Reliance on Episodic vs. Procedural Systems in Decision-Making Depends on Individual Differences in Their Relative Neural Efficiency

Yuxue Yang[1], Catherine L. Sibert[2], Andrea Stocco[1*]

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

*Corresponding author: A. Stocco.
Email: stocco@uw.edu

Author Summary

Our study investigated two possible mechanisms for learning from experience, either by relying on memorizing details of events or by learning from rewards and punishments. Do individuals show preference for using one mechanism over another? And, if so, can we predict this preference from their brain patterns? Using an experimental task that can be carried out with either mechanism and computational models to analyze behavior and neural data, we found that different individuals do show a preference for either memory-based or reward-based learning strategies. Most importantly, a machine learning model can predict an individual's preference from one's resting-state brain connectivity with high accuracy, and the predictive brain connections are in agreement with our understanding of memory-based and reward-based learning strategies. This is the first time we've seen such a clear link between brain connectivity and decision-making strategy preference. Our research suggests that human brains are rational and adaptive, and carry out tasks using the most efficient neural circuitry.

Abstract

Experiential decision-making can 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.

Introduction

Two distinct frameworks have been proposed to explain how experience influences human behavior. According to one class of theories, previous decisions and their outcomes are stored as traces in episodic long-term memory, and decision-making is controlled by which memory is easier to retrieve [1,2]: we will refer to this as the *declarative* framework. The other class of theories suggests that decision-making based on experience is governed by the basic principles of reinforcement learning, in which 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] the lateral prefrontal cortex [7], and the medial frontal and parietal regions associated with the default mode network [8,9]. In the *procedural* framework, decisions from experience rely on brain circuits for implicit reward learning and habit formation, such as the basal ganglia [10], the supplementary motor area [11], the cerebellum [12], and the specific sensorimotor regions involved in the task itself [13].

Both systems are concurrently active at any given time [10,14], so, why should individuals choose to rely on one system over the other? One likely explanation is that optimal behavior in humans is inherently limited by processing capacity [15], therefore individuals must adopt different procedures to maximize the outcome of a decision while consuming the minimum amount of cognitive resources [16–18]. Because the cost of a decision also depends on the neural resources it consumes, individuals with different

neural characteristics would make decisions in different ways. This paper puts forward the hypothesis that individuals rely on *declarative* or *procedural* processes based on the relative neural efficiency of their corresponding brain circuits.

A common indicator of neural efficiency within a circuit is their functional resting-state connectivity, that is, the degree of correlation between the spontaneous neural activities of different regions in that circuit [19]. Higher correlations at rest reflect tighter coupling of neural dynamics and greater exchange and integration of communication between regions [20]. Thus, relatively higher functional connectivity within the *declarative* or *procedural* circuit should predict greater reliance on the corresponding system when making decisions.

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 [6,14,21]. 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. Investigating the relationship between brain connectivity and preferred decision-making processes requires a task where the *procedural* system is equally as effective as the *declarative* system strategy. One such task is the Incentive Processing Task (IPT) [22], 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 decision is made 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 were 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 were 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 [23] 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 *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.

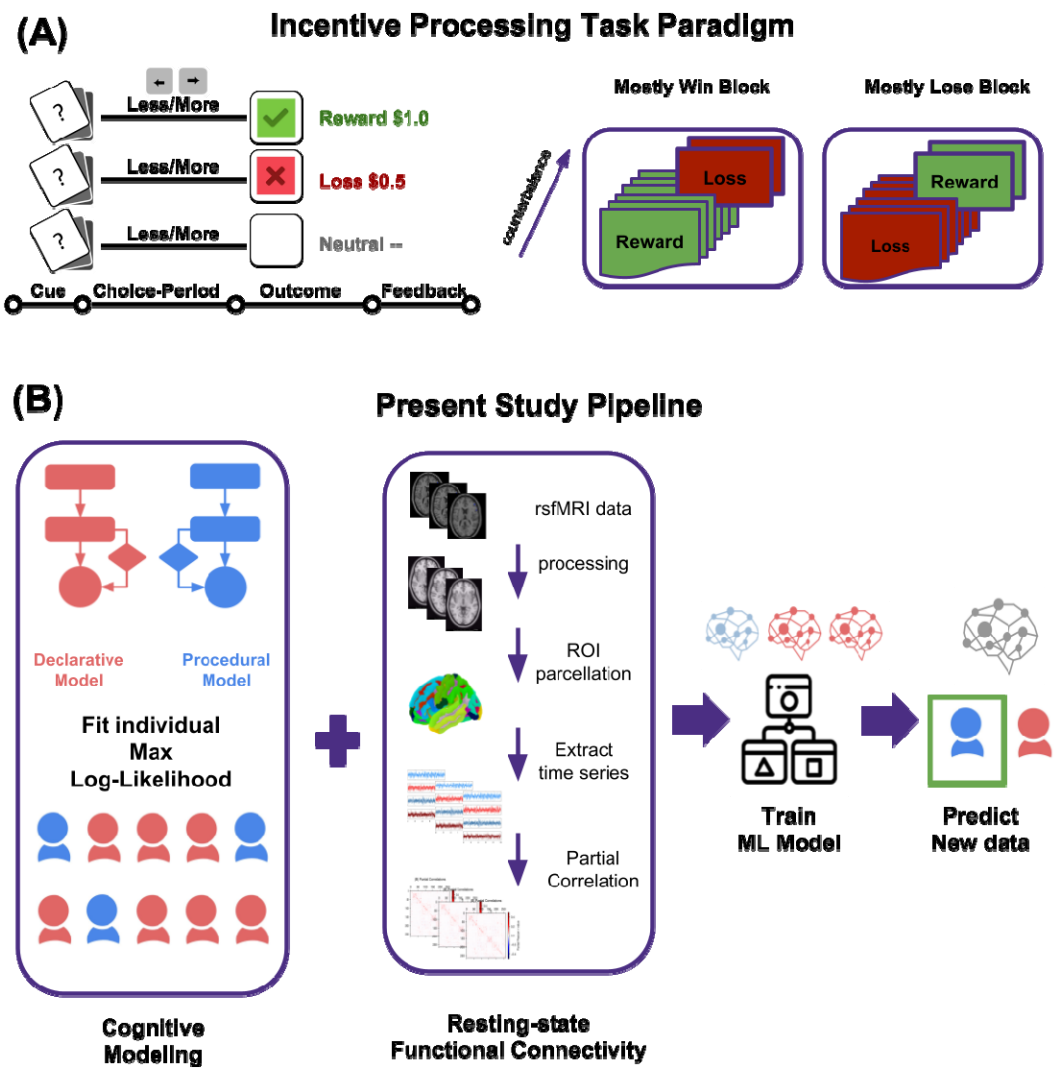


Fig 1. (A) Incentive processing task paradigm. (B) The pipeline of the present study. First, each subject's behavioral data are fit by two distinct cognitive models, *declarative* vs. *procedural* model through maximum log-likelihood. Second, each subject's neuroimaging data are processed to construct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

