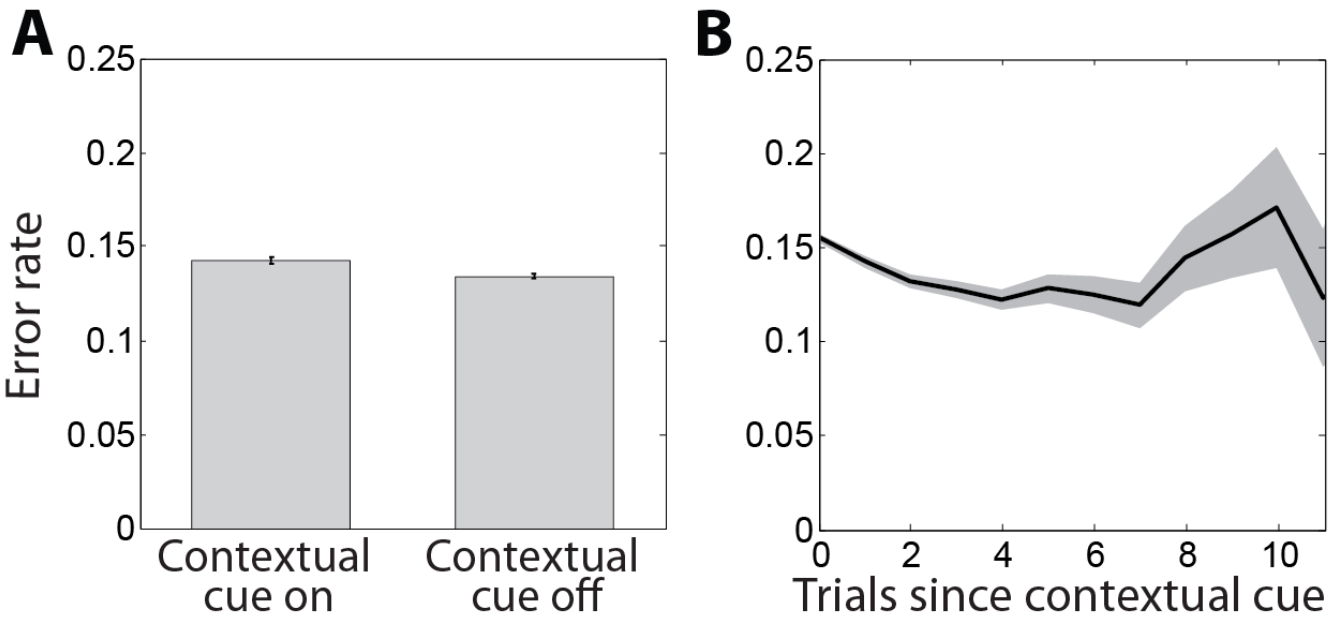

Figure S1. The presence of the contextual cue had little effect on the likelihood of error (related to Figure 1).

(A) Error rate (number of error trials divided by total number of trials) for trials with and without the contextual cue. Error trials are defined as trials in which the monkey's anticipatory licking response to the CS was not predictive of the US (see Experimental Procedures). **(B)** Error rate as a function of the number of trials experienced since the last presentation of the contextual cue. Error bars: estimated s.e.m.

