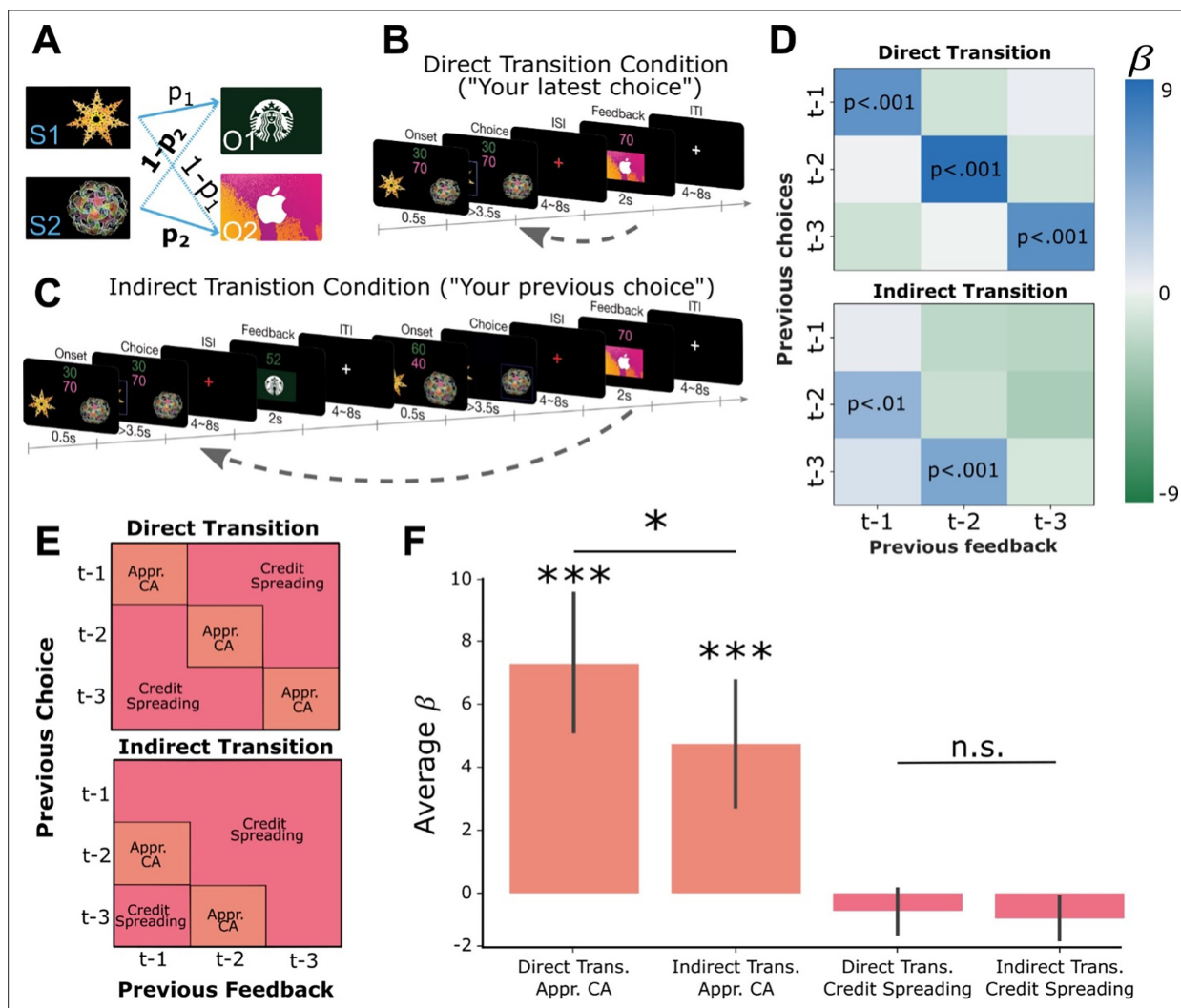


---

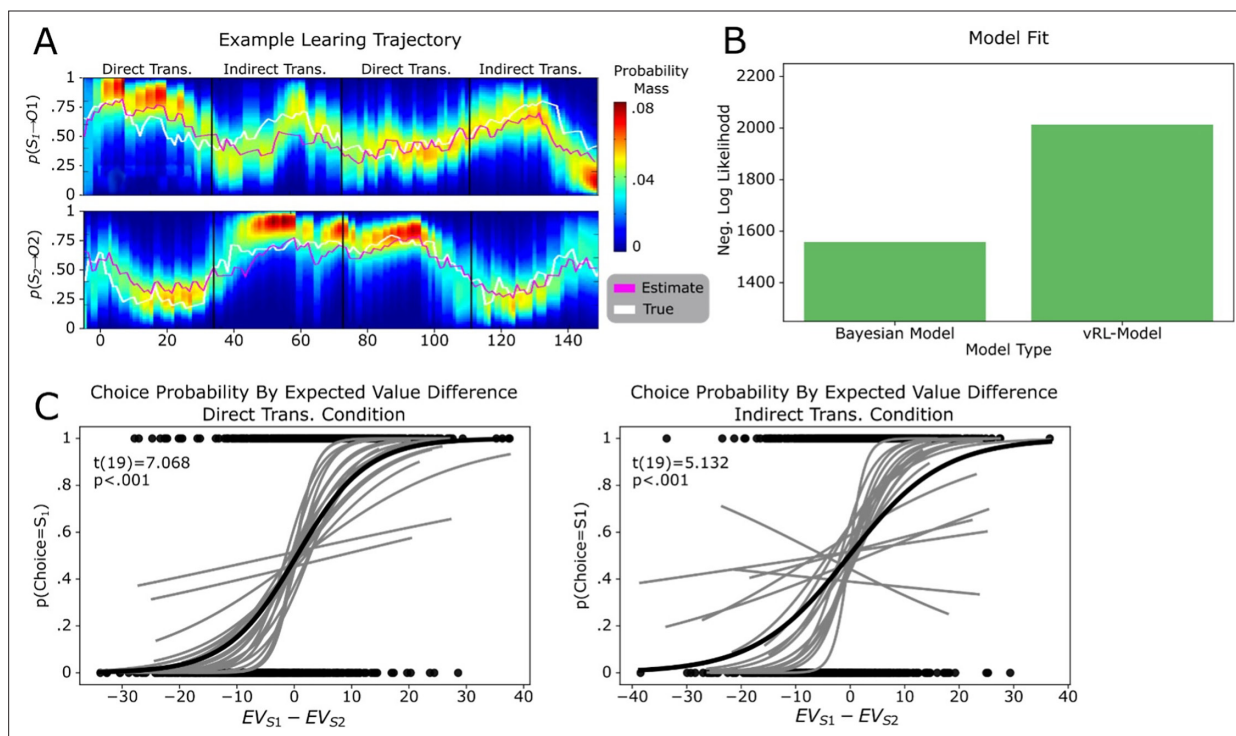
## Figures and figure supplements

Neural mechanisms of credit assignment for delayed outcomes during contingent learning

