Table 1. Amounts of endogenous GAs in wild type and *gid2* (ng per g of fresh weight).

	GA_{53}	GA_{44}	GA_{19}	GA_{20}	GA_1
Wild type					
Lot 1	5.6	2.6	20	0.5	0.3
Lot 2	4.0	2.9	28	0.3	0.5
gid2-1					
Lot 1	4.1	6.0	26	2.4	47
Lot 2	4.2	6.0	23	2.4	56

rylated SLR1 protein, we pretreated the wild type and *gid2-1* with uniconazol, an inhibitor of GA biosynthesis. We detected one faint radioactive band in uniconazol-pretreated wild type and this band disappeared after treatment with GA_3 (Fig. 3C, lanes 1 and 2). This supports our theory that the phosphorylated SLR1 protein is destabilized by bioactive GA. In contrast, we observed one strong radioactive band in *gid2-1* and GA_3 treatment increased its intensity (Fig. 3C, lanes 3 and 4). The mobility of the radioactive band corresponded to the upper band in *gid2-1*, and the intensity of the upper band observed by immunoblotting increased after treatment with GA (Fig. 3C, lanes 5 and 6). This GA-induced phosphorylation of SLR1 protein in *gid2-1* was gradually increased after GA_3 treatment (Fig. 3D). These results indicate that GA increases SLR1 phosphorylation and may lead to degradation of phosphorylated SLR1 in wild type but that degradation of the phosphorylated SLR1 in *gid2* is inhibited and consequently the protein is accumulated.

The fact that a loss of function in an F-box protein, GID2, causes accumulation of the SLR1 protein leads us to speculate that GA-dependent degradation of SLR1 protein is caused by the ubiquitin/26S proteasome pathway. To test this possibility, we examined the polyubiquitination of SLR1 protein in vivo by immunoblotting with antibody to ubiquitin (Ub). In wild type treated with a proteasome inhibitor, MG132, a low level of polyubiquitinated SLR1 was observed without GA treatment (Fig. 3E, lane 1), and GA treatment induced the accumulation of polyubiquitinated SLR1 protein (Fig 3E, lane2). In contrast, in *gid2-1*, we observed no ubiquitinated SLR1 with or without GA treatment (Fig. 3E, lanes 3 and 4). These results suggest that the SLR1 protein is degraded via the ubiquitin/26S proteasome pathway mediated by the SCF^{GID2} complex.

The F-box protein in the SCF complex functions as a receptor that selectively recruits target proteins into the complex to degrade these proteins through ubiquitination. This SCF-mediated signaling pathway is well conserved in yeast, mammals, and higher plants (*21–25*). According to recent advances in understanding SCF-mediated pathways in yeast and animals (*21–23*), modification of the target protein is a prerequisite for interaction between the target and F-box proteins, and phosphorylation is one of the most common types of modification of target proteins. Although there are no previous reports that phosphorylation of target proteins triggers SCF-mediated degradation in plants, our results indicate that GA-dependent phosphorylation of SLR1 triggers the ubiquitin-mediated degradation in a manner similar to the SCF-mediated pathway in yeast and animals.

References and Notes

1. P. J. Davies, *Plant Hormones* (Kluwer Academic, Dordrecht, Netherlands, 1995).
2. J. Peng *et al.*, *Genes Dev.* **11**, 3194 (1997).
3. A. L. Silverstone, C. N. Ciampaglio, T.-P. Sun, *Plant Cell* **10**, 155 (1998).
4. J. Peng *et al.*, *Nature* **400**, 256 (1999).
5. A. Ikeda *et al.*, *Plant Cell* **13**, 999 (2001).
6. P. M. Chandler, A. Marion-Poll, M. Ellis, F. Gubler, *Plant Physiol.* **129**, 181 (2002).
7. A. L. Silverstone *et al.*, *Plant Cell* **13**, 1555 (2001).
8. A. Dill, H.-S. Jung, T.-P. Sun, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 14162 (2001).
9. H. Itoh, M. Ueguchi-Tanaka, Y. Sato, M. Ashikari, M. Matsuoka, *Plant Cell* **14**, 57 (2002).
10. F. Gubler, P. M. Chandler, R. G. White, D. J. Llewellyn, J. V. Jacobsen, *Plant Physiol.* **129**, 191 (2002).
11. X. Fu *et al.*, *Plant Cell* **14**, 3191 (2002).
12. Materials and methods are available as supporting material on *Science* Online.
13. M. Ashikari, J. Wu, M. Yano, T. Sasaki, A. Yoshimura, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10284 (1999).
14. H. Itoh *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 8909 (2001).
15. M. Ashikari *et al.*, *Breed. Sci.* **52**, 143 (2002).
16. A. Sasaki *et al.*, data not shown.
17. C. M. Steber, S. E. Cooney, P. McCourt, *Genetics* **149**, 509 (1998).
18. E. T. Kipreos, M. Pagano, *Genome Biol.* **5**, 1 (2000).
19. R. J. Dashaies, *Annu. Rev. Cell Dev. Biol.* **15**, 435 (1999).
20. M. Yang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 11416 (1999).
21. F. N. Li, M. Jonston, *EMBO J.* **16**, 5629 (1997).
22. D. Skowyra, K. L. Craig, H. Tyers, S. J. Elledge, J. W. Harper, *Cell* **91**, 209 (1997).
23. J. T. Winston *et al.*, *Genes Dev.* **13**, 270 (1999).
24. W. M. Gray, S. Kepinski, D. Rouse, O. Leyser, M. Estelle, *Nature* **414**, 271 (2001).
25. L. Xu *et al.*, *Plant Cell* **14**, 1999 (2002).
26. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

27. We thank C. Steber for sharing SLY1 data prior to publication, A. Yoshimura for donating the rice dwarf mutant stock, and S. Hattori for excellent technical assistance. Supported by a Grant-in-Aid for the Center of Excellence, a Grant-in-Aid from the Program for the Promotion of Basic Research Activities for Innovative Bioscience (M.M.), the MAFF Rice Genome Project (M.A., M.M.), and a research fellowship from Japan Society for the Promotion of Science (H.I).

