MULTIPLE REWARD SIGNALS IN THE BRAIN

Wolfram Schultz

The fundamental biological importance of rewards has created an increasing interest in the neuronal processing of reward information. The suggestion that the mechanisms underlying drug addiction might involve natural reward systems has also stimulated interest. This article focuses on recent neurophysiological studies in primates that have revealed that neurons in a limited number of brain structures carry specific signals about past and future rewards. This research provides the first step towards an understanding of how rewards influence behaviour before they are received and how the brain might use reward information to control learning and goal-directed behaviour.

GOAL-DIRECTED BEHAVIOUR
Behaviour controlled by representation of a goal or an understanding of a causal relationship between behaviour and attainment of a goal.

REINFORCERS
Positive reinforcers (rewards) increase the frequency of behaviour leading to their acquisition. Negative reinforcers (punishers) decrease the frequency of behaviour leading to their encounter and increase the frequency of behaviour leading to their avoidance.

*Institute of Physiology and Program in Neuroscience, University of Fribourg, CH-1700 Fribourg, Switzerland.
e-mail:
Wolfram.Schultz@unifr.ch*

The fundamental role of reward in the survival and well-being of biological agents ranges from the control of vegetative functions to the organization of voluntary, GOAL-DIRECTED BEHAVIOUR. The control of behaviour requires the extraction of reward information from a large variety of stimuli and events. This information concerns the presence and values of rewards, their predictability and accessibility, and the numerous methods and costs associated with attaining them.

Various experimental approaches including brain lesions, psychopharmacology, electrical self-stimulation and the administration of addictive drugs, have helped to determine the crucial structures involved in reward processing[1–4]. In addition, physiological methods such as *in vivo* microdialysis, voltammetry[5–9] and neural imaging[10–12] have been used to probe the structures and neurotransmitters that are involved in processing reward information in the brain. However, I believe that the temporal constraints imposed by the nature of the reward signals themselves might be best met by studying the activity of single neurons in behaving animals, and it is this approach that forms the basis of this article. Here, I describe how neurons detect rewards, learn to predict future rewards from past experience and use reward information to learn, choose, prepare and execute goal-directed behaviour (FIG. 1). I also attempt to place the processing of drug rewards within a general framework of neuronal reward mechanisms.

Behavioural functions of rewards

Given the dynamic nature of the interactions between complex organisms and the environment, it is not surprising that specific neural mechanisms have evolved that not only detect the presence of rewarding stimuli but also predict their occurrence on the basis of representations formed by past experience. Through these mechanisms, rewards have come to be implicit or explicit goals for increasingly voluntary and intentional forms of behaviour that are likely to lead to the acquisition of goal objects.

Rewards have several basic functions. A common view is that rewards induce subjective feelings of pleasure and contribute to positive emotions. Unfortunately, this function can only be investigated with difficulty in experimental animals. Rewards can also act as positive REINFORCERS by increasing the frequency and intensity of behaviour that leads to the acquisition of goal objects, as described in CLASSICAL and INSTRUMENTAL CONDITIONING PROCEDURES. Rewards can also maintain learned behaviour by preventing EXTINCTION[13]. The rate of learning depends on the discrepancy between the occurrence of reward and the predicted occurrence of reward, the so-called 'reward prediction error'[14–16] (BOX 1).

Rewards can also act as goals in their own right and can therefore elicit approach and consummatory behaviour. Objects that signal rewards are labelled with positive MOTIVATIONAL VALUE because they will elicit effortful

behavioural responses. These motivational values arise either through innate mechanisms or, more often, through learning. In this way, rewards help to establish value systems for behaviour and serve as key references for behavioural decisions.
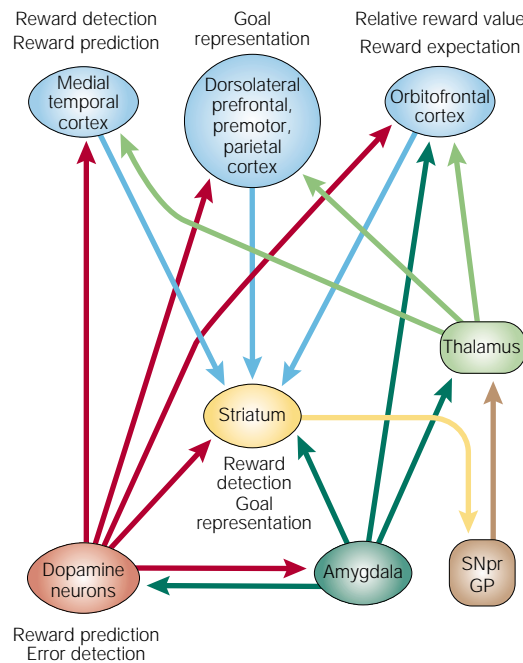

Figure 1 | **Reward processing and the brain.** Many reward signals are processed by the brain, including those that are responsible for the detection of past rewards, the prediction and expectation of future rewards, and the use of information about future rewards to control goal-directed behaviour. (SNpr, substantia nigra pars reticulata; GP, globus pallidus.)

Box 1 | Reward prediction errors

Behavioural studies show that reward-directed learning depends on the predictability of the reward[14–16]. The occurrence of the reward has to be surprising or unpredicted for a stimulus or action to be learned. The degree to which a reward cannot be predicted is indicated by the discrepancy between the reward obtained for a given behavioural action and the reward that was predicted to occur as a result of that action. This is known as the prediction error and underlies a class of error-driven learning mechanisms[33].

In simple terms, if a reward occurs unpredictably after a given action then the prediction error is positive, and learning about the consequences of the action that produced the reward occurs. However, once the consequences of that action have been learned (so that the reward that follows subsequent repetition of the action is now predicted), the prediction error falls to zero and no new information about the consequences of the action is learned. By contrast, if the expected reward is not received after repeating a learned action then the prediction error falls to a negative value and the behaviour is extinguished.