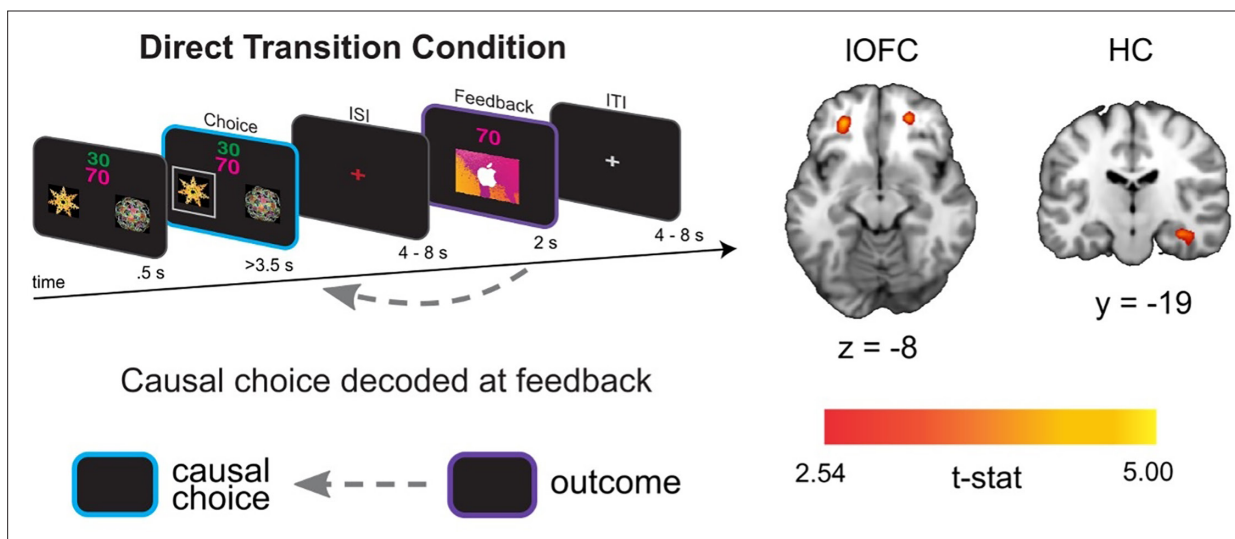
**Phillip P Witkowski and Lindsay JH Rondot *et al.***



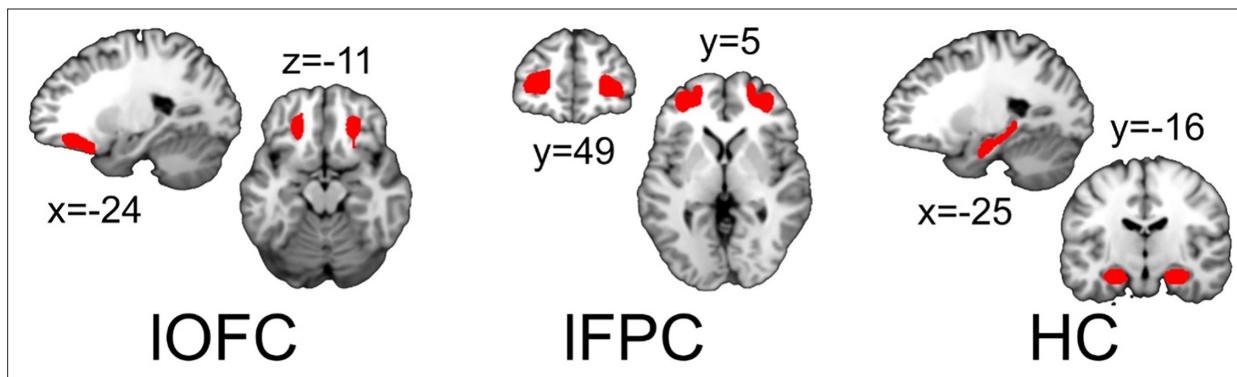
**Figure 1.** Learning Task Design and Behavioral Results. (A) Two abstract shapes were probabilistically related to each of two outcome identities by independent transition probabilities  $p_1$  and  $p_2$ . (B) Schematic of the direct transition condition. Participants chose one of the two shapes on each trial based on two pieces of information: their estimates of the probability that each would lead to either outcome identity (gift cards) and the randomly generated number of points they could potentially win if that outcome was obtained. The color of each number indicated the identity of the outcome on which that number of points could be won. In the example, green indicates the number of points for the Starbucks gift card, while pink indicates the number of points for iTunes. Next, participants observed the outcome of their choice (the gift card and amount) after a delay. (C) Schematic of the indirect transition condition. Same as (B) except that after participants made their choice they transitioned into another independent decision. After this second decision was made, participants observed the outcome of their first decision. (D) Results of logistic regression analysis predicting the current choice based on previously observed choice-outcome relationships. Each cell represents the combination of a previously observed choice with an observed outcome. The color of each cell shows the value of beta estimates for each combination of previous choice and observed outcome, averaged across participants. Positive values indicate that the choice-outcome pair predicted choosing the same shape again when that shape previously led to the currently desired outcome. (E) Theoretical decomposition of the matrix in (D) into groups of cells which reflect 'appropriate credit assignment' given the task structure (orange) and 'credit spreading' (pink). (F) Mean ( $\pm$  SEM) of beta coefficients for specific choice-outcome combinations averaged across the groupings of cells shown in E for each condition. See **Figure 1—figure supplement 1** for model outputs and Bayesian model fitting.



