Adaptive Coding of Reward Value by Dopamine Neurons

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

It is important for animals to estimate the value of rewards as accurately as possible. Because the number of potential reward values is very large, it is necessary that the brain's limited resources be allocated so as to discriminate better among more likely reward outcomes at the expense of less likely outcomes. We found that midbrain dopamine neurons rapidly adapted to the information provided by reward-predicting stimuli. Responses shifted relative to the expected reward value, and the gain adjusted to the variance of reward value. In this way, dopamine neurons maintained their reward sensitivity over a large range of reward values.

In order to select the action associated with the largest reward, it is critical that the neural representation of reward has minimal uncertainty. A fundamental difficulty in representing the value of rewards (and many other stimuli) is that the number of possible values has no absolute limits. By contrast, the representational capacity of the brain is limited, as exemplified by its finite number of neurons and the limited number of possible spike outputs of each neuron. If a neuron's limited outputs were allocated evenly to represent the large, potentially infinite number of possible reward values, then that neuron's activity would allow for little if any discrimination between rewards. However, a neuron's discriminative capacity can be improved if the neuron has access to information indicating that some reward values are more likely to occur than others and if it can allocate most of its spike outputs to representing the most probable values. Conditioned, reward-predicting stimuli could provide such information for neurons, as they do in a more general way for behavior (1–3). Here we investigate how dopamine neurons adapt to the information about reward value contained in predictive stimuli. These neurons play a major role in reward processing (4–7) and respond to rewards and reward-predicting stimuli (8–11).

We presented distinct visual stimuli that specified both the probability and magnitude of otherwise identical juice rewards to monkeys well trained in a Pavlovian procedure (12). Standard procedures were employed to extracellularly record the activity of single dopamine neurons of midbrain groups A8, A9, and A10 in two awake Macaque monkeys (12). We report data for all recorded neurons that displayed electrophysiological characteristics typical of dopamine neurons (wide impulses at low rates) (12, 13). In an attempt to accurately portray the whole population of dopamine neurons, we did not select neurons on the basis of their modulation by a reward event.

The expected value of future rewards (the sum of possible reward magnitudes, each weighted by its probability) is thought to be an important variable determining choice behavior (14–17). To test this, we trained an animal with a set of five distinct visual stimuli presented in pseudorandom alternation. Each stimulus indicated the probability that a specific liquid volume would be delivered 2 s after stimulus onset. Anticipatory licking before liquid delivery was elicited by the smallest positive expected liquid volume tested (0.05 ml at probability $p = 0.5$) and increased with expected liquid volume, suggesting that the animals had learned to use the stimuli to predict liquid delivery and that the larger liquid volumes corresponded to larger reward values (Fig. 1A). The transient activation of dopamine neurons increased monotonically with the expected liquid volume associated with each stimulus (Fig. 1, B and C). For example, the stimulus predicting 0.15 ml at $p = 1.0$ elicited significantly greater neural activation than the stimulus predicting the same magnitude reward at $p = 0.5$, but less activation than the stimulus predicting 0.50 ml at $p = 0.5$. The activation of dopamine neurons also increased with the combination of magnitude and probability when the stimuli predicted that either of two nonzero magnitudes would occur with equal probability (Fig. 1C, animal B).

To investigate whether individual neurons might be preferentially sensitive to probability or magnitude, we took independent measures of sensitivity to magnitude and probability in each neuron ($n = 57$ neurons). There was a positive correlation ($R^2 = 0.23$, $P < 0.005$), indicating that those neurons that were most sensitive to reward magnitude were also most sensitive to probability (Fig. 1D). Thus, it appears that dopamine neurons encode a combination of magnitude and probability, as expressed, for example, by the expected reward value, rather than distinguishing between the two.