The episodic control model is based on the memory model of Anderson and Schooler [24] and follows the multiple trace theory [25]. In this model, the outcome of every decision is stored as an episodic memory. After encoding a memory, without any intervention, that memory starts to be forgotten, and the availability of each memory trace decays according to a power law [26]. Multiple experiences of the same decision-outcome pair, however, result in the accumulation of traces of the same memory, increasing the *activation* of that memory. Formally, the activation of a memory *m* at time *t* reflects the balance between these two effects, and is given by below equation:

$$A_t(m) = \sum_n \log (t - t_n)^{-d} \qquad (1)$$

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 [27,28]. A decision is guided by the most active memory of past interactions. If the retrieved memory is of a "win" outcome, the model repeats its associated decision, (e.g., choosing "more" if the "win" was associated with a "more" choice). If the memory is of an associated loss, the model executes the *opposite* action (e.g., choosing "less" if the loss was associated with a "more" choice). The probability that a given memory *m* will be retrieved is controlled by a Boltzmann function with noise temperature *v*:

$$P(m) = e^{A(m)/v} / \sum_i e^{A(i)/v} \qquad (2)$$

The *procedural* model is based on the reinforcement learning framework [3]. In this model, the two possible decisions ("more" or "less") are represented as two actions

associated with the cue (in this paradigm, this is the presentation of the two cards). Each action a has an associated value $V(a)$ that is updated over time based on the difference between the expected and the actual outcome [29] For example, receiving a "win" feedback after making a "more" choice will increase the value of the "more" action, $V(more)$, making it more likely to be selected in future choices, while receiving "loss" feedback will decrease the value of the "more" action, making it less likely to be selected in future choices. Formally, after the n-th decision, the value $V_n(a)$ of an action is updated using the temporal difference method [30]:

$$V_n(a) = V_{n-1}(a) + \alpha[R - V_{n-1}(a)] \quad (3)$$

In Equation 3, R is the reward associated with an outcome, (1 for "win", -0.5 for "lose", corresponding to the dollar amount won by participants), and α is the learning rate. The probability that an action a will be selected is controlled by a Boltzmann function with noise temperature τ:

$$P(a) = e^{V(a)/\tau} / \sum_i e^{V(i)/\tau} \quad (4)$$

Thus, each model depends on one learning parameter (d or α) and one noise parameter (ν or τ).

Both models were fit to each participant using maximum likelihood estimation [31], i.e., by fitting their parameters to maximize the probability of generating the specific series of decisions made by that participant (see Methods). Each participant was then

assigned to either the *declarative* or *procedural* group based on which version of the model had the greatest likelihood of producing their observed behavioral 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, such as the hippocampus, and retrieval, such as the ventral prefrontal cortex and default mode network. Correspondingly, we expect individuals best fit by the *procedural* model to display greater activity in the circuits associated with habit formation, such as the basal ganglia, the cerebellum, and the visuo-motor regions involved in interacting with a computer task. 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 predictict 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

Reliability of Model Identification

Because our results depend on correctly identifying the model that best represents each participant, a model recovery procedure [32] was run to ensure that our identification process was reliable. In this procedure, simulated participant data was generated by instantiating a model with random parameter values (uniformly drawn within their plausible boundaries; see Methods); the model-fitting procedure was then applied, treating the simulated data as participant data, and the accuracy to which the original model could be correctly identified was recorded (see Methods). Figure 2 provides the confusion matrix for the results of the procedure.

For both models, the recovery rate is approximately 85%. To account for errors in model assignment, in all of the following analyses, observations from a given participant are weighted by the difference in log-likelihood between the fit of the two models – that is, the stronger the evidence is in favor of one model, the greater the weight associated with the corresponding observation.

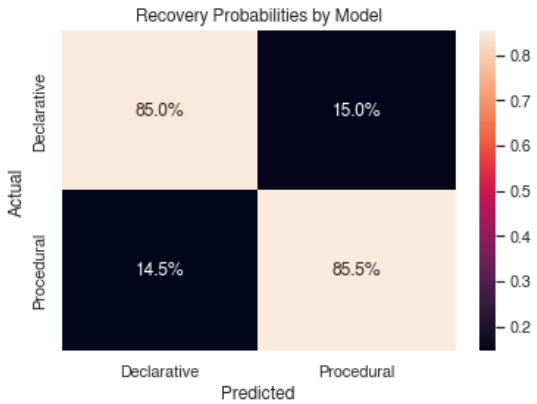


Fig 2. Confusion matrix showing the probability of correct model identification from 40,000 simulated runs. See Methods for details.

A second analysis was carried out to test whether each model's *parameters* (i.e., d and nu for the declarative model and alpha and tau for the procedural one) could also be correctly identified. This was done to make sure that any inferences about individual differences in the relative efficiency of the two systems could be grounded on reliable differences in parameters. Figure 3 illustrates the results of these simulations. All four parameters could be correctly identified from the simulated data, with correlations ranging from $r = 0.68$ (procedural temperature T, $p < 0.001$) to $r = 0.88$ (procedural learning rate alpha, $p < 0.001$).