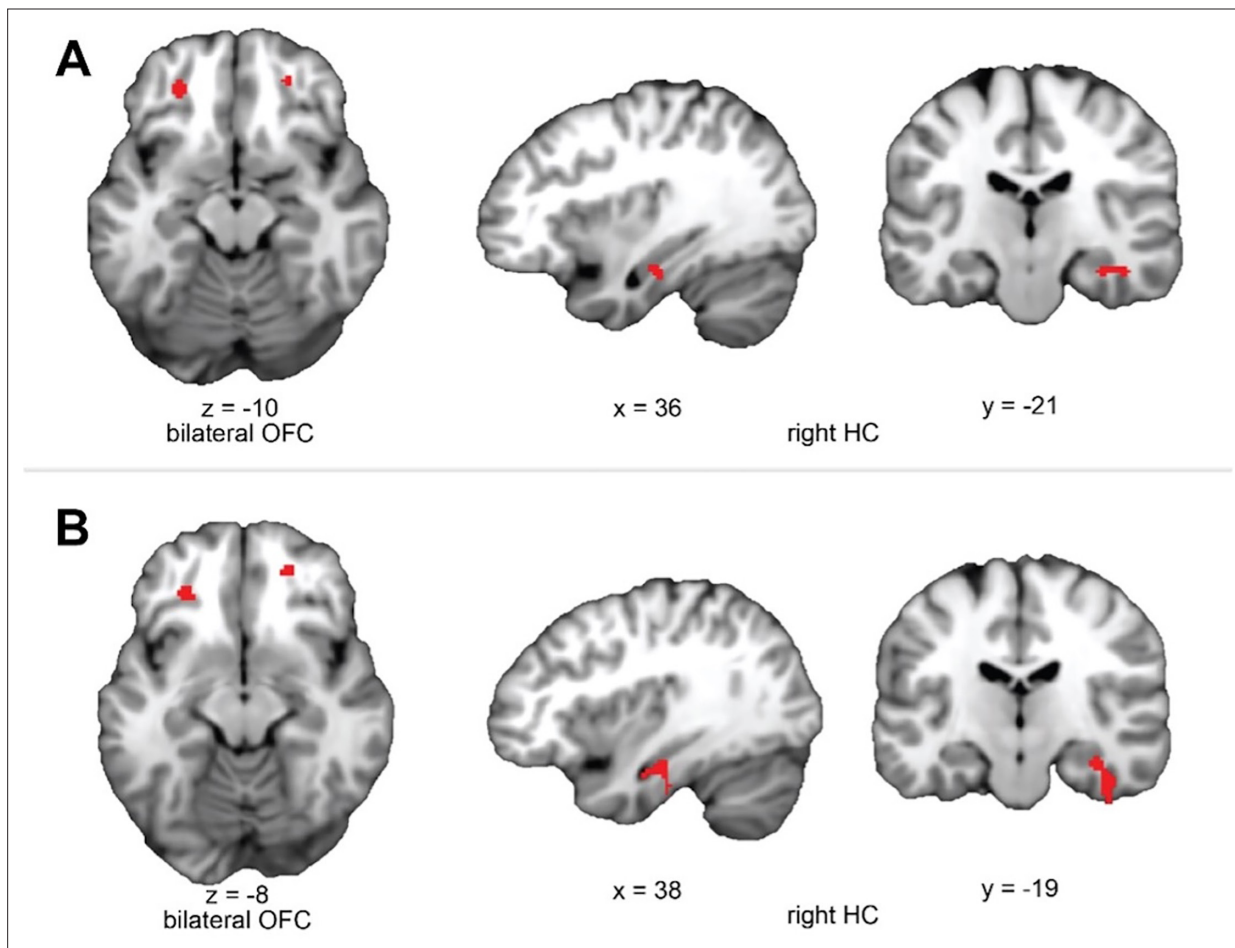
**Figure 1—figure supplement 1.** Follow up behavioral analyses. **(A)** Example trajectory across the experiment of the belief estimates generated from the Bayesian learner. Top is the trajectory of  $S_1$ , and the bottom is the trajectory of  $S_2$ . While lines represent the true probability trajectory is shown in white and the estimated belief is shown in pink. Color heatmap shows the probability mass for each possible belief in  $S_x \rightarrow O1$ . **(B)** Comparison of model fits between our Bayesian model and a value-based RL model (vRL) which used an interactive updating procedure to track the value of each shape based on the history of received rewards. The exceedance probability for the Bayesian model was 1, and 0 for the vRL model, suggesting that Bayesian model, which tracked transition probabilities between choices and outcomes, better fit participants actual choices compared to a value tracking model. **(C)** Logistic regression curves estimating the change in choice probabilities given the expected value difference between choices. Gray line shows participant specific lines, and the black line shows the effect across groups (associated t-statistics are calculated across participants). The left side shows the effect in the direct transition condition and the right side shows the indirect transition condition.



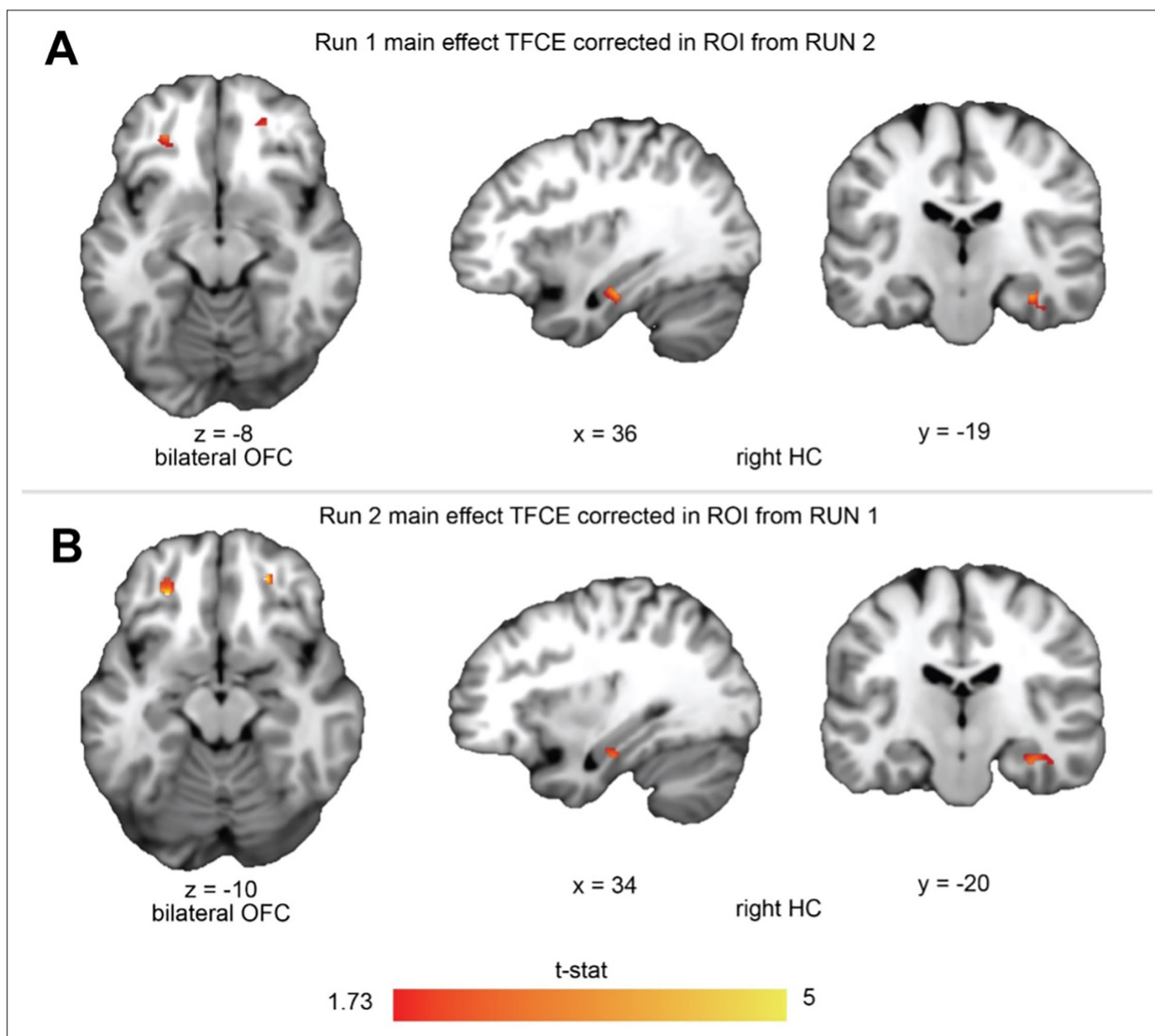
**Figure 2.** IOFC and HC carry representations of the causal choice when viewing outcomes. Left side shows the analysis scheme for decoding representations of the causal choice at feedback in the direct transition condition. An SVM decoder was used to differentiate trials at the time of the outcome (purple) based on the causal choice selected during the 'choice period' (cyan). The right side shows axial and coronal slices through a t-statistic map showing significant decoding in OFC and HC during feedback. For illustration, all maps are displayed at threshold of  $t(19) = 2.54$ ,  $p < 0.01$  uncorrected. All effects survive small volume correction in a priori defined anatomical ROIs. See **Figure 2—figure supplements 1–3** for ROI definition and **Figure 2—figure supplement 4** for power analysis.