Supporting Online Material

www.sciencemag.org/cgi/content/full/299/5614/1896/DC1
Materials and Methods
Fig. S1
References

3 December 2002; accepted 22 January 2003

Discrete Coding of Reward Probability and Uncertainty by Dopamine Neurons

Christopher D. Fiorillo,* Philippe N. Tobler, Wolfram Schultz

Uncertainty is critical in the measure of information and in assessing the accuracy of predictions. It is determined by probability *P*, being maximal at *P* = 0.5 and decreasing at higher and lower probabilities. Using distinct stimuli to indicate the probability of reward, we found that the phasic activation of dopamine neurons varied monotonically across the full range of probabilities, supporting past claims that this response codes the discrepancy between predicted and actual reward. In contrast, a previously unobserved response covaried with uncertainty and consisted of a gradual increase in activity until the potential time of reward. The coding of uncertainty suggests a possible role for dopamine signals in attention-based learning and risk-taking behavior.

The brain continuously makes predictions and compares outcomes (or inputs) with those predictions (*1–4*). Predictions are fundamentally concerned with the probability that an event will occur within a specified time period. It is only through a rich representation of probabilities that an animal can infer the structure of its environment and form associations between correlated events (*4–7*). Substantial evidence indicates that dopamine neurons of the primate ventral midbrain code errors in the prediction of reward (*8–10*). In the simplified case in which reward magnitude and timing are held constant, prediction error is the discrepancy between the probability *P* with which reward is predicted and the actual outcome (reward or no reward). Thus, if dopamine neurons code reward prediction error, their activation after reward should decline monotonically as the

probability of reward increases. However, in varying probability across its full range (P = 0 to 1), a fundamentally distinct parameter is introduced. Uncertainty is maximal at P = 0.5 but absent at the two extremes (P = 0 and 1) and is critical in assessing the accuracy of a prediction. We examined the influence of reward probability and uncertainty on the activity of primate dopamine neurons.

Two monkeys were conditioned in a Pavlovian procedure with distinct visual stimuli indicating the probability (P = 0, 0.25, 0.5, 0.75, and 1.0) of liquid reward being delivered after a 2-s delay (*11*). Anticipatory licking responses during the interval between stimulus and reward increased with the probability of reward (Fig. 1), indicating that the animals discriminated the stimuli behaviorally. However, at none of the intermediate probabilities was there a difference in the amount of anticipatory licking between rewarded and unrewarded trials (fig. S1). This suggests that the expectation of reward did not fluctuate significantly on a trial-by-trial basis as a result of the monkey learning the reward schedule (*11*).

Dopamine neurons of ventral midbrain areas A8, A9, and A10 (fig. S2) were identified solely on the basis of previously described electrophysiological characteristics, particularly the long waveform of their impulses (1.5 to 5.0 ms) (*11*). The analyses presented here are for the entire population of dopamine neurons sampled, without selection for the presence of any event-related response. Dopamine neurons (n = 188) showed little or no response to fully predicted reward (P = 1.0), but they displayed the typical phasic activations (*8–10*) when reward was delivered with $P < 1.0$, even after extensive training (Fig. 2, A and B). The magnitude of the reward responses increased as probability decreased, as illustrated by linear regression analyses (correlation coefficient r^2 = 0.97, P = 0.002 and r^2 = 0.92, P = 0.01 in monkeys A and B, respectively) (Fig. 2C and fig. S3A) (*12*). Although dopamine neurons discriminated the full range of probabilities effectively as a population, in contrast to Fig. 2A, many single neurons appeared not to discriminate across the full range (*13*). For trials in which reward was predicted with intermediate probabilities (P = 0.25 to 0.75) but did not occur, neuronal activity was significantly suppressed. The amount of suppression tended to increase with probability (r^2 = 0.65, P = 0.20 and r^2 = 0.80, P = 0.10 in monkeys A and B, respectively) (Fig. 2, B and D) although the quantification of suppression may have been limited by the

Institute of Physiology, University of Fribourg, CH-1700 Fribourg, Switzerland, and Department of Anatomy, University of Cambridge, Downing Street, Cambridge CB2 3DY, UK.

*To whom correspondence should be addressed. E-mail: cdf28@cam.ac.uk

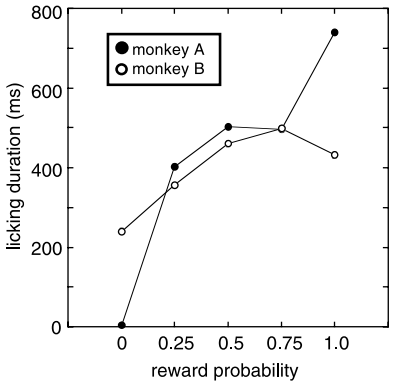
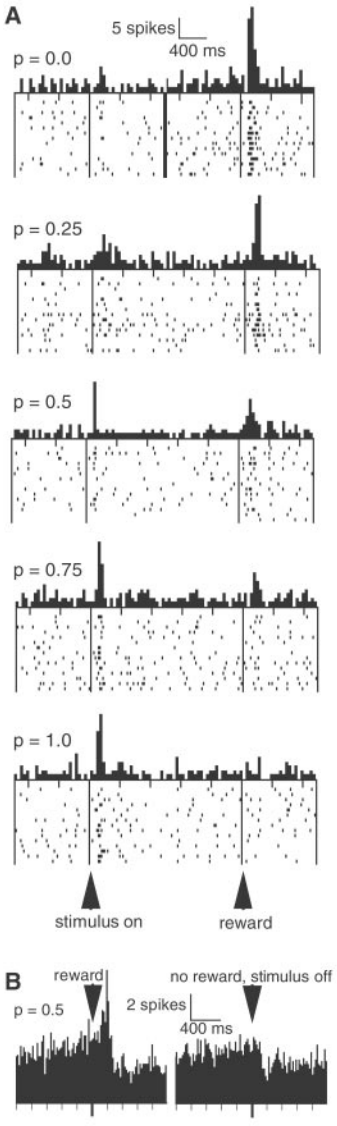


Fig. 1. Conditioned licking behavior increased with reward probability. The ordinate displays the mean duration of licking during the 2-s period from conditioned stimulus onset to potential reward. Each point represents the mean (±SEM) of 2322 to 6668 trials. Standard errors are too small to be visible. The behavioral data shown were collected between the first and last day of recordings and include data collected in the absence of physiological recordings.