Reward detection and perception

Although there are no specialized peripheral receptors for rewards, neurons in several brain structures seem to be particularly sensitive to rewarding events as opposed to motivationally neutral events that are signalled through the same sensory modalities. A prominent example are the dopamine neurons in the pars compacta of substantia nigra and the medially adjoining ventral tegmental area (groups A8, A9 and A10). In various behavioural situations, including classical and instrumental conditioning, most dopamine neurons show short, phasic activation in a rather homogeneous fashion after the presentation of liquid and solid rewards, and visual or auditory stimuli that predict reward[17–20] (FIG. 2a, b). These phasic neural responses are common (70–80% of neurons) in medial tegmental regions that project to the nucleus accumbens and frontal cortex but are also found in intermediate and lateral sectors that project to the caudate and putamen[20]. These same dopamine neurons are also activated by novel or intense stimuli that have attentional and rewarding properties[20–22]. By contrast, only a few dopamine neurons show phasic activations to punishers (conditioned aversive visual or auditory stimuli)[23,24]. Thus, the phasic response of A8, A9 and A10 dopamine neurons seems to preferentially report rewarding and, to a lesser extent, attention-inducing events.
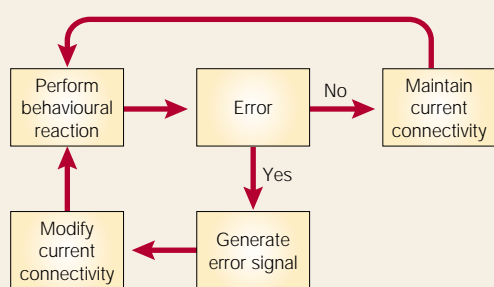
The observation of a phasic dopamine response to novel attention-inducing stimuli led to the suggestion that this response might reflect the salient attention-inducing properties of rewards rather than their positive reinforcing aspects[25,26]. However, dopamine neurons are rarely activated by strong attention-generating events such as aversive stimuli[23,24] and they are depressed rather than activated by the omission of a reward, which presumably also generates attention[20,27–29]. Of course, the possibility that the dopamine activation might encode a specific form of attention that is only associated with rewarding events cannot yet be completely ruled out.

A closer examination of the properties of the phasic dopamine response suggests that it might encode a reward prediction error rather than reward *per se*. Indeed, the phasic dopamine activity is enhanced by surprising rewards but not by those that are fully predicted, and it is depressed by the omission of predicted rewards[20,30] (FIG. 2c). The reward prediction error response also occurs during learning episodes[27]. In view of the crucial role that prediction errors are thought to play during learning[14–16], a phasic dopamine response that reports a reward prediction error might constitute an ideal teaching signal for approach learning[31]. This signal strongly resembles[32] the teaching signal used by very effective 'temporal difference' reinforcement models[33]. Indeed, neuronal networks using this type of teaching signal have learned to play backgammon at a high level[34] and have acquired serial movements and spatial delayed response tasks in a manner consistent with the performance of monkeys in the laboratory[35,36].
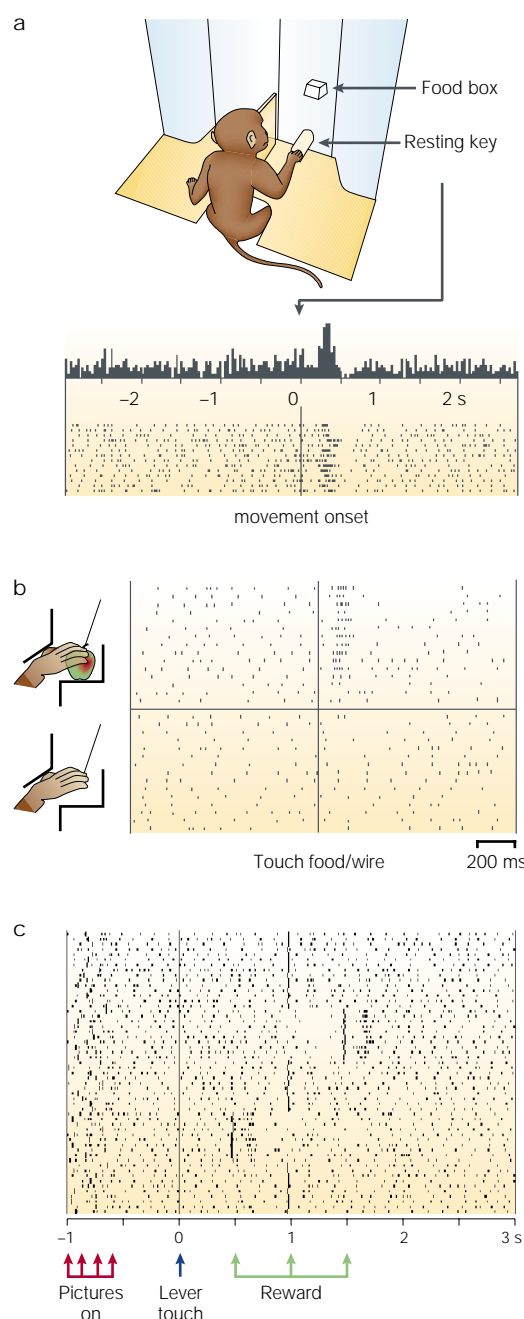
Figure 2 | **Primate dopamine neurons respond to rewards and reward-predicting stimuli.** The food is invisible to the monkey but the monkey can touch the food by placing its hand underneath the protective cover. The perievent time histogram of the neuronal impulses is shown above the raster display, in which each dot denotes the time of a neuronal impulse in reference to movement onset (release of resting key). Each horizontal line represents the activity of the same neuron on successive trials, with the first trials presented at the top and the last trials at the bottom of the raster display. **a** | Touching food reward in the absence of stimuli that predict reward produces a brief increase in firing rate within 0.5 s of movement initiation. **b** | Touching a piece of apple (top) enhances the firing rate but touching the bare wire or an inedible object that the monkey had previously encountered does not. The traces are aligned to a temporal reference point provided by touching the hidden object (vertical line). (Modified from REF. 18.) **c** | Dopamine neurons encode an error in the temporal prediction of reward. The firing rate is depressed when the reward is delayed beyond the expected time-point (1 s after lever touch). The firing rate is enhanced at the new time of reward delivery whether it is delayed (1.5 s) or precocious (0.5 s). The three arrows indicate, from left to right, the time of precocious, habitual and delayed reward delivery. The original trial sequence is from top to bottom. Data are from a two-picture discrimination task. (Figure modified with permission from REF. 27 © (1998) Macmillan Magazines Ltd.).

