Dopamine neurons report an error in the temporal prediction of reward during learning

Jeffrey R. Hollerman[1,2] and Wolfram Schultz[1]

[1] *Institute of Physiology, University of Fribourg, CH-1700 Fribourg, Switzerland*

[2] *Present address: Department of Psychology, Allegheny College, Meadville, Pennsylvania, 16335, USA*

Correspondence should be addressed to W.S. (Wolfram.Schultz@unifr.ch)

Many behaviors are affected by rewards, undergoing long-term changes when rewards are different than predicted but remaining unchanged when rewards occur exactly as predicted. The discrepancy between reward occurrence and reward prediction is termed an 'error in reward prediction'. Dopamine neurons in the substantia nigra and the ventral tegmental area are believed to be involved in reward-dependent behaviors. Consistent with this role, they are activated by rewards, and because they are activated more strongly by unpredicted than by predicted rewards they may play a role in learning. The present study investigated whether monkey dopamine neurons code an error in reward prediction during the course of learning. Dopamine neuron responses reflected the changes in reward prediction during individual learning episodes; dopamine neurons were activated by rewards during early trials, when errors were frequent and rewards unpredictable, but activation was progressively reduced as performance was consolidated and rewards became more predictable. These neurons were also activated when rewards occurred at unpredicted times and were depressed when rewards were omitted at the predicted times. Thus, dopamine neurons code errors in the prediction of both the occurrence and the time of rewards. In this respect, their responses resemble the teaching signals that have been employed in particularly efficient computational learning models.

Current theories view learning as the acquisition of specific predictions[1–4]. Humans and animals learn to predict the outcomes of their behavior, including rewards. Learning depends on the extent to which these outcomes are different than predicted, being governed by the discrepancy or 'error' between outcome and prediction. Outcomes that affect learning in this way are termed 'reinforcers'. Learning proceeds when outcomes occur that are not fully predicted, then slows down as outcomes become increasingly predicted and ends when outcomes are fully predicted. By contrast, behavior undergoes extinction when a predicted outcome fails to occur. (In the laboratory, predictions may fail either because the subject made an error or because the experimenter withholds the reward for correct behavior.) Recent learning algorithms employ errors in the prediction of outcome as teaching signals for changing synaptic weights in neuronal networks[5]. In these models, an unpredicted outcome leads to a positive signal, a predicted outcome to zero signal and the absence of a predicted outcome to a negative signal. The most efficient models capitalize on the observation that a key component of predictions concerns the exact time of reinforcement[6,7]. Their teaching signals use errors in the temporal prediction of reinforcement and compute the prediction error over consecutive time steps in individual trials ('temporal difference' algorithm[8]). Thus, teaching signals come to report progressively earlier reinforcement-related events and thus predict the outcome rather than simply reporting that it has occurred. They are particularly efficient for learning, as they can influence the behavioral reaction before it is executed. Reinforcement models that use predictive teaching signals can learn a wide variety of behavioral tasks, from balancing a pole on a cart wheel[9] to playing world-class backgammon[10]. It is therefore of interest to determine whether real nervous systems might process rewards in a similar manner during learning.

Results from lesioning and psychopharmacological experiments indicate a role of dopamine systems in behavior driven by rewards and in reward-based learning[12–14]. We have studied the neural mechanisms underlying this role of dopamine in monkeys and have previously reported that midbrain dopamine neurons show responses to food and liquid rewards that depend on their predictability[15,16]. The present study investigated whether these responses could have the formal characteristics of teaching signals. We found that the magnitude of dopamine responses to a juice reward reflected the degree of reward predictability during individual learning episodes. An unexpected reward evoked a strong response in dopamine neurons. As the monkeys' performance improved (i.e. as they learned to predict which response would trigger a reward), the neuronal response to the reward progressively decreased. Moreover, by varying the timing of reward, we found that dopamine neurons signal not only its occurrence but also its timing relative to expectations. Thus dopamine neurons seem to track the reward prediction error and emit a signal that has all the typical characteristics of a positive reinforcing signal for learning.

