SUPPORTING ONLINE MATERIAL

Materials and Methods

Animals. Two adult female *Macaca fascicularis* monkeys were maintained under the Swiss Animal Protection Law and the supervision of the Fribourg Cantonal Veterinary Office.

Experimental Design. A classical conditioning procedure was performed with visual stimuli presented on a computer monitor. The head was fixed in place in front of the monitor. The present data were obtained from five separately trained sets of visual stimuli, two in monkey A and three in monkey B, each presented with a distinct background on the monitor. In each set, five stimuli were presented in random alternation. Pictures were chosen to have similar physical salience but to be easily discriminated. To aid discrimination, each stimulus was presented at a unique location. Stimuli of 2 s duration were followed by a fixed amount of liquid (0.15 – 0.20 ml of diluted, raspberry-flavored syrup) delivered from a spout immediately in front of the animals mouth. Licking behavior was monitored with an infrared detector.

Each stimulus was associated with a specific probability of reward. To prevent large, random fluctuations, the program specified that the pre-assigned probabilities were precise after a block of eight consecutive trials of a specific trial type. After those eight trials the counter was reset so that the next trial occurred with precisely the stated probability. The counter was also reset if the experimenter interrupted the recording for more than a few seconds. All trials were presented with an inter-stimulus interval that averaged 9 s, consisting of a fixed 4 s plus an interval determined by a Poisson process with a rate constant of 0.02 per 100 ms. Unpredicted rewards were given in a separate block of trials with the same intertrial interval, and thus occurred with a rate constant of $p = 0.02$ per 100 ms. The

relevant probability for dopamine neurons is presumably low for these 'unpredicted' rewards but is unknown, as we don't know the unit of time for which predictions are made.

Task training consisted of 100–200 trials of each stimulus per day, five days per week, for about five weeks. Recordings began after at least five days of training and emergence of discriminative conditioned licking responses.

In experiments concerning reward magnitude, the small, medium, and large rewards were 0.05 ml in 40 ms, 0.15 ml in 100 ms, and 0.50 ml in 240 ms, respectively. Anticipatory licking responses preceded all reward magnitudes. Thus even the small reward was a sufficiently strong reinforcer for conditioning.

Histology. Recording sites were marked with small electrolytic lesions and reconstructed from 40 μm thick, stereotaxically oriented coronal brain sections, stained with cresyl violet or antibodies to tyrosine hydroxylase. No significant correlations were found between neuronal position and responses. In all cases, the data was pooled. Hisological reconstructions of the position of recorded neurons are shown in figure S2.

Electrophysiological Recordings. Single unit recordings were performed as previously described (*S1*). An attempt was made to record a representative sample of the entire population of dopamine neurons; thus the presence of phasic or sustained responses to conditioned stimuli or reward was not a criterion for selecting neurons to record. Rather, dopamine neurons were identified solely by their discharge characteristics, including long waveforms (1.5 – 5.0 ms) and slow, fairly regular basal firing rates (0.1 – 8.0 Hz). Prior studies in primates have shown that ventral midbrain neurons having these properties are antidromically activated by stimulation of the striatum (*S2*), and their firing is suppressed by

systemic administration of dopamine D2 agonists (*S3*), thus fitting long established criteria for the identification of ventral midbrain dopamine neurons.

Data analysis. Typically, at least 15 trials of each trial type were performed per cell; the minimum accepted for analysis was 7. Responses were measured in standard windows and compared to the control period (1 s before stimulus onset) to calculate the percent change in spike rate. The standard windows for phasic stimulus and reward responses were chosen to cover about 60% of the duration of the response, centered on the average maximum. Standard windows varied depending on the phasic response being measured and differed slightly between monkeys; they were fixed across trial types and across neurons. The latency and duration (milliseconds) of standard windows in monkeys A and B, respectively, were 90, 90 and 110, 130 following conditioned stimulus onset, 120, 100 and 120, 100 following reward onset, and 150, 100 and 150, 100 following no reward, conditioned stimulus off. For sustained activation, the standard window was the 500 ms before the potential reward or neutral stimulus.