Having examined responses to reward-predicting stimuli of differing values, we investigated the extent to which dopamine neurons discriminated between different volumes of unpredicted liquid. We delivered three distinct liquid volumes (0.05, 0.15, and 0.50 ml) in pseudorandom alternation with a variable intertrial interval (18) and in the absence of any explicit predictive stimuli. Both individual dopamine neurons (43 of 55 neurons tested; $P < 0.01$, Wilcoxon test) and the population as a whole (55 neurons) showed greater activation for the large than for the small liquid volume (Fig. 2). Thus, the activation of dopamine neurons increased with the reward value of unpredicted liquids, similar to the responses to reward-predicting visual stimuli.

Although these results suggest that dopamine neurons encode the reward value in a monotonically increasing fashion, past work indicates that they do not represent absolute value. Rather, they appear to encode value as a prediction error by representing at each moment in time the difference between the

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

*Present address: Department of Biology, Stanford University, Stanford, CA 94305–5125, USA.
†To whom correspondence should be addressed. E-mail: ws234@cam.ac.uk

Fig. 1. Behavioral and neuronal responses to conditioned stimuli increase with expected reward value. (**A**) Anticipatory licking responses during the 2-s delay between the conditioned stimuli and liquid delivery. Each point shows the mean (± SEM) of at least 1835 trials (animal A) and is significantly different from all other points (t tests). Similar results were obtained from animal B, although the mean licking durations varied over a smaller range. (**B**) Single-neuron (top) and population responses (bottom) ($n = 57$ neurons) from the experiment in (A). Visual conditioned stimuli with their expected magnitude of reward are shown above the rasters. Expected values (probability × magnitude) were, from left to right, 0 ml (1.0 probability × 0.0 ml magnitude), 0.025 ml (0.5 × 0.05 ml), 0.075 ml (0.5 × 0.15 ml), 0.15 ml (1.0 × 0.15 ml), and 0.25 ml (0.5 × 0.50). Bin width is 10 ms in histograms of all figures. (**C**) (Left) Population responses as a function of expected liquid volume. Measurements were taken 90 to 180 ms (animal A) and 110 to 240 ms (animal B) after the onset of visual stimuli. The median (±95% confidence intervals) percent change in firing rates within the population was calculated after normalization of responses within each neuron to the response evoked after onset of the stimulus associated with the largest expected value. This stimulus elicited a median activation of 167% in animal A ($n = 57$ neurons) and 40% in animal B ($n = 53$ neurons). For animal A (squares), stimuli indicated probability and magnitude as in (B). For animal B (circles), one stimulus was never followed by liquid, whereas each of the other three stimuli was associated with two volumes of equal probability (0.05 or 0.15 ml, 0.05 or 0.50 ml, and 0.15 or 0.50 ml). In each animal, the population of neurons discriminated among each expected value tested, except for 0.0 versus 0.025 ml in animal A. (Right) An alternative analysis, illustrating the sensitivity (spikes/s/ml) of a typical dopamine response to expected liquid volume. For each individual neuron,
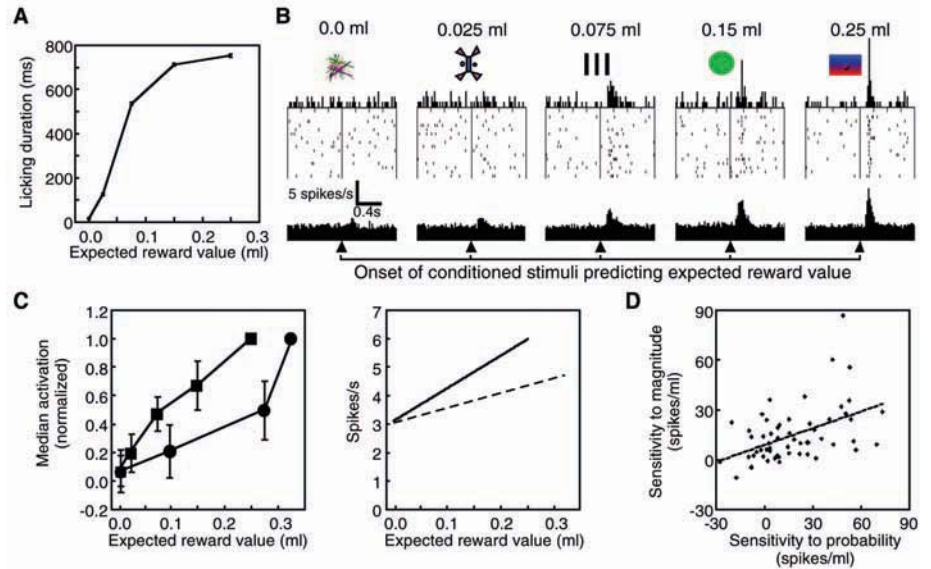
the number of impulses after stimulus onset was plotted as a function of expected magnitude, and a line was fit. The lines shown are the median lines of each population of neurons (animal A, solid line, spikes/s = 11.5 × magnitude + 3.1, $R^2 = 0.51$; animal B, spikes/s = 5.2 × magnitude + 3.0, $R^2 = 0.69$). (**D**) Positive correlation between the sensitivity of individual neurons to reward probability and magnitude ($R^2 = 0.23$, $P < 0.005$). For the data from animal A in (C), responses in each neuron ($n = 57$ neurons) are plotted both as a function of expected value, as determined both by reward probability (0.15 ml at $p = 0.0$, 0.5, and 1.0) and by liquid volume (0.05, 0.15, and 0.50 ml at $p = 0.5$). A line was fit in each case, and the slopes provided independent estimates of the sensitivity of that neuron to reward probability and magnitude. For each neuron, the slopes are plotted against each other.