Results

Dopamine neurons in pars compacta of the substantia nigra and the ventral tegmental area were studied while monkeys learned to associate visual stimuli with liquid reward. Dopamine neurons in these two different midbrain groups showed similar respons-

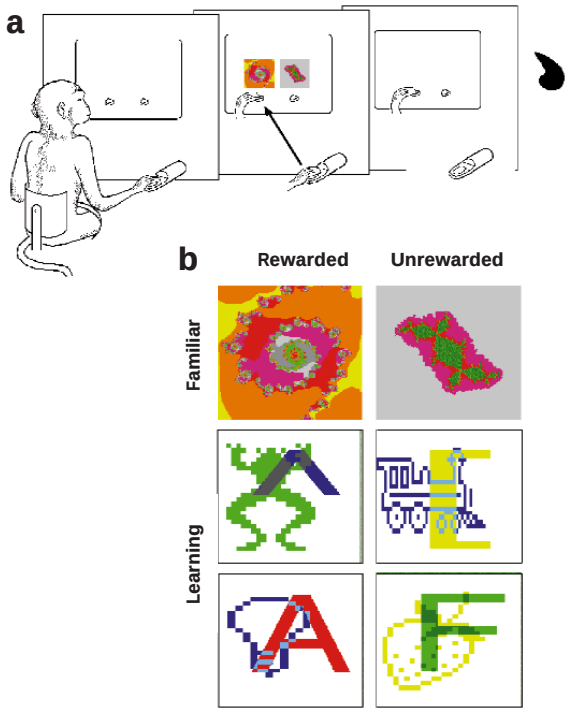

Fig. 1. The discrimination learning task. **(a)** Animals released a resting key when a pair of pictures appeared, touched the lever below the rewarded picture and received a drop of liquid. **(b)** Pictures used in the task. The same pair of two fractal pictures was used in all familiar trials (top). A new pair of two pictures was used in each block of learning trials (middle and bottom).

es and are therefore not distinguished in this report. We first tested 211 dopamine neurons during the learning of a visual-discrimination task. Animals were simultaneously presented with two pictures. If they touched a lever below one of the pictures, they received a drop of liquid, whereas the other picture was not rewarded (Fig. 1). During the initial presentations, 75% of dopamine neurons were activated when the reward occurred, comparable to other learning situations[15,17,18]. The same two pictures were presented repeatedly (varying randomly between left and right positions), and as the task was learned, the reward gradually ceased to activate dopamine neurons; instead, these neurons became responsive to presentation of the reward-predicting pictures, consistent with previous findings[15,17].

We then studied the reward responses of individual dopamine neurons during complete learning episodes. For each new episode, a novel pair of pictures was presented, whereas all other task components remained unchanged. Animals learned by trial and error to associate one of the novel pictures with reward. A learning criterion served to indicate the deviation from chance performance; this was defined as the second correct response in the first series of four consecutive correct responses. The learning criterion corresponded to the loss of chance performance and was reached on average after 12 trials, 5 of them being correctly performed (Fig. 2; see also Fig. 5a). The frequent errors during the initial learning period indicate that the pictures themselves did not contain any unintended predictive information.

Dopamine neurons showed strong activation (increase in firing rate) in response to unpredicted free liquid delivered outside the task (45 of 61 neurons, 74%), consistent with previous findings[15] (Fig. 3b). Strikingly, many neurons in the present study (33 of 66, 50%) were also activated by the reward during initial learning trials (Fig. 3). By contrast, reward activations were rare in trials with familiar pictures (8 of the 66 neurons, 12%) ($p < 0.0001$, ANOVA on magnitudes in initial ten free-liquid, learning and familiar trials; all three situations $p < 0.05$, Fisher test). Most neurons also showed activations following the novel pictures (not shown), resembling novelty responses and response generalization from learned, reward-predicting stimuli observed previously[17–22].

During individual learning episodes, each responding neuron progressively lost its reward activation after the criterion was attained, and activations approached the low levels typical for familiar trials. The rate of decrease in responsiveness was related to the duration of the learning period before criterion. Some stimulus pairs were learned more quickly than others; in cases where only a few trials were required to reach criterion, reward-related activations decreased rapidly (Fig. 3a and b), whereas in cases where learning occurred more slowly (Fig. 3c), reward responses persisted even after tens of trials. In the population of 66 dopamine neurons, mean activations following reward delivery increased nearly threefold in the first two trials with novel pictures (193% above background) and declined rapidly afterwards (to 90–110% above background), thus mirroring the learning curve of Fig. 2. Similarly, when results were analyzed with reference to criterion, reward activations were highest in the trials before the animal reached criterion (i.e. when the error rate was highest) and declined gradually thereafter (Figs 4 and 5). Differences were significant for learning versus familiar performance (Fig. 5b) and, in particular, for trials prior to reaching criterion versus subsequent learning blocks (Fig. 5b). This was also found when the learning criterion was redefined as the fourth correct response in the first series of four consecutive correct trials, which was attained after a mean of seven correct trials. Thus, the observed effect seems to be robust with respect to the criterion chosen.

