1386 (1991).
7. D. Badre, A. D. Wagner, Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia*. **45**, 2883–2901 (2007).
8. H. H. Yin, B. J. Knowlton, The role of the basal ganglia in habit formation. *Nat. Rev. Neurosci.* **7**, 464–476 (2006).
9. S. Nieuwenhuis, C. B. Holroyd, N. Mol, M. G. H. Coles, Reinforcement-related brain potentials from medial frontal cortex: origins and functional significance. *Neurosci. Biobehav. Rev.* **28**, 441–448 (2004).
10. K. Foerde, B. J. Knowlton, R. A. Poldrack, Modulation of competing memory systems by distraction. *Proceedings of the National Academy of Sciences*. **103**, 11778–11783 (2006).
11. R. A. Poldrack, M. G. Packard, Competition among multiple memory systems: converging evidence from animal and human brain studies. *Neuropsychologia*. **41**, 245–251 (2003).
12. R. A. Poldrack, J. Clark, E. J. Paré-Blagoev, D. Shohamy, J. Creso Moyano, C. Myers, M. A. Gluck, Interactive memory systems in the human brain. *Nature*. **414**, 546–550 (2001).
13. H. A. Simon, A behavioral model of rational choice. *Models of man, social and rational: Mathematical essays on rational human behavior in a social setting*, 241–260 (1957).
14. J. W. Payne, J. R. Bettman, E. J. Johnson, *The Adaptive Decision Maker* (Cambridge University Press, 1993).
15. G. E. Gigerenzer, R. E. Hertwig, T. E. Pachur, Heuristics: The foundations of adaptive behavior. **844** (2011).
16. X. Shen, E. S. Finn, D. Scheinost, M. D. Rosenberg, M. M. Chun, X. Papademetris, R. T. Constable, Using connectome-based predictive modeling to predict individual behavior from brain connectivity. *Nat. Protoc*. **12**, 506–518 (2017).
17. M. P. van den Heuvel, H. E. Hulshoff Pol, Exploring the brain network: a review on resting-state fMRI functional connectivity. *Eur. Neuropsychopharmacol*. **20**, 519–534 (2010).
18. M. J. Frank, L. C. Seeberger, R. C. O'reilly, By carrot or by stick: cognitive reinforcement learning in Parkinsonism. *Science*. **306**, 1940–1943 (2004).
19. B. J. Knowlton, L. R. Squire, M. A. Gluck, Probabilistic classification learning in amnesia. *Learn. Mem*. **1**, 106–120 (1994).
20. M. R. Delgado, L. E. Nystrom, C. Fissell, D. C. Noll, J. A. Fiez, Tracking the hemodynamic responses to reward and punishment in the striatum. *J. Neurophysiol*. **84**, 3072–3077 (2000).

21. D. C. Van Essen, S. M. Smith, D. M. Barch, T. E. J. Behrens, E. Yacoub, K. Ugurbil, WU-Minn HCP Consortium, The WU-Minn Human Connectome Project: an overview. *Neuroimage*. **80**, 62–79 (2013).
22. Newell, Rosenbloom, Mechanisms of skill acquisition. In J. R. Anderson, *Cognitive skills and their acquisition* (1981)
23. P. Zhou, F. Sense, H. van Rijn, A. Stocco, Reflections of idiographic long-term memory characteristics in resting-state neuroimaging data. *Cognition*. **212**, 104660 (2021).
24. R. S. Sutton, Learning to predict by the methods of temporal differences. *Mach. Learn.* **3**, 9–44 (1988).
25. J. R. Anderson, *How Can the Human Mind Occur in the Physical Universe?* (Oxford University Press, 2009).
26. D. M. Barch, G. C. Burgess, M. P. Harms, S. E. Petersen, B. L. Schlaggar, M. Corbetta, M. F. Glasser, S. Curtiss, S. Dixit, C. Feldt, D. Nolan, E. Bryant, T. Hartley, O. Footer, J. M. Bjork, R. Poldrack, S. Smith, H. Johansen-Berg, A. Z. Snyder, D. C. Van Essen, WU-Minn HCP Consortium, Function in the human connectome: task-fMRI and individual differences in behavior. *Neuroimage*. **80**, 169–189 (2013).
27. J. D. Power, A. L. Cohen, S. M. Nelson, G. S. Wig, K. A. Barnes, J. A. Church, A. C. Vogel, T. O. Laumann, F. M. Miezin, B. L. Schlaggar, S. E. Petersen, Functional network organization of the human brain. *Neuron*. **72**, 665–678 (2011).
28. R. Tibshirani, Regression shrinkage and selection via the lasso. *J. R. Stat. Soc*. **58**, 267–288 (1996).
29. R. W. Cox, AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* **29**, 162–173 (1996).
30. R Core Team, *R: A Language and Environment for Statistical Computing* (2020), (available at https://www.R-project.org/).
31. N. C. Silver, W. P. Dunlap, Averaging correlation coefficients: Should Fisher's z transformation be used? *J. Appl. Psychol.* **72**, 146–148 (1987).
32. I. Kotseruba, J. K. Tsotsos, 40 years of cognitive architectures: core cognitive abilities and practical applications. *Artificial Intelligence Review*. **53**, 17–94 (2020).
33. A. Stocco, A Biologically Plausible Action Selection System for Cognitive Architectures: Implications of Basal Ganglia Anatomy for Learning and Decision-Making Models. *Cogn. Sci.* **42**, 457–490 (2018).
34. T. Haile, C. S. Prat, A. Stocco, "One Size Doesn't Fit All: Idiographic Computational Models Reveal Individual Differences in Learning and Meta-Learning Strategies" in *Proceeding of the 18th International Conference on Cognitive Modeling* (2020), pp. 75–81.
35. Y. C. Yang, A. M. Karmol, A. Stocco, Core Cognitive Mechanisms Underlying Syntactic Priming: A Comparison of Three Alternative Models. *Front. Psychol.* **12**, 662345 (2021).
36. A. Vabalas, E. Gowen, E. Poliakoff, A. J. Casson, Machine learning algorithm validation with a limited sample size. *PLoS One.* **14**, e0224365 (2019).