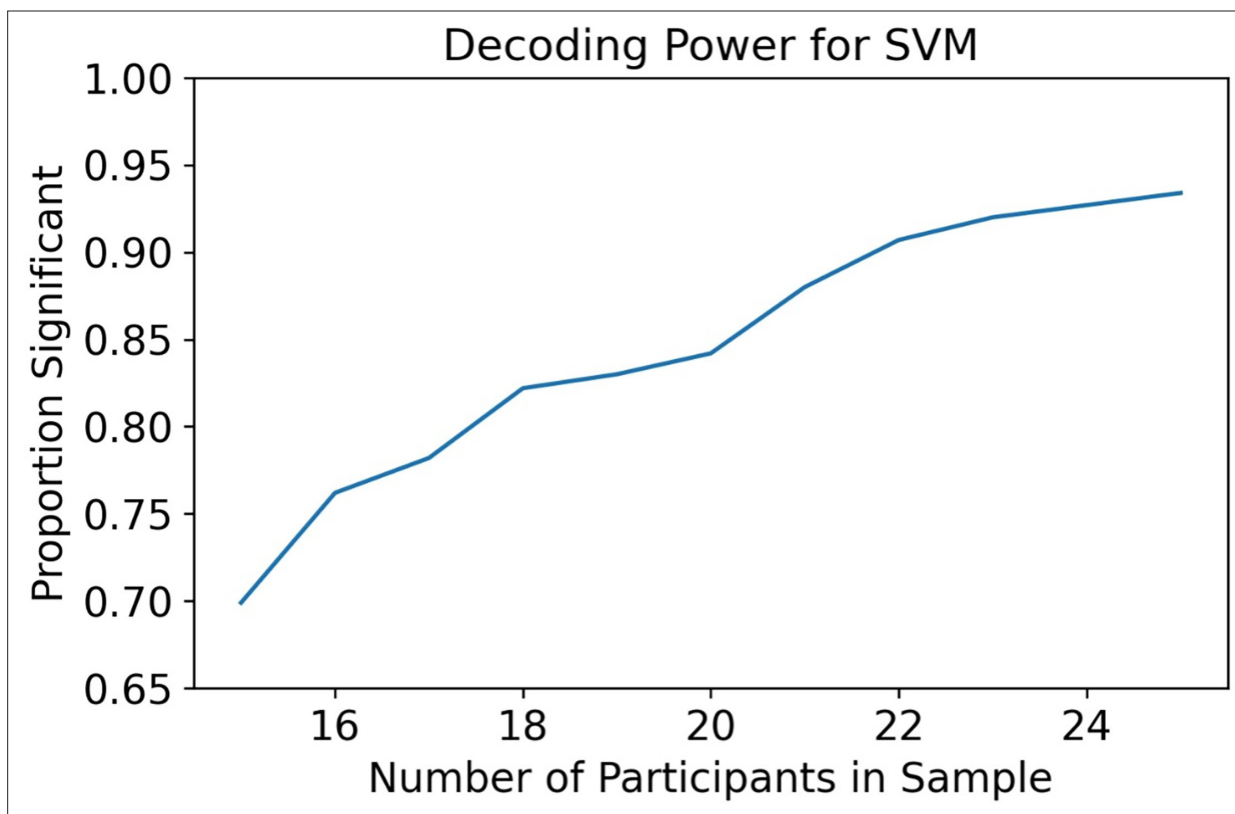
**Figure 2—figure supplement 1.** Pre-selected anatomical ROIs. Illustrations of pre-selected anatomical ROIs taken from *Neubert et al., 2015*. The IOFC ROI corresponds to index 9 and 30, IFPC corresponds to indexes 14 and 35. The HC ROI was defined in *Yushkevich et al., 2015*.



**Figure 2—figure supplement 2.** Functionally defined ROIs for in the direct transitions condition. (A) Despite having a priori defined anatomical ROIs for our decoding analysis of the causal choice, we wanted to test whether our results depended on these ROI definitions by using a data-driven approach. Here, we trained an SVM classifier to decode representations of the causal choice in run 1 of the direct transition condition, then tested the decoder on run 2 to find regions of the orbitofrontal cortex (OFC) and hippocampus (HC) that significantly decoded causal choice representations at a significance level of  $t(19) > 2.54$ ,  $p < 0.01$ , uncorrected. We then used these regions as ROIs for a separate analysis which trained the classifier in run 1 and tested the classifier in run 2. (B) Shows ROIs generated from the same procedure as described in A, but the use of each run for training and testing are switched.