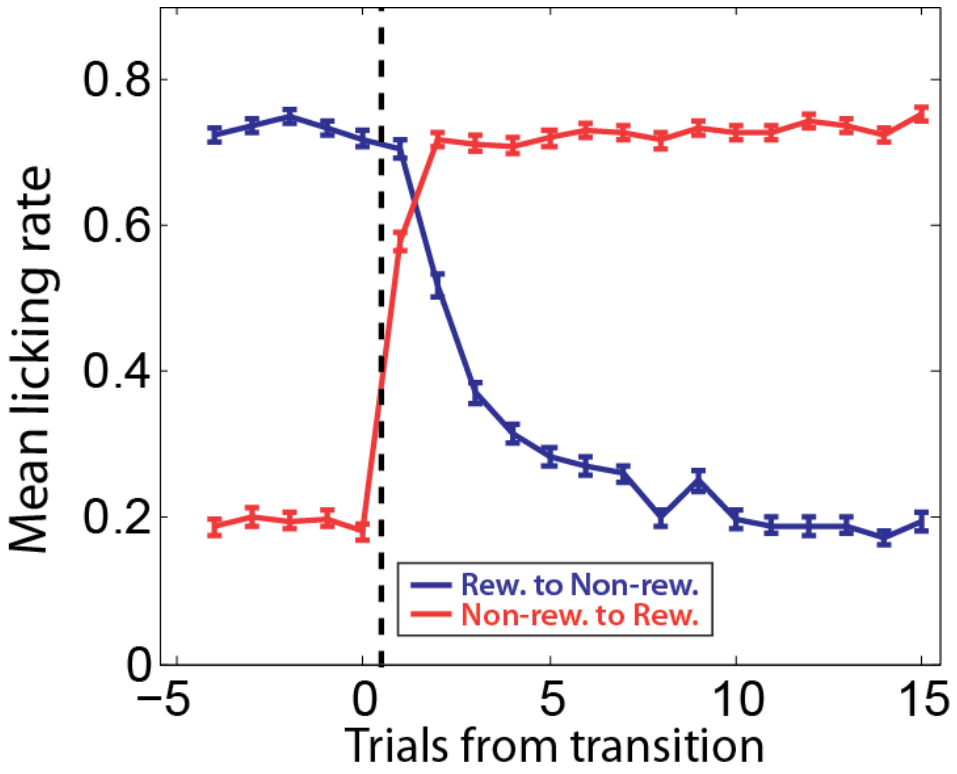

Figure S2: Average licking rate around a block switch (related to Figure 2B).

Mean licking rate, across all block switches for each trial type, plotted as a function of the trial's position relative to the switch, where 1 denotes the first trial of the new block. Error bars: s.e.m.

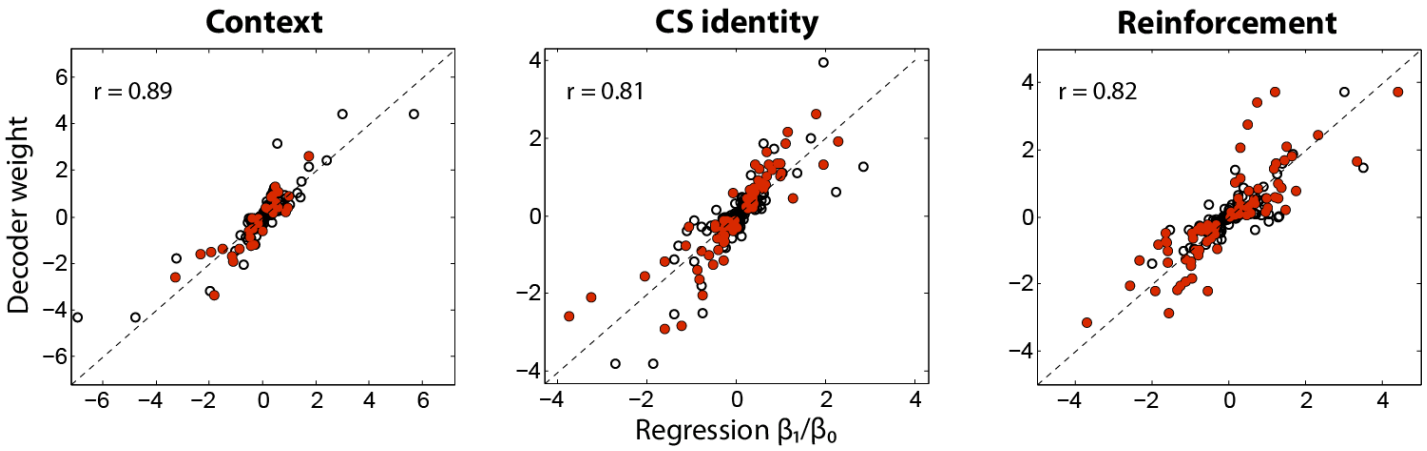

Figure S3. Weights assigned to neurons by the decoding algorithm correlate with regression coefficients from analysis of single neurons (related to Figure 5).
Relationship between the weight assigned to each neuron by the decoding algorithm and the coefficient found by fitting its firing rate with the simple regression model $FR = \beta_0 + \beta_1 X$, where X is the independent variable. Left: X = Context; Center: X= CS identity; Right: X = Reinforcement expectation. The weight of each cell (y-axis), as determined by the linear decoder is plotted against the value of β_1 normalized by the intercept β_0 (x-axis). Both axes are in units of z-score. Red dots indicate cells for which the linear regression fit was significant. Firing rates were taken during the 500-ms time window considered in Figures 6A-B, 7 and 8 (0.1 to 0.6 sec relative to CS onset). The correlation coefficient r between the weights and values of β_1/β_0 is indicated in each panel.

