Supplemental material for "Dissociating hippocampal and striatal contributions to sequential prediction learning." (*European Journal of Neuroscience 35:1011-1023*).

Methods

Transition matrix generation

Images were selected according to a set of first-order contingencies specified by three session-unique, pseudorandomly generated transition matrices for each subject. These matrices were designed to minimize the influence of response biases arising from an overall unbalance between the unconditional image probabilities — even ephemerally — and also to allow us to observe reaction time across a wide range of first-order conditional probabilities. The specific procedure used is as follows. Each of the four rows was generated, one at a time, in a pseudorandom order. For the first row, four values were selected in a range of 0.05 to 0.95, and then normalized to sum to one. For ensuing rows, each value was pseudorandomly selected from a range with an upper bound of one minus the sum of the appropriate column, and then normalized across the row. For the final row, each column was set to one minus the sum of the other rows in that column, obviating any further normalization. In this way we ensured that each set of image contingencies was a probability distribution (rows sums to one) and also ensured uniform overall presentation frequency in the steady state (uniform stationary distribution: all columns sum to one).

Among matrices generated this way, we retained only those that achieved the uniform stationary distribution quickly (fast *mixing time*, see below) and selected for a wide range of transition entropy (SD of row entropy > 0.25), to allow for a broad sample in this measure as well. The transition probabilities in the matrices ultimately used in our task ranged between 0.0019 and 0.7803 (mean 0.2503, standard deviation 0.1878), and transition entropy ranged between 0.6415 to 1.3854 (mean 1.1157, standard deviation 0.1859).

Mixing time. Mixing time is a measure of how many steps a Markov process can be expected to proceed before the distribution of states visited approximates the stationary distribution (which our case, uniform). Fast mixing was ensured by a two-step process. First, we normalized each row and column of the transition matrix, as detailed above. This property generates matrices that are likely to have a high difference between the first and second eigenvalues, ensuring a low value for the conductance of the resulting graph. The conductance is heuristically related to mixing time (Sinclair, 1993). Second, a measure of worst-case mixing time was computed by beginning with any individual image and counting the number

of successive applications of the transition matrix required to obtain expected observation frequencies over images within one percent of the uniform distribution (25% per image). Probability matrices which did not mix in fewer than ten steps were discarded.

Learning rate analysis

We sought to identify activations in each of our brain regions of interest as indicative of one of two distinct learning processes, both of which influence behavior. These processes are distinguished by their learning rate, and so we sought to statistically validate that BOLD signal in each region is best explained by regressors generated by one rate rather than the other. This analysis was complicated by the fact that these regressors are correlated with one another, and so rather than employing a traditional regression analysis, we used a method designed to infer which value of the learning rate parameter which would optimally describe activation in each region. Specifically, we constructed the forward entropy and conditional probability regressors as estimated by a single process learning at the rate α_0 — which we set to the average of the two behaviorally identified rates — and included two additional regressors measuring how these regressors would change if they had been generated from the model with a different learning rate. Technically, we defined these additional regressors as the partial derivatives of the probability and entropy timeseries with respect to the learning rate parameter, evaluated at α_0 (Holmes and Friston, 1998). This analysis allows us to estimate the change in learning rate, relative to the reference point α_0, that would best explain BOLD in an area, by using a regression to estimate coefficients for the first two terms in the Taylor expansion of the dependence of the regressor on the learning rate. This takes the following form:

$$F(\alpha_{BOLD}) \approx F(\alpha_0) + (\alpha_{BOLD} - \alpha_0)^{\partial F(\alpha)/\partial \alpha} \qquad [S1]$$