Fig. 2. Neural discrimination of liquid volume. (**A**) (Top) Rasters and histograms of activity from a single dopamine neuron. (Bottom) Population histograms of activity from all neurons tested ($n = 55$ neurons). Three volumes of liquid were delivered in pseudorandom alternation in the absence of any explicit predictive stimuli. The intertrial interval ensured that the expected volume at any given moment was low (18). Thick horizontal bars above the rasters indicate the time of reward delivery, and thin horizontal bars indicate the single standard time window that was used for measuring the magnitude of all responses in all neurons, as summarized in (B). Similar windows were used for all analyses and plots (supporting text). (**B**) Neural response as a function of liquid volume. Median (±95% confidence intervals) percentage
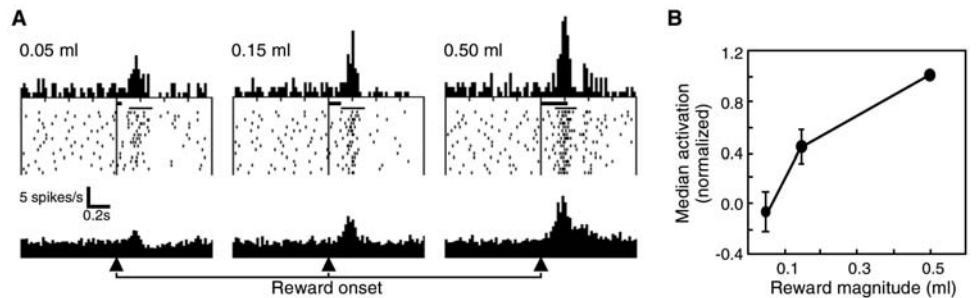
change in activity for the population of neurons ($n = 55$ neurons) was calculated for responses to each volume after normalization in each neuron to the response after delivery of 0.5 ml, which itself elicited a median activation of 159% above baseline activity.

reward value (the sum of current and future rewards) and its expected value (before observation of current sensory input). Recent work demonstrates that, when signaling prediction errors, dopamine neurons are able to use contextual information in addition to information from explicitly conditioned stimuli (19). In the experiments shown in Figs. 1 and 2, all visual stimuli and liquid volumes were delivered in a context in which the expected reward value at each moment in time was low and invariant across trial types because of the intertrial interval (18). In our next set of experiments, we delivered differ-

ent volumes of liquid in the presence of explicit predictions indicated by conditioned stimuli, allowing us to systematically vary the expected value and range of reward.