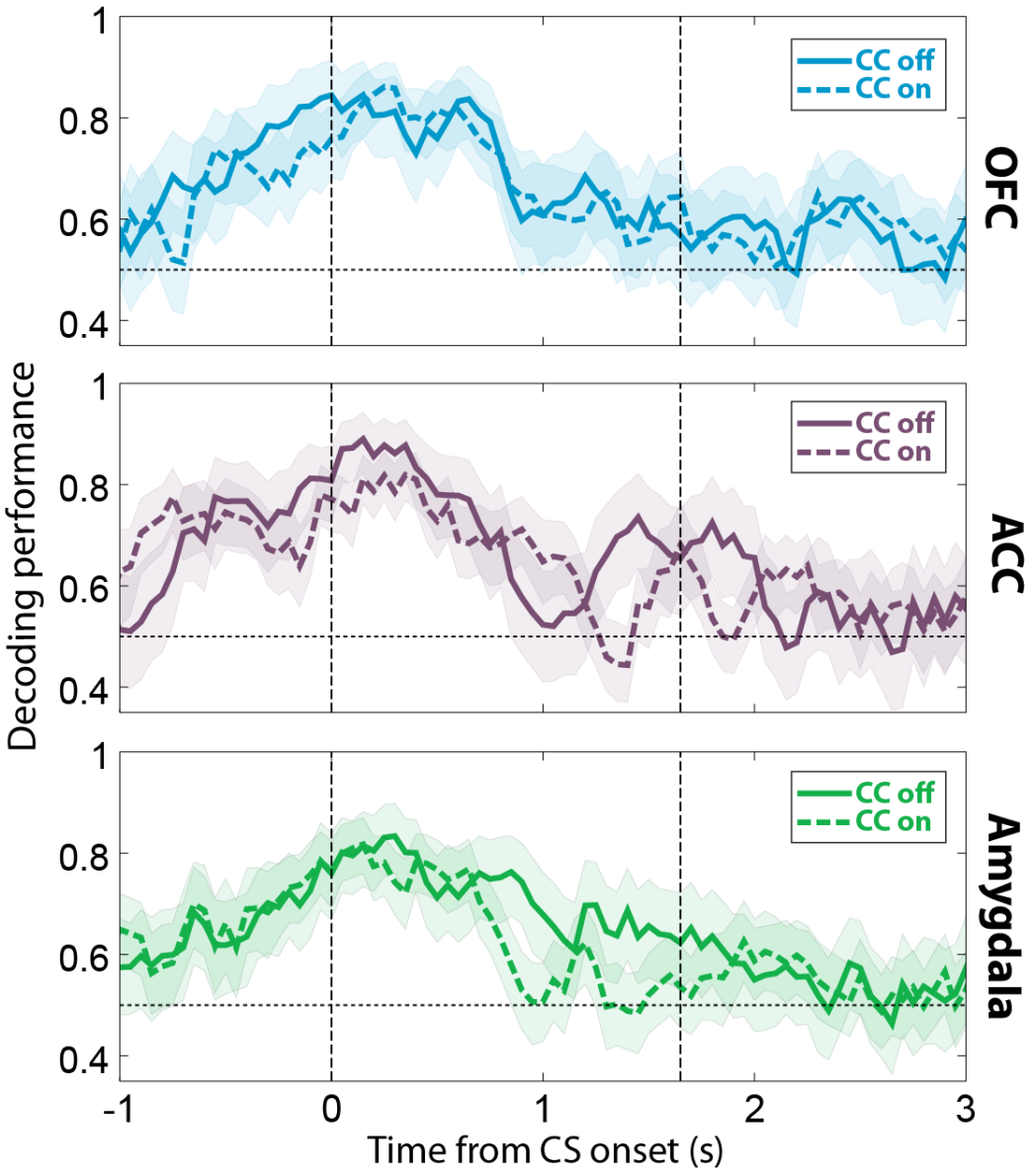

Figure S4: The presence of the contextual cue had little effect on the neural representation of context (related to Figure 5A). The linear decoder was trained to decode context on a subset of trials that did not contain a contextual cue. The decoder was then tested on held-out trials that either did (dashed curve) or did not (solid curve) contain a contextual cue. The decoding of context was very similar in all 3 brain areas regardless of whether a contextual cue was presented.

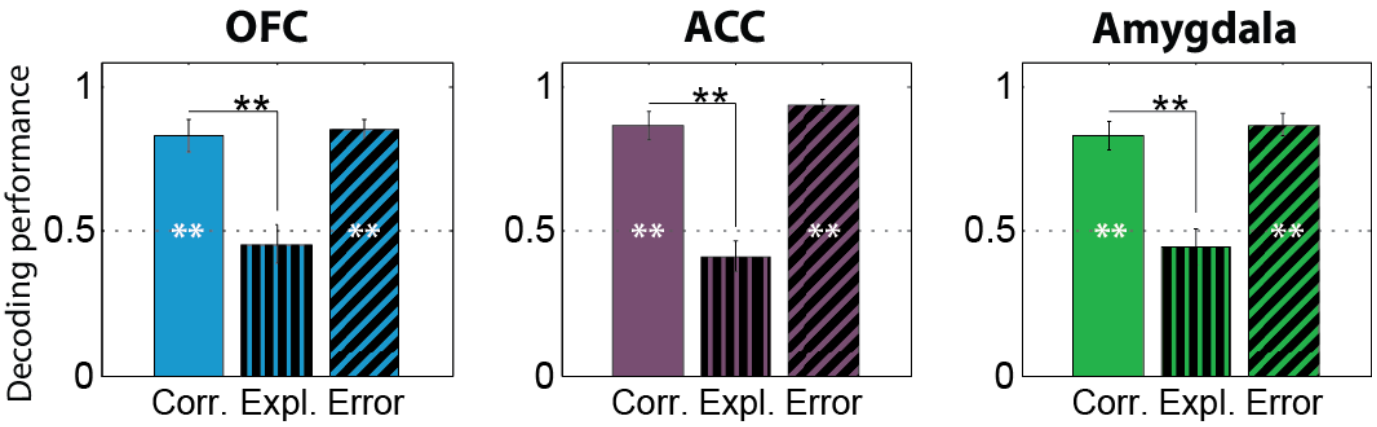

Figure S5: Decoding performance for the context signal before CS presentation (related to Figure 7B).