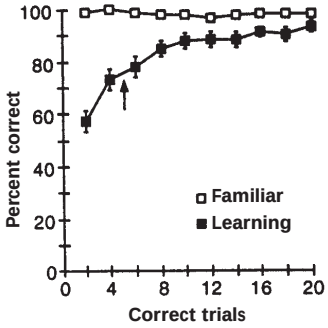

Fig. 2. Learning curves. Performance increased rapidly over successive trials with each pair of novel pictures but was stable with the familiar pictures. Percentage correct is calculated using the number of trials required to attain each two correct trials. Arrow indicates the mean number of trials required to reach the behavioral criterion for learning (second correct trial of first series of four consecutive correct trials). Data are means from 54 familiar and learning blocks in which 20 or more correct trials were performed and dopamine neurons were recorded (45 blocks) or only learning was studied (9 blocks). Each of the 54 blocks used a novel picture pair. Bars show standard errors and are not visible for familiar trial data because of their small size (0–1%).

The relation to reward prediction became further evident when the predicted reward failed to occur because of behavioral errors (Fig. 6a). Activity was significantly depressed in 70% of neurons (28 of 40) that were recorded during at least six error trials during learning. Depressions began at 99 ± 29 ms after reward would have been delivered upon correct behavioral response and lasted 401 ± 36 ms (mean ± standard

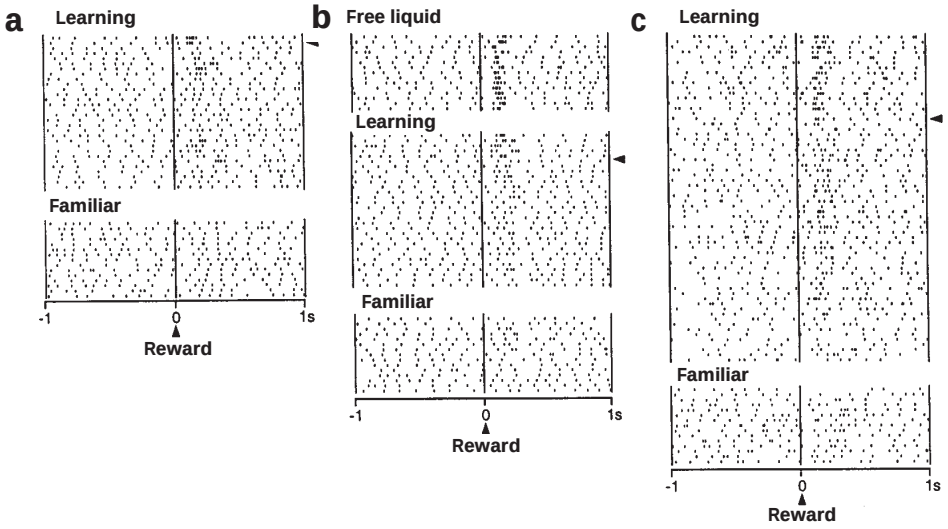

Fig. 3. Reward responses of three dopamine neurons **(a–c)** during learning of pairs of novel pictures. Reward responses decreased after the learning criterion (second of four correct responses) was reached (arrowhead). The panels (a), (b) and (c) show activity during fast, medium and slow learning episodes, in which criterion was reached after 2, 6 and 16 correct trials respectively. The bottom panels in each case show the absence of response to a predicted reward in blocks of trials with familiar pictures. The top panel in (b) shows the response to a free reward outside the context of the task. The response lever is touched at 1.0 s before reward was delivered (-1 s), except in free liquid trials. Dots denote neuronal impulses, aligned to the electric pulse that opened the liquid reward valve (center vertical lines). Individual lines represent consecutive correct trials in sequence from top to bottom. Error trials are omitted.

error). This decrease in firing suggests that the neurons signal the (negative) discrepancy between reward occurrence and reward prediction. The depression occurred at the exact time of habitual reward, suggesting that the prediction to which the neurons respond is specific with respect to the timing of the reward. In order to investigate this, we changed the time of reward delivery, using the familiar pictures to provide stable predictions over consecutive trials. The reward was normally delivered 1.0 s after a correct lever touch and did not lead to a change in firing rate. When reward was suddenly delayed by 0.5 s, i.e. 1.5 s after the lever touch, however, dopamine neurons exhibited significant depressions at the usual time of reward (9 of 14 neurons) and in addition were activated by reward at the new time (8 of the 14 neurons) (Fig. 6b). When reward was delivered at 0.5 s earlier than usual, i.e. 0.5 s after lever touch, dopamine neurons were activated by reward at that time, but showed no depression at the usual time of reward (6 of 8 neurons).