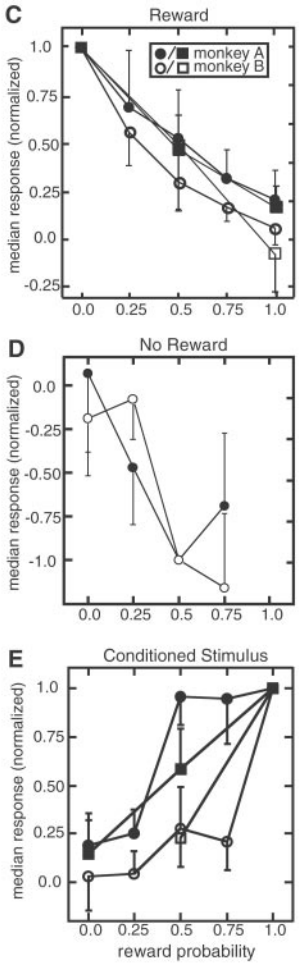


Fig. 2. Dependence of phasic neuronal responses on reward probability. (**A**) Rasters and histograms of activity in a single cell, illustrating responses to the conditioned stimuli and reward at various reward probabilities, increasing from top to bottom. The thick vertical line in the middle of the top panel (P = 0) indicates that the conditioned stimulus response to the left and the reward response to the right were not from a single trial type as in other panels but were spliced together. Reward at P = 0.0 was given in the absence of any explicit stimulus at a rate constant of 0.02 per 100 ms and thus presumably occurred with a low subjective probability (*11*). Only rewarded trials are shown at intermediate probabilities. Bin width = 20 ms. (**B**) Population histograms of rewarded (left) and unrewarded (right) trials at P = 0.5 (n = 39, monkey A, set 1). Bin width = 10 ms. (**C** to **E**) The median response (n = 34 to 62) measured in fixed standard windows, along with symmetric 95% confidence intervals (bars) (*11*). Circles and squares represent data from analogous experiments, with the squares representing a subsequent replication of the prior "circle" data but with distinct visual stimuli and only two or three probabilities tested. Error bars represent standard errors. In (C), the median magnitude of reward responses as a function of probability is shown, normalized in each neuron to the response to unpredicted reward. Unpredicted reward caused a median increase in activity that ranged from 76 to 270% above baseline for the four picture sets. Analogous to (C), fig. S3A shows means (±SEM) for a subset of responsive neurons (*11*). In (D), the median magnitude of responses to no reward as a function of probability is shown, normalized in each neuron to the response at P = 0.5. Median decreases in activity at P = 0.5 ranged from –22 to –55% below baseline. Symbols represent picture sets as shown in (C). At reward probability P = 0 for monkey B, a neutral visual stimulus was predicted (P = 0.5) by the conditioned stimulus. The data point shows the response after the neutral stimulus failed to occur. In (E), responses to conditioned stimuli are shown, normalized in each neuron to the response to the stimulus predicting reward at P = 1.0. The median response to this stimulus ranged from 67 to 194% above baseline. Symbols represent picture sets as shown in (C). The stimuli with P = 0 for monkey A, set 2, and for monkey B, set 1, predicted the subsequent occurrence of a neutral visual stimulus with P = 0.5.

low spontaneous activity levels. Conditioned stimuli elicited the typical phasic activations (*8–10*), with their magnitude increasing with increasing reward probability ($r^2 = 0.80$, $P = 0.04$ and $r^2 = 0.69$, $P = 0.08$ in monkeys A and B, respectively) (Figs. 2, A and E, and 3, A and B). In summary, the phasic activations varied monotonically with reward probability, although further conclusions about the quantitative relations are not warranted (*13*).

The present work revealed an additional, previously unreported activation of dopamine neurons. There was a sustained increase in activity that grew from the onset of the conditioned stimulus to the expected time of reward (Fig. 3, A and B). At $P = 0.5$, 29% of 188 neurons showed significant increases in activity before potential reward, whereas 3% showed decreases ($P < 0.05$, Wilcoxon test). By contrast, at $P = 1.0$, only 9% showed significant increases, and 5% showed significant decreases. For the population response, the sustained activation was maximal at $P = 0.5$, less pronounced at $P = 0.25$ and 0.75, and absent at $P = 0.0$ and 1.0 (Fig. 3C and fig. S3B). Statistical analysis revealed a significant effect of uncertainty on the population response ($P < 0.005$ in each of four data sets) (*11*), indicating that the sustained activation codes uncertainty (*14*). Furthermore, the peak of the sustained activation occurs at the time of potential reward, which corresponds to the moment of greatest uncertainty (*15*). The particular function of uncertainty signaled by dopamine neurons is not known (*13*), but we note that common measures of uncertainty (variance, standard deviation, and entropy) are all maximal at $P = 0.5$ and have highly nonlinear relations to probability, being very sensitive to small changes in probability near the extremes ($P = 0$ or 1).

The phasic and sustained activations differed not only in timing and relation to reward probability, but also in their occurrence in single neurons. In Fig. 3D, the magnitude of the phasic and sustained activation is shown for each neuron ($n = 241$). First, a substantial number of neurons had little or no response of either type (*13*); however, the magnitudes of each type of response fell along a continuum, with no evidence for subpopulations among dopamine neurons. Second, the magnitude of the sustained activation showed no consistent relation to the magnitude of phasic activation across neurons. This was the case both for the phasic response to conditioned stimuli ($r = 0.095$, $P > 0.10$) and for the response to unpredicted reward ($r = -0.024$) (Fig. 3D). In contrast, there was a significant positive correlation of phasic responses between conditioned stimuli and reward ($r = 0.196$, $P < 0.01$) (fig. S4). Thus, the phasic and sustained activations appear to occur independently and within a single population of dopamine neurons.