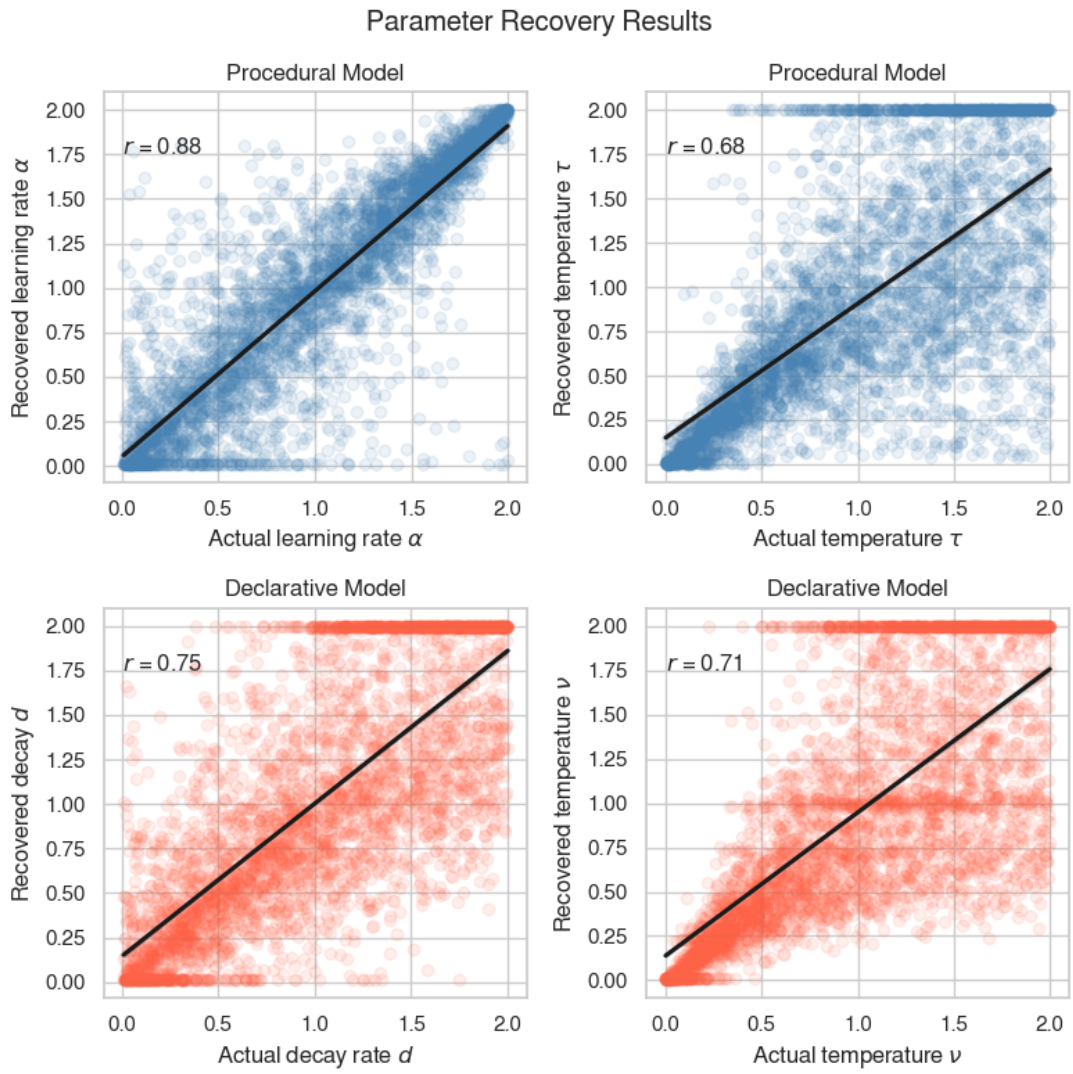


Fig 3. Results of the parameter recovery analysis for the individual parameters of the procedural (blue, top row) and the declarative (red, bottom row) models from 40,000 simulated runs. See Methods for details.

Decision-Making Process Identification

Having examined the reliability of the models, we then proceeded with categorizing participants based on which model fit them best. A total of 199 participants

were fit by the two models. Of these, 35 (17%) were best fit by the declarative model and thus were included in the declarative group.

The remaining 164 individuals were best fit by the procedural model and included in the procedural group. Figure 4 illustrates the distributions of the log-likelihood differences (Procedural - Declarative) across participants. Individuals for which the difference was < 0 were categorized as procedural, and those whose difference was > 0 were categorized as declarative.

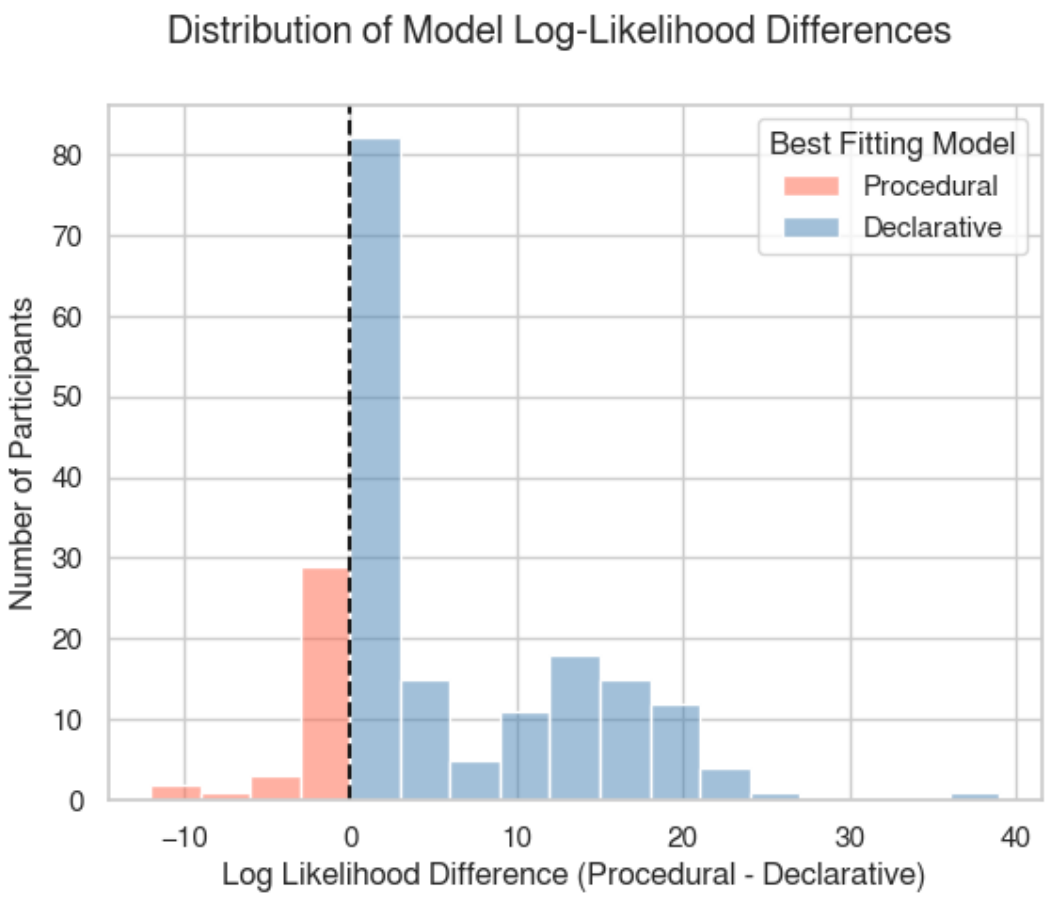


Fig 4. The distributions of the log-likelihood differences (Procedural - Declarative) across participants. Red bars represent the count of participants who are best fit by the procedural model, whereas blue bars represent the count of participants who are best fit by the declarative model.

Group Differences In Behavioral Responses During The Task

Our prediction is that, because of the different dynamics in the two models, individuals best fitted by the procedural model would exhibit greater sensitivity to trial feedback and a greater probability of switching from their previous decision than individuals best fitted by the procedural model. To test this prediction, a dummy variable was created to code each trial after the very first as either a response switch (1) or not (0). The effect of the previous trial feedback on the probability of a response switch was then analyzed using a logistic mixed-effects model implemented in the lme4 package in R [33]. In the model, Group (declarative vs. procedural) and Feedback (Win vs. Loss vs. Neutral) were treated as fixed effects, and individual subjects were treated as random effects, as shown in Figure 5.