Discussion

The activations of individual dopamine neurons by reward were inversely related to the progress of learning in a given learning episode, progressively decreasing as stimuli were learned and vanishing as performance was consolidated. The rapid change of responsiveness to reward observed in these discrete learning episodes indicates that dopamine neurons have a considerable degree of flexibility. Learning curves reflect the degree of reinforcer unpredictability[1–4], and dopamine neuron activations seem to reflect this, by encoding the unpredictability of reward occurrence and tracking its decrease as learning progresses. The progressive decrease in responsiveness suggests that dopamine neurons may also code reward unpredictability in a quantitative way. Consistent with this possibility, we observed that reward activations during early trials in individual learning episodes were not as great as responses to unpredicted free liquid; this may reflect the partial (50%) reward prediction inherent in the repeated two-choice learning situation. The activations by unpredicted free liquid that was given independent of the task suggest that unpredicted rewards are reported irrespective of learning context.

The results from reward omission at predicted times suggest that the predictions influencing dopamine neurons include not only the occurrence but also the time of reward. Dopamine neurons show a positive response (activation) when a reward is not predicted or when it occurs at an unpredicted time, they

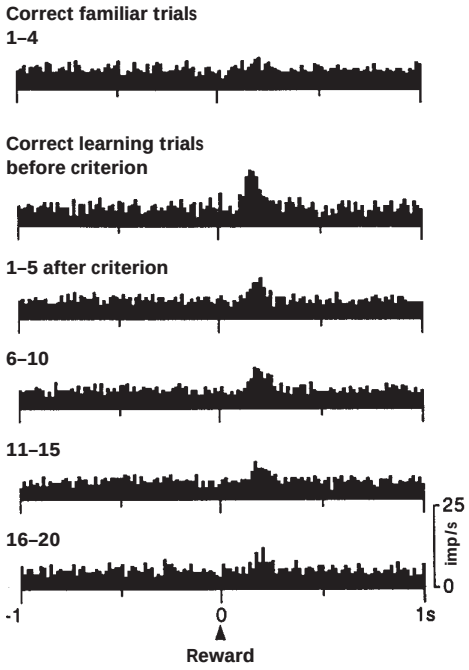

Fig. 4. Changes of average population response (54 neurons tested) to reward during learning. Trials with familiar pictures are also shown for comparison. Note the absence of response to reward with familiar pictures, strong activations during initial learning trials before reaching the criterion and progressive decrease after the criterion. Learning data are from episodes with at least 20 correct trials. Average population activity is shown for the first five trials (familiar pictures, top panel), for the total set of correct trials before criterion (second panel) or for sets of 5 consecutive correct trials at different stages (first to fifth, sixth to tenth, etc.) after criterion was reached (bottom four panels).

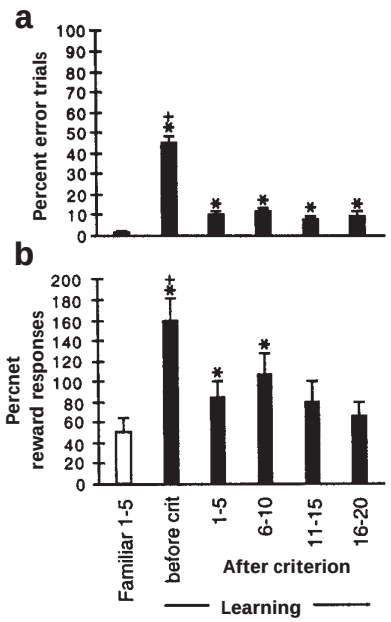

Fig. 5. Comparison between progress of learning and neuronal responses to reward. **(a)** Behavioral performance, at chance during the initial learning period before reaching criterion and improving rapidly after criterion is reached. **(b)** Responses of dopamine neurons during the same sessions. Firing rate is expressed as a percentage increase over baseline activity for the same neuron (thus, 200% represents 3 x baseline firing), and is averaged for 54 neurons during the time interval 130–220 ms after reward. In (a) and (b), differences were significant between blocks of five consecutive correct trials at $p < 0.025$, ANOVA before criterion versus subsequent blocks; * $p < 0.05$ familiar versus all learning blocks, post-hoc Fisher test; + $p < 0.05$ learning before criterion versus all subsequent blocks.