Acknowledgments

Funding

This work was supported by award FA9550-19-1-0299 from the Air Force Office of Scientific Research to AS.

Author contributions

Conceptualization: AS, CS, YY
Methodology: YY, AS
Formal analysis: YY, AS
Visualization: YY, AS
Funding acquisition: AS
Project administration: AS
Supervision: AS
Writing – original draft: AS, YY, CS
Writing – review & editing: N/A

Competing interests

Authors declare that they have no competing interests.

Data and materials availability

All data and codes are available in the OSF repository: https://osf.io/wf4my/

Materials and Methods

This study analyzed both behavioral and neuroimaging data obtained from a subset of the Human Connectome Project (HCP) dataset (2013). Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research and by the McDonnell Center for Systems Neuroscience at Washington University. A total of 199 participants (111 females, 85 males, and 3 did not disclose) who completed both sessions of the task-based fMRI gambling game were included in this study. All participants were healthy adults with no neurodevelopmental or neuropsychiatric disorders. The experimental protocol, subject recruitment procedures, and consent to share de-identified information were approved by the Institutional Review Board at Washington University.

The Incentive Processing Task in the HCP

This incentive decision making task was adapted from the gambling paradigm developed by Delgado and colleagues *(20)*. Participants were asked to guess if the number on a mystery card (represented by a "?", and ranging from 1 to 9) was more or less than 5. After making a guess, participants were given feedback, which could take one of three forms, Win (a green up arrow and $1), Loss (a red down arrow and -$0.50), or Neutral (a gray double-headed arrow and the number 5). The feedback did not depend on the subject's response, but was determined in advance; the sequence of pre-defined feedback was identical for all participants. The task was presented in two runs, each of which contains 64 trials divided into eight blocks. Blocks could be Mostly Loss (6 loss trials pseudo-randomly interleaved with either 1 neutral and 1 reward trial, 2 neutral trials, or 2 reward trials) or Mostly Win (6 win trials pseudo-randomly interleaved with either 1 neutral and 1 loss trial, 2 neutral trials, or 2 loss trials). In each of the two runs, there were two Mostly Win and two Mostly Loss blocks, interleaved with 4 fixation blocks (15 seconds each). All participants received money as compensation for completing the task, and the amount of reward is standard across subjects.

fMRI Data Processing and Analysis

This study employed the "minimally preprocessed" version of resting-state fMRI data and incentive processing task fMRI data, which has already undergone a minimal number of standard preprocessing steps including artifact removal, motion correction, normalization, and registration to the standard MNI ICBM152 template. Additional preprocessing steps were performed using the AFNI software (*29*), including despiking, spatial smoothing with an isotropic Gaussian 3D filter FWHM of 8 mm, and removal of linear components related to the six motion parameters and their first-order derivatives.

Functional connectivity measures were constructed from the HCP resting-state data using Power et al.'s whole-brain parcellation (*27*). This parcellation was used to construct a 264 Region of Interest (ROI) functional atlas, with each ROI containing 81 voxels. This parcellation atlas is defined in the MNI space and was applied to all participants in the HCP dataset. The extraction of the time series and calculation of the connectivity matrices was performed using R (*30*) and Python. Pearson correlation coefficients and partial correlation coefficients between the time series of each brain region were calculated for each participant, resulting in a 264 × 264 symmetric connectivity matrices for each session for each subject. The average correlation coefficients across subjects were calculated by first transforming each r value into a Z-value, and

then retransforming the average *Z* value back into an equivalent *r* value using the hyperbolic tangent transformation (*31*).

For task-based fMRI data analysis, we specified the first-level analysis model and estimated the parameters corresponding to the difference between Mostly Win and Mostly Lose blocks, as in (*26*). The resulting contrast maps from were for each subject were then used in a second-level weighted t-test between-Declarative and Procedural groups. The test was implemented using AFNI's 3dttest++ software, and its weights corresponded to the absolute difference in log-likelihood between the best-fitting Declarative and best-fitting Procedural model. This way, the contribution of each observation was proportionally scaled to the evidence favoring each participant's assignment to their groups. The statistical significance level was set at a significance level of $q < 0.05$ corrected for multiple comparisons using a False Discovery Rate (FDR) procedure.