Here F(α) is the regressor of interest (i.e., the probability or entropy timeseries), viewed as a function of the learning rate α, and α_{BOLD} is some other learning rate for which the regressor would best fit the BOLD signal. To encode learning rates in this analysis we used a change of variables by which the original Rescorla-Wagner learning rate was transformed by an inverse sigmoid, so that it ranged throughout the real numbers and estimates of it could be treated with Gaussian statistics. Thus F(α) above first maps the unconstrained argument α through the logistic sigmoid function to produce a learning rate in original (0-1 constrained) units, and the partial derivative accordingly includes the derivative of this transform by the chain rule. This linear approximation to the (nonlinear) relationship between

the regressor and the learning rate parameter allows the use of a GLM to approximately estimate the learning rates that would best explain BOLD correlates to the regressor. In particular, the weight estimated for the partial derivative regressor corresponds to α_{BOLD} α_0 (or, more particularly, $k \cdot [\alpha_{BOLD}\alpha_0]$, if the net effect of the regressor on BOLD is scaled by multiplying both sides of Equation S1 by some factor k). This is just the degree to which the best-fit (inverse-sigmoid transformed) learning rate for explaining the BOLD response differs from α_0, the value used to calculate our regressor of interest and its derivative. We thus computed estimates of α_{BOLD} for each regressor (entropy or probability) at a voxel by first extracting the regression weights for the partial derivative regressor for each subject. To normalize these coefficients to a common scale in units of transformed learning rate (even if they originated from different regions or underlying regressors, i.e. probability or entropy), we divided these weights by the average, across subjects, of the regression weights for the corresponding regressor F(α_0) at the voxel, this corresponding to the overall scale factor k mentioned above. Lastly, we added the reference value α_0, converting the result into the range of our behaviorally-obtained rates. Our statistical analyses were all performed on the learning rate estimates in the transformed units, taken across the population. Specifically, we test whether the computed α_{BOLD} is statistically distinguishable from learning rate values obtained by fitting behavior, via t-tests against each (transformed) fit rate. We also test whether α_{BOLD} differs between regions, by comparing the estimates in paired-sample t-tests. For Figure 5 (main text), we mapped the mean estimates and their confidence intervals through the sigmoid to depict them in units of Rescorla-Wagner learning rate. To maximize power, to examine learning-rate effects at areas where there was learning- related activity, and to identify areas to allow between-region comparisons, we first selected voxels at which to perform these analyses of learning rates using contrasts on the main effect of the conditional probability and entropy regressors (not their derivatives). This was one motivation for choosing α_0 to be the midpoint of the fast and slow rates i.e., that it is roughly equally suited to detect activity related to either. Additionally, the linear approximation to α_{BOLD} is most accurate when the difference α_{BOLD} - α_0 is small, suggesting placing α_0 equally close to both relevant learning rates. We selected the voxels of peak group activation within each of our a priori regions of interest. Differences between parameters in the subsequent tests were considered reliable at a level of $p < 0.05$. Note that the various tests on these estimates do not require correction for bias introduced by the (region-wise) multiple comparisons used in voxel selection. This is because the estimate α_{BOLD} is computed from the estimate for the partial derivative, scaled by the estimate for the main effect. The partial derivative regressors are by definition orthogonal to the main effects, and thus the estimate for the partial derivative is essentially independent of the selecting contrast. In particular, after whitening and

filtering, which in principle can re-correlate orthogonal regressors (Kriegeskorte et al., 2009), the average correlation across subjects between regressors representing the derivative and the original variable was 0.026 for conditional probability and 0.076 for forward entropy. For testing α_{BOLD} against the average learning rate α_0 (equivalent to testing the estimate for the partial derivative against zero), the scale factor is a constant, and only the partial derivative affects the significance of the test. For the other tests on implied learning rates, the selecting contrast does introduce a bias, but it is a bias that works against the hypotheses we test. In particular, the selecting contrast extremizes the scale factor (the parameter estimate for the main effect of entropy or probability), which (since the estimate is divided by this factor) minimizes the absolute estimate of α_{BOLD} α_0, or equivalently biases the estimate of α_{BOLD} toward α_0. In other words, selecting voxels for best correlation with the probability or entropy computed for learning rate α_0 tends to bias the estimated learning rate toward the selecting learning rate in our case, the midpoint – which biases the estimate against the hypotheses that α_{BOLD} differs between regions or that it corresponds to the fast or slow values of the learning rate. Thus, observing a significant effect for the derivative regressors is a strong test of the hypothesis that either of the extreme, behaviorally-obtained, learning rates, is a superior explanation of neural activity compared to the midpoint value used for voxel selection.