The phasic activation of dopamine neurons has a time course of tens of milliseconds. It is possible that this facet of the physiology of these neurons describes only one aspect of dopamine function in the brain. Feeding, drinking, punishment, stress and social behaviour result in a slower modification of the central dopamine level, which occurs over seconds and minutes as measured by electrophysiology, *in vivo* dialysis and voltammetry[5,7,37,38]. So, the dopamine system might act at several different timescales in the brain from the fast, restricted signalling of reward and some attention-inducing stimuli to the slower processing of a range of positive and negative motivational events. The tonic gating of a large variety of motor, cognitive and motivational processes that are disrupted in Parkinson's disease are also mediated by central dopamine systems.

Neurons that respond to the delivery of rewards are also found in brain structures other than the dopamine system described above. These include the striatum (caudate nucleus, putamen, ventral striatum including the nucleus accumbens)[39–44], subthalamic nucleus[45], pars reticulata of the substantia nigra[46], dorsolateral and orbital prefrontal cortex[47–51], anterior cingulate cortex[52], amygdala[53,54], and lateral hypothalamus[55]. Some reward-detecting neurons can discriminate between liquid and solid food rewards (orbitofrontal cortex[56]), determine the magnitude of rewards (amygdala[57]) or distinguish between rewards and punishers (orbitofrontal cortex[58]). Neurons that detect rewards are more common in the ventral striatum than in the caudate nucleus and putamen[40]. Reward-discriminating neurons in the lateral hypothalamus and the secondary taste area of the orbitofrontal cortex decrease their response to a particular food upon satiation[59,60]. By contrast, neurons in the primary taste area of the orbitofrontal cortex continue to respond during satiety and thus seem to encode taste identity rather than reward value[61].

Most of the reward responses described above occur in well-trained monkeys performing familiar tasks, regardless of the predictive status of the reward. Some neurons in the dorsolateral and orbital prefrontal cortex respond preferentially to rewards that occur unpredictably outside the context of the behavioural task[50,51,62,63] or during the reversal of reward associations to visual stimuli[58]. However, neurons in these structures do not project in a widespread, divergent fashion to multiple postsynaptic targets and thus do not seem to be able to exert a global reinforcing influence similar to that described for the dopamine neurons[29].

Other neurons in the cortical and subcortical structures mentioned above respond to conditioned reward-predicting visual or auditory stimuli[41,51,53,58,64–66], and discriminate between reward-predicting and non-reward-predicting stimuli[27,51]. Neurons within the orbitofrontal cortex discriminate between visual stimuli that predict different liquid or food rewards but show few relationships to spatial and visual stimulus features[56]. Neurons in the amygdala differentiate between the visual aspects of foods and their responses decreasing with selective satiety[53].
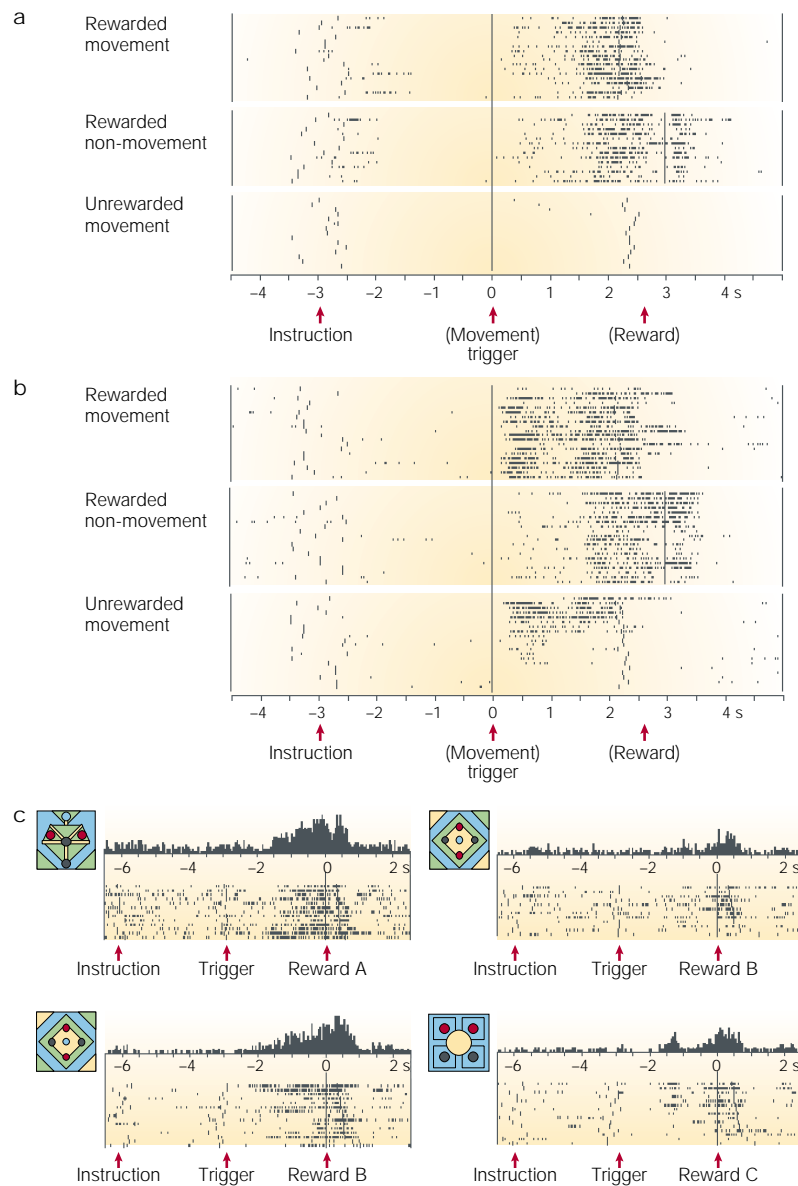
Figure 3 | **Neuronal activity in primate striatum and orbitofrontal cortex related to the expectation of reward. a** | Activity in a putamen neuron during a delayed go–no go task in which an initial cue instructs the monkey to produce or withhold a reaching movement following a trigger stimulus. The initial cue predicts whether a liquid reward or a sound will be delivered following the correct behavioural response. The neuron is activated immediately before the delivery of reward, regardless of whether a movement is required in rewarded trials, but is not activated before the sound in unrewarded movement trials (bottom). The three trial types alternated randomly during the experiment and were separated here for clarity. (Modified from REF. 71.) **b** | Adaptation of reward-expectation activity during learning in the same putamen neuron. A novel instructional cue is presented in each trial type on a computer screen. The monkey learns by trial and error which behavioural response and which reinforcer is associated with each new instructional cue. The neuron shows typical reward-expectation activity during the initial unrewarded movement trials. This activity disappears when learning advances (bottom graph; original trial sequence from top to bottom). (Modified from REF. 72.) **c** | Coding of relative reward preference in an orbitofrontal neuron. Activity increased during the expectation of the preferred food reward (reward A was a raisin; reward B was a piece of apple). Although reward B was physically identical in the top and bottom panels, its relative motivational value was different (low in the top panel, high in the bottom panel). The monkey's choice behaviour was assessed in separate trials (not shown). Different combinations of reward were used in the two trial blocks: in the top trials, A was a raisin and B was a piece of apple; in the bottom trials, B was a piece of apple and C was cereal. Each reward was specifically predicted by an instruction picture, which is shown above the histograms. (Figure modified with permission from REF. 56 © (1999) Macmillan Magazines Ltd.)