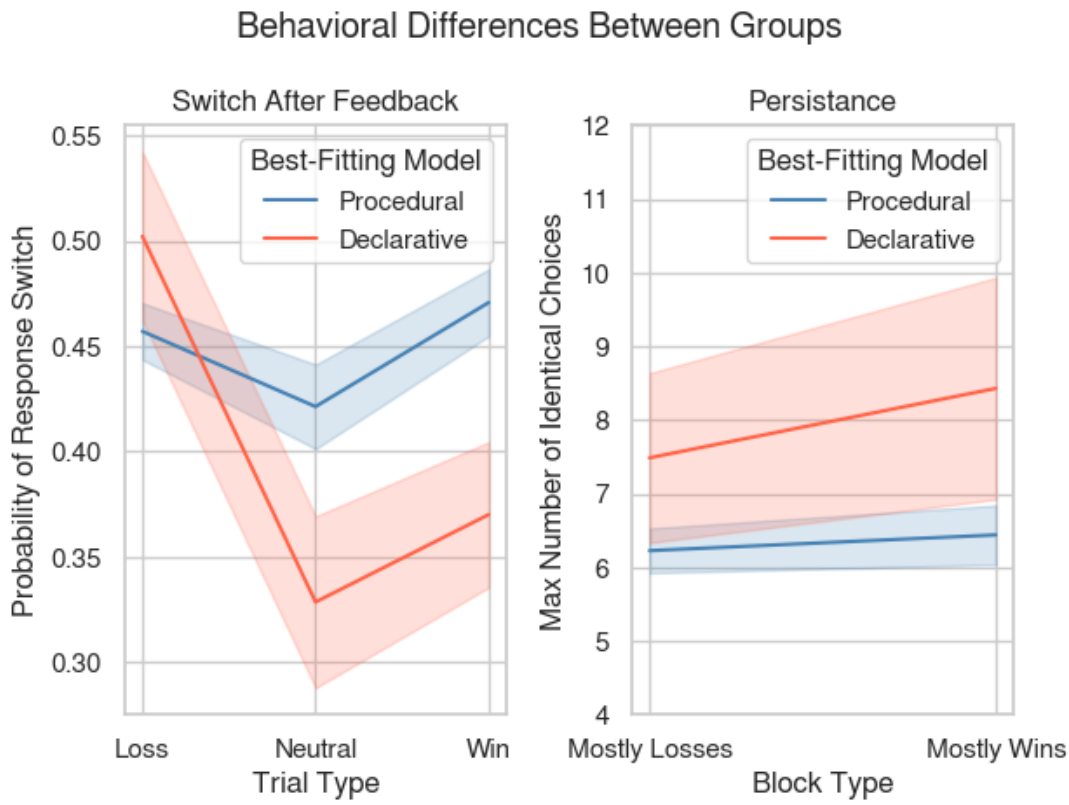


Fig 5. Main behavioral differences between participants best fit by procedural (blue) or declarative models (red). Left: the procedural group shows a greater tendency to change responses after and neural and a "win" feedback. Right: the declarative group shows a greater tendency to repeat the same choice within a block.

As expected, we found a significant main effect of model (β = 0.87, t = 2.21, p < 0.02), a significant main effect of trial when the feedback was a loss (β = 1.13, t = 2.92, p = 0.003), and a significant interaction between the two factors (β = -0.95, t = -2.44, p = 0.01). In summary, participants in the procedural group showed a significantly greater probability of shifting their behavior than participants in the declarative group, and, although both groups were more likely to change their responses after a "loss" than a

"win", this difference was magnified in the declarative group (Figure 5, left). 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.

We also reasoned that the declarative group would exhibit longer streaks of identical choices than the procedural group. Furthermore, since both models are more likely to change their responses after a "loss" than a "win", we expected that for both models the persistence would be larger in the "Mostly Win" blocks. To test these predictions, the maximum number of consecutive identical choices was computed for each block and individual, and the data were analyzed with a linear mixed model implemented in the *glmer* package in R. In the model, the persistence variable was modeled with a Poisson distribution, and both Block ("Mostly Lose" vs. "Mostly Win") and Model (declarative vs. procedural) were included as fixed-effects. Because of the small number of observations per participant, no random effect was included. The results showed a significant main effect of both block (blocks: $\beta = 0.25$, $t = 5.56$, $p < 0.0001$) and group (β -0.70, $t = -19.64$, $p < 0.0001$), as well as their interaction ($\beta = -0.13$, $t = -2.78$, $p = 0.005$). As expected, the declarative group had higher persistence than the procedural group and this effect was magnified in the "Mostly Win" block (Figure 5, right).

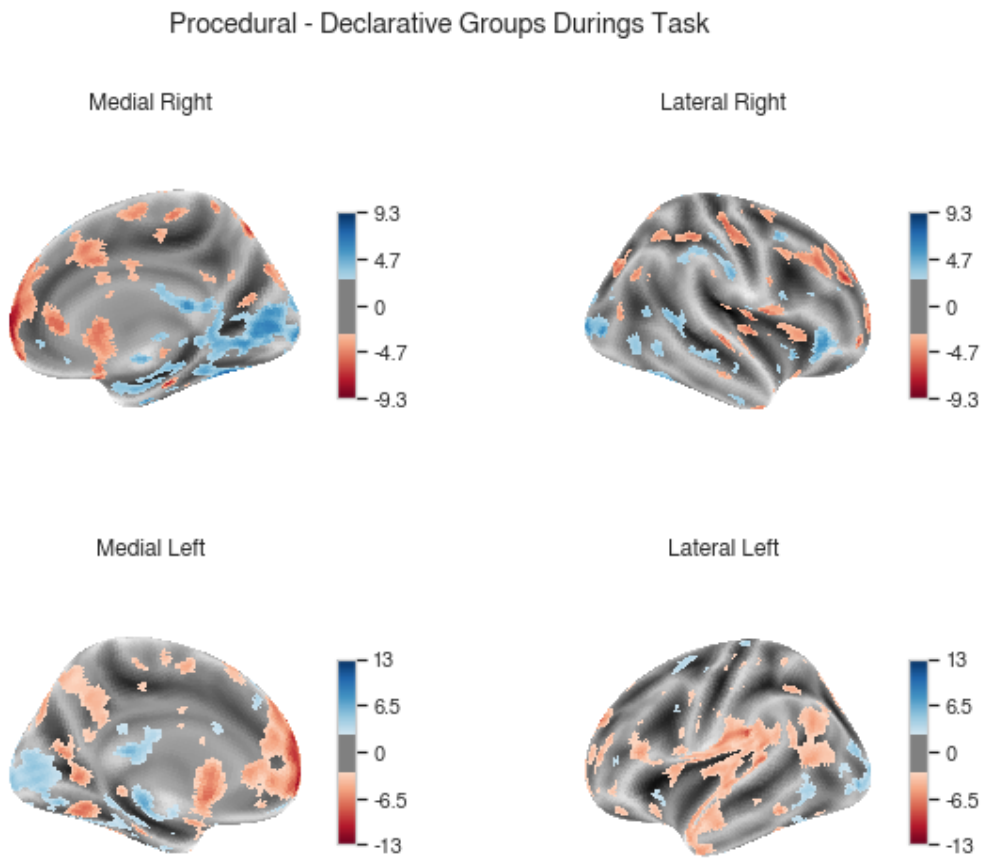
Group Differences in Brain Activity During the Task

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. This contrast is particularly important in this analysis, since participants are more likely to

keep engaging their preferred decision-making system during “Mostly Win” blocks (as implied by Figure 5, right).