Neural activity from -0.4 to 0.1s relative to CS onset was analyzed to determine decoding accuracy for context on correct, exploration and error trials**.** Grey asterisks inside bars indicate a performance significantly different from chance level (0.5). Black asterisks indicate significant differences in performance between conditions. * $p<0.05$, ** $p<0.01$, bootstrap.

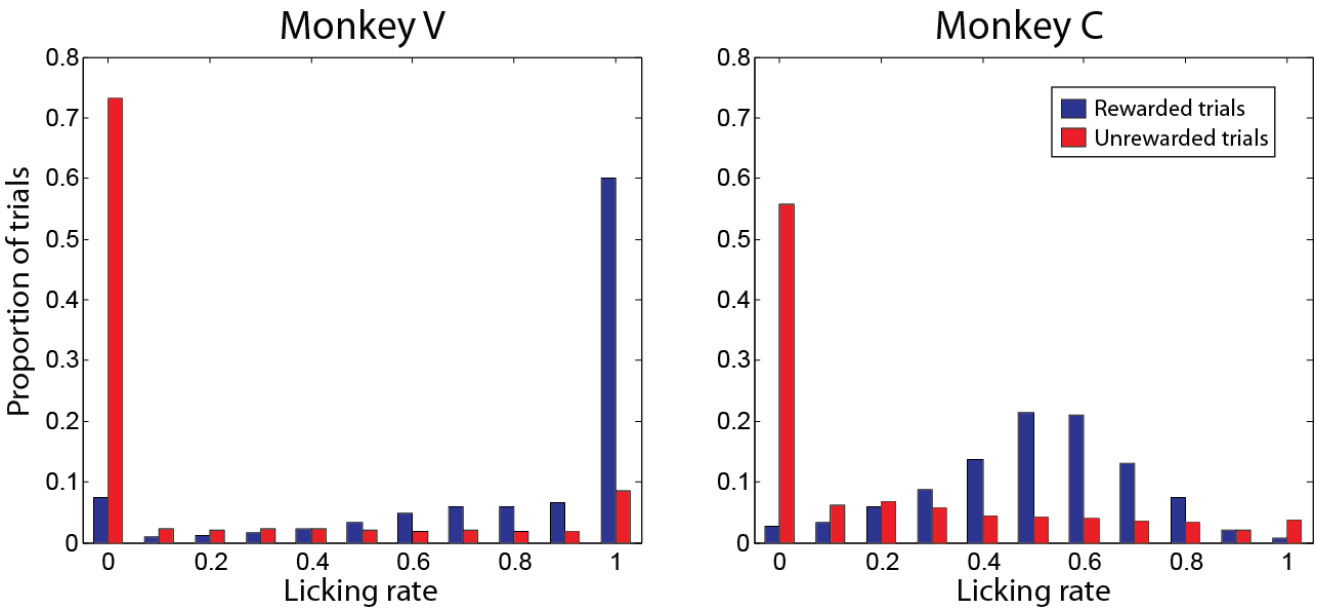

Figure S6. Licking patterns of the two monkeys (related to Experimental Procedures).

Distribution of licking rates of all trials during rewarded and non-rewarded trials for each monkey. The licking rate falls into one of two well separated distributions of values according to whether the trial is rewarded or not. This separation is the basis for "binarizing" the licking rate into "licks" and "no licks". The first 5 trials of each block are excluded due to the very different licking strategy adopted by monkeys during the exploration phase.

	Exploration trials		Error trials	
	Rewarded	Non-rew.	Rewarded	Non-rew.
Monkey V	5.2%	3.3%	5.6%	7.5%
	8.5%		13.1%	
Monkey C	5.8%	3.3%	6.4%	8.4%
	9.1%		14.8%	
Both monkeys	5.5%	3.2%	5.9%	7.9%
	8.7%		13.8%	

Table S1. Percentage of exploration phase and error trials for each monkey (related to Experimental Procedures).
Trials were further divided into rewarded and non-rewarded trials.

Supplemental Experimental Procedures

Animals, surgical procedures and training

Two rhesus monkeys (*Macaca mulatta*, one female, 5 kg, one male, 10 kg) were used in these experiments. Each animal was surgically implanted with a plastic head post and a single recording chamber giving access to both the prefrontal cortex and the amygdala. Surgery was conducted using aseptic techniques under general anesthesia (isoflurane), and analgesics and antibiotics were provided during postsurgical recovery. Structural magnetic resonance imaging (MRI) was used to guide the positioning of the recording chamber prior to surgery and to verify its position after surgery. Electrode placements for recordings were planned from MRI images using Brainsight 2 software (Rogue Research). After surgical implantation of the head post and the recording chamber, animals were trained on the task until their anticipatory licking behavior consistently reflected the CS-US associations for each context across multiple context switches within an experimental session. Training took 8 weeks for monkey C and 15 weeks for monkey V.