Expectation of future rewards

A well-learned reward-predicting stimulus evokes a state of expectation in the subject. The neuronal correlate of this expectation of reward might be the sustained neuronal activity that follows the presentation of a reward-predicting stimulus and persists for several seconds until the reward is delivered (FIG. 3a). This activity seems to reflect access to neuronal representations of reward that were established through previous experience.

Reward-expectation neurons are found in monkey and rat striatum[39,44,67,68], orbitofrontal cortex[51,56,69,70] and amygdala[69]. Reward-expectation activity in the striatum and orbitofrontal cortex discriminates between rewarded and unrewarded trials[51,71] (FIG. 3a) and does not occur in anticipation of other predictable task events, such as movement-eliciting stimuli or instruction cues, although such activity is found in other striatal neurons[39,67]. Delayed delivery of reward prolongs the sustained activity in most reward-expectation neurons and early delivery reduces the duration of this activity[51,67,68]. Reward-expectation neurons in the orbitofrontal cortex frequently distinguish between different liquid and food rewards[56].

Expectations change systematically with experience. Behavioural studies suggest that, when learning to discriminate rewarded from unrewarded stimuli, animals initially expect to receive a reward on all trials. Only with experience do they adapt their expectations[72,73]. Similarly, neurons in the striatum and orbitofrontal cortex initially show reward-expectation activity during all trials with novel stimuli. With experience, this activity is progressively restricted to rewarded rather than unrewarded trials[72,73] (FIG. 3b). The important point here is that the changes in neural response occur at about the same time as, or a few trials earlier than, the adaptation in behaviour.

An important issue is whether the neurons that seem to encode a given object are specific for that particular object or whether they are tuned to the rewards that are currently available or predicted. When monkeys are presented with a choice between two different rewards, neurons in the orbitofrontal cortex discriminate between the rewards on the basis of the animal's preference, as revealed by its overt choice behaviour[56]. For example, if a given neuron is more active when a preferred reward (such as a raisin) is expected rather than a non-preferred reward (such as a piece of apple) then the neuron will fire in preference to the previously non-preferred reward if the same non-preferred reward (the apple) is presented with an even less preferred reward (such as cereal) (FIG. 3c). Results such as these indicate that these neurons may not discriminate between rewards on the basis of their (unchanged) physical properties but instead encode the (changing) relative motivational value of the reward, as shown by the animal's preference behaviour. Similar preference relationships are found in neurons that can detect stimuli that predict reward[56]. Neurons that encode the relative motivational values of action outcomes in given behavioural situations might provide important information for neuronal mechanisms in the frontal lobe that underlie goal-directed behavioural choices.
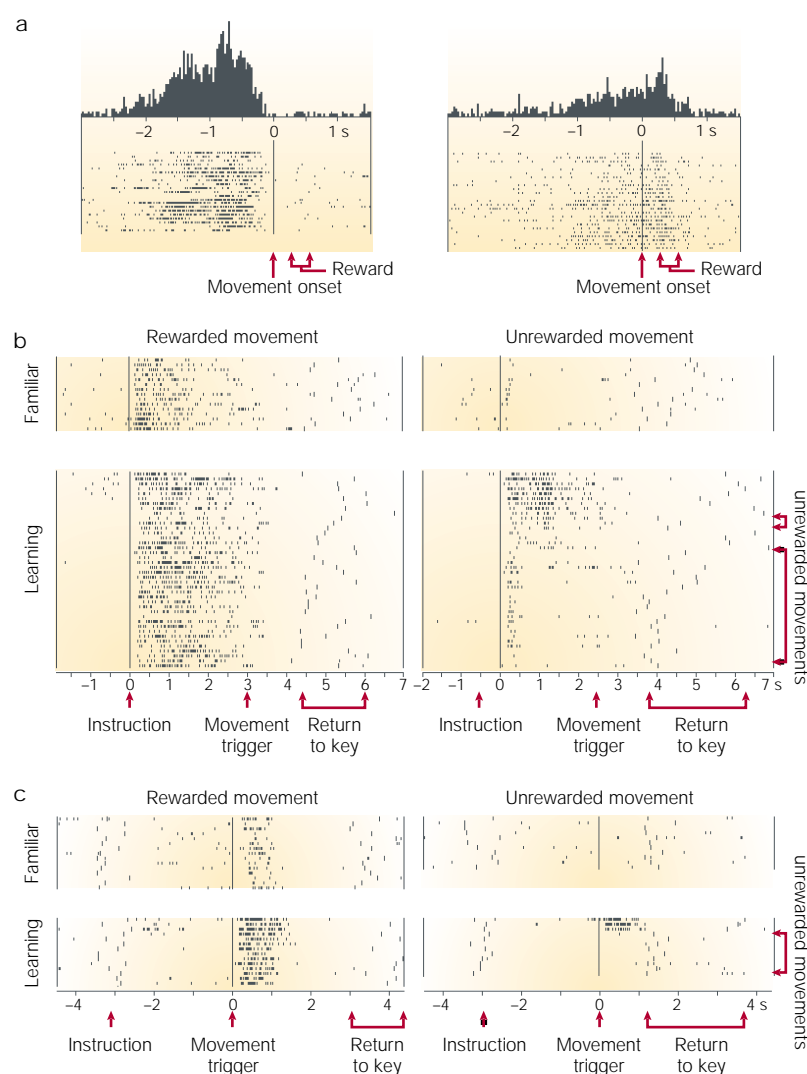
Figure 4 | **Behaviour-related activity in the primate caudate reflects future goals. a** | Self-initiated arm movements in the absence of movement triggering external stimuli. Activation stops before movement starts (left) or, in a different neuron (right), on contact with the reward. (Modified from REF. 77.) **b** | Movement preparatory activity in a delayed go–no go task during trials with familiar stimuli (top) and during learning trials with novel instruction cues (bottom). The neuron shows selective, sustained activation during familiar trials when reward is predicted in movement trials (left) but no activation in unrewarded movement trials (right) and rewarded non-movement trials (not shown). By contrast, during learning trials with unrewarded movement, the activation initially occurs when movements are performed that are reminiscent of rewarded movement trials and subsides when movements are performed appropriately (arrows to the right). **c** | Reward-dependent movement-execution activity showing changes during learning that are similar to the preparatory activity in (**b**). (Modified from REF. 71.)