Response Switch Analysis

Because in the Incentive Processing task the feedback is scheduled in advance and does not depend on actions taken by participants, all participants have exactly the same raw performance (in terms of "correct" choices), and is thus impossible to analyze their behavior in terms of either accuracy or learning. This poses a challenge when trying to analyze behavior since the most common metrics used in decision-making from experience (accuracy and learning rates) cannot be used. Instead, the most meaningful way to examine participants' behavior in response to feedback is by analyzing their Win-Stay, Lose-Shift (WSLS) probabilities. Thus, our main behavioral measurement was the tendency to switch responses after Loss feedback and after Win feedback. The response switch is coded as 0 if the current response is the same as the next response, and coded as 1 if the current response is not the same as the next response. Because the response switch is a binary variable, the analysis was conducted with logistic mixed-effects models using orthogonal contrast coding as implemented in the "lme4" package in R. Given that Neutral trials make up only a small proportion of total trials, they were excluded in statistical tests. In the mixed-effect model, Block Type (Mostly Win or Mostly Loss) and Trial Type (Win or Loss) were treated as fixed effects, and individual participants were treated as random effects. The parameters were estimated based on the maximum likelihood.

On the group level, there is no significant effect of feedback nor Block Type on the probability of switching responses. However, and critically for this study, participants do exhibit different behavioral response profiles on an individual level. **Fig. S1.** shows the mean probability of response switching as a function of Trial Type (feedback received) and Block Type, each line representing the mean performance of a single individual.

We also examined whether the response times change as a function of previous feedback (the Trial Type of the previous trial) and Block Type. Excluding neutral trials in the statistical analysis, on average, participants tend to take longer when making decisions in Mostly Reward blocks than in Mostly Loss blocks ($\beta = 15.21$, $SE = 6.39$, $p = 0.017$), regardless of previous feedback. Compared to the probability of response switching, however, the pattern of RTs was found to be noisier and less consistent across individuals and was therefore not included in the following modeling analysis.

Computational Models

While the behavioral data does not reveal major effects across subjects, it offers an exciting opportunity from a modeling perspective. There exist two competing explanations of how decision-making occurs in a repeated choice paradigm, one based on episodic memory of previous choices (1) and one based on reinforcement learning (3). Each explanation is dependent

on different mechanisms, and, ultimately, reliant on different strategies. Both explanations were implemented as two computational models in the ACT-R cognitive architecture *(25)*: a Declarative model, reliant on memory retrieval, and a Procedural model, which makes use of reinforcement learning.

ACT-R is the most prominent and successful cognitive architecture in psychology and neuroscience (*32*), which provides a comprehensive framework to understand a wide variety of cognitive functions and learning processes. In ACT-R, decision-making processes are represented in two fundamental ways: chunks and production rules. A Chunk is a vector-like structure that stores semantic or episodic memories. A Production rule is a basic action unit that represents procedural knowledge as an "IF-THEN" conditional statement. Productions and chunks interact through a set of modules which represent different cognitive processes. Two distinct cognitive representations in ACT-R make it the ideal modeling architecture for the present study.

Declarative Model

The Declarative model relies on the Declarative module to retrieve a memory of prior choices and their corresponding outcomes (**Fig. S2. A**). When presented with a mystery card, the model selects one of the possible choices, LESS or MORE, for evaluation, and attempts a retrieval of a previous episodic memory in which that action was used. If the retrieved memory contains a WIN result associated with that choice, it will execute that action, but if the history contains a LOSE or NEUTRAL result, the model will execute the alternate action. After being presented with feedback, the mode the model encodes as new memory associating the action with its outcome. Memories are retrieved based on their *activation*, *a* noisy quantity that depends on the frequency and recency with which the decision-outcome episodes have been experienced (Eq. 1).

Procedural Model

By contrast, the Procedural model (**Fig. S2. B**) represents the possible actions of the decision-making processes as competing rules, and reinforcement learning is used to increase the use of the rule that leads to the best outcomes. Instead of encoding each trial as a memory of action and associated feedback, the model has two competing rules that execute the MORE and LESS actions. When presented with the mystery card, the model chooses one of the rules to execute based on its expected value. Initially, both rules have equal value, and one will be chosen at random. After each decision, the model is presented with a WIN, LOSE, or NEUTRAL response, and this feedback is encoded as the reward term in the reinforcement learning equation (Eq. 2) (+1 for a WIN result, -1 for a LOSE result, and 0 for a NEUTRAL result). Positive rewards will encourage the model to repeat the associated action, while a sequence of losses will decrease the value of an action and encourage the selection of the alternate action.

Individual Fit and Model Evaluation

To examine the predictions of our model, we used a grid-search approach to find the best possible parameters within the parameter space shown in Table S3. Each model simulates 64 trials, the same as the experimental paradigm for participants, repeated over 50 runs. The simulated stimuli were presented in the same order as the real experimental stimuli to avoid any potential noise from sequence effects in the simulation. Following the six conditions (Reward, Loss, Neutral trials in Mostly Reward Block and Reward, Loss, Neutral trials in Mostly Loss

Block), the mean probability of response switching, P(Switch), and its standard deviation are computed.