Behavioral task

Monkeys performed a serial-reversal trace-conditioning task in which they initiated a trial by looking at a central fixation point for 1 sec. While maintaining fixation, they were then presented one of two possible conditioned stimuli (CS) for 0.35 sec (monkey V) or for 0.15 sec (monkey C). A shorter CS presentation was used for monkey C to prevent systematic fixation breaks upon seeing the non-rewarded CS. The CSs consisted of abstract images (fractal patterns) that were unique to each experiment. After CS offset,

monkeys waited a trace period of 1.5 sec during which fixation was no longer required. Then, monkeys received an unconditioned stimulus (US) consisting of either a liquid reward (rewarded trials) or nothing (non-rewarded trials) depending upon which CS was presented. Rewarded and non-rewarded trials followed a pseudo-random schedule ensuring that 3 rewarded trials and 3 non-rewarded trials were experienced on every 6-trial set. The association between CS and US depended on the context in which the trial occurred. In Context 1, CS1 was paired with reward and CS2 was paired with no-reward, while in Context 2 the associations were reversed. Each context was maintained for a whole block of 30, 40 or 50 trials (with equal probabilities) for monkey V and of 20, 30 or 40 trials for monkey C. Context 1 and Context 2 blocks were alternated many times (ranging from 12 to 30) during an experiment. In addition, a contextual cue indicating the context in effect was randomly shown on 40 percent of the trials. The contextual cue consisted of a color frame at the periphery of the screen, yellow for Context 1 and blue for Context 2, and was displayed from the fixation point onset until US onset. The contextual cue was always shown on the first trial of a new context block, and in total occurred randomly on 40% of the trials. The inter-trial interval between the end of the US delivery and the next fixation point onset was 3 sec.

Behavioral measures

The reward delivery spout was placed a few millimeters away from the animal's mouth, so monkeys needed to lick at the spout before reward delivery to avoid missing it. Anticipatory licking behavior was measured by detecting the interruption of an infrared laser beam passing between the monkey's lips and the reward delivery spout. Licking rate

was defined as the proportion of time spent licking during the 500 ms immediately preceding the onset of the US (reward or no reward) and thus ranged between 0 and 1. Identifying exploration phases, error trials and the "behavioral switch" (Fig. 2B) involved applying a threshold to the licking rate in order to separate trials with high licking from trials with low licking. This procedure was justified by the fact we observed that licking rates were close to a bimodal variable, with two levels of licking (low and high) typically corresponding to the two reinforcement outcomes (reward and no reward, respectively) (Fig. S6 and Fig. 1B). We applied a local threshold that could change across an experiment to account for possible animal satiation or demotivation that could vary between the beginning and the end of the session. On each trial, the threshold was defined as the mean licking rate for the three blocks (the previous, the current and the next) around the trial. This approach relies on the fact that rewarded and non-rewarded trials are equally frequent within a block. However, we excluded from this average the first 5 trials of each of the three blocks due to the high values of the licking rate on non-rewarded trials (in addition to during rewarded trials) that characterized the exploration phase.

Although we chose to use a local threshold for categorizing licking levels, 95% of trials are classified identically when applying a unique threshold for all sessions and both monkeys. We also compared our local threshold method to the more optimal approach of fitting a logistic model to the licking rate in order to predict whether a trial is rewarded or not. The two methods achieved very similar results.

With the licking rate "binarized" into below- and above-threshold trials, we defined the exploration phase as the set of trials at the beginning of the block that ends on the trial preceding the first non-rewarded trial in which the monkey's licking rate was below threshold (see Table S1 for proportions of error and exploration phase trials). We defined error trials as the ones after the exploration phase in a given block in which the binary licking rate (a reflection of the monkey's reward expectation) did not match the reinforcement delivered at the end of the trial (*i.e.* licking in a non-rewarded trial or not licking in a rewarded trial). Because the transition out of the exploration phase is not necessarily immediately followed by a non-rewarded trial, it is likely that the real exploratory period (the one corresponding to a mental state of the animal) is effectively shorter than the one identified by our method. Given the statistics of the sequence of rewarded and unrewarded trials in the task, we may be overestimating the length of the real exploration phase by an average of 0.66 trials. Error trials happening during the exploration phase are only counted as "Exploration trials" (not "Error trials"). "Correct trials" refer to non-error trials outside of exploration phases. Therefore, "Correct", "Exploration" and "Error" trials are non-overlapping sets.

Population decoding

Pseudo-simultaneous population response vectors:

The decoding algorithm was based on a population decoder trained on pseudo-simultaneous population response vectors (Meyers et al., 2008). The components of these vectors corresponded to the spike counts of the recorded neurons in specific time bins computed in the following way.