Although the sustained activation occurs in response to reward uncertainty, it is important to know whether it is specific to motivationally relevant stimuli or generalizes to all uncertain events. We conditioned two visual stimuli in a series, with the second following the first in only half of the trials ($P = 0.5$). The stimuli were distinct but entirely analogous to the other stimuli used for conditioning. Dopamine neurons showed neither sustained (Figs. 3C and 4A) nor phasic responses (Fig. 2, D and E) to either the first or second of these stimuli. Thus, the sustained activation seems to be related to uncertainty about motivationally relevant stimuli.

If the sustained dopamine activation is related to the motivational properties of uncertain rewards, it should vary with reward magnitude. We used distinct visual stimuli to predict the magnitude of potential reward at $P = 0.5$ and found that the sustained activation of dopamine neurons increased with increasing reward magnitude ($n = 84$, $P < 0.02$ in each monkey) (Fig. 4A) (*11*). The sustained activation could reflect the discrepancy in potential reward rather than absolute reward magnitude. To address this issue, we performed an additional experiment (53 neurons in monkey B) in which reward was delivered in each trial but varied between two magnitudes at $P = 0.5$. One stimulus predicted a small or medium reward, another predicted a small or large reward, and a third predicted a medium or large reward. The sustained activation was maximal after the stimulus predicting the largest variation (small versus large reward) ($P < 0.01$) (Fig.

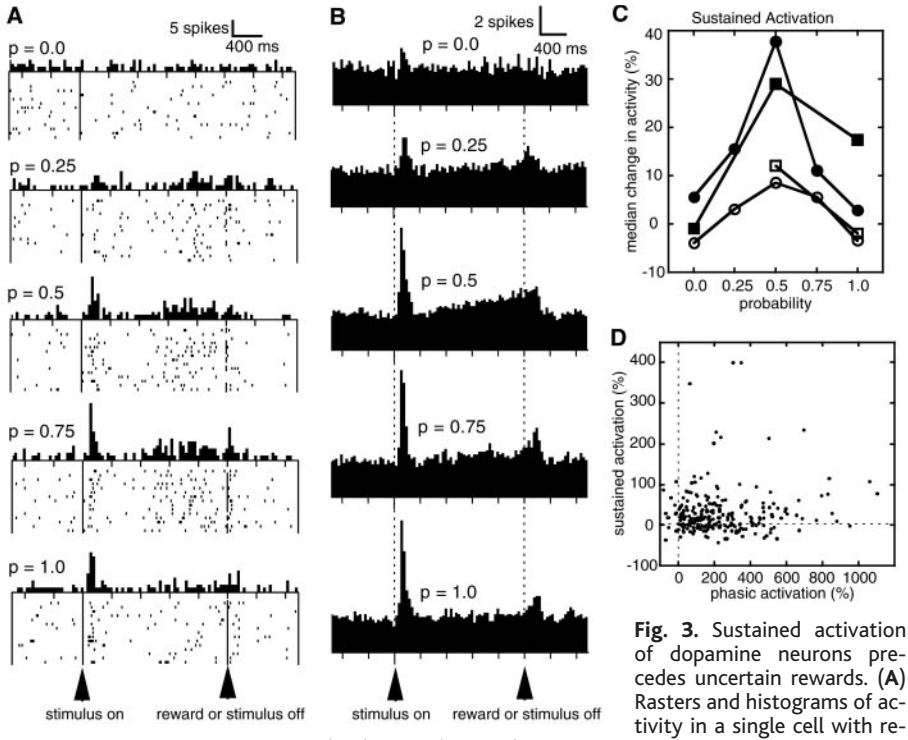


Fig. 3. Sustained activation of dopamine neurons precedes uncertain rewards. (**A**) Rasters and histograms of activity in a single cell with reward probabilities ranging from 0.0 (top) to 1.0 (bottom). This neuron showed sustained activation before potential reward at all three intermediate probabilities. Both rewarded and unrewarded trials are shown at intermediate probabilities; the longer vertical marks in the rasters indicate the occurrence of reward. Bin width = 20 ms. (**B**) Population histograms at reward probabilities ranging from 0.0 (top) to 1.0 (bottom). Histograms were constructed from every trial in each neuron in the first picture set in monkey A (35 to 44 neurons per stimulus type; 638 total trials at $P = 0$ and 1200 to 1700 trials for all other probabilities). Both rewarded and unrewarded trials are included at intermediate probabilities. At $P = 0.5$, the mean (±SD) rate of basal activity in this population was 2.5 ± 1.4 impulses per second before stimulus onset and 3.9 ± 2.7 in the 500 ms before potential reward. (**C**) Median sustained activation of dopamine neurons as a function of reward probability. In analogy, means (±SEM) are shown in fig. S3B for a subset of responsive neurons (*11*). Symbols have the same meaning as in Fig. 2C. For monkey A, set 1, the points at $P = 0.25$ and 0.75 may underestimate the amount of sustained activation, as 11 cells with unusually high levels of sustained activity at $P = 0.5$ (median activation of 72%) were not tested at $P = 0.25$ or 0.75. This was because, at the time of those experiments, the novel form of activation cast doubt on the dopaminergic identity of the neurons. For $P = 0$ in monkey A, set 2, and in monkey B, set 1, there was a 50% chance of a neutral stimulus following the conditioned stimulus. (**D**) Sustained responses (at $P = 0.5$) plotted against phasic responses to unpredicted reward ($P = 0$) for all neurons recorded in both monkeys (188 neurons, with an additional 53 neurons tested with different reward magnitudes as in Fig. 4B; five outlying neurons, in both dimensions, are not shown).