The calculation of the 95% confidence intervals shown in figure 2 was done as recommended for simple approximation by Iglewicz (S4), multiplying the appropriate t value by the interquartile range and dividing by 1.075 times the square root of the number of observations.

Statistical analyses of the sustained activations shown in figure 3C were performed as follows. For the two data sets in which five probabilities were tested, the percent change in activity in the 500 ms before reward was ranked across the five probabilities for each neuron. The ranked values were then subjected to a Kruskal-Wallis test with three groups defined by the degree of uncertainty (p=0.0 and 1.0; p=0.25 and 0.75; p=0.5). The initial ranking of the data points accounted for the paired nature of the data from each cell, while the Kruskal-

Wallis test is appropriate for multiple comparisons of nonparametric data. For the one data set with three probabilities and two levels of uncertainty, the responses were ranked and then tested with a Mann-Whitney Test. For the data set with only two probabilities examined (p=0.5 and 1.0), the unranked data was subjected to a Wilcoxon Signed Rank test. The data concerning reward magnitude (Fig. 4) were analyzed in an analogous manner, with data sets having two or three levels of magnitude. The data shown in Fig. 4B revealed a significant effect ($P<0.01$) when analyzed by either Wilcoxon tests, or Kruskal-Wallis or Mann-Whitney tests after ranking.

For the correlation analysis carried out for figures 3D and S4, correlation coefficients (r) were derived from a partial correlation matrix of activity observed in each cell during four periods: the control period and standard windows (see above) for sustained activation (at p=0.5), phasic reward (at p=0), and phasic conditioned stimulus (at p=1.0) responses.

Additional Data

Analysis of conditioned responses on rewarded vs. unrewarded trials

The question arises as to whether or not the animals predictions varied on a trial by trial basis dependent on the probability schedule. As discussed in the methods, the reward probabilities were not truly random, but structured so that the actual probabilities matched the pre-assigned probabilities after a block of 8 consecutive trials of a given trial type. Because there were as many as five trial types (one for each conditioned stimulus) randomly interleaved, it would appear difficult to count rewarded vs. unrewarded trials for a given trial type. Nonetheless, with extensive experience the animal (or the neurons) might learn the negative correlation between consecutive trials of a given trial type ("since that stimulus was followed by reward last time, it is less likely to be rewarded this time"). If this occurred, it would reduce the average amount of uncertainty at all intermediate probabilities, and could

cause a significant skew in the measured probability functions. Another possibility, not requiring such sophisticated cognition, is that the animal bases its predictions simply on a weighted average of past trials. In this case, the animal's prediction would assume a positive correlation between consecutive trials ("if this stimulus was rewarded last time, it probably will be this time"). The simplest way to assess the extent to which either of these processes might have influenced reward expectations is to compare behavioral and physiological responses on rewarded vs. unrewarded trials at intermediate probabilities. In the first scenario outlined above, in which the animal has learned something about the structure of the probability schedule, one would expect behavioral and neuronal responses to the conditioned stimulus to correspond to higher reward probabilities on rewarded trials as compared to unrewarded trials. In the second scenario, if the animal simply adjusts its predictions based on a weighted average of past trials (with sufficiently high weight given to the most recent trials), then one would expect behavioral and neuronal responses to the conditioned stimulus to correspond to lower reward probabilities on rewarded trials as compared to unrewarded trials. Figures S1A and S1B show that both licking behavior and neuronal responses to conditioned stimuli failed to discriminate rewarded from unrewarded trials. This suggests that neither the animals nor the neurons learned the probability schedule to a significant extent, and that their predictions were probably based on a weighted average of more than just the last few trials.