Consistent with past work, a reward occurring exactly at the expected value (0.15 ml) elicited no response. However, when liquid volume was unpredictably smaller (0.05 ml) or larger (0.50 ml) in a minority of trials, dopamine neurons were suppressed or activated, respectively, compared to both the prestimulus baseline and the response to the expected volume delivered in the majority of trials ($P < 0.01$, Mann-Whitney test) (Fig. 3,

A and B). In an additional experiment, one stimulus predicted that either the small or medium volume would be delivered with equal probability, whereas another stimulus predicted either the medium or large volume with equal probability. In both cases, delivery of the larger of the two potential volumes elicited an increase in activity, whereas the smaller volume elicited a decrease (Fig. 3C). Thus, the identical medium volume had opposite effects on activity depending on the prediction ($P < 0.01$ in 19 of 53 neurons, Mann-Whitney; $P < 0.0001$ for the population of 53 neurons, Wilcoxon test) (Fig. 3D). These

results show how dopamine neurons process reward magnitude relative to a predicted magnitude and that a reward outcome that is positive on an absolute scale can nonetheless suppress the activity of dopamine neurons.

Although these results suggest that dopamine responses shift relative to the predicted reward magnitude, it is not known how their activity scales with the difference between actual and expected reward. To this end, we analyzed the dopamine responses at the time of the reward in the experiment shown in Fig. 1. Each of three distinct visual stimuli, presented on pseudorandomly alternating trials, predicted that one of two potential liquid volumes would be delivered with equal probability. Animals discriminated behaviorally between the three reward-predicting stimuli (Fig. 1A). Confirming the data described above, the larger of the two volumes always elicited an increase in activity at the time of the reward, and the smaller a decrease. However, the magnitude of activation or suppression appeared to be identical in each case, despite the fact that the absolute difference between actual and expected volume varied over a 10-fold range (Fig. 4, A and B). Thus, the responses of dopamine neurons did not appear to scale according to the absolute difference between actual and expected reward. Rather, the sensitivity or gain of the neural responses appeared to adapt according to the discrepancy in volume between the two potential outcomes.

To document this result further, we plotted the median neural responses as a function of liquid volume and drew a straight line to connect the data points representing the larger and smaller outcomes after each visual stimulus (Fig. 4C). The slope of these lines provided an estimate of the neurons' gain or sensitivity with respect to liquid volume. When the discrepancy was large, the sensitivity of dopamine neurons was low, and when the discrepancy was small, sensitivity was high. As a result of this adaptation, the neural responses discriminated between the two likely outcomes equally well, regardless of their absolute difference in magnitude. The present data are not sufficient to determine precisely to which aspect of the reward prediction the neuron's sensitivity adapted, but further analysis provided limited evidence that sensitivity adapted to the discrepancy between potential liquid volumes (such as the difference or variance) rather than to their expected value (12) (fig. S2).

Our results suggest that the activity of dopamine neurons carries information on the magnitude of reward. In representing reward magnitude, neural activity displayed two forms of adaptation that depended on the prediction that was in place at the time of the reward. First, the activity increased or de-