The activity of each neuron in every trial was aligned to CS onset, the event that in the analysis defines time zero. Within each trial we then computed the spike count over time bins of 250 ms (Fig. 5A-B), 50 ms (Fig. 5C) or 500 ms (Figs. 6, 7, 8) that we displaced in steps of 50 ms (Fig. 5A-B) or 5 ms (Fig. 5C). All analyses were then carried out independently in every time bin as explained in the following.

For every neuron and every time bin, we z-scored the distribution of spike counts across all trials and all conditions (the mean and standard deviations used were always computed on the training data). Given a task condition *c* (for instance, CS1 is shown) and a time bin *t* (for instance, between 0 and 250 ms after CS presentation) we generated pseudo-simultaneous population response vectors by sampling, for every neuron *i*, the z-scored spike count in a trial in condition *c*, that we indicate by $n^c_i(t)$. This procedure resulted in the single trial population response vector $n^c(t) = (n^c_1(t), n^c_2(t), \ldots, n^c_N(t))$, where *N* is the number of recorded neurons in the area under consideration.

The visual stimuli used as the two CSs were fractal images replaced in every experimental session. By working with pseudo-simultaneous population response vectors composed of the activity of neurons recorded in different sessions, we assume that the neural activity patterns evoked by the presentation of any CS are statistically equivalent. Specifically, by combining multiple recording sessions we implicitly assume that the visual stimuli are equivalent and that, had all neurons been recorded simultaneously with the same two stimuli, we would have obtained similar responses to the ones in this study. Combining the activity recorded in separate sessions also neglects possible effects of

correlated firing rate variability across the neuronal population that could influence the decoding of the response vectors (Abbott and Dayan, 1999).

Analogous to the analysis by Britten *et al.* (Britten et al., 1992), we also sample the same neuron twice, after switching the labels of the two features being decoded. For instance, for every neuron with a particular response profile to CS1 and CS2, we introduce an "anti-neuron" that has that same response profile but to CS2 and CS1 instead. This effectively allows us to train and test our decoding algorithm on larger population response vectors and increase our statistical power. However, all statistically significant results reported remain valid without the inclusion of "anti-neurons".

In order to meaningfully compare the performance of the population decoder across the different brain areas, we used the same number of neurons for all areas, randomly discarding excess neurons (at every iteration of the training and testing procedures, we discarded a random set of excess neurons). We also discarded neurons for which we had fewer than 10 trials of "correct" behavior in each condition (*i.e.* in each context-CS combination). In total (including "anti-neurons"), we were left with 286 cells in every one of the three brain areas.

Training and testing the population decoder:

In the task under consideration, every trial was indexed by one of 4 conditions given by the combination of CS identity (CS1 or CS2) and the reinforcement outcome associated with the presented stimulus (either rewarded or non-rewarded). In all decoding analyses

we discarded trials where the contextual cue was presented, so as to discard contextual-cue selectivity as a possible explanation for apparent context selectivity. We also distinguish three behavioral states of the monkey: correct behavior, exploration phase and errors (as defined in the Behavioral measures section).

All decoding analyses consisted in training a population decoder to discriminate between population response vectors belonging to two distinct classes that corresponded to two sets of experimental conditions, and then testing the performance of the trained decoder in discriminating between the two classes. This was done on each time bin independently. Training and testing were performed on distinct sets of population response vectors, obtained by sampling from a training set of trials or from a testing set of trials, respectively.

We trained the decoder to discriminate between conditions where the CS was either one of the two possible images (decoding of CS identity), between conditions where the CS was either rewarded or not (decoding of reinforcement expectation), and between conditions where the CS-US associations corresponded to either one of the two contexts (decoding of context). We then tested its performance at discriminating between the same two conditions as during training, but on a different set of trials. The training of the decoder was always done on trials where the licking behavior was "correct". The testing of the decoder performance, on the other hand, was done on correct, error and exploration trials. This allowed us to assess how the neural patterns of activity differ between the state in which the monkey's behavioral prediction is correct and the states in which the

monkey is engaged in exploratory behavior or fails to predict the reinforcement accurately.

On average, our data set contained ~130 trials per neuron for each of the 4 conditions. For a given neuron and a given condition, we used around 80% of the correct trials for training the decoder and the rest for testing. We chose 80% because it resulted in a testing set of correct trials that was comparable in number (~15% of the total) to the number of trials in the testing set for the other two categories (exploration and error). Population response vectors for training and testing were generated by randomly sampling (with replacement) these trial pools for each neuron 1,000 times.

The training procedure consists in finding a threshold and a set of weights (one per neuron) that maximally discriminate between the population activity evoked in distinct classes of conditions. In order to do that we first obtained trial-averaged firing rate activity patterns corresponding to all task conditions. In other words, for every time bin t and every condition c we generated a population vector $v^c(t)$ by averaging the recorded spike counts across trials belonging to the training set of trials: $v^c(t) = \langle n^c(t) \rangle$, where the triangular brackets denote a trial average across the training set. For any given time bin t, we then find the weights that maximally discriminate the firing rate patterns of the conditions that belong to the first class from those of the conditions in the second class.