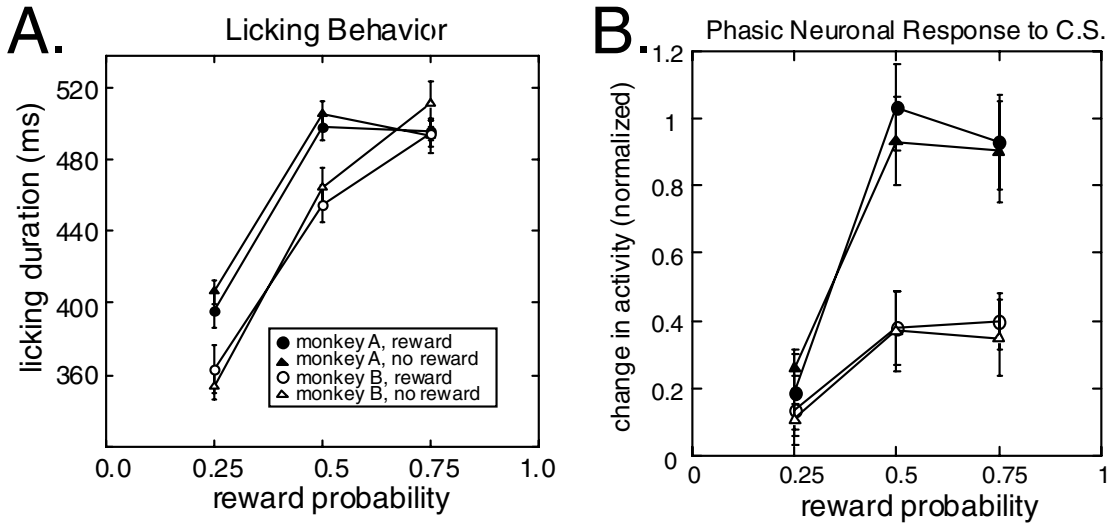

Fig. S1. Conditioned behavioral and neuronal responses failed to discriminate rewarded from unrewarded trials, though both responses were sensitive to reward probability. **A.** The data shown is the same as in figure 1, except rewarded and unrewarded trials have been analyzed separately. Conditioned licking responses are quantified as the duration of licking in the 2 s interval between stimulus onset and potential reward. Each point represents the mean (±s.e.m) duration of licking of 905 – 4966 trials. **B.** The data shown represents a subset of the data in figure 2E, now with rewarded and unrewarded trials analyzed separately. Responses were normalized in each neuron to the response (percent change in activity) following the conditioned stimulus predicting reward at $p = 1.0$, and the mean (±s.e.m.) of these values is shown. Only neurons showing greater than 50% increases in activity following onset of the stimulus with $p = 1.0$ were used in this analysis (n=27-36). By selecting neurons in this way, the data became more parametric; hence the standard error is used here but not in figure 2.

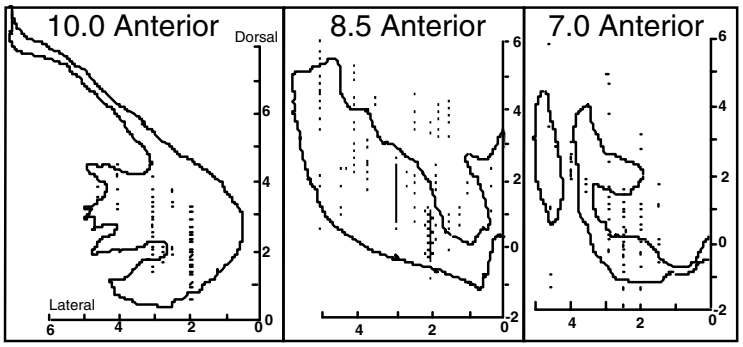

Fig. S2 The figure above displays histological reconstructions of the positions of recorded cells. Each outline represents the area of dense staining for the dopamine-synthesizing enzyme tyrosine hydroxylase in the ventral midbrain. The sections depicted were taken at 7.0, 9.5, and 10.0 mm anterior to the interaural line in monkey A. Neurons from both hemispheres in both monkeys are shown, each recorded within ± 0.5 mm anterior-posterior of the section displayed. All neurons included in this study are shown, except 22 neurons from monkey A that were at the level of 5.5 or 6.0. All neurons at 10.0 were from monkey B.