In order to evaluate the goodness-of-fit for individual fitting, we estimated maximum Log-Likelihood across the parameter space. The likelihood function of a particular model *m* with parameters **θ** *given data* ***x***, indicated as $L(\mathrm{m}, \theta \mid x)$, is the probability that the parameterized model would produce that data, that is $L(\mathrm{m}, \theta \mid x) = P(x \mid \mathrm{m}, \theta)$. Common model comparison metrics, such as the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), are both based on likelihood, but rely on closed-form likelihood functions. While it is possible to derive such functions for simple models (such as logistic models or linear models), they can be incredibly difficult to derive for more complex models and impossible for arbitrarily complex models based on ACT-R and other high-level architectures. Some attempts have been made to evaluate complex models with basic likelihood metrics like BIC *(33, 34)*. However, the equation used to estimate BIC is a closed-form approximation that is based on Residual Sum of Squares and was originally derived for linear models; as such, it does not necessarily hold for ACT-R.

In this paper, we followed the computationally expensive but more accurate solution of empirically calculating the likelihood function by simulating each model and set of parameters multiple times and calculating the empirical probability distribution of each set of results (*35*). Knowing the mean and standard deviation of this distribution, the value of $P(x|m, \boldsymbol{\theta})$ can then be calculated directly. If a model is designed to predict n data points (corresponding, for instance, to different experimental conditions), the likelihood can be expressed as the joint probability that any of those data points can be produced. For simplicity, and assuming independence, this can be expressed as the product of the probability of observing each individual data point in the empirical data, i.e., $L(\mathrm{m}, \boldsymbol{\theta} \mid x_1,\ x_2, \dots, x_n) = \prod_i L(m, \boldsymbol{\theta} \mid x_i) = \prod_i P(x_i|m, \boldsymbol{\theta})$. The probability that a model *m* with parameters **θ** would generate the observed data x_i can be calculated directly from the mean $\mu_{i,m,\theta}$ and standard deviation $\sigma_{i,m,\theta}$ of the model's output for the *i*-th variable, that is $Z\left(x_i - \mu_{i,m,\theta}\right)/\sigma_{i,m,\theta}$. Finally, to avoid computational problems with vanishing small probabilities, it is common to use *log*-likelihoods, so that:

$$\log\ = \log \prod_i L(m,\ \theta \mid x_i) = \sum_i \log P(x_i \mid m, \theta) = \sum_i \log\left[Z\left(x_i - \mu_{i,m,\theta}\right)/\sigma_{i,m,\theta}\right]$$

After model fitting, an examination of the resulting parameters for the best-fitting Declarative and Procedural models for each participant shows that differences in the two noise parameters are highly predictive of which model would best fit the behavioral data (**Fig. S3**). A logistic regression model similarly found that noise parameter values in the two models were jointly significant in predicting which model would best fit each participant (Table S4), with higher levels of Declarative noise predicting a greater likelihood of being best fit by the Procedural model ($p < 0.001$), and vice versa ($p < 0.001$). Note that this does not mean that our models are fitting noise, instead, we argue that noise parameters reflect the degree to which choice is dependent on the information stored by the respective memory system (declarative or procedural), and are thus important in separating the two decision-making strategies.

<u>Supervised Classification Model</u>

To explore if individuals' behavioral differences between Declarative and Procedural strategies are indicated by an individual's underlying brain structure, we trained three most commonly used supervised classification models (Logistic regression model, Decision Tree model, and Random Forest Model), using resting-state functional connectivity as input its variable, and predicted the probability of a participant being labeled as either preferring

Declarative or Procedural strategy. Considering the equally high accuracy (> 0.8) among these classification models (Random Forest Model (accuracy = 0.92) is slightly outperformed than other two models), we are confident to say that our Machine Learning models work very well in predicting the strategy selection from an individual's resting state neuroimaging data.

In addition to the predictive power of machine learning models, another very important dimension we need to carefully consider in Machine Learning related research is the model's interpretability. While some ML models are excellent at predicting outcome variables, as complexity grows exponentially (Deep neural networks), they become a black box that is fundamentally difficult to interpret. Therefore, choosing an appropriate ML model with reasonable predictability and interpretability is critical in research. We chose the logistic regression model because it is the most simple and powerful binary classification model that has been widely used in many fields of research. It could be used to predict the likelihood of an event happening or a choice being made. Rather than fitting data as a straight line, the logistic regression model uses the logistic function to squeeze the output of a linear equation between 0 and 1. The equation of the logistic function to estimate the probability of an event x is shown below:

$$\mathrm{P}(x) = \frac{1}{1 + e^{-x}}$$

In order to handle an imbalanced dataset with unequal target labels, up-sampling was applied by randomly adding data from the minority class. We also applied individualized weights to each training sample, which is the absolute difference of maximum log-likelihood between two models. Specifically, for subjects who are better distinguished by two models (Declarative vs. Procedural), we increased the weights of these data points in later ML model training, while for those who had a very close fit maximum log-likelihood between two models, the training procedure was less reliant on these samples. Having 69,696 (264 ROI × 264 ROI) connections, we want to select only the most important connections contributing to the prediction, therefore, Lasso regularization was applied to the Logistic Model. Lasso is a machine learning regression analysis technique that performs both variable selection and regularization in order to improve the prediction accuracy and interpretability of the computational model. It can reduce model complexity by penalizing large numbers of coefficients and also prevents overfitting which may result from simple linear regression. Lasso minimization is calculated using Eq 5 where the tuning parameter λ controls the degree of penalty: for greater values of λ, more coefficients are forced to become 0.