are uninfluenced when rewards occur at a predicted time, and they display a negative response (depression) when a reward fails to occur at the predicted time, unless the reward has already been delivered earlier than the predicted time. Like rewards themselves, a conditioned stimulus that has been associated with reward will evoke a response, provided that its occurrence is unpredicted. No response occurs, however, if the conditioned stimulus is predicted[18]. This suggests that the coding of temporal prediction errors may apply to both unconditioned and conditioned appetitive events. Dopamine neurons therefore seem to be sensitive and flexible detectors of errors in reward prediction, signaling not that an appetitive outcome has occurred but that this outcome is different than predicted at this exact moment. Moreover, this signal is updated very rapidly in response to changing expectations and reward occurrence. This computation is presumably the result of synaptic inputs to dopamine neurons, possibly from the striatum[23].

The temporal aspects of predictions reflected in dopamine responses correspond well to the temporal components of predictions observed in animal learning experiments[6,7]. Dopamine responses show several essential characteristics of teaching signals of temporal-difference models of learning[23–26]. They are increased when the primary reinforcement is unpredicted, they are also increased in response to secondary reinforcers that become associated with the primary reward as a result of learning, and they are reduced when a predicted reinforcement fails to occur. The potential influence of dopamine on synaptic plasticity[27–29] would allow the dopamine neurons of the substantia nigra and ventral tegmental area to provide a teaching signal for modifying synaptic transmission in the striatum and frontal cortex. This may constitute a neuronal mechanism contributing to the well established role of dopamine in reward-driven behavior and in appetitive learning[12–14].

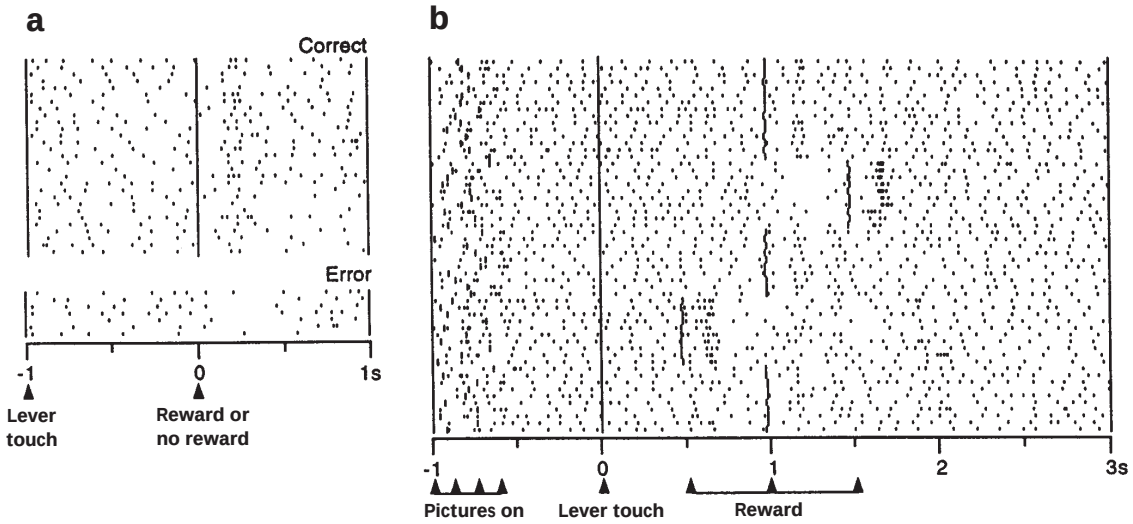

Fig. 6. Responses of dopamine neurons related to errors in the temporal prediction of reward. **(a)** Comparison of correct and error trials, showing effect of behaviorally determined reward occurrence. Correct learning trials lead to reward and to neuronal activation (top panel). Behavioral errors lead to reward omission and induce a depression in the same dopamine neuron at the time that reward is normally given (bottom panel). Activity in error trials is aligned to lever touch, which in correct trials would be followed after 1.0 s by reward. **(b)** Effects of reward timing during familiar trials. Following a correct response, the reward was delivered after 1.0 s (as expected), 1.5 s (delayed) or 0.5 s (early) Activity of a dopamine neuron was depressed when reward was delayed and increased at the new time of delayed or precocious reward (familiar trials). Reward delivery is marked by a longer line; the slight jitter reflects the fact that the interval between the lever press (to which the traces are aligned) and the reward delivery was not controlled with absolute precision (timing varies ± 8 ms). 'Pictures on' indicates the various times of appearance of pictures, as indicated by small vertical line in each raster. In each panel, original trial sequence is from top to bottom.