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 heavily (See Materials and Methods). The False Discovery Rate [34] procedure was used to correct for multiple comparisons at a level of $q < 0.01$. The analysis identified several brain regions that show significant BOLD signal differences between the procedural and declarative groups (Figure 6; Table 1). Specifically, participants in the procedural group (Figure 6, blue) showed greater task activation in the primary (inferior and superior occipital cortex and lingual gyrus) and secondary (left and right inferior temporal cortex) visual regions, and in the primary (M1) and secondary (supplementary motor area) motor regions, together with greater task activation in the bilateral insula. This pattern of activation is consistent with the acquisition of habitual stimulus-response associations.

On the other hand, participants in the declarative group (Figure 6, red) showed greater brain activity in the canonical nodes of the default mode network (the bilateral medial frontal cortex, the bilateral precuneus, and the bilateral inferior parietal lobules) as well as regions involved in the encoding (right hippocampus, left anterior temporal lobe), and the retrieval (left and right lateral ventral prefrontal cortex) of declarative memories.



Fig

Fig. 6. Differences between procedural and declarative participants during the incentive processing task. In the overlays, the color indicates the direction of the difference (blue: Procedural > Declarative; Red: Declarative > Procedural) and color intensity indicates the magnitude of the associated *T* statistics. Regions are thresholded at a significance level of at $q < 0.01$, corresponding to a voxel-wise threshold of $T >= 2.70$.

Table 1: Results from the task fMRI comparison. For clarity, only clusters spanning > 50 voxels and locations spanning > 1% of the cluster and > 10 voxels within the cluster are shown. L = Left, R = Right; the word "gyrus" is omitted from the location names.

Contrast	MNI (Peak)	Automated Anatomical Labeling (AAL) Locations	Size (voxels)	Peak *T* value
Procedural > Declarative	22, -82, -48	R Cerebellum (4.44%); L Cerebellum (5.22%); R Fusiform (3.15%); L Calcarine (2.99%); R Calcarine (2.94%); R Lingual (2.76%); L Middle Occipital (2.64%); R Middle Temporal (2.14%); Left Fusiform (2.12%); L Lingual (1.62%); R Inferior Frontal (1.35%); R Inferior Temporal (1.27%); R Middle Occipital (1.21%); R Hippocampus (1.13%); L Inferior Occipital (1.12%); L Inferior Temporal (1.01%)	30,662	14.09
	62, -26, 46	R Postcentral (34.73%); R Supramarginal (34.27%); R Inferior Parietal(17.60%); R Rolandic (11.89%)	858	7.48
	-14, -34, 58	L Paracentral (43.00%); L Postcentral (6.11%); L Precuneus (4.07%); L Precental (1.53%)	393	5.55
	-22, 4, 48	L Superior Frontal (36.72%); L Middle Frontal (35.94%); L Supplementary Motor Area (3.52%)	256	4.41
	-10, 36, 42	L Superior Frontal (90.04%); ; L Middle Frontal (2.90%)	241	5.23
	36, 14, 60	R Middle Frontal (96.94%);	196	6.75
	12, 24, 20	R Anterior Cingulate (57.05%);	149	5.49
	-6, 18, 26	L Anterior Cingulate (40.74%); L Median Cingulate (10.37%	135	4.9
	-40, 16, -4	L Insula (92.06%)	126	5.1
	10, -22, 66	R Paracentral (43.93%); R Supplementary Motor Area (29.91%); R Precental (26.17%)	107	4.73
	-24, 68, 16	L Middle Frontal (24.14%); L Superior Frontal (18.39%)	87	4.61
	-40, -8, 52	L Precental (91.89%)	74	4.15

	28, 62, 24	R Middle Frontal (56.16%); R Superior Frontal (39.73%)	73	5.47
	-32, 0, 2	L Putamen (56.72%)	67	4.33
Declarative > Procedural	-4, 66, -2	L Superior Frontal (7.26%); L Superior Temporal (3.29%); L Middle Temporal (3.25%); R Superior Frontal (5.56%); L Precuneus (2.12%); L Middle Frontal (2.12%); L Postcentral (2.00%); L Inferior Temporal gyrus (2.74%); L Inferior Frontal part (1.58%); L Angular (1.54%); L Rolandic (1.47%); L Anterior Cingulate (1.43%); L Supramarginal (1.35%); L Caudate (1.35%); R Median Cingulate (1.27%); R Superior Temporal (1.24%); L Precental (1.17%); L Medial Frontal Cortex (1.13%); R Insula (1.10%); L Insula (1.06%); R Anterior Cingulate (1.05%); R Supplementary Motor Area (1.04%); L Temporal Pole (1.03%); L Supplementary Motor Area (1.01%); R Precuneus (0.99%)	38,651	-15.92
	22, -44, -54	R Cerebellum (86.87%)	545	-15.02
	22, -24, -26	R Parahippocampal (26.42%); R Cerebellum (33.97%); R Fusiform (18.24%)	159	-12.17
	64, -4, 32	R Inferior Frontal (38.31%); R Postcentral (31.17%); R Precental (30.52%);	154	-9.32
	40, 0, -44	R Inferior Temporal (89.47%); R Temporal Pole (6.58%)	152	-7.13
	2, -88, -16	L Calcarine (25.49%); L Lingual (17.65%); L Crus (7.84%);	102	-8.99
	6, -20, 76	R Paracentral Lobule (59.60%); R Supplementary Motor Area (12.12%);	99	-7.52
	52, -38, 56	R Postcentral (89.53%)	86	-8.25
	0, -4, 32	R Median Cingulate (52.56%); L Median Cingulate (39.74%)	78	-4.21
	-42, 20, 50	L Middle Frontal (98.68%)	76	-6
	14, -26, -8	R Lingual (15.71%); R Hippocampus (8.57%);	70	-4.72
	40, 24, -26	R Temporal Pole (98.33%);	60	-5.2

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 [35]. 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: [36]) regression model, which reduces a large number of potential predictors while retaining the most predictive features. The LASSO model was implemented using the *glmnet* package in R [37]. To ensure generalization, we chose the optimal LASSO hyperparameter λ_{nested} (0.0059) using a nested cross-validation method (See Supplementary Materials). Having the optimal λ_{nested}, we refit the LASSO model using the Leave-One-Out (LOO) cross-validation approach and obtained the correlation coefficient betas (□) for subsequent brain connectivity analysis.

The prediction performance was evaluated by regular LOO cross-validation and nested cross-validation approaches. The training and testing accuracy with regular cross-validation and nested cross-validation approaches are shown in Table 2. 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.

Table 2. Summary of the LASSO Model Performance

	Best λ	Testing AUC	Number of Non-Zero β
Regular CV	0.0027	1.000	92
Nested CV	0.0059	0.911	91

Finally, we examined which functional connectivity features were predictive of individual preferences in decision making. After LASSO regularization, 92 functional connections (approximately 0.11% of the total) had non-zero β parameters, suggesting a very sparse neurofunctional connectivity. Fig 7 shows the non-zero β weights of connection between networks and within networks as extracted from the fit LASSO model. 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 examined the corrected weight matrix by multiplying the β matrix with the sign of the functional connectivity matrix, obtaining a corrected weight correlation matrix W. The weight 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. Fig 7A shows these different functional connections.