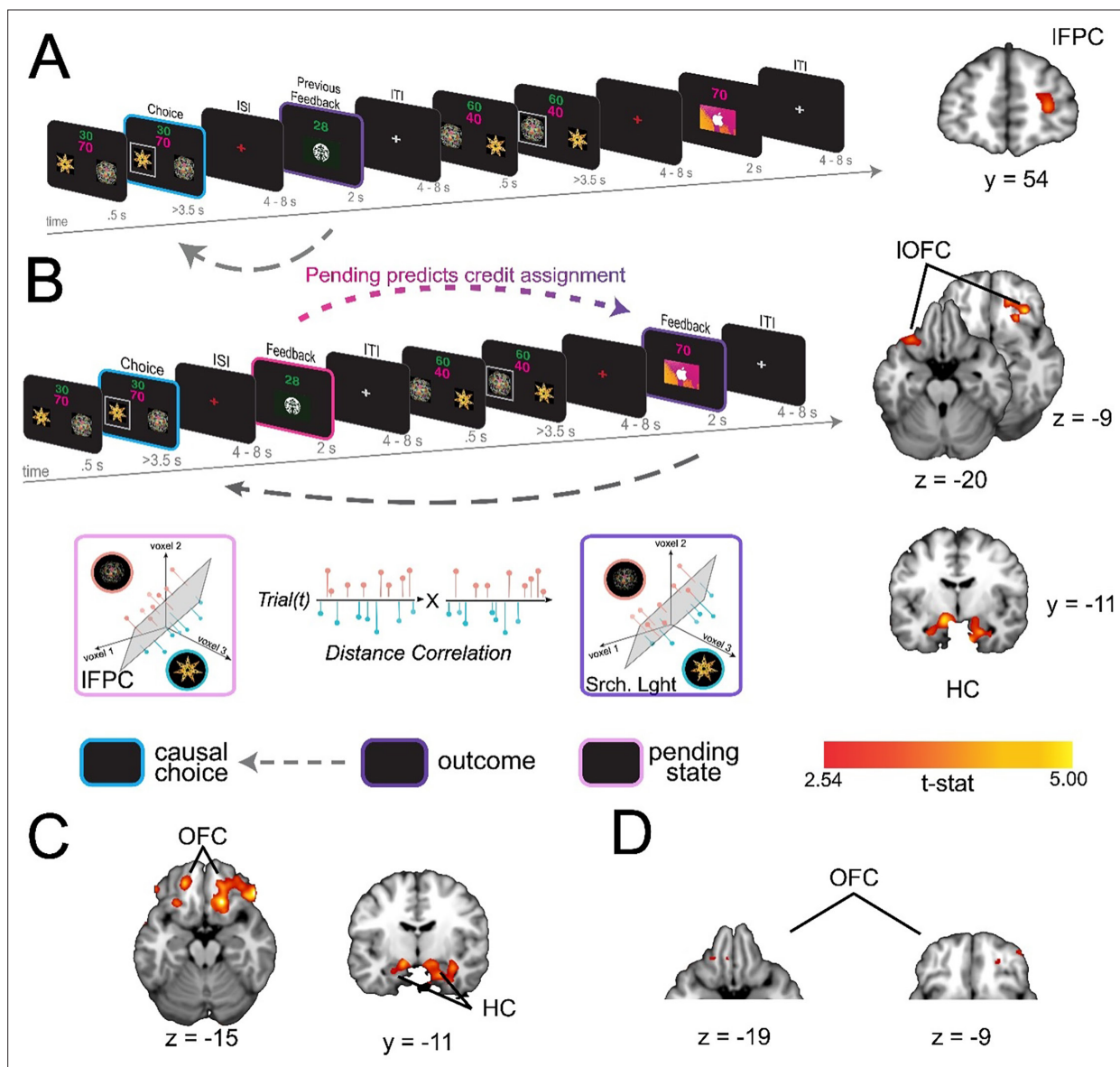


**Figure 2—figure supplement 3.** Main effect of choice decoding accuracy at the time of feedback TFCE corrected in each run of the direct transition condition. **(A)** Regions of the OFC showing significant decoding of the causal choice in run 1 of the direct transition condition. Significance was tested using TFCE correction over voxels with the ROI generated from run 2, using the procedure described above (**Figure 2—figure supplement 2**). For illustration, we show voxels that survive at threshold to  $t(19)=1.73$ ,  $p<0.05$  uncorrected. **(B)** Shows the same as A but for voxels in run 2, using the ROI generated from run 1.