$$\text{Loss} = \text{Error}(Y - \hat{Y}) + \lambda \sum_{1}^{n} w_i^2$$

To account for the large disparity between the number of participants and the number of predictors, we performed a Grid search cross-validation using glmnet package in R (2010) to determine the best value for the fit hyper-parameter $\lambda_{regular}$ (0.015). To alleviate the potential problems of small sample size in neuroimaging studies, Vabalas and colleagues (*36*) used Nested CV approaches producing robust and unbiased model performance regardless of sample size. Following their suggestions, we fit the model with *n*-iteration nested cross-validation (*n* = 200) to determine the optimal hyperparameter λ. For each iteration, the dataset was randomly split into training and testing (the ratio of training to testing is 1:3). Instead of adopting computational expensive Leave-One-Out (LOO), we adopted the k-folds cross-validation (*k* = 20) method. The whole dataset was randomly split into 20 folds and trained on 19 folds of samples and the prediction was made on the remaining one-fold of samples to obtain the best hyperparameter

value λ_k, which gave the lowest classification error. With each best lambda λ_k, the model was restrained and made predictions on testing datasets. Then the process was repeated k times. To guarantee maximum generalizability, we chose the median of lambdas ($\lambda_{nested} = 0.0239$).

The mean accuracy score, true positive rate (TPR), true negative rate (TNR), false positive rate (FPR), and false negative rate (FNR) were calculated across all folds to evaluate the overall performance of the model, taking sample weights into account. By definition, the receiver operating characteristic curve (ROC) demonstrates the performance of a classification model by plotting the relationship between TPR vs. FPR at different classification thresholds. We calculated the AUC (Area under the curve), which is one of the most important metrics for evaluating a classification model's performance; as the AUC of a model approaches 1, the model approximates an ideal, perfect classifier. It provides information about how well a classification model is capable of distinguishing between classes.

Fig. S1. The mean probability of response switching as a function of feedback and block type. Each color dot and gray line represents the mean probability of response switching of a single participant, and the black dot represents the mean and 95% confidence interval across participants.

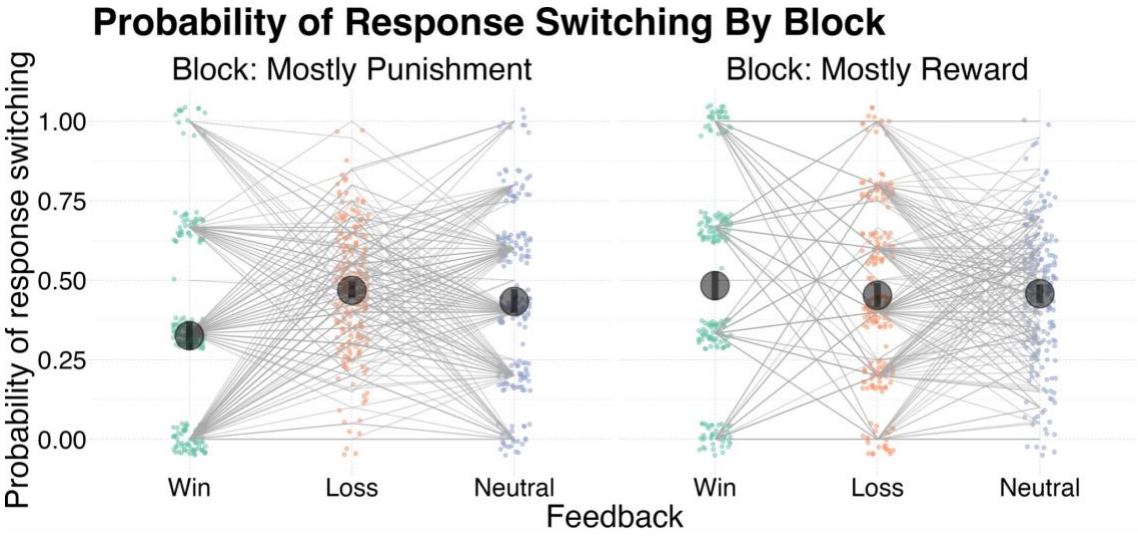


Fig. S2. (A) The flowchart of the Procedural model in (B) The flowchart of the Declarative model. Both models are implemented in the ACT-R cognitive architecture, which includes basic cognitive resources and visuo-motor capabilities that are common to both models. Arrows represent the flow of information through the model components. Rounded rectangles represent fundamental cognitive modules (e.g., visual, motor, long-term memory); squared rectangles represent atomic mental actions ("production rules" in the ACT-R framework); document icons represent episodic traces ("chunks" in ACT-R jargon) and include incoming sensory information.