Connections predictive of declarative processes involve clearly different networks than connections predictive of procedural processes, as shown in Fig 7(B) and Fig 7(C). As expected, reliance on declarative processes in decision-making was predicted by greater connectivity in the networks of regions associated with task control (task control networks and subcortical networks) as well as episodic memory (memory retrieval networks). Reliance on procedural learning processes, however, was predicted by greater connectivity in sensorimotor and salience regions. As Figures 7(B) and 7(C) show, the predictive functional connectivity also largely overlaps with the task-fMRI results, indicating that greater functional connectivity between two regions at rest predicts a greater likelihood that these two regions would be used to support the decision-making processes during the task.

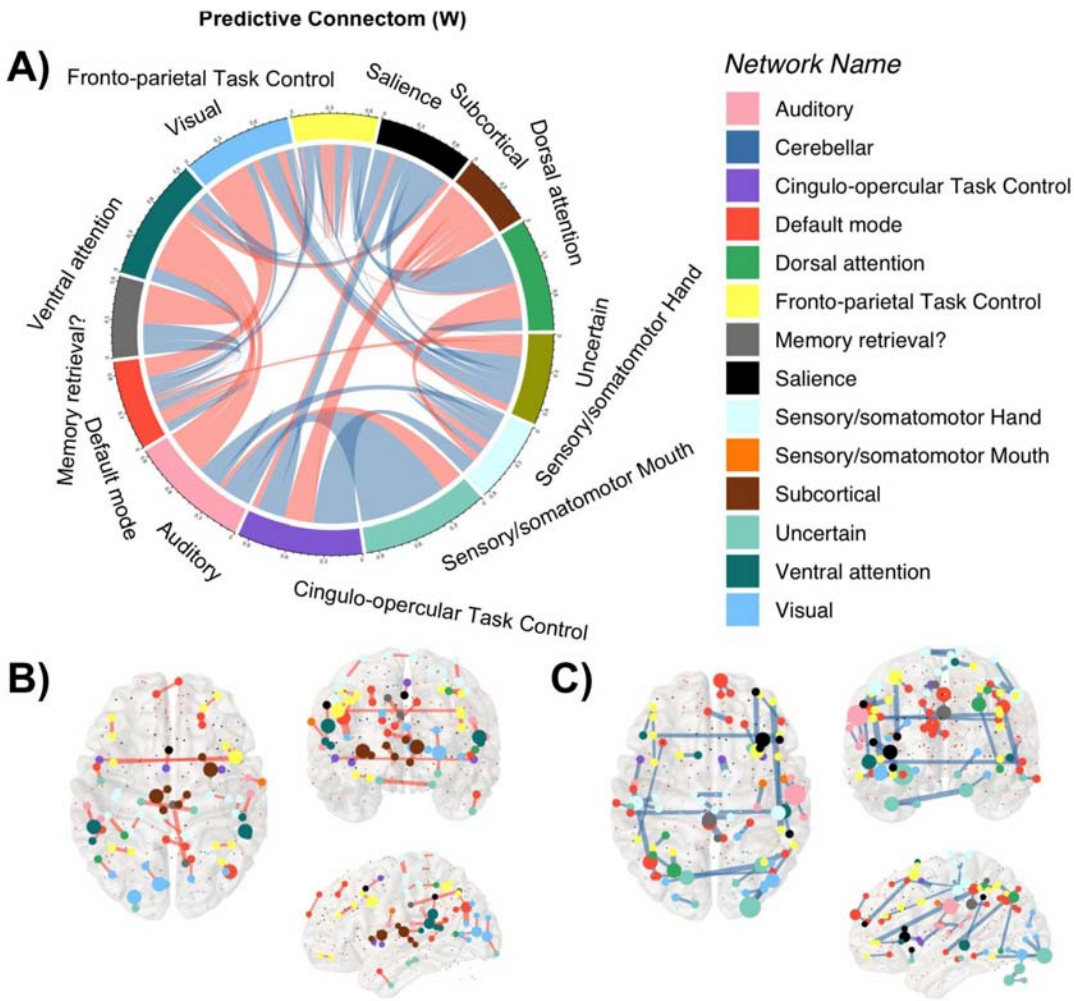


Fig 7. The distribution of predictive connections in declarative and procedural functional networks. (A) Predictive Connectom across the Power parcellation networks. Red connections denote positive weight values ($W > 0$), predicting the procedural Group, and Blue connections denote negative weights values ($W < 0$), predicting the procedural Group. (B) The anatomical location of the function connection is predictive of declarative group assignment; nodes represent regions, node colors indicate the network they belong to; node size indicates the degree of connectivity (the number of adjacent edges), , and line width indicates the strength of connectivity between two nodes. (C) Anatomical location of the functional connections predictive of procedural group assignment.

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, in fact, successfully be 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 (Welch Two Sample *T*-test: $t(50.04) = 3.37$, $p = 0.001$, see Figure 8).

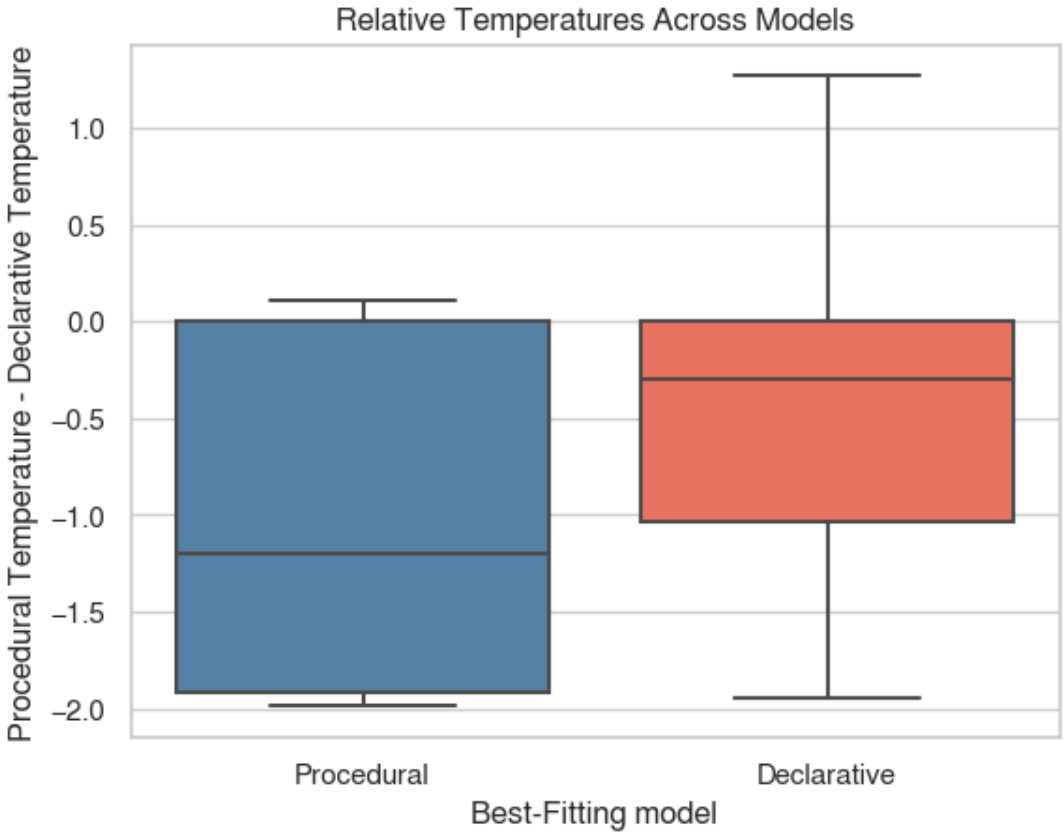


Fig 8. Estimated procedural and declarative temperature parameters between the two groups.