From reward expectation to goal direction

Expected rewards might serve as goals for voluntary behaviour if information about the reward is present while the behavioural response towards the reward is being prepared and executed[74]. This would require the integration of information about the expectation of reward with processes that mediate the behaviour leading to reward acquisition. Recent experiments provide evidence for such mechanisms in the brain.

Neurons in the striatum, in the supplementary motor area and in the dorsolateral premotor cortex show enhanced activity for up to 3 s before internally initiated movements towards expected rewards in the absence of external movement-triggering stimuli[75–78]. In some of these neurons, the enhanced activity continues until the reward is collected (FIG. 4a), possibly reflecting an expectation of reward before the onset of movement. In other neurons, the enhanced activity stops when movement starts or slightly earlier and is therefore probably related to movement initiation. The finding that some of these neurons are not activated when the movement is elicited by external stimuli suggests a role in the internal rather than the stimulus-dependent control of behaviour.

Delayed response tasks are assumed to test several aspects of working memory and the preparation of movement. Many neurons in the dorsolateral prefrontal cortex show sustained activity during the delay period between the sample being presented and the choice phase of the task. This delay seems to reflect information about the preceding cue and the direction of the upcoming movement. The activity might vary when different food or liquid rewards are predicted by the cues[79] or when different quantities of a single liquid reward are predicted[80]. Neurons in the caudate nucleus also show directionally selective, delay-dependent activity in oculomotor tasks. This activity seems to depend on whether or not a reward is predicted and occurs preferentially for movement towards the reward[81].

Some neurons in the anterior striatum show sustained activity during the movement preparatory delay. The activity stops on movement onset and is not observed during the period preceding the reward. Interestingly, this activity is much more commonly observed in rewarded than in unrewarded trials (FIG. 4b, top). A similar reward dependency is seen during the execution of movement[71] (FIG. 4c, top). When novel instructional cues that predict a reward rather than non-reward are learned in such tasks, reward-dependent movement preparation and execution activity develop in parallel with the behavioural manifestation of reward expectation[72] (FIG. 4b, c, bottom). These data show that an expected reward has an influence on movement preparation and execution activity that changes systematically during the adaptation of behaviour to new reward predictions. So, both the reward and the movement towards the reward are represented by these neurons. The reward-dependent activity might be a way to represent the expected reward at the neuronal level and influence neuronal processes underlying the behaviour towards the acquisition of that reward. These activities are in general compatible with a goal-directed account.

In many natural situations, a reward might be available only after the completion of a specific sequence of individual movements. Neurons in the ventral striatum and one of its afferent structures, the perirhinal cortex, report the positions of individual stimuli within a behavioural sequence and thus signal progress towards the reward. By contrast, neurons in the higher order visual area TE report the visual features of cues irrespective of reward imminence[43,44,82].

In situations involving a choice between different rewards, agents tend to optimize their choice behaviour according to the motivational values of the available rewards. Indeed, greater efforts are made for rewards with higher motivational values. Neurons in the parietal cortex show stronger task-related activity when monkeys choose larger or more frequent rewards over smaller or less frequent rewards[83]. The reward preference neurons in the orbitofrontal cortex might encode a related process[56]. Together, these neurons appear to encode the basic parameters of motivational value that underlies choice behaviour.

In addition, agents tend to modify their behaviour towards a reward when the motivational value of the reward for that behaviour changes. Neurons in the cingulate motor area of the frontal cortex might provide a neural correlate of this behaviour because they are selectively active when monkeys switch to a different movement in situations in which performing the current movement would lead to less reward[84]. These neurons do not fire if switching between movements is not associated with reduced reward or when rewards are reduced without movement switches. This might suggest a mechanism that selects movements that maximize reward.

Multiple reward signals

Information about rewards seems to be processed in different forms by neurons in different brain structures, ranging from the detection and perception of rewards, through the expectation of imminent rewards, to the use of information about predicted rewards for the control of goal-directed behaviour. Neurons that detect the delivery of rewards provide valuable information about the motivational value and identity of encountered objects. This information might also help to construct neuronal representations that permit subjects to expect future rewards according to previous experience and to adapt their behaviour to changes in reward contingencies.

Only a limited number of brain structures seem to be implicated in the detection and perception of rewards and reward-predicting stimuli (FIG. 1). The midbrain dopamine neurons report the occurrence of reward relative to its prediction and emit a global reinforcement signal to all neurons in the striatum and many neurons in the prefrontal cortex. The dopamine reward response might be the result of reward-detection activity in several basal ganglia structures, the amygdala and the orbitofrontal cortex[28]. Some neurons in the dorsal and ventral striatum, orbitofrontal cortex, amygdala and lateral hypothalamus discriminate well between different rewards, which might contribute to the perception of rewards, be involved in identifying particular objects and/or provide an assessment of their motivational value.

Reward-expectation activity might be an appropriate signal for predicting the occurrence of rewards and might thereby provide a suitable mechanism for influencing behaviour that leads to the acquisition of rewards. Although rewards are received only after the completion of behavioural responses, they could influence behaviour before reward acquisition through this mechanism. This retrograde action might be mediated by the reward-predicting response of the dopamine neurons and by the reward-expectation neurons in the striatum and orbitofrontal cortex.