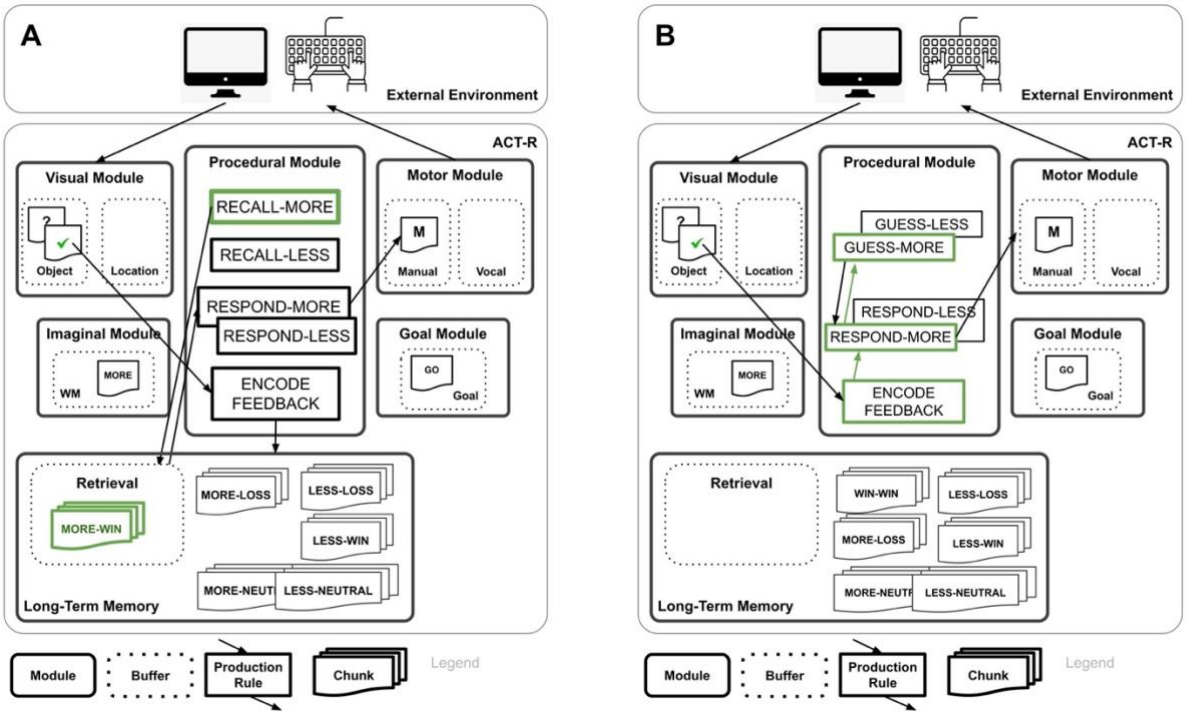
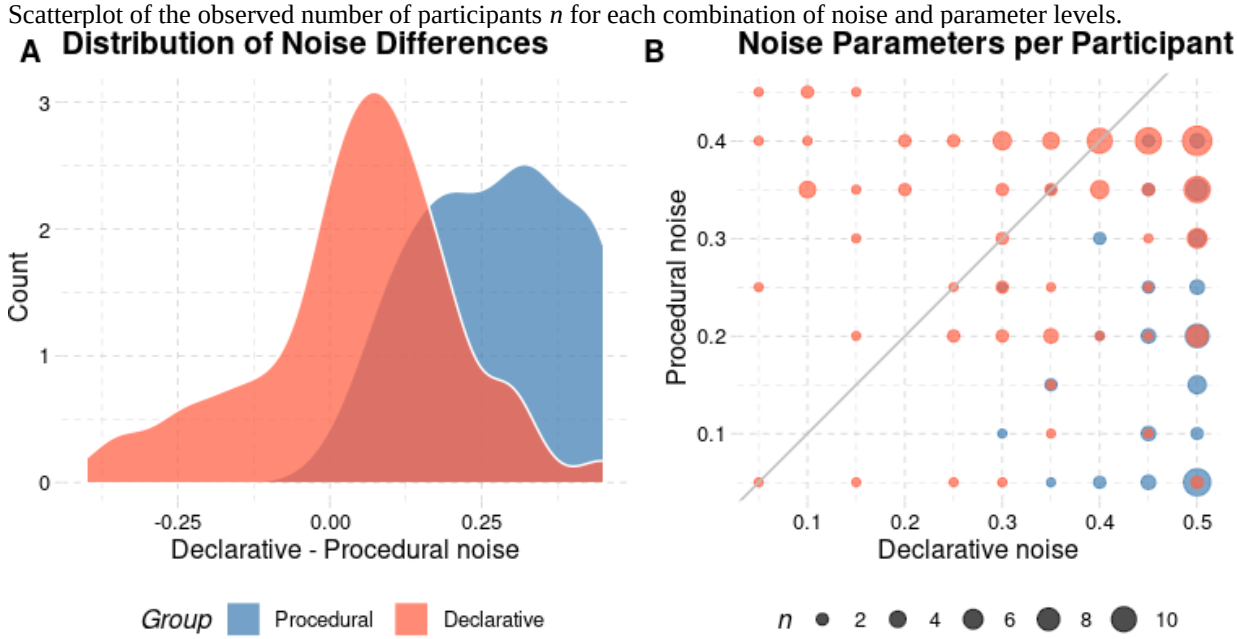


Fig. S3: Noise parameters for the best-fitting Declarative and Procedural models are significantly different across Declarative and Procedural groups. (A) Distribution of the differences between noise parameters (B) Scatterplot of the observed number of participants *n* for each combination of noise and parameter levels.