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 it is in principle conceivable to use 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.

Methods

Materials

This study analyzed both behavioral and neuroimaging data obtained from a subset of the Human Connectome Project (HCP) dataset [23]. 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 and fit by the two ACT-R models. By excluding 23 participants who missed resting-state fMRI scanning, a total of 176 participants were fit by predictive LASSO model. 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 [22]. 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 Lose" (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 Lose" 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 [38,39], 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 [35]. 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 [40] 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 matrix 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 [41].

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 [42]. The resulting contrast maps 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 $q < 0.01$ corrected for multiple comparisons using a False Discovery Rate (FDR [34]) procedure.

Computational Models

Episodic Memory Model

The episodic memory model relies on the declarative model of Anderson and Schooler [24], currently implemented as part of the influential ACT-R cognitive architecture [43]. When presented with a mystery card, the model retrieves a previous episodic memory in which the number was guessed successfully. After being presented with feedback, the model encodes a 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 represents the possible actions of the decision-making processes as competing stimulus-response 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 actions that implement the "More" and "Less" decisions. 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, -0.5 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.

Model Fitting

Model parameters were fit to maximize their likelihoods given the data. The likelihood of a model M with parameters θ given a vector of data $\mathbf{x}$, indicated as $L(M, \theta \mid \mathbf{x})$, is the probability of observing the data, given the model and its parameter, that is $L(M, \theta \mid \mathbf{x}) = P(\mathbf{x} \mid M, \theta)$. The data $\mathbf{x}$ consists of a vector of decisions $x_1, x_2, \ldots, x_N$:

$$L(M, \theta \mid \mathbf{x}) = P(x_1 \mid M, \theta) \cdot P(x_2 \mid M, \theta) \cdot \ldots \cdot P(x_N \mid M, \theta) = \prod_i P(x_i \mid M, \theta)$$

Since the product of conditional probabilities becomes vanishingly small, it is common to use *log*-likelihood:

$$\log L(M, \theta \mid \mathbf{x}) = \log \prod_i P(x_i \mid M, \theta) = \sum_i \log P(x_i \mid M, \theta)$$

Note that, because both models are *memory* models, the probability of decision x_i depends on the *history* of previous decisions $x_{i-1}, x_{i-2}, \ldots, x_1$ and their outcomes, i.e. $P(x_i \mid M, \theta) = P(x_i \mid M, \theta, x_{i-1}, x_{i-2}, \ldots, x_1)$. In both models, the history of previous decisions is implicitly represented in their respective memory representations, that is, the activations

of episodic memories and the values of actions. Thus, the probabilities of generating a response x_i can be derived by Equations 2 and 4, provided that the history of previous decisions and outcomes is recorded in the corresponding activations and values.

To do so, each model is run forward, first calculating the probability of producing the participant's response in the current trials; next updating the cumulative log-likelihood; and finally adjusting the activation of episodic traces and the value of the decision based on Equations 1 and 3 before moving to the next trial. For each participant, the model simulates the exact series of 64 trials as they experienced them, in the exact order and with the same inter-trial and inter-block intervals.

In order to evaluate the goodness-of-fit for individual fitting, we estimated maximum Log-Likelihood across the parameter space, i.e. we selected the value of parameters θ^* such that $\theta^* = \text{argmax}_\theta L(M, \theta \mid \mathbf{x})$. Because no closed-form solutions exist to derive the maximum likelihood of Equations 1 and 3, the best-fitting parameter values were identified using Powell's optimization algorithm [44], as implemented in Python's SciPy package [45]. This method was chosen because it does not require explicit derivatives and allows us to specify meaningful bounds for parameter values. All parameters were constrained within the bounds [0, 2]; they are all meaningless when negative and lead to instability and essentially random choices when > 2.

The starting parameter values for the algorithm were $\alpha = 0.1$, $\tau = 0.1$ for the procedural model, and $d = 0.5$, $\nu = 0.1$ for the declarative model. The starting values of α and d were chosen as the most common values found in their respective modeling frameworks; they are also the default values of the corresponding parameters in ACT-R.

Supervised Classification Model

To explore if individuals' behavioral differences between declarative and procedural strategies are indicated by an individual's underlying brain structure, we trained three of the most commonly used supervised classification models (Logistic regression model, Decision Tree model, and Random Forest model), using resting-state functional connectivity as the input variable, and predicted the probability of a participant being labeled as either preferring the declarative or procedural strategy. Considering the equally high accuracy (> 0.8) among these classification models (Random Forest Model (accuracy = 0.92) slightly outperformed the 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 (as in 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. We chose the logistic regression model because it is a 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 constrain a value between 0 and 1. Specifically, the output value x of a linear model is transformed into a corresponding probability $P(x)$ by the function:

$$P(x) = 1/(1 + e^{-x})$$

In order to handle an imbalanced dataset with unequal target labels, upsampling 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 either of the 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 $\times$ 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 a simple linear regression analysis. In LASSO, the best values associated with each regressor are chosen as to minimize a loss function that includes the ordinary squared errors as well as a penalty term for each regressor's β weight:

$$Loss = (y - \hat{y})^2 + \lambda \sum_i^n |\beta_i|$$

where the tuning parameter λ controls the degree of penalty: for greater values of λ, more coefficients are forced to become 0.

To account for the large disparity between the number of participants and the number of predictors, we performed a grid search cross-validation using the *glmnet* package in R [37] to determine the best value for the fit hyper-parameter $\lambda_{regular}$. To alleviate the potential problems of small sample size in neuroimaging studies, Vabalas and colleagues [46] used Nested CV approaches to produce 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 1:4). Instead of adopting a computationally expensive Leave-One-Out (LOO) method, 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. 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 (λ_{nested}). This approach allowed us to obtain a reliable and unbiased assessment of the LASSO model's performance while accounting for the variability in the data and the selection of hyperparameters.

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.

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: AS,YY

Formal analysis: AS,YY

Visualization: AS,YY

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 code are available in the OSF repository: https://osf.io/wf4my/

References

1. Gonzalez C, Dutt V. Instance-based learning: integrating sampling and repeated decisions from experience. Psychol Rev. 2011;118: 523–551.