Fig. 4. Neural sensitivity to liquid volume adapts in response to predictive stimuli. (A) Activity of a single neuron showing nearly identical responses to three liquid volumes spanning a 10-fold range. Each of three pseudorandomly alternating visual stimuli (shown at left) was followed by one of two liquid volumes at $p = 0.5$ (top, 0.0 or 0.05 ml; middle, 0.0 or 0.15 ml; bottom, 0.0 or 0.5 ml). Responses after onset of visual stimuli increased with their associated expected reward values. Only rewarded trials are shown. (B) Population histograms for different liquid volumes from the experiment in (A) (57 neurons, animal A). (C) Each line connects responses occurring in the context of a specific conditioned stimulus, and its slope provides a measure of gain or sensitivity. Each point represents the median (±95% confidence intervals) response of the population taken after normalizing the percentage change in activity in each neuron to the response after unpredicted liquid (0.15 ml) delivered in a separate block of trials (which elicited an activation of 266% above baseline in animal A, $n = 57$ neurons, and 97% in animal B, $n = 53$ neurons). (Left) The experiment in (A) and (B). (Right) The same experiment, but performed in animal B with two nonzero liquid volumes per conditioned stimulus at equal probability ($p = 0.5$) (stimulus 1: 0.05 versus 0.15 ml, stimulus 2: 0.15 versus 0.5 ml, stimulus 3: 0.05 versus 0.5 ml).
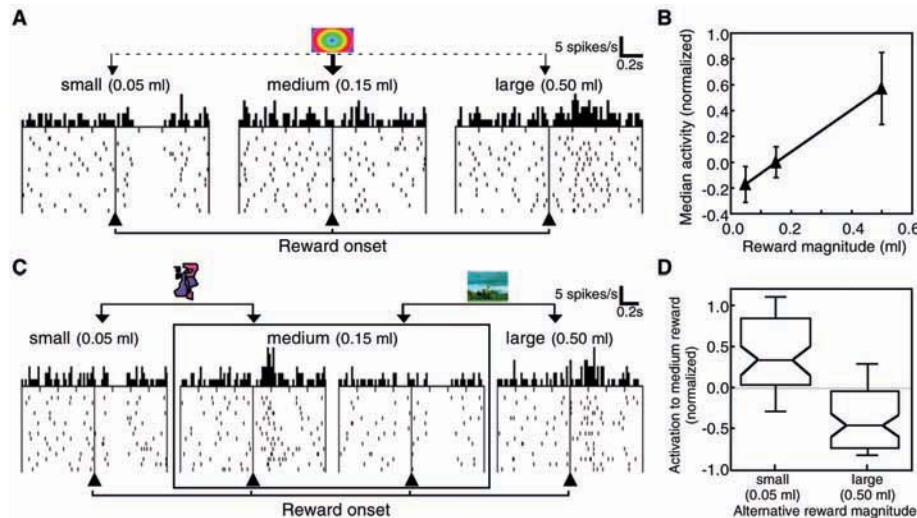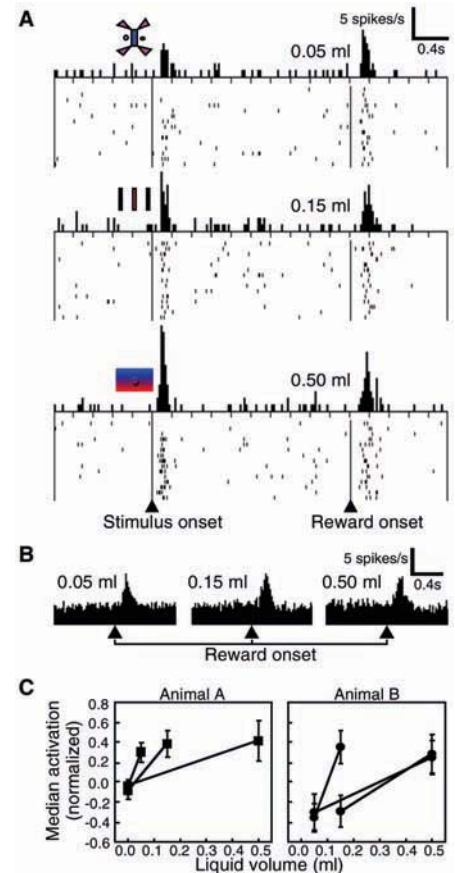