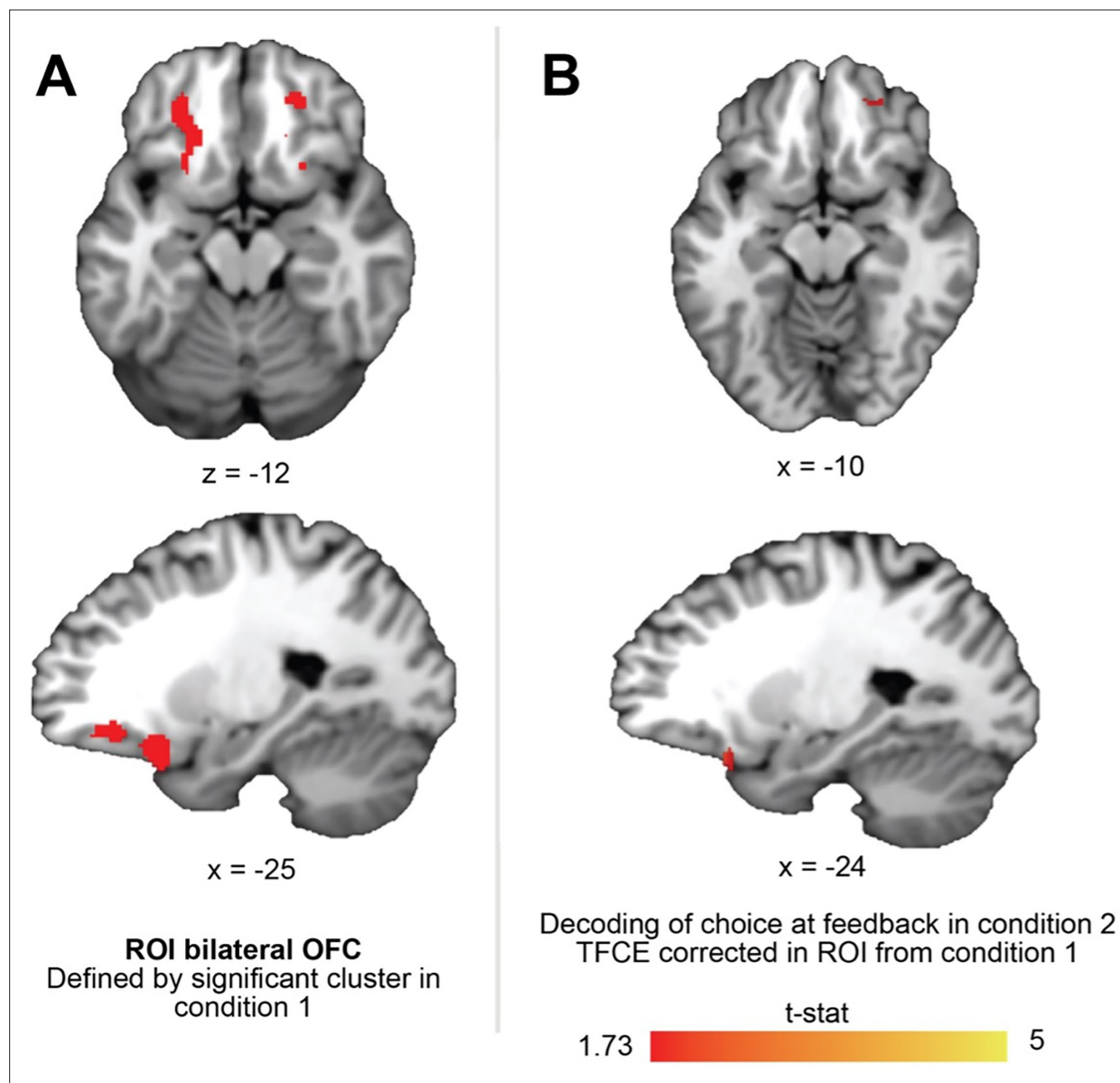


**Figure 2—figure supplement 4.** Power analysis for Reinstatement Effect in the IOFC. Power analysis using an independent data set. Twenty-eight participants completed an associative learning task, in which they learned the causal associations between four different choices, and two food rewards. We estimated voxel activity at the time of the outcome for each trial and tested for multivariate patterns of the causal choice in the IOFC, using the same procedures described in the main text (see Methods). We began by drawing 1000 samples of participants of size N, with replacement, for values of N ranging from 15 to 25. We then tested for significant decoding of the causal choice within each subset using small-volume TFCE correction. Finally, we calculated the proportion of these samples that were at or below a significance level of  $p_{TFCE} < 0.05$ .