Only a few neuronal populations other than dopamine neurons are so far known to process reward information. These include the dorsal and ventral striatum[30–36], subthalamic nucleus[37], amygdala[38], dorsolateral prefrontal cortex[39,40], orbitofrontal cortex[41] and anterior cingulate cortex[42]. Some neurons in these structures are activated during the expectation of rewards and in response to rewards and reward-predicting stimuli. Also, noradrenaline neurons of locus coeruleus respond to a wide variety of attention-inducing stimuli, including rewards and punishments[43–45], and cholinergic neurons of nucleus basalis are activated by many rewarding and punishing events[46,47]. Nevertheless, despite the observation that some of these activities depend on reward unpredictability[36] or behavioral errors[39,42], they do not seem to code a reward prediction error in the way that dopamine neurons do; moreover, noradrenergic and cholinergic neurons differ from dopamine neurons in that rewards are not their strongest stimuli. Thus dopamine neurons display unique response characteristics, which can be conceptualized as the coding of temporal-prediction errors according to formal learning algorithms.

Methods

Experiments were performed on two monkeys (*Macaca fascicularis*), which were moderately fluid deprived to increase their motivation to work for juice rewards. In the discrimination task, two color pictures ($13 \times 13°$) appeared on a computer screen to the left and right of center (Fig. 1a). The monkey released a key, touched a small lever below one picture within 1 s and, if it selected the correct picture, received 0.15 ml of apple juice 1.0 s later. No reward occurred if no response or an incorrect response occurred. Pictures varied randomly between left and right positions and extinguished upon lever touch. Liquid arrived at the mouth 55 ms after the electronic pulse activated the liquid valve. Trials lasted 4–6 s; inter-trial intervals were 4–6 s. In free-liquid trials, animals received 0.15 ml of apple juice at irregular intervals outside of any specific task.

Animals first learned the discrimination task with fractal pictures as the familiar stimuli. Initially, animals reacted to a single, rewarded picture presented ipsilateral to the working hand, then contralaterally and subsequently in random alternation. Finally, the unrewarded picture was added at gradually increasing sizes until it matched the rewarded picture. After the fractal picture pair was fully learned, animals successively learned many pairs of novel pictures, similar to previous learning studies[48]. These pairs of pictures were presented together at equal size, varying randomly between left and right locations. The animal had to learn by trial and error which was the correct picture in each pair. Each learning picture was randomly chosen from 65536 possibilities, consisting of one of 64 yellow, red, green or blue alphanumeric symbols superimposed on one of 64 simple forms in yellow, red, green or blue (Fig. 1b). Dopamine neurons were first studied during learning of the discrimination task using fractal pictures. Subsequently, every dopamine neuron was studied in separate, randomly alternating blocks of trials with the familiar picture pair, a novel picture pair and free liquid.

As described[15,17,18], activity from single dopamine neurons was recorded extracellularly during 20–60 minutes in two monkeys using standard electrophysiological techniques. Dopamine neurons discharged polyphasic, initially negative or positive impulses with relatively long durations (1.8–5.5 ms) and low frequencies (2.0–8.5 impulses per s). Impulses contrasted with those of the non-dopaminergic pars reticulata neurons of substantia nigra (70–90 impulses per s and < 1.1 ms duration) and neighboring fibers (< 0.4 ms duration). Neuronal activity changes were compared against a 500-ms control period before the first event in each trial by using a Wilcoxon procedure with a constant time window of 130–220 ms following liquid reward, comprising 80% of onset and offset times of statistically significant increases[15,17,18] ($p < 0.01$). Magnitudes of activation were calculated in the constant time window for every neuron tested, independent of response significance. Recording sites of dopamine neurons randomly sampled from groups A8, A9 and A10 were marked with small electrolytic lesions and reconstructed from 40-μm thick, tyrosine hydoxylase-immunoreacted or cresyl violet-stained coronal brain sections. Experimental protocols conformed to the Swiss Animal Protection Law and were supervised by the Fribourg Cantonal Veterinary Office.

Acknowledgements

We thank Anthony Dickinson, David Gaffan and P. Read Montague for helpful discussions and advice, and B. Aebischer, J. Corpataux, A. Gaillard, A. Pisani, A. Schwarz and F. Tinguely for expert technical assistance. Supported by Swiss NSF, Roche Research Foundation and NIMH postdoctoral fellowship to J.R.H.