4B). These data indicate that the amount of sustained activation by reward uncertainty in dopamine neurons increases with the discrepancy between potential rewards.

The present results demonstrate two distinct response types in dopamine neurons. Brief, phasic activations changed monotonically with increasing reward probability, whereas slower, more sustained activations developed with increasing reward uncertainty. These sustained activations were not observed in previous studies in which predictions had low uncertainty. Thus, the activity of dopamine neurons carries information about two intimately related but fundamentally distinct statistical parameters of reward. A potentially analogous coding scheme was identified in neurons of the fly visual system, in which the visual stimulus and uncertainty about that visual stimulus appeared to be coded independently in single neurons (*16*).

By systematically varying reward probability, we show that the phasic activity of dopamine neurons matches the quantitative definition of reward prediction error. Responses to reward decreased with increasing reward probability, and, conversely, responses to the predictive stimulus increased. Furthermore, reward always elicited responses when it occurred at $P < 1$, even after thousands of pairings between stimulus and reward. By always coding prediction error over the full range of probabilities, dopamine neurons could provide a teaching signal in accord with the principles of learning originally described by Rescorla and Wagner (*17*–*19*).

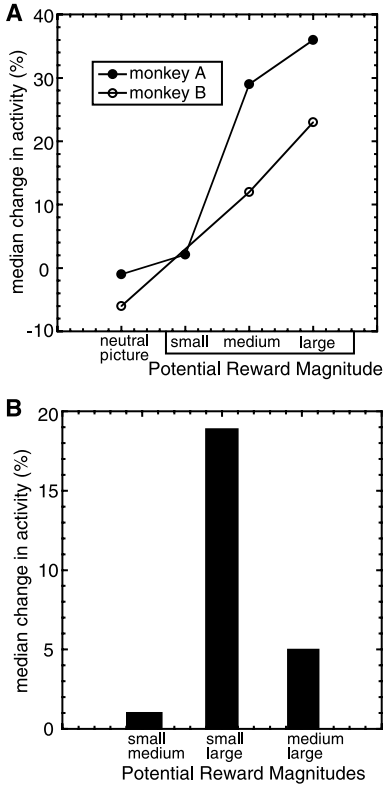


Fig. 4. Sustained activation is dependent on the discrepancy in potential reward magnitude. (**A**) All stimuli predicted potential reward (0.05, 0.15, or 0.5 ml of liquid) or a neutral picture at $P = 0.5$. Data are from 35 cells in monkey A and 49 cells in monkey B. (**B**) Each stimulus predicted that reward would be one of two potential magnitudes, each at $P = 0.5$, as indicated on the abscissa. Every trial was rewarded with one of the two potential reward magnitudes. Data are from 53 cells in monkey B.

In addition to those principles described by Rescorla-Wagner, other basic intuitive principles of associative learning have been described, focusing in particular on the importance of attention (*20*, *21*). It is generally accepted that no single principle alone is sufficient to explain all observations of animal learning, and the various theories are thus considered to be complementary (*6*, *7*). The Pearce-Hall theory proposes that attention (and thus learning) is proportional to uncertainty about reinforcers (*21*, *22*). As dopamine neurons are activated by reward uncertainty, dopamine could facilitate attention and learning in accord with the Pearce-Hall theory. This raises the possibility that two fundamental principles of learning are embodied by two distinct types of response in dopamine neurons (*23*).

The link between uncertainty, attention, and learning has two related aspects [another aspect is given in (*24*)]. The goal of learning can be seen as finding accurate predictors for motivationally significant events. Subjective uncertainty indicates that the animal lacks an accurate predictor and thus indicates the utility of identifying a more accurate predictor (*25*). Similarly, and as indicated by mathematical principles of information (*26*), only in the presence of uncertainty is it anticipated that there will be information available in the outcome. If reward ($P = 1$) or no reward ($P = 0$) occurs exactly as predicted, that event contains no information beyond that already given by the conditioned stimulus; that is, it is redundant. However, when the prediction of reward is uncertain, the outcome (reward or no reward) always contains information. The outcome at $P = 0.5$ contains, on average, the maximal amount of information (one bit) of any probability. The processing of this reward information is demonstrated by the fact that prediction error signals are always generated in dopamine neurons when reward outcomes occur under conditions of uncertainty. Thus, subjective reward uncertainty corresponds both to the utility of identifying more accurate predictors and to the expectation of reward information. Through its widespread influence, dopamine could control a nonselective form of attention or arousal, which is dependent on uncertainty and designed to aid the learning of predictive stimuli and actions.

Although dopaminergic signals may promote a particular form of attention, an extensive literature has already established the critical importance of dopamine in reward and reinforcement. Whereas the phasic response of dopamine neurons to reward prediction error fits remarkably well with dopamine's presumed role in appetitive reinforcement (*10*, *17*, *18*), the activation by reward uncertainty may appear inconsistent with a reinforcing function. This apparent discrepancy would be resolved to the extent that postsynaptic neurons can discriminate the two forms of activity. However, it seems unlikely that the two patterns of activity can be discriminated perfectly, especially given the slow time course of dopamine transmission. Rather than arguing against a role for the activity of dopamine neurons in reinforcement, one might ask whether reward uncertainty itself has rewarding and reinforcing properties. Indeed, gambling behavior is defined by reward uncertainty and is prevalent throughout many cultures. Animals display a potentially related behavior, preferring variable over fixed reward schedules [for discussion, see (*27*) and (*28*)]. The present results suggest that dopamine is elevated during gambling in a manner that is dependent on both the probability and magnitude of potential reward. This uncertainty-induced increase in dopamine could contribute to the rewarding properties of gambling, which are not readily explained by overall monetary gain or dopamine's corresponding role in prediction error (as losses tend to outnumber gains) (*29*). The question arises as to why a reward signal would be produced by reward uncertainty. Although risk-taking behavior may be maladaptive in a laboratory or casino, where the probabilities are fixed and there is nothing useful to learn, it could be advantageous in natural settings, where it would be expected to promote learning of stimuli or actions that are accurate predictors of reward (*25*). Thus, the sustained, uncertainty-induced increase in dopamine could act to reinforce risk-taking behavior and its consequent reward information, whereas the phasic response after prediction error could mediate the more dominant reinforcement of reward itself.