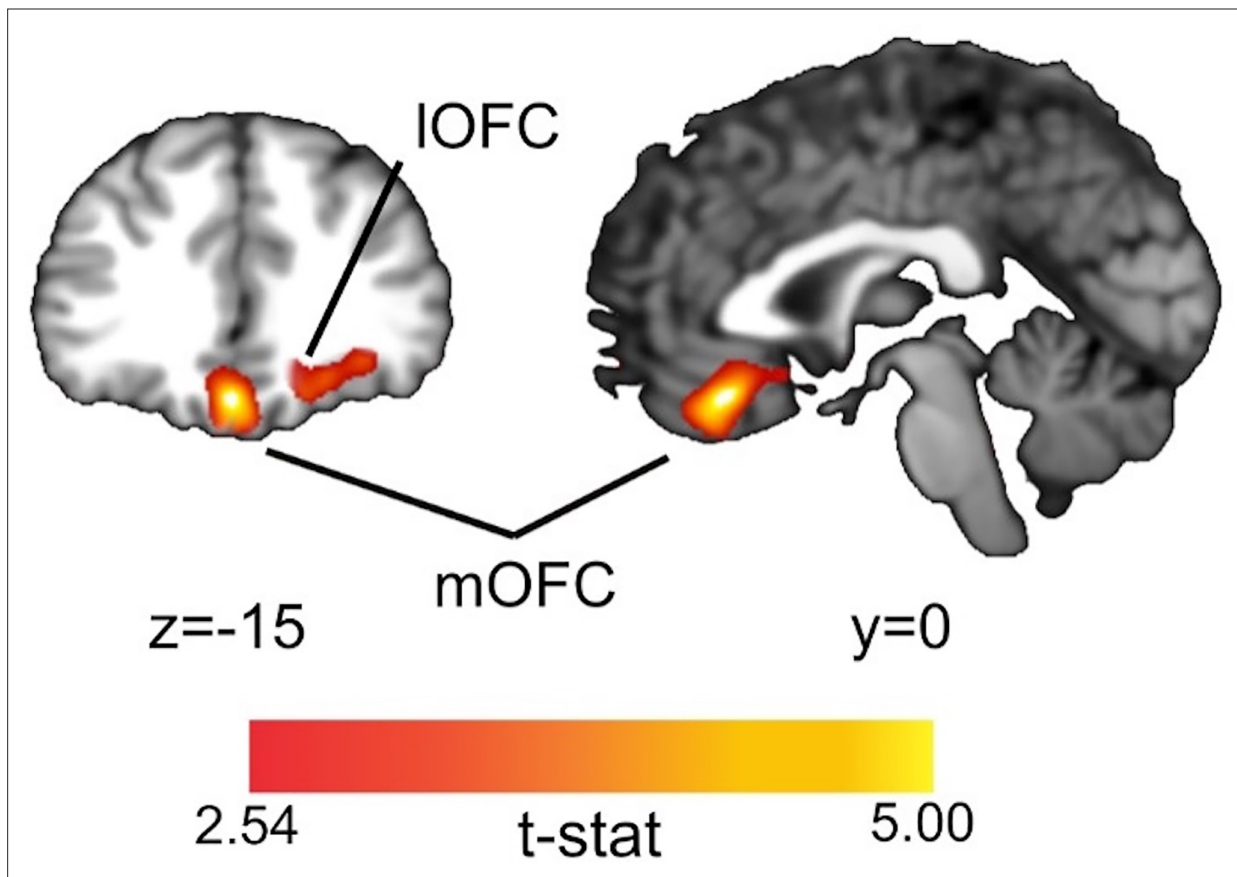




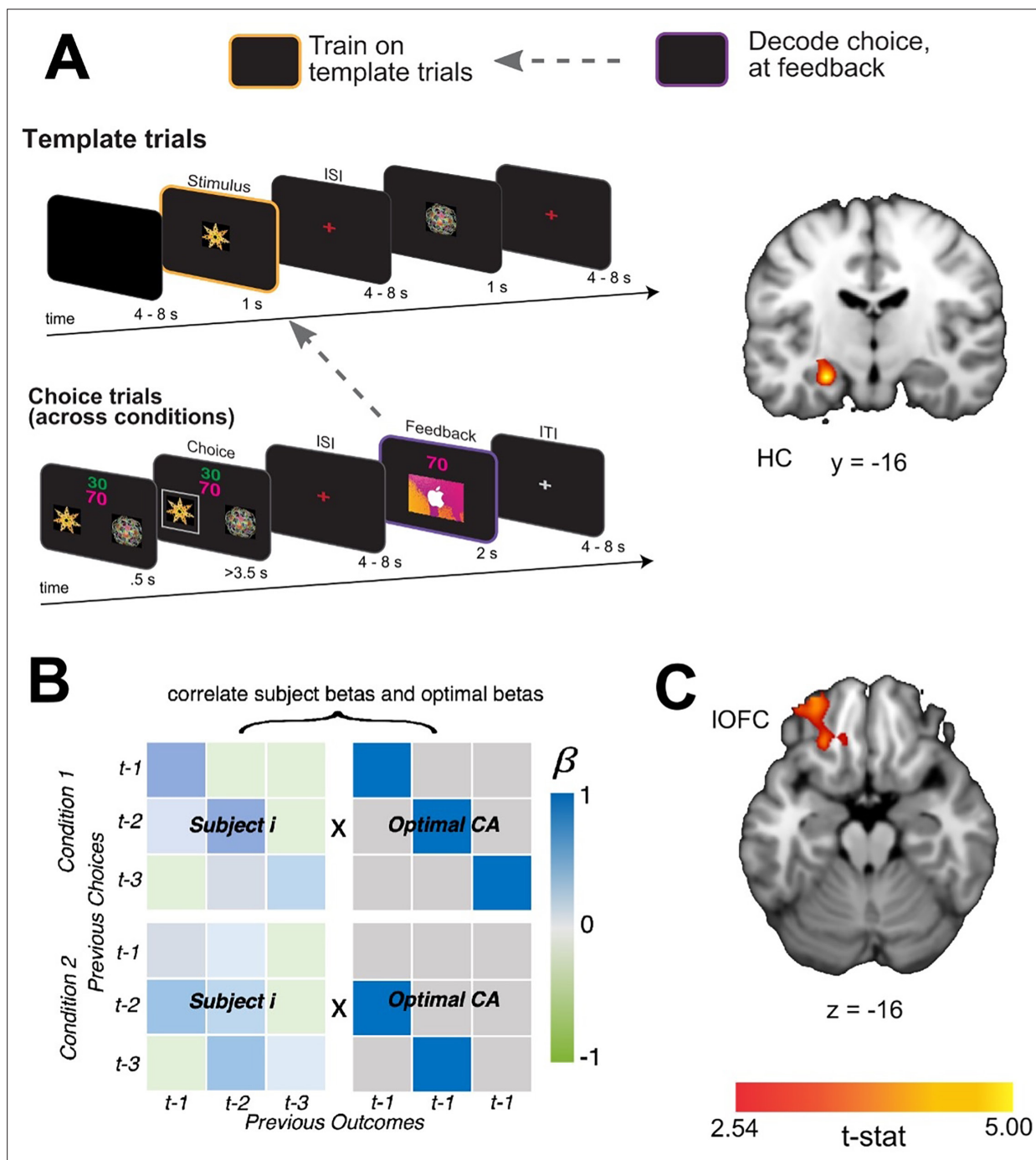
**Figure 3.** IFPC carries representations of the pending causal choice during indirect transitions and predicts credit assignment in the IOFC and HC. (A) Left side shows the analysis scheme for decoding information about the causal choice in 'pending state' (pink) in the indirect transition condition. We decoded information about the previous choice during the feedback period, during which the causal choice should be 'pending' credit assignment in the next trial. The image on the right shows a coronal slice through a t-statistic map, showing significant decoding in IFPC. (B) The analysis scheme for the information connectivity analysis which uses the trial-by-trial fidelity of causal choice representations in the 'pending state' (pink) to predict the fidelity of these same choices when the outcome is observed (purple). The right side shows axial and coronal slices of a t-statistic map showing effects in IOFC and HC. All maps are displayed using the same conventions as **Figure 2** and all effects survive small volume correction in a priori defined anatomical ROIs (for whole brain analysis, see **Figure 3—figure supplement 2**). (C) Axial (left) and coronal (right) slices through a t-statistic map showing the results of a control analysis in which we test the proportion of correct classifications of causal choice information in OFC and HPC at the time of the outcome for trials in which the IFPC showed correct classification for the causal choice during pending trials. The proportion of correct trials was compared to a permuted baseline of randomly drawn trials for each participant then combined over participants to create a t-statistic. (D) Secondary control analysis in which we reran the classification analysis for causal choice information at the time of outcome, but only on trials where IFPC was found to correctly decode pending causal choice information. Note that this test is different from A because we allowed the classifier to create a new hyperplane separating categories for only those trials in which the IFPC decoding was 'correct'. For illustration, all maps are displayed at a threshold of  $t(19)=2.54$ ,  $p<0.01$  uncorrected. All effects survive small volume correction in a priori and functionally defined anatomical ROIs. See **Figure 3—figure supplements 1–2** for ROI definition and whole brain searchlight.



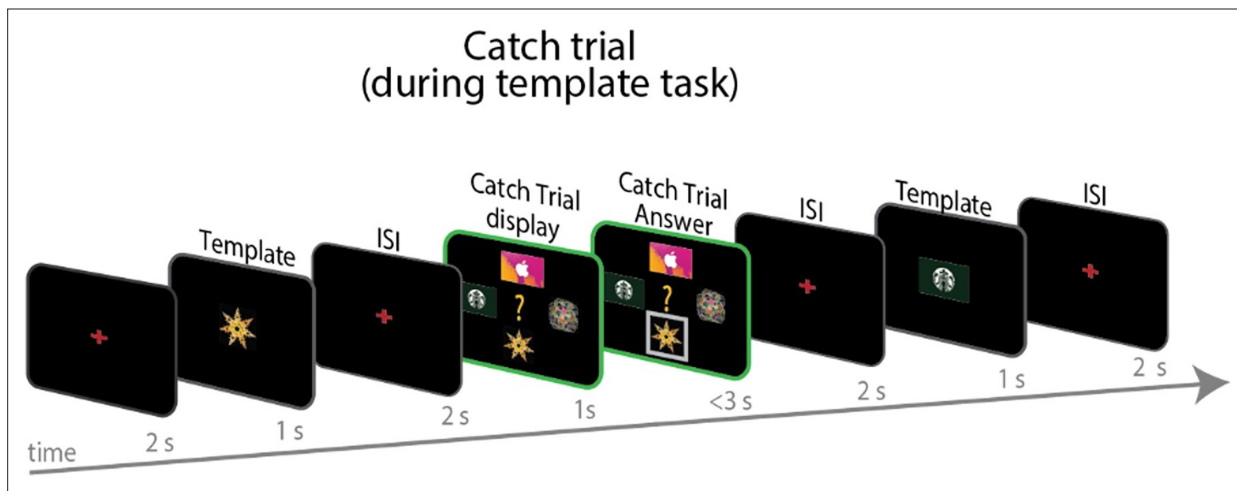
**Figure 3—figure supplement 1.** Significant information connectivity between IFPC and OFC in functionally defined ROI from direct transition condition. **(A)** We did not observe significant decoding of the causal choice in a bilateral OFC ROI defined by significant cluster in the indirect transition condition. Thus, we used the accuracy map for decoding choices at feedback during the direct transition condition ( $t(19) > 1.73$ ;  $p < 0.05$ ) in the OFC, averaged across runs. **(B)** We then used those clusters as ROI for TFCE correction for regions of the IOFC that showed significant information connectivity with IFPC. We did this by testing for significant correlations between the trial-by-trial fidelity of pending representations in the IFPC and causal choice representation during feedback in IOFC (see Methods).