Parametric analysis of a subset of the data shown in figures 2 and 3

In the main text, the median of the entire population was used as the measure of responsiveness. This allowed an unbiased measure of the entire population, while being insensitive to the nonparametric nature of the data. The median is also relatively insensitive to outliers, which are produced inappropriately when normalizing to values that are negative or close to zero, as was done for figure 2. An alternative approach is to select responsive neurons, which makes the data more parametric and may provide a more sensitive measure of relative responses as a function of probability. The results of this analysis are shown in figures S3A and S3B for subsets of the neurons that contributed to figures 2C and 3C. The results are similar. The only potentially meaningful distinction is that in figure S3B, but

perhaps not in 3C, there is clearly more sustained activation at $p = 1.0$ than at $p = 0.0$ ($P<0.01$, Wilcoxan signed rank test). It is important to recognize that although analyzing only responsive neurons may give a more accurate measure, it could also lead to a skewed measure of the overall population response (if the neglected neurons have distinct properties and don't merely contribute random noise). For at least two of the three data sets shown in figure 3C, the medians of the whole populations were not different between $p = 0.0$ and $p = 1.0$. Nonetheless, figure S3B could indicate a meaningful asymmetry in the relationship of the sustained activation to probability (though no difference is apparent between probabilities of 0.25 and 0.75). An alternative explanation is that the sustained activation at $p = 1.0$ resulted from a context-dependent generalization effect of the uncertainty that was associated with the other stimuli which were present on alternating trials. If this is the case, there should be no sustained activation at $p = 1.0$, and no difference in activity between $p = 0.0$ and $p = 1.0$, in a context in which all stimuli predict reward ($p=1.0$) or no reward ($p=0.0$) with certainty. Such experiments were performed in 37 neurons in monkey A and 48 neurons in monkey B. These experiments used distinct picture sets, and none of these neurons were among those reported in the main text. The mean (±s.e.m.) activation in monkey A at $p = 0.0$ was 9.0 ± 5.7% and the median was 5%, while at $p = 1.0$ the mean was 0.4 ± 4.2% and the median was 0%. In monkey B, the mean activation at $p = 0.0$ was 0.0 ± 2.3% and the median was 0.0%, while at $p = 1.0$ the mean was –6.5 ± 4.9% and the median was –13%. In monkey B, the amount of activity was marginally but significantly less at $p = 1.0$ than at $p = 0.0$ ($P < 0.05$, Wilcoxan signed rank test). The same trend is apparent in monkey A, though this was not significant. Thus the discrepancy between $p = 0.0$ and $p = 1.0$ in figure S3B appears to arise either from the general context of uncertainty created by the frequent, interleaved presentation of stimuli predicting reward at intermediate probabilities, or from a skew introduced by the selection of highly responsive neurons.

