Supplemental information

Novelty and uncertainty regulate the balance between exploration and exploitation through distinct mechanisms in the human brain

Jeffrey Cockburn, Vincent Man, William A. Cunningham, and John P. O'Doherty
Figure S 1: Related to Figure 3. Learning in the forgetful Bayesian agent. A) The forgetting curve for three values of $\eta$. When $\eta = 0$, values are integrated equally regardless of how many trials have passed, reflecting optimal Bayesian integration with perfect memory (red). As $\eta$ increases, recently observed outcomes are weighted more heavily than those observed in the past (blue and green). This implies that the expected value and uncertainty will reflect more recent outcomes as $\eta \to 1$. B) The effect of $\eta$ on the regression model’s predicted effect of expected value as a function of trial number. When $\eta = 0$ both the fmUCB and the regression model agree on the expected value derived for each option, resulting in a consistent effect of EV across trials (i.e. $\beta_{EV} = 0$). As $\eta \to 1$, the expected values derived by the computational and regression models grow increasingly divergent, resulting in a decreasing effect of EV.

Figure S 2: Related to Figure 3. Parameter and model identifiability in the fmUCB model. A) Correlation between generative and estimated parameters ($\beta$: $r = 0.96$; $\eta$: $r = 0.95$; $N_I$: $r = 0.85$; $U_I$: $r = 0.97$; $U_T$: $r = 0.97$). x-axis: generative parameters; y-axis: estimated parameters; red dashed line: identity line; gray line: best fit linear correlation. B) Model identification showing that the generative model is reliably captured as the best fit model across a range of candidate model structures. Panel names indicate the generative model variant; x-axis: models fit to simulated data; y-axis: negative log evidence for each model fit indicating error and accounting for additional parameters. The generative model can be identified by light grey fill, and the best fit model by a + symbol.
Figure S 3: Related to Figure 4. GLM regressor correlations. 

A) Mean correlation among regressors for GLM 1 across participants, consisting of selected and rejected option utility (sU-S, sU-R respectively), selected and rejected option estimation uncertainty (u-S, u-R), and selected and rejected option novelty (n-S, n-R). 

B) Mean correlation among regressors for GLM 2 across participants, consisting of selected and rejected option expected value (q-S, q-R respectively), selected and rejected option uncertainty bias (uB-S, uB-R), selected and rejected option estimation uncertainty (u-S, u-R), and selected and rejected option novelty (n-S, n-R). Color bars depict correlation coefficients ranging from [1, -1].

C) Neural correlates associated with the uncertainty bias value (selected + rejected option value) from the GLM reported in Figure 4B with stimulus novelty terms removed. Reported voxels have t-score > 2.
Figure S 4: Related to Figure 4. Significant clusters related to GLM effects reported in Figure 4. A) Significant clusters identified for GLM 1. B) Significant clusters identified for GLM 2. Significant clusters were identified using whole-brain cluster-level FWE correction at p < 0.05 and cluster-forming threshold at p < 0.001 uncorrected. Peak voxel t-scores are reported for each cluster (MNI coordinates). C) Mean beta estimates from each of the clusters reported in Figure 4 (Sel = Selection option, Rej = Rejected option). Mean estimates were computed by extracting the beta estimates for each voxel in the cluster across participants, then averaging over voxels and participants. Bars depict the average cluster beta estimate across participants, and error bars depict the standard error.
Figure S 5: Related to Figure 5. Voxels correlated with expected value and reward prediction errors across mechanisms of novelty integration. **A**) A cluster of voxels in vmPFC correlated with q-values estimated by the policy bias mechanism (blue), and voxels correlated with q-values as estimated by the optimistic initialization mechanism (green bordered). Mask comprising 15mm sphere centered on independent meta-analysis peak voxel in vmPFC bordered in white. **B**) A cluster of voxels in ventral striatum correlated with reward prediction errors estimated by the policy bias mechanism (blue), and voxels correlated with reward prediction errors as estimated by the optimistic initialization mechanism (green bordered). Whole-brain maps for signal of interest were tested with a cluster-forming threshold of $p < 0.001$ uncorrected, followed by cluster-level FWE correction at $p < 0.05$.

Figure S 6: Related to Figure 5. Voxels correlated with the uncertainty bias. A cluster of voxels in mPFC correlated with the uncertainty bias absent familiarity modulation (blue), but only at a liberal threshold of $p < 0.01$. This region overlaps the cluster associated with the familiarity modulated uncertainty bias (green border), which was tested with a cluster-forming threshold of $p < 0.001$ uncorrected, followed by cluster-level FWE correction at $p < 0.05$. 

43