Neurons in the ventral striatum have more reward-related responses and reward-expectation activity than neurons in the caudate nucleus and putamen[40,68]. Although these findings might provide a neurophysiological correlate for the known motivational functions of the ventral striatum, the precise characterization of the neurophysiology of the various subregions of nucleus accumbens that have been suggested by behavioural studies[85] remains to be explored.

The striatal reward activities and dopamine reward inputs might influence behaviour-related activity in the prefrontal, parietal and perirhinal cortices through the different cortex–basal-ganglia–cortex loops. The orbitofrontal sectors of the prefrontal cortex seem to be the most involved in reward processing[56,60] but reward-related activity is found also in the dorsolateral area[79,80]. The above analysis suggests that the orbitofrontal cortex is strongly engaged in the details of the detection, perception and expectation of rewards, whereas the dorsolateral areas might use reward information to prepare, plan, sequence and execute behaviour directed towards the acquisition of rewarding goals.

The different reward signals might also cooperate in reward-directed learning. Dopamine neurons provide an effective teaching signal but do not seem to specify the individual reward, although this signal might modify the activity of synapses engaged in behaviour leading to the reward. Neurons in the orbitofrontal cortex, striatum and amygdala might identify particular rewards. Some of these neurons are sensitive to reward prediction and might be involved in local learning mechanisms[29]. The reward-expectation activity in cortical and subcortical structures adapts during learning according to changing reward contingencies. In this way, different neuronal systems are involved in different mechanisms underlying the various forms of adaptive behaviour directed towards rewards.

Drug rewards

Is it possible to use this expanding knowledge of the neural substrates of reward processing to gain an understanding of how artificial drug rewards exert their profound influence on behaviour? Psychopharmacological studies have identified the dopamine system and the ventral striatum, including the nucleus accumbens, as some of the critical structures on which drugs of abuse act[2]. These very structures are also implicated in the processing of rewards outlined above. One important question is thus whether drugs of abuse modify existing neuronal responses to natural rewards or constitute rewards in their own right and, as such, engage existing neuronal reward mechanisms.

There is evidence that the firing pattern of some neurons in the nucleus accumbens can be enhanced or depressed by the self-administration of cocaine[86,87].

Box 2 | **Cognitive rewards**

In contrast to basic rewards, which satisfy vegetative needs or relate to the reproduction of the species, the evolution of abstract, cognitive representations facilitates the use of a different class of rewards. Objects, situations and constructs such as novelty, money, challenge, beauty, acclaim, power, territory and security constitute this different class of rewards, which have become central to our understanding of human behaviour.

  Do such higher order, cognitive rewards use the same neuronal reward mechanisms and structures as the more basic rewards? If rewards have common functions in learning, approach behaviour and positive emotions, could there be common reward-processing mechanisms at the level of single neurons, neuronal networks or brain structures? Do the neurons that detect or expect basic rewards show similar activities with cognitive rewards? Although cognitive rewards are difficult to investigate in the animal laboratory, it might be possible to adapt higher rewards such as novel or interesting images to the study of brain function with psychopharmacological or neurophysiological methods. Neuroimaging might offer the best tools for investigating the neural substrates of cognitive rewards.

Neurons can also discriminate between cocaine and liquid rewards[43,88], or between cocaine and heroin[89], possibly even better than between natural rewards[90]. There is also evidence that some of these responses indicate drug detection regardless of the action of the animal[86,91,92].

Other neurons in the nucleus accumbens show activity that is sustained for several seconds when the subject approaches a lever that can be pressed to receive a cocaine injection and stops rapidly when the lever is pressed[86,87,92]. Although some neurons might be activated by lever pressing regardless of drug injection[91,92], others are activated only by pressing a lever for specific drug or natural rewards[88–90]. These activations might reflect drug-related approach behaviour or preparation for lever pressing but could also be related to the expectation of drug reward.

Existing behaviour-related activity may also be affected by drugs. The administration of amphetamine increases the phasic movement-related activity of neurons in the striatum[93]. This reveals an important influence of dopamine reuptake-blocking drugs that might also hold for cocaine.

Drugs might also exert long-lasting effects on sequences of behaviour. Some neurons in the nucleus accumbens are depressed for several minutes after cocaine administration[94,95]. This depression follows the time course of an increased dopamine concentration in this area and might reflect cocaine-induced blockade of dopamine reuptake. This effect is too slow to be related to individual movements but could provide a more general influence on drug-taking behaviour.

From the data reviewed above, some neurons do indeed seem to treat drugs as rewards in their own right. Neuronal firing is activated by drug delivery and some neurons also show anticipatory activity during drug expectation periods, as has been observed with natural rewards. Behaviour-related activity also seems to be influenced by drugs and natural rewards in a somewhat similar manner, possibly reflecting a goal-directed mechanism by which the expected outcome influences the neuronal activity related to the behaviour leading to that outcome. Drug rewards might thus exert a chemical influence on postsynaptic neurons of the reward-processing structures that in some way mimics

or corresponds to the influence of neurotransmitters released by natural rewards, including dopamine. This would be consistent with the observation that many drugs of abuse modulate dopamine neurotransmission. Drugs of abuse that mimic or boost the phasic dopamine reward prediction error might generate a powerful teaching signal and might even produce lasting behavioural changes through synaptic modifications. The reward value of drugs of abuse might also be assessed in other, currently unknown, areas that modulate executive centres via existing reward pathways, including dopamine neurons. Drugs and natural rewards might give rise to a final, convergent reward predicting message that influences behaviour-related activity.

Conclusions

A limited number of brain structures process reward information in several different ways. Neurons detect reward prediction errors and produce a global reinforcement signal that might underlie the learning of appropriate behaviours. Future work will investigate how such signals lead to synaptic modifications[96–99]. Other neurons detect and discriminate between different rewards and might be involved in assessing the nature and identity of individual rewards, and might thus underlie the perception of rewards.

However, the brain not only detects and analyses past events; it also constructs and dynamically modifies predictions about future events on the basis of previous experience[29,31]. Neurons respond to learned stimuli that predict rewards and show sustained activities during periods in which expectations of rewards are evoked. They even estimate future rewards and adapt their activity according to ongoing experience[72,73].