This maximization procedure can be formulated as a constrained quadratic programming problem (see (Anlauf and Biehl, 1989; Barak and Rigotti, 2011)), an optimization method

that is used to improve generalization performance in the Structural Risk Minimization and Support Vector Machines literature (Cortes and Vapnik, 1995) and is equivalent to a maximal margin perceptron algorithm (Barak and Rigotti, 2011; Krauth and Mezard, 1987; Rosenblatt, 1962).

To slightly increase the generalization performance of our decoder, we used a simple model averaging method known as "bootstrap resampling" or "bagging" (Breiman, 1996). For all decoding analyses, we used a 25-fold bagging.

Statistical significance and estimate of confidence intervals:

On every iteration of the training and testing procedure, we partitioned the trials into training and testing sets by random sampling without replacement. The average decoding performance, and its statistical significance and confidence intervals were estimated by repeating the partitioning procedure either 1,000 times (Fig. 5A, B) or 10,000 times (Figs. 5C, 6, 7, 8) (Golland et al., 2005). The confidence intervals shown throughout this work correspond to the 95th percentiles of the distributions obtained by repeated partitioning. The significance of the decoding performance was determined by the percentage of partitions yielding a performance above 0.5 (chance level). Similarly, significant decreases in performance (Fig. 7) were defined by the percentage (95% or 99%) of partitions yielding a performance difference above 0.

In order to compare the timing of the rise of the CS identity signal to that of the reinforcement expectation signal (Fig. 5C), we built a null population of the decoding

performance for each signal by shuffling the labels of the two conditions (CS1 and CS2 in the case of the CS identity signal; reward and no-reward in the case of the reinforcement expectation) before training and testing the decoder. Then, for each partition of the unshuffled trial pool, we identified the first time bin for which the decoding performance rose above the 95 percentile of the null population and remained above this level for the 10 subsequent bins. By doing this, we obtained two populations of "first significant time bins", one for each signal. The CS identity signal was said to significantly precede the reinforcement expectation signal if a randomly picked sample from the CS population preceded a randomly picked sample from the reinforcement population more than 95% of times. Our results (Fig. 5C) were not affected by changing the criterion used for defining the first significant time bin; requiring 5 or 20 consecutive bins instead of 10 yields the same results.

Definition and quantification of encoding signal (Fig. 8):

We quantified the signal collectively encoded in the neural population regarding a particular feature of the task (e.g. "Context") by simply summing the contribution of each neuron to the decoding performance in discriminating the feature's values (e.g. "Context 1" vs "Context 2"). This is simply done by weighting the preference (difference in trial average activity) of each neuron by the weight that the decoder attributes to each neuron as a consequence of the training procedure. For example, to quantify the population encoding signal related to "context" in a given time bin, we train the decoder to discriminate between single trial population response vectors sampled during "Context 1" and "Context 2" trials (as previously explained). This will result for every neuron *i* in a

weight w_i (obtained by averaging 25 weights obtained by the 25-fold bagging procedure) and a threshold term b. Given a trial population response vector $n^c(t) = (n^c_1(t), n^c_2(t), \ldots, n^c_N(t))$ sampled during condition c (e.g. CS 1 in Context 1), these weights allow us to compute a trial-to-trial total contribution to decoding condition c by computing the following weighted sum over spike counts:

$h^c(t) = w_1 n^c_1(t) + w_2 n^c_2(t) + \ldots + w_N n^c_N(t) - b$. Notice again that this is simply a weighted sum of recorded spike counts in time bin t. Notice also that since the decoder training is performed on z-scored spike counts, h^c is expressed in units of standard deviation of the spike counts across trials.

We define the *encoding signal* as the average difference of $h^c(t)$ in conditions c' that have to be classified as belonging to one class by the decoder (e.g. "Context 1") and conditions c'' that belong to the other class (e.g. "Context 2"), where the average is over: $S = \langle h^{c'}(t) \rangle_+ - \langle h^{c''}(t) \rangle_-$, where $\langle \rangle_+$ and $\langle \rangle_-$ indicate sample average over trials sampled in conditions c' and conditions c'', respectively. According to its definition, the encoding signal quantifies the (weighted) average selectivity across the population with regards to the two conditions. The encoding signal can thereby be used as a global measure of the strength of the signal that our particular decoder can use in order to discriminate between the two different contexts (or to discriminate between two conditions for any other variable).

Notice that an equivalent definition of S is to simply compute the preference of every neuron (i.e. average firing rate in "Context 1" minus average firing rate in "Context 2") and weighting this quantity by corresponding decoder weights. However, our previous

definition based on trial sampling of h^c has the utility of straight-forwardly extending to a procedure to compute its variability. In fact, analogously to how we compute the encoding signal *S,* we can now compute the corresponding *encoding variance*: $N = var(h^{c'}(t))_+ + var(h^{c''}(t))_-$, where $var(h^{c'}(t))_+$ indicates the sample variance of $h^{c'}(t)$ in conditions *c'* (and analogously for *c''* conditions). Encoding variance can be used as an inverse measure of reliability of the encoding signal, since it quantifies the trial-to-trial noise of the signal read out by our decoder.