Additional GLM analyses

In order to investigate alternative approaches to the fMRI data, we conducted two additional GLM analyses, which were similar to those described in the main text except for the definition of the parametric regressors of interest. For the first, (combined process) we generated probability and entropy estimates from the net transition matrix (the weighted sum of both fast and slow transition matrices, from Eq. 2), rather than either the fast or slow separately. For the second, we included regressors from both fast and slow processes in a single GLM, without orthogonalizing one to the other.

Results

fMRI: Combined Process

The analyses in the main text are premised on the assumption that neural correlates relate to either of the hypothesized learning processes separately. As a partial test of this assumption, we conducted an additional exploratory analysis using regressors generated from the combined learning process (Figure 2).

Conditional Probability When generating regressors according to the estimates of the combined process (the behaviorally-observed weighted sum of expectations learned by the fast and slow LR processes) as fit to behavior, we observed correlates to the conditional probability regressor in striatum (specifically, right medial caudate [12, 8, 6 ; $p < 0.0005$ unc.]).

Forward Entropy Examining the forward entropy regressor in the combined process revealed negative correlates in caudate [left -14, 4, 12 $p < 0.0003$ unc.; right 12, 8, 0 $p < 0.0006$ unc.]. In our other major region of interest, MTL, no activity was observed reaching even an exploratory uncorrected reporting threshold of .001. Overall, that these results are weaker than those reported in the main text (in particular, that none of these results survive correction for multiple comparisons), is consistent with our interpretation that neural activity is related to separate rather than combined processes. Also, that the results, in Figure 2, more closely resemble the correlates of the fast process (Figures 3a and 4, main text) which we speculate may be due to the slightly greater weight of this process in the combined probabilities.

Figures

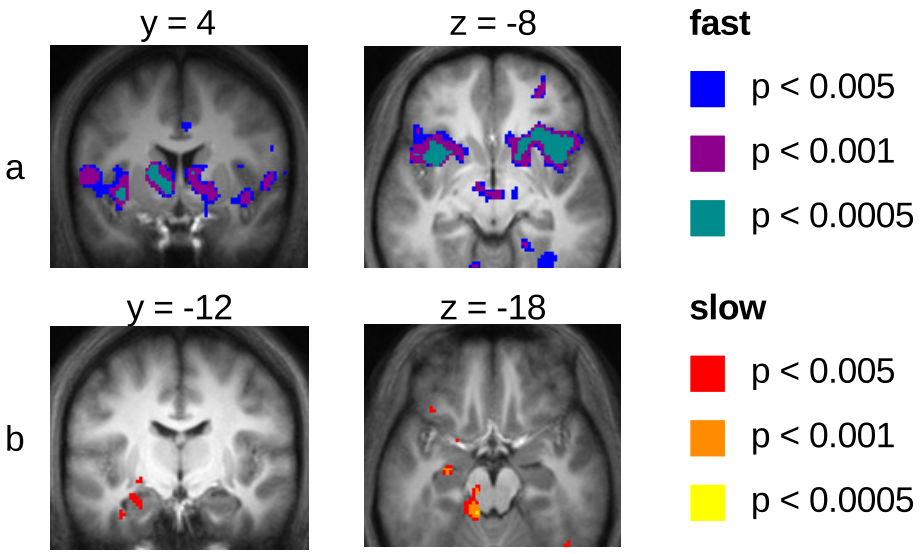

Figure 1: **Non-orthogonalized SPMs.** To observe the results of our analyses uncontaminated by any residual correlation that might be imparted by SPM's implicit orthogonalization, this GLM was run with orthogonalization turned *off*, for all regressors. The regressors of interest (probability, entropy) from both fast and slow processes were used to construct a single GLM. The entropy results, shown here, reflect similar patterns of activity as in the main text (Figure 3a and b), albeit at a less-stringent statistical threshold.