Once we understand more about how rewards are processed by the brain, we can begin to investigate how reward information is used to produce behaviour that is directed towards rewarding goals. Neurons in structures that are engaged in the control of behaviour seem to incorporate information about predicted rewards when coding specific behaviours that are directed at obtaining the rewards. Future research might permit us to understand how such activity leads to reward preferences, choices and decisions that incorporate both the cost and the benefit of behaviour. This investigation should not exclude the 'higher' or more cognitive rewards that typify human behaviour (BOX 2). The emerging imaging methods should produce a wealth of important information in this domain[10–12]. The ultimate goal should be to obtain a comprehensive understanding of the interconnected brain systems that deal with all aspects of reward that are relevant for explaining goal-directed behaviour. This would contribute to our understanding of neuronal mechanisms underlying the control of voluntary behaviour and help to reveal how individual differences in brain function explain variations in basic behaviour.

1. Fibiger, H. C. & Phillips, A. G. in *Handbook of Physiology—The Nervous System* Vol. IV (ed. Bloom, F. E.) 647–675 (Williams and Wilkins, Baltimore, Maryland, 1986).
2. Wise, R. A. & Hoffman, D. C. Localization of drug reward mechanisms by intracranial injections. *Synapse* **10**, 247–263 (1992).
3. Robinson, T. E. & Berridge, K. C. The neural basis for drug craving: an incentive-sensitization theory of addiction. *Brain Res. Rev.* **18**, 247–291 (1993).
4. Robbins, T. W. & Everitt, B. J. Neurobehavioural mechanisms of reward and motivation. *Curr. Opin. Neurobiol.* **6**, 228–236 (1996)
5. Louilot, A., LeMoal, M. & Simon, H. Differential reactivity of dopaminergic neurons in the nucleus accumbens in response to different behavioural situations. An *in vivo* voltammetric study in free moving rats. *Brain Res.* **397**, 395–400 (1986).
6. Church, W. H., Justice, J. B. Jr & Neill, D. B. Detecting behaviourally relevant changes in extracellular dopamine with microdialysis. *Brain Res.* **41**, 397–399 (1987).
7. Young, A. M. J., Joseph, M. H. & Gray, J. A. Increased dopamine release *in vivo* in nucleus accumbens and caudate nucleus of the rat during drinking: a microdialysis study. *Neuroscience* **48**, 871–876 (1992).
8. Wilson, C., Nomikos, G. G., Collu, M. & Fibiger, H. C. Dopaminergic correlates of motivated behaviour: importance of drive. *J. Neurosci.* **15**, 5169–5178 (1995).
9. Fiorino, D. F., Coury, A. & Phillips, A. G. Dynamic changes in nucleus accumbens dopamine efflux during the Coolidge effect in male rats. *J. Neurosci.* **17**, 4849–4855 (1997).
10. Thut, G. *et al.* Activation of the human brain by monetary reward. *NeuroReport* **8**, 1225–1228 (1997).
11. Rogers, R. D. *et al.* Choosing between small, likely rewards and large, unlikely rewards activates inferior and orbital prefrontal cortex. *J. Neurosci.* **19**, 9029–9038 (1999)
12. Elliott, R., Friston, K. J. & Dolan, R. J. Dissociable neural responses in human reward systems. *J. Neurosci.* **20**, 6159–6165 (2000)
13. Pavlov, P. I. *Conditioned Reflexes* (Oxford Univ. Press, London, 1927).
14. Rescorla, R. A. & Wagner, A. R. in *Classical Conditioning II: Current Research and Theory* (eds Black, A. H. and Prokasy, W. F.) 64–99 (Appleton Century Crofts, New York, 1972).
15. Mackintosh, N. J. A theory of attention: variations in the associability of stimulus with reinforcement. *Psychol. Rev.* **82**, 276–298 (1975).
16. 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).
17. Schultz, W. Responses of midbrain dopamine neurons to behavioural trigger stimuli in the monkey. *J. Neurophysiol.* **56**, 1439–1462 (1986).
18. Romo, R. & Schultz, W. Dopamine neurons of the monkey midbrain: contingencies of responses to active touch during self-initiated arm movements. *J. Neurophysiol.* **63**, 592–606 (1990).
19. Schultz, W. & Romo, R. Dopamine neurons of the monkey midbrain: contingencies of responses to stimuli eliciting immediate behavioural reactions. *J. Neurophysiol.* **63**, 607–624 (1990).
20. Ljungberg, T., Apicella, P. & Schultz, W. Responses of monkey dopamine neurons during learning of behavioural reactions. *J. Neurophysiol.* **67**, 145–163 (1992).
21. Strecker, R. E. & Jacobs, B. L. Substantia nigra dopaminergic unit activity in behaving cats: effect of arousal on spontaneous discharge and sensory evoked activity. *Brain Res.* **361**, 339–350 (1985).
22. Horvitz, J. C. Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience* **96**, 651–656 (2000).
23. Mirenowicz, J. & Schultz, W. Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature* **379**, 449–451 (1996).
24. Guarraci, F. A. & Kapp, B. S. An electrophysiological characterization of ventral tegmental area dopaminergic neurons during differential Pavlovian fear conditioning in the awake rabbit. *Behav. Brain Res.* **99**, 169–179 (1999).
25. Schultz, W. Activity of dopamine neurons in the behaving primate. *Semin. Neurosci.* **4**, 129–138 (1992).
26. Redgrave, P., Prescott, T. J. & Gurney, K. Is the short-latency dopamine response too short to signal reward? *Trends Neurosci.* **22**, 146–151 (1999).
27. Hollerman, J. R. & Schultz, W. Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature Neurosci.* **1**, 304–309 (1998).
28. Schultz, W. Predictive reward signal of dopamine neurons. *J. Neurophysiol.* **80**, 1–27 (1998).
29. Schultz, W. & Dickinson, A. Neuronal coding of prediction errors. *Annu. Rev. Neurosci.* **23**, 473–500 (2000).