References and Notes

1. R. P. N. Rao, D. H. Ballard, *Nature Neurosci.* **2**, 79 (1999).
2. D. M. Wolpert, Z. Ghahramani, *Nature Neurosci.* **3** (suppl.), 1212 (2000).
3. E. K. Engel, P. Fries, W. Singer, *Nature Rev. Neurosci.* **2**, 704 (2001).
4. R. P. N. Rao, B. A. Olshausen, M. S. Lewicki, Eds., *Probabilistic Models of the Brain* (MIT Press, Cambridge, MA, 2002).
5. C. R. Gallistel, *The Organization of Learning* (MIT Press, Cambridge, MA, 1990).
6. A. Dickinson, *Contemporary Animal Learning Theory* (Cambridge Univ. Press, Cambridge, 1980).
7. J. M. Pearce, *An Introduction to Animal Cognition* (Lawrence Erlbaum, Hove, UK, 1987).
8. W. Schultz, P. Apicella, T. Ljungberg, *J. Neurosci.* **13**, 900 (1993).
9. P. Waelti, A. Dickinson, W. Schultz, *Nature* **412**, 43 (2001).

10. W. Schultz, *J. Neurophysiol.* **80**, 1 (1998).
11. Materials and methods are available as supporting material on *Science* Online.
12. Simple linear regression coefficients for each type of phasic response were calculated for each set of data for which all probabilities were tested ($P = 0.0$, 0.25, 0.5, 0.75, and 1.0 in Fig. 2, C and E; $P = 0.0$, 0.25, 0.5, and 0.75 in Fig. 2D). This was done only as an approximation and does not imply linearity in the response functions. In addition to the nonlinear factors discussed in (*13*), there is imprecision in the subjective timing of the 2-s interval between stimulus onset and potential reward (*15*). This probably accounts for the small but significant activation to "fully" predicted reward in monkey A (Figs. 2C and 3B).
13. Unpublished data (*30*), as well as Figs. 3 and 4, suggest that the responses of dopamine neurons multiplicatively combine the probability and magnitude of reward. Thus, it is not necessarily the case that the maximal responses observed in this study for a given reward magnitude (those at $P = 0.0$, 0.5, or 1.0, depending on the type of response) are actually the maximal evoked responses of a given neuron. One would expect that, like other neurons coding the intensity of a signal, dopamine neurons have a stimulus-response function that is sigmoid, being insensitive to values above or below a particular range. The likelihood that individual neurons have distinct thresholds has critical implications for understanding the shape of the probability-response functions presented in Figs. 2 and 3 and could explain why many neurons shown in Fig. 3D appear to be unresponsive. The shape of the probability functions that we measured would depend on the range of values to which most of the neurons are sensitive. Because these ranges are unknown, the only interpretation that should be given to the data at this time is that dopamine neuronal responses follow probability or uncertainty in a monotonic fashion.
14. The present experiments were performed with a standard delay conditioning procedure, meaning that the conditioned stimulus remained on for the full 2-s delay until the potential time of reward. In a separate experiment, a smaller number of neurons ($n = 22$) were tested with trace conditioning in which the conditioned stimulus indicating the probability of reward was on for 1 s, and potential reward occurred following an additional 1-s interval after stimulus offset. Although there may have been some sustained activation in the trace condition at $P = 0.5$ ($P < 0.1$), the activity preceding potential reward (during either 250- or 500-ms periods) was significantly less than that in experiments with delay conditioning ($P < 0.05$, Mann-Whitney test). Furthermore, a distinct behavioral pattern emerged with trace conditioning; the likelihood of licking increased before stimulus offset, decreased subsequently, and then increased again before reward. The explanation for the apparent discrepancy between trace and delay conditioning is unclear, but it could be related to the presence of temporal information provided by the continued presence of the delay stimulus; that is, as long as the delay stimulus is present, the time of potential reward must not have passed, and this information could suppress incoming inhibitory signals that are (imprecisely) timed to coincide with potential reward (*15*).
15. Objectively, potential reward always occurred after a 2-s delay. However, it is known that subjective timing is imprecise. Thus, the time course of the slowly developing sustained activation could reflect the increasing likelihood that the interval is nearing completion. Unpublished data (*30*) on the phasic activation of dopamine neurons to the delivery of reward earlier or later than predicted suggest a similar degree of temporal imprecision in the prediction. It is therefore reasonable to hypothesize that dopamine neurons code the uncertainty in reward in the subsequent moment (the very near future).
16. A. L. Fairhall, G. D. Lewin, W. Bialek, R. R. de Ruyter van Steveninck, *Nature* **412**, 787 (2001).
17. R. R. Montague, P. Dayan, T. J. Sejnowski, *J. Neurosci.* **16**, 1936 (1996).
18. W. Schultz, P. Dayan, R. R. Montague, *Science* **275**, 1593 (1997).
19. R. A. Rescorla, A. R. Wagner, in *Classical Conditioning II: Current Research and Theory*, A. H. Black, W. S. Prokasy, Eds. (Appleton-Century-Crofts, New York, 1972), pp. 64–69.
20. N. J. Mackintosh, *Psychol. Rev.* **82**, 276 (1975).
21. J. M. Pearce, G. A. Hall, *Psychol. Rev.* **87**, 532 (1980).
22. H. Kaye, J. M. Pearce, *J. Exp. Psychol. Anim. Behav. Process.* **10**, 90 (1984).
23. The fact that there are two distinct dopamine signals, each with unique properties, suggests two distinct functions for dopamine. However, this does not necessarily imply that the two signals must be processed independently. Thus, each signal may contribute to the performance of two or more functions. Furthermore, questions concerning the functions of dopamine in target areas (such as reinforcement and attention) are distinct from questions about the qualitative nature of stimuli (rewarding versus attention-inducing) that contribute to the activation of dopamine neurons (*31*).
24. P. Dayan, S. Kakade, P. R. Montague, *Nature Neurosci.* **3** (suppl.), 1218 (2000).
25. In the artificial, impoverished conditions of a laboratory setting or a casino, the probabilities associated with particular stimuli or actions are fixed, and there is nothing else useful to be learned. However, the natural environment contains a high degree of correlation between a multitude of events; this is implicit in the adaptive utility of associative learning. Thus, an animal should not assume that uncertainty signals the objective absence of accurate predictors but rather that it is ignorant of those predictors. Although accurate predictors of reward may not always be present in the environment, one would not expect the learning machinery of the brain to assume their absence. In fact, there is a period of uncertainty about all rewards before accurate predictors are found. If subjective uncertainty is assumed to result from ignorance of predictors rather than absence of predictors, then it would be appropriate for subjective uncertainty to have attention-inducing and reinforcing properties that would ultimately enhance learning and reduce uncertainty.
26. C. E. Shannon, *Bell Syst. Tech. J.* **27**, 379 (1948).
27. C. R. Gallistel, J. Gibbon, *Psychol. Rev.* **107**, 289 (2000).
28. J. E. Mazur, *Psychol. Rev.* **108**, 96 (2001).
29. Alternative attempts to explain gambling behavior focus on the fact that people (particularly those with prefrontal deficits) may misperceive reward probabilities or magnitudes or combine them in an inappropriate manner. This "cognitive" hypothesis fails to explain why gambling is appealing (and sometimes addictive) to a large number of otherwise healthy people, most of whom are aware that the odds are against them and that they have lost and will continue to lose money. In addition, any attempt to explain gambling behavior must address the fact that gambling is common at all probabilities (except $P = 0$ or 1, by definition) and all reward magnitudes. The present work suggests that activation of dopamine neurons may occur to a comparable extent during the expectation of a small reward at intermediate probabilities or a large reward at low probabilities. Thus, dopamine could contribute to the appeal of gambling in general. A behavior as prevalent as gambling must be explained in terms that are consistent with natural selection. The present hypothesis does so by pointing out that risk-taking promotes learning in natural environments.
30. C. D. Fiorillo, P. N. Tobler, W. Schultz, unpublished data.
31. P. Redgrave, T. J. Prescott, K. Gurney, *Trends Neurosci.* **22**, 146 (1999).