In all cases, *S* and *N* are computed by sampling trials in the test set of trials, i.e. recorded trials that were not used to train the decoder and obtain the weights w_i. The variability of both quantities is estimated in the same way we estimate decoding accuracy confidence intervals, that is by repeating the training and sampling procedure over test trials after having a different random partition between training and testing sets of trials.

The methods just described enable us to quantify encoding signal and encoding variance which in turn allows us to determine whether a decrease in decoding performance between two experimental conditions is due to a decrease in signal (that quantifies the neural population selectivity or to an increase in variance (that quantifies the spike count variability across all neurons).

Comparison with alternative population decoder algorithms:

The results presented using the linear decoder just described differ little from those obtained using a Fisher discriminant method (Fisher, 1936). The Fisher discriminant

method is equivalent to a multidimensional ROC analysis on the firing rates of all neurons simultaneously and is therefore very directly related to the activity of individual neurons. The decoding accuracy simply represents differences in mean spike counts between different conditions, averaged over the population by weighting each neuron with the strength of its selectivity.

The linear decoder approach we took for identifying task-related signals in the neural activity is tightly linked to more conventional approaches such as regression analyses of single unit data. More specifically, given a trained decoder we can compute an average weight that the decoder attributes to each neuron by averaging the readout weights across bagging folds. Figure S3 shows that the resulting average weight that the trained decoder attributes to each neuron is strongly correlated with the coefficient found when fitting the firing rate of each individual neuron with a simple linear regression model (see Fig. S3 legend), so that the weights reflect the selectivity of individual neurons. The advantage of the decoder method over linear regression, however, is that it offers a principled way to combine all neurons and evaluate the selectivity of the whole population. In particular, this measure does not exclude neurons whose firing rate is not significantly fit by the regression model. Instead, the decoder technique translates population selectivity into a unique cross-validated quantity, which corresponds to an estimate of the probability of correctly predicting the variable being decoded in future trials. Nonetheless, the weighting of each neuron is strongly correlated with its selectivity.

We use the performance of a population decoder as a measure of how strongly different task-related signals are encoded in the brain. However, it is noteworthy that our approach can only give a lower bound on the total information encoded. This is a fundamental limitation of any parametric or model-driven analysis method. An analysis that relies on single-cell activity (such as t-test or ROC) is typically poor at quantifying single-trial decoding accuracy and population-wide selectivity. On the other hand, a decoder that is more sophisticated than our linear classifier might be able to achieve higher decoding performance, but at the expense of requiring a level of computational complexity that goes beyond what could be reasonably ascribed to local neural processing. Our choice of decoder is meant to strike a balance between statistical power and physiological relevance. Another practical benefit of using a linear decoder is that it can be thought of as a principled way of combining individual single-cell selectivity analyses across neurons.

Supplemental References

Abbott, L.F., and Dayan, P. (1999). The effect of correlated variability on the accuracy of a population code. Neural Comput. *11*, 91–101.

Anlauf, J.K., and Biehl, M. (1989). The AdaTron: An Adaptive Perceptron Algorithm. Europhys. Lett. EPL *10*, 687–692.

Barak, O., and Rigotti, M. (2011). A Simple Derivation of a Bound on the Perceptron Margin Using Singular Value Decomposition. Neural Comput. *23*, 1935–1943.

Breiman, L. (1996). Bagging predictors. Mach. Learn. *23*, 123–140.

Britten, K.H., Shadlen, M.N., Newsome, W.T., and Movshon, J.A. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. J. Neurosci. Off. J. Soc. Neurosci. *12*, 4745–4765.

Cortes, C., and Vapnik, V. (1995). Support-Vector Networks. Mach. Learn. *20*, 273–297.

Fisher, R.A. (1936). The Use of Multiple Measurements in Taxonomic Problems. Ann. Eugen. *7*, 179–188.

Golland, P., Liang, F., Mukherjee, S., and Panchenko, D. (2005). Permutation Tests for Classification. In Learning Theory, P. Auer, and R. Meir, eds. (Springer Berlin Heidelberg), pp. 501–515.

Krauth, W., and Mezard, M. (1987). Learning algorithms with optimal stability in neural networks. J. Phys. Math. Gen. *20*, L745–L752.

Meyers, E.M., Freedman, D.J., Kreiman, G., Miller, E.K., and Poggio, T. (2008). Dynamic population coding of category information in inferior temporal and prefrontal cortex. J. Neurophysiol. *100*, 1407–1419.

Rosenblatt, F. (1962). Principles of neurodynamics: perceptrons and the theory of brain mechanisms (Spartan Books).