Fig. 3. Bidirectional dopamine responses to reward outcomes reflect deviations from predictions. (A) A single conditioned stimulus was usually followed by an intermediate volume of liquid (0.15 ml) that elicited no change in the neuron's activity (center). However, on a small minority of trials, smaller (0.05 ml) or larger (0.50 ml) volumes were unpredictably substituted, and neural activity decreased (left) or increased (right), respectively. Neural responses to the large liquid volume were relatively long-lasting (supporting online text). (B) Median responses (±95% confidence intervals) from the population as a function of liquid volume for the experiment in (A) (12 neurons from animal A, 17 neurons from animal B). Responses in each neuron were normalized to the response after the unpredicted delivery of liquid (0.15 ml) in a separate block of trials and in the absence of any explicit reward-predicting stimulus. (C) Responses of a single neuron to three liquid volumes, delivered in the context of two different predictions. One stimulus predicted small or medium volume with equal probability, whereas another stimulus predicted medium or large volume. The medium volume activated the neuron in one context, but suppressed activity in the other. (D) Population responses ($n = 53$ neurons, animal B) to medium reward in the experiment in (C). The plot shows the median, the ±95% confidence intervals (notches corresponding to obtuse angles), the 25th and 75th percentiles (boundaries corresponding to right angles), and the 10th and 90th percentiles (bars). In each neuron, percentage change in activity was normalized to the response to unpredicted liquid (0.15 ml, which elicited a median increase in activity of 97%).

creased depending on whether the reward outcome was larger or smaller, respectively, than an intermediate reference point such as expected value. A second, unanticipated form of adaptation was the change in sensitivity or gain of neural activity that appeared to depend on the range of likely reward magnitudes (Fig. 4). Thus, the larger of two potential rewards always elicited the same increase in activity and the smaller of the two elicited the same decrease in activity, regardless of absolute magnitude. The identical responses to liquid volumes spanning a 10-fold range were not due to an insensitivity of the dopamine neurons, which were capable of greater activations (Fig. 4C, note normalization of data points) and discriminated well among these same liquid volumes when delivered in the absence of explicit predictive stimuli (Fig. 2). Rather, the gain of neural activity with respect to liquid volume appeared to adapt in proportion to the range or standard deviation of the predicted reward outcomes, so that neural discrimination between the two reward outcomes that were most probable from the animal's perspective was robust regardless of their absolute difference in magnitude.

The efficiency and accuracy with which neural activity can code the value of a stimulus (such as liquid volume) can be greatly increased if neurons make use of information about the probabilities of potential reward values. Neural activity can then be devoted to representing probable values at the expense of improbable values. Our evidence suggests that the transient dopamine response to conditioned stimuli may carry information on expected reward value, and previous work shows that the more sustained activity of dopamine neurons reflects a measure of reward uncertainty such as variance (10). If the system possesses prior information consisting of the expected value and variance of reward, then this information need not be represented redundantly at the time of re-

ward. Discarding this old information may be achieved by subtracting the expected value from the absolute reward value and then dividing by the variance. Analogous normalization processes appear to occur in early visual neurons (20–22). It is not known to what extent the normalization processes observed in dopamine neurons are actually performed in dopamine neurons as opposed to their afferent input structures (23). Because the new information is by definition precisely the information that the system needs to learn, the activity of dopamine neurons would be an appropriate teaching signal (24).

Adaptation appears to be a nearly universal feature of neural activity. There is substantial evidence, particularly from the early visual system, that adaptation contributes to the efficient representation of stimuli (20–22, 25–28). We have extended the principles of efficient representation to the study of reward. Reward is central to processes underlying behavior, such as reinforcement learning and decision-making, and consideration of limitations and efficiency in the neural representation of reward may yield insights into these processes.