5

Table S1. Results of the Logistic Mixed Effects Model of the Probability of Response Switch.

Statistical Test	**odds ratio**	***se***	***z***	***p***
(Intercept)	0.88*	0.05	-2.30	**0.022**
Model Group	0.71***	0.04	-6.11	**<0.001**
Block Type	0.98	0.03	-0.80	0.423
Trial Type	1.08**	0.03	2.84	**0.005**
Model Group by Block Type	1.07*	0.03	2.46	**0.014**
Model Group by Trial Type	0.8***	0.02	-8.10	**<0.001**
Block Type by Trial Type	1.02	0.03	0.78	0.434
Model Group by Block Type by Trial Type	1.10***	0.03	3.47	**0.001**
Random Effect				
σ^2	3.29			
ICC	0.12			
N HCPID	199			
observation	9746			
Marginal R^2/Conditional R^2	0.030/0.144			
Log-Likelihood	-4010.323			

** p<0.05 ** p<0.01 *** p<0.001*

Table S2. Results from Task fMRI. For clarity, only clusters spanning > 50 voxels, and anatomical locations spanning > 5% of the cluster, are shown.

Contrast	Peak MNI (*x*, *y*, *z*)	Automated Anatomical Labeling (AAL) Locations	Size (voxels)	Peak *T*-value
Declarative > Procedural	26, -84, 12	Right Calcarine fissure and surrounding cortex (23.50%); Right Lingual gyrus (11.80%); Right Superior occipital gyrus (7.16%); Right Middle occipital gyrus (6.66%); Right Cuneus (5.24%)	1,983	6.96
	22, -26, 46	Right Median cingulate and paracingulate gyri (14.16%); Right Supplementary motor area (10.19%); Right Paracentral lobule (7.21%); Right Precuneus (7.15%)	1,511	4.20
	-24, 42, -16	Left Superior frontal gyrus, medial orbital (29.44%); Left Inferior frontal gyrus, orbital part (26.60%); Left Superior frontal gyrus, orbital part (20.08%); Left Orbitofrontal (11.04%)	951	5.95
	-24, -58, -8	Left Lingual gyrus (34.98%); Left Fusiform gyrus (22.79%); Left Calcarine fissure and surrounding cortex (13.07%); Left Cuneus (12.19%); Left Inferior occipital gyrus (7.24%)	566	6.26
	-22, 22, 28	Left Middle frontal gyrus (50.09%); Left Superior frontal gyrus, dorsolateral (12.74%); Left Inferior frontal gyrus (5.66%)	565	3.73
	-34, -42, 34	Left Inferior parietal, but supramarginal and angular gyri (13.40%); Left Postcentral gyrus (7.92%); Left Precental gyrus (5.28%);	530	5.54
	18, -26, -22	Right Parahippocampal gyrus (24.35%); Right Cerebelum 3 (10.73%); Right Cerebelum 4 (5.76%); Right Hippocampus (5.24%)	382	5.42
	8, -62, -44	Right Cerebelum_8_R (23.49%); Vermis_8 (13.17%); Right Cerebelum_9_R (12.46%); Right Cerebelum_Crus1_R (9.25%)	281	4.64
	-16, 32, 10	Left Superior frontal gyrus, dorsolateral (17.24%); Left Middle frontal gyrus (11.49%)	261	3.91
	-46, -10, -36	Left Inferior temporal gyrus (72.92%); Left Middle temporal gyrus (21.67%);	240	5.07
	-14, -42, 16	Left Posterior cingulate gyrus (20.38%); Right Posterior cingulate gyrus (7.64%); Left Precuneus (5.10%);	157	5.85
	14, 36, 42	Right Superior frontal gyrus, dorsolateral (60.96%); Right Superior frontal gyrus, medial (24.66%); Right Median cingulate and paracingulate gyri (12.33%)	146	3.47
	-20, 4, 4	Left putamen (68.64%); Left pallidum (16.10%)	118	4.09
	2, 32, -12	Right Frontal_Med_Orb_R (33.91%); Left Olfactory cortex (24.35%); Right Anterior cingulate and paracingulate gyri (13.04%); Left Anterior cingulate and paracingulate gyri (13.04%); Left Gyrus rectus (5.22%)	115	3.99
	-66, -38, 8	Left Middle temporal gyrus (55.96%); Left Superior temporal gyrus (40.37%)	109	6.71
	-36, -82, 26	Left Middle occipital gyrus (97.96%)	98	5.77
	30, -64, 60	Right Superior parietal gyrus (81.63%); Right Angular gyrus (11.22%); Right Inferior parietal and angular gyrus (7.14%);	98	6.47
	-34, -70, 54	Left Inferior parietal, but supramarginal and angular gyri (43.96%); Left Superior parietal gyrus (34.07%); Left Angular gyrus (13.19%)	91	5.43