32. We thank A. Dickinson, S. Baker, R. Moreno, K. Tsutsui, I. Hernadi, P. Dayan, and two anonymous reviewers for helpful comments on the manuscript. Funding was provided by the Human Frontiers Science Program (C.D.F.), Swiss National Science Funds (W.S. and P.N.T.), and Wellcome Trust (W.S.).

Supporting Online Material
www.sciencemag.org/cgi/content/full/299/5614/1898/DC1
Materials and Methods
SOM Text
Figs. S1 to S4
References and Notes

14 August 2002; accepted 12 February 2003

Identified Sources and Targets of Slow Inhibition in the Neocortex

Gábor Tamás,* Andrea Lőrincz, Anna Simon, János Szabadics

There are two types of inhibitory postsynaptic potentials in the cerebral cortex. Fast inhibition is mediated by ionotropic γ-aminobutyric acid type A ($GABA_A$) receptors, and slow inhibition is due to metabotropic $GABA_B$ receptors. Several neuron classes elicit inhibitory postsynaptic potentials through $GABA_A$ receptors, but possible distinct sources of slow inhibition remain unknown. We identified a class of GABAergic interneurons, the neurogliaform cells, that, in contrast to other GABA-releasing cells, elicited combined $GABA_A$ and $GABA_B$ receptor–mediated responses with single action potentials and that predominantly targeted the dendritic spines of pyramidal neurons. Slow inhibition evoked by a distinct interneuron in spatially restricted postsynaptic compartments could locally and selectively modulate cortical excitability.

Gamma-aminobutyric acid (GABA) is the major inhibitory transmitter in the cerebral cortex (*1*). Extracellular stimulation of afferent cortical fibers elicits biphasic inhibitory postsynaptic potentials (IPSPs) in cortical cells. The early phase is due to the activation of $GABA_A$ receptors resulting in Cl^- conductance, and the late phase is mediated by K^+ channels linked to $GABA_B$ receptors through heterotrimeric GTP-binding proteins (*2–6*). Although dual recordings revealed several classes of interneurons evoking fast $GABA_A$ receptor–mediated responses in the postsynaptic cells, it is not clear whether distinct groups of inhibitory cells are responsible for activating $GABA_A$ and $GABA_B$ receptors. GABAergic neurons terminate on separate subcellular domains of target cells (*7*,

Department of Comparative Physiology, University of Szeged, Közép fasor 52, Szeged H-6726, Hungary.

*To whom correspondence should be addressed. E-mail: gtamas@bio.u-szeged.hu

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.