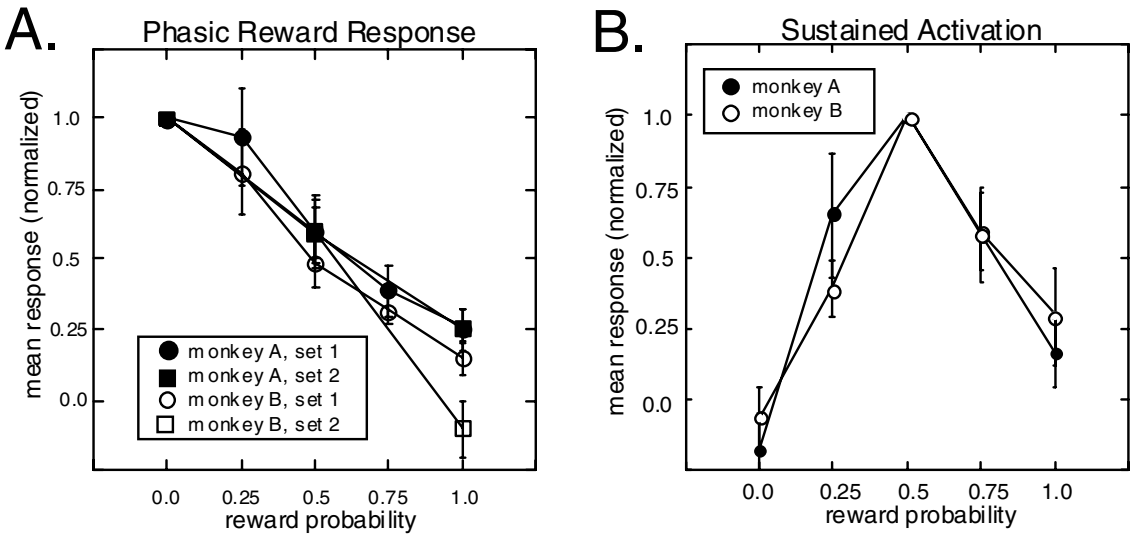

Fig. S3. The relationships of the phasic reward response and sustained activation to reward probability. Whereas figures 2C and 3C show median responses for the entire neuronal populations sampled, these figures show means (±s.e.m.) for selected groups of responsive neurons. **A.** The mean response was calculated following normalization within each neuron to the response to unpredicted reward (p=0.0). This figure includes a subset of neurons from figure 2C in which the phasic reward response at p = 0.0 exceeded 50% above basal activity. Each point represents the mean value for 26 –54 neurons. **B.** The mean sustained activation was calculated following normalization within each neuron to the response at p = 0.5. This figure is based on a subset of neurons from figure 3C in which the sustained activation at p = 0.5 exceeded 30% above basal activity, and sufficient data was obtained at all probabilities. The mean (±s.e.m.) increase in activity at p = 0.5 was 220 ± 102% in monkey A (n=16) and 88 ± 17% in monkey B (n=14). The activation at p = 1.0 appears to result either from a contextual generalization effect due to the uncertainty associated with the other stimuli, or to

a skew introduced by selecting highly responsive neurons, as discussed in the supplementary text above.

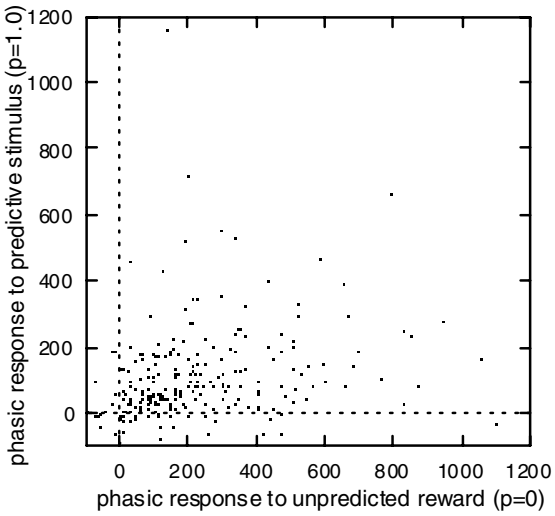

Fig. S4 The magnitude of the phasic activation to reward is correlated across neurons with the magnitude of the phasic activation to a conditioned stimulus (r=0.196, P<0.01, n=241). This is in contrast to figure 3D, which shows no correlation between the sustained activation and the phasic activation to reward. The conditions eliciting the largest average responses are shown (*p*=0 for reward, *p*=1.0 for conditioned stimulus). Each point represents a single dopamine neuron. Response values are given as percent change from basal activity. Five outliers are not shown. Correlation coefficients were not derived directly from the data shown, but rather from a partial correlation matrix of firing rates that took into account the covariance of each measure with basal firing rate.

References for supplementary online material.

S1. W. Schultz, P. Apicella, T. Ljungberg, *J. Neurosci.* **13**, 900 (1993).

S2. W. Schultz, *J Neurophysiol* **56**, 1439 (1986).

S3. W. Schultz, R. Romo, *J Neurophysiol* 57, 201 (1987).

S4. B. Iglewicz, in *Understanding Robust and Exploratory Data Analysis*, D.C. Hoaglin, F. Mosteller, J.W. Tukey (Wiley, New York, 1983), pp. 404-430.