RECEIVED 1 APRIL; ACCEPTED 16 JUNE 1998

1. Rescorla, R. A. & Wagner, A. R. in *Classical Conditioning II: Current Research and Theory* (eds Black, A. H. & Prokasy, W. F.) 64–99 (Appleton Century Crofts, New York, 1972).
2. Dickinson, A. *Contemporary Animal Learning Theory* (Cambridge Univ. Press, Cambridge, 1980).
3. Mackintosh, N. J. A theory of attention: Variations in the associability of stimulus with reinforcement. *Psychol. Rev*. 82, 276–298 (1975).
4. Pearce, J. M. & Hall, G. A model for Pavlovian conditioning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol. Rev*. 87, 532–552 (1980).
5. Sutton, R. S. & Barto, A. G. Toward a modern theory of adaptive networks: expectation and prediction. *Psychol. Rev*. 88, 135–170 (1981).
6. Smith, M. C. CS-US interval and US intensity in classical conditioning of the rabbit's nictitating membrane response. *J. Comp. Physiol. Psychol*. 66, 679–687 (1968).
7. Dickinson, A., Hall, G. & Mackintosh, N. J. Surprise and the attenuation of blocking. *J. Exp. Psychol. Anim. Behav. Proc*. 2, 313–322 (1976).
8. Sutton, R. S. Learning to predict by the method of temporal difference. *Machine Learning* 3, 9–44 (1988).
9. Barto, A. G., Sutton, R. S. & Anderson, C. W. Neuronlike adaptive elements that can solve difficult learning problems. *IEEE Trans Syst. Man Cybernet. SMC* 13, 834–846 (1983).
10. Tesauro, G. TD-Gammon, a self-teaching backgammon program, achieves master-level play. *Neural Comp*. 6, 215–219 (1994).
12. Wise, R. A. Neuroleptics and operant behavior: The anhedonia hypothesis. *Behav. Brain Sci*. 5, 39–87 (1982).
13. Fibiger, H. C. & Phillips, A. G. in *Handbook of Physiology - The Nervous System*, vol IV (ed. Bloom, F. E.) 647–675 (Williams & Wilkins, Baltimore, 1986).
14. Robbins, T. W. & Everitt, B. J. Neurobehavioural mechanisms of reward and motivation. *Curr. Opin. Neurobiol*. 6, 228–236 (1996).
15. Mirenowicz, J. & Schultz, W. Importance of unpredictability for reward responses in primate dopamine neurons. *J. Neurophysiol*. 72, 1024–1027 (1994).
16. Schultz, W., Dayan, P. & Montague, R. R. A neural substrate of prediction and reward. *Science* 275, 1593–1599 (1997).
17. Ljungberg, T., Apicella, P. & Schultz, W. Responses of monkey dopamine neurons during learning of behavioral reactions. *J. Neurophysiol*. 67, 145–163 (1992).
18. Schultz, W., Apicella, P. & Ljungberg, T. Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *J. Neurosci*. 13, 900–913 (1993).
19. Steinfels, G. F., Heym, J., Strecker, R. E & Jacobs, B. L. Behavioral correlates of dopaminergic unit activity in freely moving cats. *Brain Res*. 258, 217–228 (1983).
20. Schultz, W. & Romo, R. Dopamine neurons of the monkey midbrain: Contingencies of responses to stimuli eliciting immediate behavioral reactions. *J. Neurophysiol*. 63, 607–624 (1990).
21. Mirenowicz, J. & Schultz, W. Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature* 379, 449–451 (1996).
22. Horvitz, J. C., Stewart, T. & Jacobs, B. L. Burst activity of ventral tegmental dopamine neurons is elicited by sensory stimuli in the awake cat. *Brain Res*. 759, 251–258 (1997).
23. Houk, J. C., Adams, J. L. & Barto, A. G. in *Models of Information Processing in the Basal Ganglia* (eds Houk, J. C., Davis, J. L. & Beiser, D. G.) 249–270 (MIT Press, Cambridge, 1995).
24. Montague, P. R., Dayan, P., Person, C. & Sejnowski, T. J. Bee foraging in uncertain environments using predictive hebbian learning. *Nature* 377, 725–728 (1995).
25. Montague, P. R., Dayan, P. & Sejnowski, T. J. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci*. 16, 1936–1947 (1996).
26. Suri, R. E. & Schultz, W. Learning of sequential movements with dopamine-like reinforcement signal in neural network model. *Exp. Brain Res*. (in press).