2. Stewart N, Chater N, Brown GDA. Decision by sampling. Cogn Psychol. 2006;53: 1–26.

3. Niv Y. Reinforcement learning in the brain. J Math Psychol. 2009;53: 139–154.

4. Lee D, Seo H, Jung MW. Neural basis of reinforcement learning and decision making. Annu Rev Neurosci. 2012;35: 287–308.

5. Chelian SE, Paik J, Pirolli P, Lebiere C, Bhattacharyya R. Reinforcement learning and instance-based learning approaches to modeling human decision making in a prognostic foraging task. 2015 Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob). ieeexplore.ieee.org; 2015. pp. 116–122.

6. Knowlton BJ, Squire LR, Gluck MA. Probabilistic classification learning in amnesia. Learn Mem. 1994;1: 106–120.

7. Badre D, Wagner AD. Left ventrolateral prefrontal cortex and the cognitive control of memory. Neuropsychologia. 2007;45: 2883–2901.

8. Raichle ME, Snyder AZ. A default mode of brain function: a brief history of an evolving idea. Neuroimage. 2007;37: 1083–90; discussion 1097–9.

9. Sestieri C, Corbetta M, Romani GL, Shulman GL. Episodic memory retrieval, parietal cortex, and the default mode network: functional and topographic analyses. J Neurosci. 2011;31: 4407–4420.

10. Foerde K, Knowlton BJ, Poldrack RA. Modulation of competing memory systems by distraction. Proceedings of the National Academy of Sciences. 2006;103: 11778–11783.

11. Nieuwenhuis S, Holroyd CB, Mol N, Coles MGH. Reinforcement-related brain potentials from medial frontal cortex: origins and functional significance. Neurosci Biobehav Rev. 2004;28: 441–448.

12. Taylor JA, Ivry RB. Cerebellar and prefrontal cortex contributions to adaptation, strategies, and reinforcement learning. Prog Brain Res. 2014;210: 217–253.

13. Hill NM, Schneider W. Brain changes in the development of expertise: Neuroanatomical and neurophysiological evidence about skill-based adaptations. The Cambridge handbook of expertise and expert performance. 2006; 653–682.

14. Poldrack RA, Clark J, Paré-Blagoev EJ, Shohamy D, Creso Moyano J, Myers C, et al. Interactive memory systems in the human brain. Nature. 2001;414: 546–550.

15. Simon HA. A behavioral model of rational choice. Models of man, social and rational: Mathematical essays on rational human behavior in a social setting. 1957; 241–260.

16. Payne JW, Payne JW, Bettman JR, Johnson EJ. The Adaptive Decision Maker. Cambridge University Press; 1993.

17. Payne JW, Bettman JR, Johnson EJ. Adaptive strategy selection in decision making.

J Exp Psychol Learn Mem Cogn. 1988;14: 534–552.

18. Gigerenzer G. Why Heuristics Work. Perspect Psychol Sci. 2008;3: 20–29.

19. Shen X, Finn ES, Scheinost D, Rosenberg MD, Chun MM, Papademetris X, et al. Using connectome-based predictive modeling to predict individual behavior from brain connectivity. Nat Protoc. 2017;12: 506–518.

20. van den Heuvel MP, Hulshoff Pol HE. Exploring the brain network: a review on resting-state fMRI functional connectivity. Eur Neuropsychopharmacol. 2010;20: 519–534.

21. Frank MJ, Seeberger LC, O'reilly RC. By carrot or by stick: cognitive reinforcement learning in parkinsonism. Science. 2004;306: 1940–1943.

22. Delgado MR, Nystrom LE, Fissell C, Noll DC, Fiez JA. Tracking the hemodynamic responses to reward and punishment in the striatum. J Neurophysiol. 2000;84: 3072–3077.

23. Van Essen DC, Smith SM, Barch DM, Behrens TEJ, Yacoub E, Ugurbil K, et al. The WU-Minn Human Connectome Project: an overview. Neuroimage. 2013;80: 62–79.

24. Anderson JR, Schooler LJ. Reflections of the Environment in Memory. Psychol Sci. 1991;2: 396–408.

25. Nadel L, Samsonovich A, Ryan L, Moscovitch M. Multiple trace theory of human memory: computational, neuroimaging, and neuropsychological results. Hippocampus. 2000;10: 352–368.

26. Newell, Rosenbloom. Mechanisms of skill acquisition. Cognitive skills and their acquisition. 1981. Available:

https://books.google.com/books?hl=en&lr=&id=7oEtr1KvMbgC&oi=fnd&pg=PA1&dq=+practice++newell+rosenbloom+1981&ots=1Nb-lfK7Zn&sig=gQdE9Ifcd1i0JBe6E8MABdrZmTU

27. Zhou P, Sense F, van Rijn H, Stocco A. Reflections of idiographic long-term memory characteristics in resting-state neuroimaging data. Cognition. 2021;212: 104660.

28. Sense F, Behrens F, Meijer RR, van Rijn H. An individual's rate of forgetting is stable over time but differs across materials. Top Cogn Sci. 2016;8: 305–321.

29. Rescorla RA, Wagner AR, Others. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. Classical conditioning II: Current research and theory. 1972;2: 64–99.

30. Sutton RS. Learning to predict by the methods of temporal differences. Mach Learn. 1988;3: 9–44.

31. Myung IJ. Tutorial on maximum likelihood estimation. J Math Psychol. 2003;47: 90–100.

32. Wilson RC, Collins AG. Ten simple rules for the computational modeling of behavioral data. Elife. 2019;8. doi:10.7554/eLife.49547

33. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. J Stat Softw. 2015;67: 1–48.

34. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Series B Stat Methodol. 1995;57: 289–300.

35. Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA, et al. Functional network organization of the human brain. Neuron. 2011;72: 665–678.

36. Tibshirani R. Regression shrinkage and selection via the lasso. J R Stat Soc. 1996;58: 267–288.

37. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Softw. 2010;33: 1–22.

38. Cox RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res. 1996;29: 162–173.

39. Cox RW. AFNI: what a long strange trip it's been. Neuroimage. 2012;62: 743–747.

40. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2020. Available: https://www.R-project.org/

41. Silver NC, Dunlap WP. Averaging correlation coefficients: Should Fisher's z transformation be used? J Appl Psychol. 1987;72: 146–148.

42. Barch DM, Burgess GC, Harms MP, Petersen SE, Schlaggar BL, Corbetta M, et al. Function in the human connectome: task-fMRI and individual differences in behavior. Neuroimage. 2013;80: 169–189.

43. Anderson JR. How Can the Human Mind Occur in the Physical Universe? Oxford University Press; 2009.

44. Powell MJD. An efficient method for finding the minimum of a function of several variables without calculating derivatives. Comput J. 1964;7: 155–162.

45. Virtanen P, Gommers R, Oliphant TE, Haberland M. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nature. 2020. Available: https://www.nature.com/articles/s41592%E2%80%90019%E2%80%900686%E2%80%902

46. Vabalas A, Gowen E, Poliakoff E, Casson AJ. Machine learning algorithm validation with a limited sample size. PLoS One. 2019;14: e0224365.