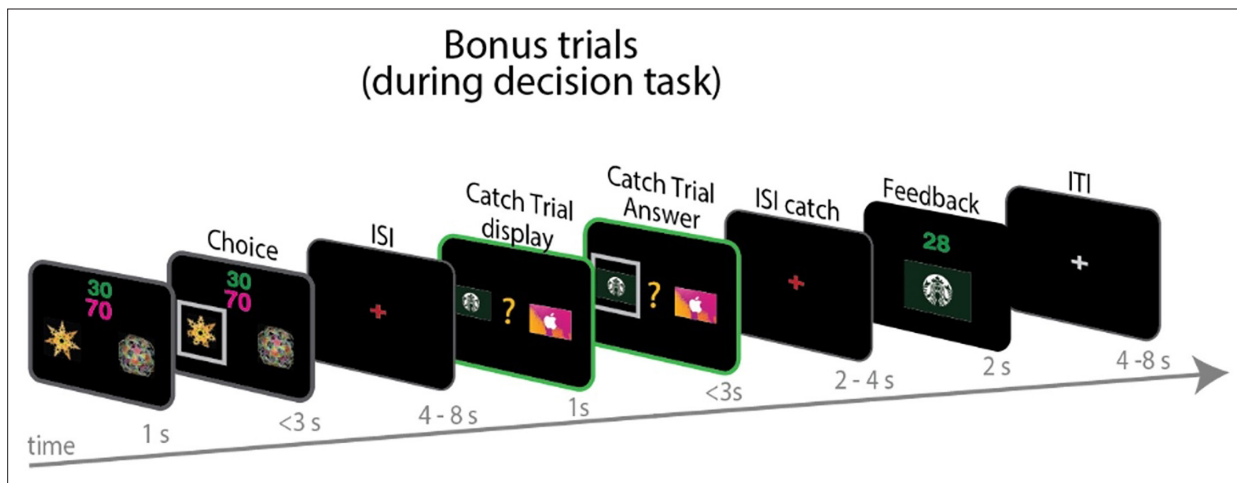
**Figure 3—figure supplement 2.** Exploratory information connectivity analysis for 'Indirect transition condition'. To ascertain whether additional regions maybe involved in credit assignment beyond those that formed the focus of our study, we repeated the analysis described in **Figure 3** but using a whole brain search light procedure. All aspects of the analyses were the same as those previously conducted except that we corrected for multiple comparisons at the whole brain level using TFCE. For the 'direct transition condition', we found no additional regions that showed high decoding for the causal choice at the time of outcome. However, for the 'indirect transition condition' we identified a region of medial OFC (mOFC) which showed information about the causal choice that was predicted by pending representation in IFPC (pTFCE <.05). The left panel shows a coronal slice through a t-statistic map, thresholded using the same conventions as **Figure 3**; the right panel shows a sagittal slice through the same map. These results suggest a potential role for mOFC in credit assignment uniquely during the 'indirect transition condition'.



**Figure 4.** Task-independent representations of causal stimuli in HC at feedback. (**A - left**) Schematic of the decoding procedure. In task-independent 'template trials', participants passively viewed images corresponding to the two choice stimuli and two outcome stimuli in the main task (for more information see **Figure 4—figure supplement 1**). We used these trials to train an SVM to differentiate stimuli outside the task context and then tested for representations of the causal choice stimulus at the time of feedback during the learning task. (**A - right**) A coronal slice through a t-statistic map showing regions of the HC with significantly above chance decoding for the causal choice stimulus identity at the time of feedback, across conditions. In this figure, 'CA' refers to 'credit-assignment'. (**B**) Analysis scheme for generating each participant's overall credit assignment precision.  $\beta$ -values for each participant were taken from the behavioral model predicting current choices given all combinations of the previous three choices and outcomes (**Equation 1**). Each participant's pattern of  $\beta$ -values (left side matrices) were correlated with a matrix representing an optimal pattern of regression betas given the task structure (right side matrices). The optimal matrix was a binary matrix with ones where credit should be assigned for a given outcomes and zeros everywhere else. (**C**) Axial slice through a t-statistic map showing regions where decoding of the stimulus identity was significantly correlated with estimates of credit assignment precision. All maps are displayed using the same conventions as **Figure 2** and all effects survive small volume correction in a priori defined anatomical ROIs. See **Figure 4—figure supplements 1 and 2** for catch and bonus trials definition.



**Figure 4—figure supplement 1.** Depiction of catch trials. To ensure that participants where we included valuable *catch trials* in the passive observing 'template task'. Participants were asked to report which image out of the four (2 gift cards and 2 stimuli) was the last one presented on the screen. They were endowed an extra £10 from which we removed £1 for every incorrect response. There were four catch trials per template run.



**Figure 4—figure supplement 2.** Depiction of bonus trials. To ensure that participants where we included valuable *catch trials* in the passive observing 'template task'. Participants were asked to report which image out of the four (2 gift cards and 2 stimuli) was the last one presented on the screen. They were endowed an extra £10 from which we removed £1 for every incorrect response. There were four catch trials per template run. The decision task included 'bonus trials' in which participants could predict which gift card they expected to see on the subsequent feedback screen given their choice. They were given 3£ extra on the final gift card that was given to them for every correct answer. The first run of the direct transition condition had two catch trials; the second run had one. Both runs of the indirect transition condition had one catch trial each.