	20, -20, -4	Right Thalamus (13.64%)	88	4.51
	16, -14, 22	Right Caudate nucleus (40.28%); Right Thalamus (18.06%);	72	4.31
	-16, 0, 66	Left Supplementary motor area (72.46%); Left Superior frontal gyrus, dorsolateral (24.64%)	69	3.68
	-2, 50, 10	Left Anterior cingulate and paracingulate gyri (58.93%); Left Superior frontal gyrus, medial (35.71%); Right Anterior cingulate and paracingulate gyri (5.36%);	56	3.31
	-4, -14, -28	Left Parahippocampal gyrus (19.23%)	52	4.13
Procedural > Declarative	36, -56, -54	Right Cerebelum 8 (11.43%); Right Cerebelum 9 (10.06%); Right Cerebelum 10 (5.86%)	1,382	-6.57
	-10, -8, 20	Left Thalamus (25.14%); Left Inferior temporal gyrus (7.14%); Left Hippocampus (6.03%)	1,261	-5.51
	22, -36, 76	Right Middle frontal gyrus (25.70%); Right Superior frontal gyrus, dorsolateral (24.53%); Right Precental gyrus (22.43%); Right Postcentral gyrus (15.19%)	428	-5.22
	58, 24, 24	Right Inferior frontal gyrus, opercular part (47.40%); Right Inferior frontal gyrus, triangular part (40.62%); Right Rolandic operculum (8.33%);	192	-5.85
	-42, -48, -40	Left Crus 1 (47.37%); Left Cerebelum 7b (16.32%); Left Cerebelum 8 (10.53%); Left Cerebelum 6 (9.47%); Left Cerebelum 4 (7.89%); Left Crus 2 (7.37%);	190	-5.55
	52, 28, -4	Right Inferior frontal gyrus, triangular part (54.50%); Right Inferior frontal gyrus, orbital part (42.86%)	189	-4.89
	66, -18, 20	Right Rolandic operculum (45.21%); Right Postcentral gyrus (28.08%); Right Supramarginal gyrus (13.01%); Right Superior temporal gyrus (11.64%);	146	-4.03
	-6, -70, -8	Left Cerebelum 6 (33.04%); Left Lingual gyrus (30.36%); Vermis 6 (16.07%); Left Cerebelum 4 (8.04%)	112	-5.06
	-14, -80, -34	Left Crus 2 (94.55%); Left Crus 1 (5.45%);	110	-5.52
	18, 12, -4	Right Lenticular nucleus, putamen (52.27%); Right Lenticular nucleus, pallidum (10.23%)	88	-4.45
	-38, 20, 6	Left Inferior frontal gyrus, triangular part (59.09%); Left Insula (40.91%);	88	-5.07
	64, -48, 16	Right Middle temporal gyrus (59.30%); Right Superior temporal gyrus (40.70%);	86	-4.92
	4, -54, 68	Right Precuneus (80.25%); Left Precuneus (16.05%);	81	-5.18
	4, -32, -4	Right Lingual gyrus (13.92%)	79	-4.38
	-52, 18, -6	Left Temporal pole: superior temporal gyrus (38.36%); Left Inferior frontal gyrus, orbital part (19.18%); Left Inferior frontal gyrus, triangular part (16.44%)	73	-5.15
	-16, -10, 56	Left Superior frontal gyrus, dorsolateral (45.21%); Left Supplementary motor area (8.22%)	73	-3.26
	-30, -54, -18	Left Cerebelum 6 (50.00%); Left Fusiform gyrus (41.67%); Left Cerebelum 4 (8.33%);	72	-4.70

	44, 32, 42	Right Middle frontal gyrus (81.94%); Right Superior frontal gyrus, dorsolateral (8.33%);	72	-4.64
	0, -26, 24	Left Posterior cingulate gyrus (1.47%);	68	-3.97
	42, -12, -34	Right Inferior temporal gyrus (62.50%); Right Fusiform gyrus (37.50%);	56	-4.24
	-2, 10, -6	Left Caudate nucleus (31.25%); Left Olfactory cortex (12.50%); Left Lenticular nucleus, pallidum (10.42%); Left Lenticular nucleus, putamen (6.25%);	48	-4.18

Table S3. Model parameter space in the simulations.

Models	Parameter	Value	Meaning
Declarative	ε	0 - 0.5	activation noise
	d	0.2 - 0.85	memory decay
Procedural	s	0 - 0.5	utility noise
	α	0.05 - 0.5	learning rate

Table S4. Results of logistic regression analysis of the effects of noise parameters on group assignment.

Factor	**Estimate**	**Standard Error**	**T-value**	**p value**
Intercept	3.34	1.21	2.75	0.006
Declarative noise	12.86	2.76	4.66	**<0.001**
Procedural noise	-10.56	1.78	-5.94	**<0.001**