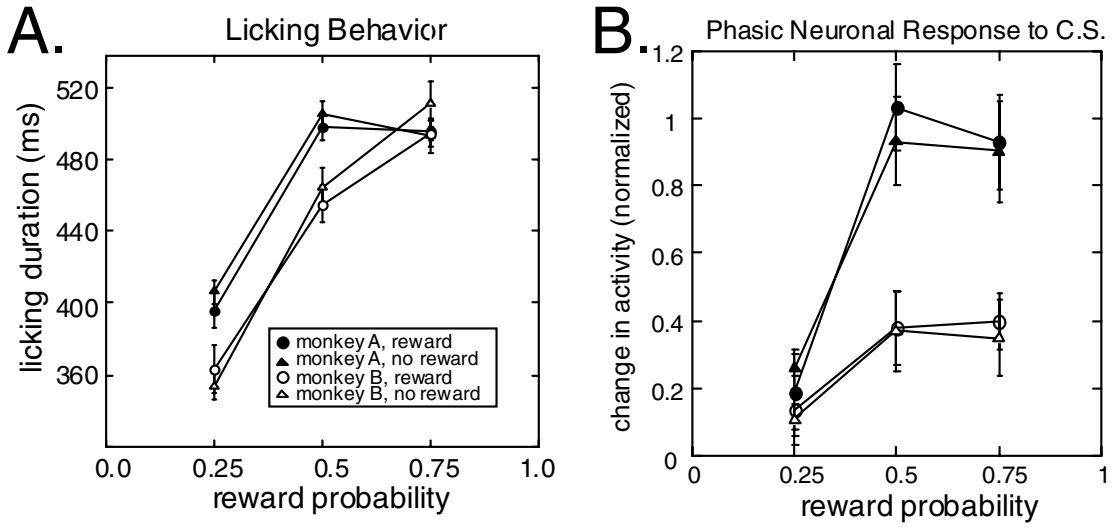


Fig. S1. Conditioned behavioral and neuronal responses failed to discriminate rewarded from unrewarded trials, though both responses were sensitive to reward probability. **A.** The data shown is the same as in figure 1, except rewarded and unrewarded trials have been analyzed separately. Conditioned 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.

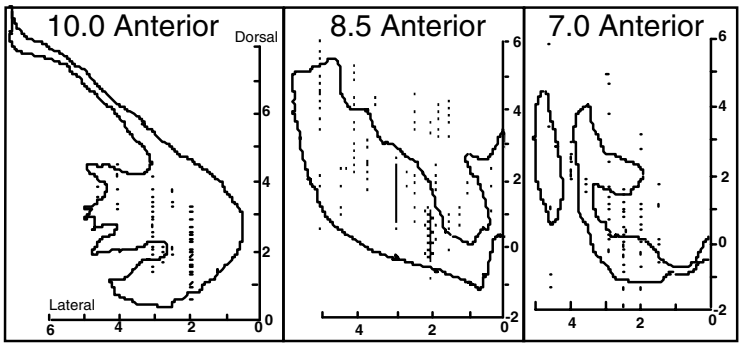


Fig. S2 The figure above displays histological reconstructions of the positions of recorded cells. Each outline represents the area of dense staining for the dopamine-synthesizing enzyme tyrosine hydroxylase in the ventral midbrain. The sections depicted were taken at 7.0, 9.5, and 10.0 mm anterior to 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.

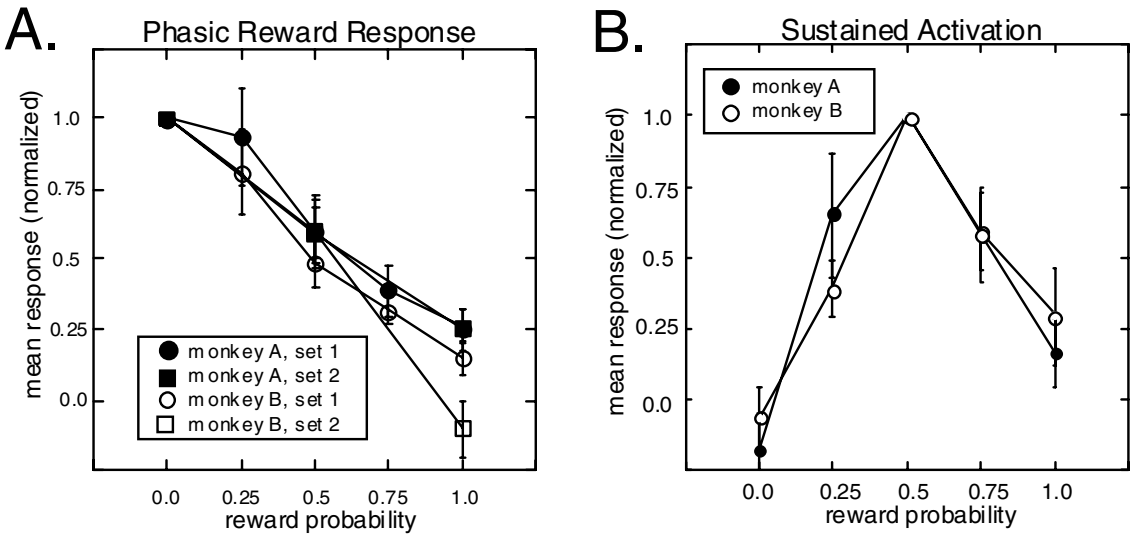


Fig. S3. The relationships of the phasic reward response and sustained activation to reward probability. Whereas figures 2C and 3C show median responses for the entire neuronal populations sampled, these figures show means (±s.e.m.) for selected groups of responsive neurons. **A.** The mean response 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.

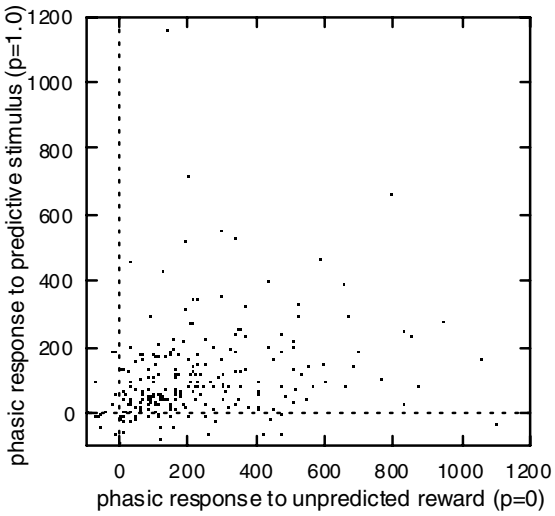


Fig. S4 The magnitude of the phasic activation to reward is correlated across neurons with the magnitude of the phasic activation to a conditioned stimulus (r=0.196, P<0.01, n=241). This is in contrast to figure 3D, which shows no correlation between the sustained activation and the phasic activation 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.