27. Calabresi, P., Maj, R., Pisani, A., Mercuri, N. B. & Bernardi, G. Long-term synaptic depression in the striatum: Physiological and pharmacological characterization. *J. Neurosci*. 12, 4224–4233 (1992).
28. Wickens, J. R., Begg, A. J. & Arbuthnott, G. W. Dopamine reverses the depression of rat corticostriatal synapses which normally follows high-frequency stimulation of cortex in vitro. *Neuroscience* 70, 1–5 (1996).
29. Calabresi, P. *et al.* Abnormal synaptic plasticity in the striatum of mice lacking dopamine D2 receptors. *J. Neurosci*. 17, 4536–4544 (1997).
30. Hikosaka, O., Sakamoto, M. & Usui, S. Functional properties of monkey caudate neurons. III. Activities related to expectation of target and reward. *J. Neurophysiol*. 61, 814–832 (1989).
31. Apicella, P., Ljungberg, T., Scarnati, E. & Schultz, W. Responses to reward in monkey dorsal and ventral striatum. *Exp. Brain Res*. 85, 491–500 (1991).
32. Schultz, W., Apicella, P., Scarnati, E. & Ljungberg, T. Neuronal activity in monkey ventral striatum related to the expectation of reward. *J. Neurosci*. 12, 4595–4610 (1992).
33. Williams, G. V., Rolls, E. T., Leonard, C. M. & Stern, C. Neuronal responses in the ventral striatum of the behaving monkey. *Behav. Brain Res*. 55, 243–252 (1993).
34. Aosaki, T. *et al.* Responses of tonically active neurons in the primate's striatum undergo systematic changes during behavioral sensorimotor conditioning. *J. Neurosci*. 14, 3969–3984 (1994).
35. Bowman, E. M., Aigner, T. G. & Richmond, B. J. Neural signals in the monkey ventral striatum related to motivation for juice and cocaine rewards. *J. Neurophysiol*. 75, 1061–1073 (1996).
36. Apicella, P., Legallet, E. & Trouche, E. Responses of tonically discharging neurons in the monkey striatum to primary rewards delivered during different behavioral states. *Exp. Brain Res*. 116, 456–466 (1997).
37. Matsumura, M., Kojima, J., Gardiner, T. W. & Hikosaka, O. Visual and oculomotor functions of monkey subthalamic nucleus. *J. Neurophysiol*. 67, 1615–1632 (1992).
38. Nishijo, H., Ono, T. & Nishino, H. Topographic distribution of modality-specific amygdalar neurons in alert monkey. *J. Neurosci*. 8, 3556–3569 (1988).
39. Watanabe, M. The appropriateness of behavioral responses coded in post-trial activity of primate prefrontal units. *Neurosci. Lett*. 101, 113–117 (1989).
40. Watanabe, M. Reward expectancy in primate prefrontal neurons. *Nature* 382, 629–632 (1996).
41. Thorpe, S. J., Rolls, E. T. & Maddison, S. The orbitofrontal cortex: neuronal activity in the behaving monkey. *Exp. Brain Res*. 49, 93–115 (1983).
42. Niki, H. & Watanabe, M. Prefrontal and cingulate unit activity during timing behavior in the monkey. *Brain Res*. 171, 213–224 (1979).
43. Aston-Jones, G. & Bloom, F. E. Norepinephrine-containing locus coeruleus neurons in behaving rats exhibit pronounced responses to nonnoxious environmental stimuli. *J. Neurosci*. 1, 887–900 (1981).
44. Sara, S. J. & Segal, M. Plasticity of sensory responses of locus coeruleus neurons in the behaving rat: implications for cognition. *Prog. Brain Res*. 88, 571–585 (1991).
45. Aston-Jones, G., Rajkowski, J., Kubiak, P. & Alexinsky, T. Locus coeruleus neurons in monkey are selectively activated by attended cues in a vigilance task. *J. Neurosci*. 14, 4467–4480 (1994).
46. Richardson, R. T. & DeLong, M. R. Nucleus basalis of Meynert neuronal activity during a delayed response task in monkey. *Brain Res*. 399, 364–368 (1986).
47. Wilson, F. A. W. & Rolls, E. T. Neuronal responses related to reinforcement in the primate basal forebrain. *Brain Res*. 509, 213–231 (1990).
48. Gaffan, E. A., Gaffan, D. & Harrison, S. Disconnection of the amygdala from visual association cortex impairs visual reward association learning in monkeys. *J. Neurosci*. 8, 3144–3150 (1988).