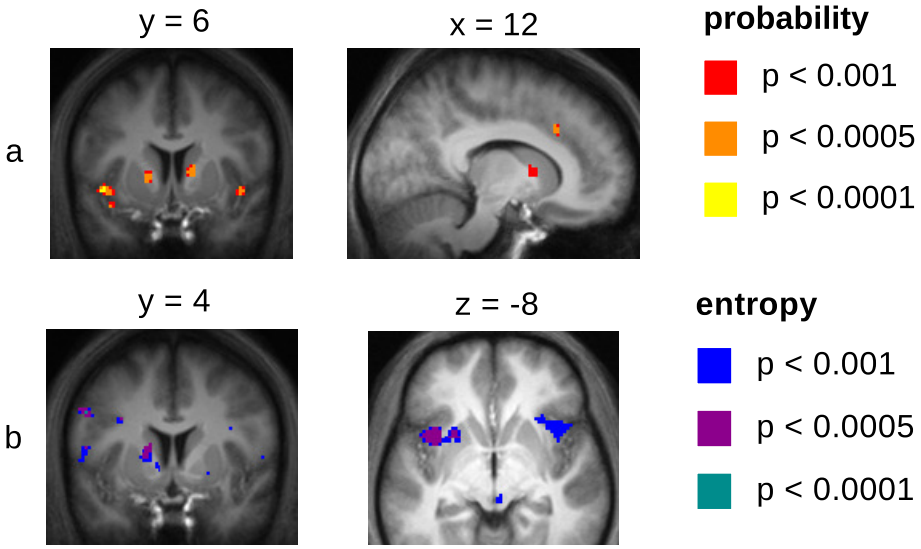

Figure 2: **Combined process SPMs.** These maps were generated using the two process regressors, with learning rate and weight parameters as fit to behavior. The probability and entropy results, shown here, reflect a similar pattern of activity as in the fast process (Figures 3a and 4, main text), perhaps due to the slightly greater weight applied to this process in effecting behavior

Tables

Region	Hem	MNI	Extent	T	p (unc.)
Putamen	R	18 6 -6	84	4.67	0.000014 *
Inferior Frontal Gyrus	R	36 16 -4	1008	9.56	0.000000 ++
Thalamus	L	-14 -18 4	198	9.34	0.000000 ++
Right Calcarine Sulcus	R	22 -56 4	1004	8.24	0.000000 ++
Left Anterior Insula	L	-40 14 -4	923	8.03	0.000000 ++

Table 1: Areas of *negative* correlation with the forward entropy regressor in the fast process GLM.
** = $p < 0.05$ after FWE correction for search in an anatomically defined mask of the ventral striatum and hippocampus.
* = $p < 0.1$ after correction.
++ = $p < 0.05$ after whole-brain correction for FWE.
+ = $p < 0.1$ after whole-brain correction.

Region	Hem	(xyz)	Extent	T	p_{unc}
Hippocampus (Anterior)	L	-26 -10 -18	36	5.71	0.000021 **
Parahippocampal Cortex	L	-16 -34 -18	35	4.84	0.000109 *

Table 2: Regions of significant correlation with the forward entropy regressor in the slow process GLM.
** = $p < 0.05$ after FWE correction for search in an anatomically defined mask of the ventral striatum and hippocampus.
* = $p < 0.1$ after FWE correction.
Note: Parahippocampal cluster was significantly active outside of the hippocampal mask, and therefore was disadvantaged by correction on the hippocampus proper.

Region	Hem	MNI	Extent	T	p_{unc}
Putamen	R	18 14 -4	25	4.52	0.000204 *

Table 3: Regions of significant correlation with the conditional probability regressor in the fast process GLM.
** = $p < 0.05$ after FWE correction for search in an anatomically defined mask of the bilateral ventral striatum and hippocampus.
* = $p < 0.1$ after correction.

Region	Hem	MNI	Extent	T	p_{unc}
Postcentral Gyrus	L	-58 -6 20	40	5.47	0.000032
Anterior Cingulate	L	0 28 24	81	5.12	0.000063
FrontalInfTriR (aal)	R	54 20 0	19	4.25	0.000347

Table 4: Regions of significant correlation with the conditional probability regressor in the slow process GLM.
No clusters survived whole-brain correction for FWE.

References

Holmes, A. P. and Friston, K. J. (1998). Generalisability , Random Effects & Population Inference. *Neuroimage*, 7:S754.

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S. F., and Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nature Neuroscience*, 12(5):535–40.

Sinclair, A. (1993). *Algorithms for random generation and counting: a Markov chain approach*. Springer.