References and Notes

1. I. P. Pavlov, *Conditional Reflexes* (Oxford Univ. Press, Oxford, 1927).
2. E. L. Thorndike, *Animal Intelligence: Experimental Studies* (Macmillan, New York, 1911).
3. R. A. Rescorla, A. R. Wagner, in *Classical Conditioning II: Current Research and Theory*, A. Black, W. F. Prokasy, Eds. (Appleton-Century-Crofts, New York, 1972), pp. 64–99.
4. H. C. Fibiger, A. G. Phillips, in *The Nervous System*, vol. 4 of *Handbook of Physiology*, F. E. Bloom, Ed. (Williams & Wilkins, Baltimore, MD, 1986), pp. 647–675.
5. R. A. Wise, D. C. Hoffman, *Synapse* **10**, 247 (1992).
6. T. E. Robinson, K. C. Berridge, *Brain Res. Rev.* **18**, 247 (1993).
7. T. W. Robbins, B. J. Everitt, *Curr. Opin. Neurobiol.* **6**, 228 (1996).
8. P. Waelti, A. Dickinson, W. Schultz, *Nature* **412**, 43 (2001).
9. T. Satoh, S. Nakai, T. Sato, M. Kimura, *J. Neurosci.* **23**, 9913 (2003).
10. C. D. Fiorillo, P. N. Tobler, W. Schultz, *Science* **299**, 1898 (2003).
11. R. Kawagoe, Y. Takikawa, O. Hikosaka, *J. Neurophysiol.* **91**, 1013 (2004).
12. Materials and methods are available as supporting material on *Science* Online.
13. No significant correlations were found between neuronal position in areas A8, A9, and A10 and the different types of responses in all the experiments reported; thus, the data were pooled.
14. A. Arnauld, P. Nicole, *La logique, ou, L'art de penser* (Librairie philosophique J. Vrin, Paris, 1981/1662).
15. P. W. Glimcher, *Decisions, Uncertainty and the Brain* (MIT Press, Cambridge, MA, 2003).
16. P. Shizgal, *Curr. Opin. Neurobiol.* **7**, 198 (1997).
17. M. I. Leon, C. R. Gallistel, *J. Exp. Psychol. Anim. Behav. Process.* **24**, 265 (1998).
18. All trials were separated by an average intertrial interval of 9 s, consisting of a fixed 4 s plus an interval drawn from a truncated Poisson-like distribution with a mean of 5 s. Thus, the probability that a trial would begin at any given moment was low.
19. H. Nakahara, H. Itoh, R. Kawagoe, Y. Takikawa, O. Hikosaka, *Neuron* **41**, 269 (2004).
20. M. V. Srinivasan, S. B. Laughlin, A. Dubs, *Proc. R. Soc. London Ser. B.* **216**, 427 (1982).
21. N. Brenner, W. Bialek, R. R. de Ruyter van Steveninck, *Neuron* **26**, 695 (2000).
22. A. L. Fairhall, G. D. Lewen, W. Bialek, R. R. de Ruyter van Steveninck, *Nature* **412**, 787 (2001).
23. W. Schultz, *J. Neurophysiol.* **80**, 1 (1998).
24. W. Schultz, P. Dayan, R. R. Montague, *Science* **275**, 1593 (1997).
25. H. B. Barlow, in *Sensory Communication*, W. A. Rosenblith, Ed. (MIT Press, Cambridge, MA, 1961), pp. 217–234.
26. S. B. Laughlin, *Z. Naturforsch. Teil C Biochem. Biophys. Biol. Virol.* **36**, 910 (1981).
27. I. Ohzawa, G. Sclar, R. D. Freeman, *Nature* **298**, 266 (1982).
28. S. M. Smirnakis, M. J. Berry, D. K. Warland, W. Bialek, M. Meister, *Nature* **386**, 69 (1997).
29. We thank K. I. Tsutsui, I. Hernadi, A. Dickinson, and S. B. Laughlin for helpful comments. Supported by the Wellcome Trust, Swiss National Science Foundation (W.S. and P.N.T.), Roche Research Foundation (P.N.T.), Janggen-Poehn Foundation (P.N.T.), and Human Frontiers Science Program (C.D.F.).