30. Mirenowicz, J. & Schultz, W. Importance of unpredictability for reward responses in primate dopamine neurons. *J. Neurophysiol.* **72**, 1024–1027 (1994).
31. Schultz, W., Dayan, P. & Montague, R. R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
32. 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).
33. Sutton, R. S. & Barto, A. G. Toward a modern theory of adaptive networks: expectation and prediction. *Psychol. Rev.* **88**, 135–170 (1981).
    This paper proposed a very effective 'temporal difference' reinforcement learning model that computes a prediction error over time. The teaching signal incorporates primary reinforcers and conditioned stimuli, and resembles in all aspects the response of dopamine neurons to rewards and conditioned, reward-predicting stimuli, although dopamine neurons also report novel stimuli.
34. Tesauro, G. TD-Gammon, a self-teaching backgammon program, achieves master-level play. *Neural Comp.* **6**, 215–219 (1994).
35. Suri, R. E. & Schultz, W. Learning of sequential movements by neural network model with dopamine-like reinforcement signal. *Exp. Brain Res.* **121**, 350–354 (1998).
36. Suri, R. & Schultz, W. A neural network with dopamine-like reinforcement signal that learns a spatial delayed response task. *Neuroscience* **91**, 871–890 (1999).
37. Schultz, W. & Romo, R. Responses of nigrostriatal dopamine neurons to high intensity somatosensory stimulation in the anesthetized monkey. *J. Neurophysiol.* **57**, 201–217 (1987).
38. Abercrombie, E. D., Keefe, K. A., DiFrischia, D. S. & Zigmond, M. J. Differential effect of stress on *in vivo* dopamine release in striatum, nucleus accumbens, and medial frontal cortex. *J. Neurochem.* **52**, 1655–1658 (1989).
39. 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).
40. Apicella, P., Ljungberg, T., Scarnati, E. & Schultz, W. Responses to reward in monkey dorsal and ventral striatum. *Exp. Brain Res.* **85**, 491–500 (1991).
41. Apicella, P., Scarnati, E. & Schultz, W. Tonically discharging neurons of monkey striatum respond to preparatory and rewarding stimuli. *Exp. Brain Res.* **84**, 672–675 (1991).
42. Lavoie, A. M. & Mizumori, S. J. Y. Spatial-, movement- and reward-sensitive discharge by medial ventral striatum neurons of rats. *Brain Res.* **638**, 157–168 (1994).
43. 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).
44. Shidara, M., Aigner, T. G. & Richmond, B. J. Neuronal signals in the monkey ventral striatum related to progress through a predictable series of trials. *J. Neurosci.* **18**, 2613–2625 (1998).
45. Matsumura, M., Kojima, J., Gardiner, T. W. & Hikosaka, O. Visual and oculomotor functions of monkey subthalamic nucleus. *J. Neurophysiol.* **67**, 1615–1632 (1992).
46. Schultz, W. Activity of pars reticulata neurons of monkey substantia nigra in relation to motor, sensory and complex events. *J. Neurophysiol.* **55**, 660–677 (1986).
47. Niki, H., Sakai, M. & Kubota, K. Delayed alternation performance and unit activity of the caudate head and medial orbitofrontal gyrus in the monkey. *Brain Res.* **38**, 343–353 (1972).
48. Niki, H. & Watanabe, M. Prefrontal and cingulate unit activity during timing behaviour in the monkey. *Brain Res.* **171**, 213–224 (1979).
49. Rosenkilde, C. E., Bauer, R. H. & Fuster, J. M. Single cell activity in ventral prefrontal cortex of behaving monkeys. *Brain Res.* **209**, 375–394 (1981).
50. Watanabe, M. The appropriateness of behavioural responses coded in post-trial activity of primate prefrontal units. *Neurosci. Lett.* **101**, 113–117 (1989).
51. Tremblay, L. & Schultz, W. Reward-related neuronal activity during go–no go task performance in primate orbitofrontal cortex. *J. Neurophysiol.* **83**, 1864–1876 (2000).
52. Niki, H. & Watanabe, M. Cingulate unit activity and delayed response. *Brain Res.* **110**, 381–386 (1976).
53. Nishijo, H., Ono, T. & Nishino, H. Single neuron responses in amygdala of alert monkey during complex sensory stimulation with affective significance. *J. Neurosci.* **8**, 3570–3583 (1988).
    A paper written by one of the few groups investigating neurons in the primate amygdala in relation to reward-related stimuli. They reported a number of different responses to the presentation of natural rewards.

54. Nakamura, K., Mikami, A. & Kubota, K. Activity of single neurons in the monkey amygdala during performance of a visual discrimination task. *J. Neurophysiol.* **67**, 1447–1463 (1992).
55. Burton, M. J., Rolls, E. T. & Mora, F. Effects of hunger on the responses of neurons in the lateral hypothalamus to the sight and taste of food. *Exp. Neurol.* **51**, 668–677 (1976).
56. Tremblay, L. & Schultz, W. Relative reward preference in primate orbitofrontal cortex. *Nature* **398**, 704–708 (1999).
57. Pratt, W. E. & Mizumori, S. J. Y. Characteristics of basolateral amygdala neuronal firing on a spatial memory task involving differential reward. *Behav. Neurosci.* **112**, 554–570 (1998).
58. Thorpe, S. J., Rolls, E. T. & Maddison, S. The orbitofrontal cortex: neuronal activity in the behaving monkey. *Exp. Brain Res.* **49**, 93–115 (1983).
59. Rolls, E. T., Murzi, E., Yaxley, S., Thorpe, S. J. & Simpson, S. J. Sensory-specific satiety: food-specific reduction in responsiveness of ventral forebrain neurons after feeding in the monkey. *Brain Res.* **368**, 79–86 (1986).
60. Rolls, E. T., Sienkiewicz, Z. J. & Yaxley, S. Hunger modulates the responses to gustatory stimuli of single neurons in the caudolateral orbitofrontal cortex of the macaque monkey. *Eur. J. Neurosci.* **1**, 53–60 (1989).
61. Rolls, E. T., Scott, T. R., Sienkiewicz, Z. J. & Yaxley, S. The responsiveness of neurons in the frontal opercular gustatory cortex of the macaque monkey is independent of hunger. *J. Physiol.* **397**, 1–12 (1988).
62. Apicella, P., Legallet, E. & Trouche, E. Responses of tonically discharging neurons in the monkey striatum to primary rewards delivered during different behavioural states. *Exp. Brain Res.* **116**, 456–466 (1997).
63. Apicella, P., Ravel, S., Sardo, P. & Legallet, E. Influence of predictive information on responses of tonically active neurons in the monkey striatum. *J. Neurophysiol.* **80**, 3341–3344 (1998).
64. Apicella, P., Legallet, E. & Trouche, E. Responses of tonically discharging neurons in monkey striatum to visual stimuli presented under passive conditions and during task performance. *Neurosci. Lett.* **203**, 147–150 (1996).
65. 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).
66. Aosaki, T. *et al.* Responses of tonically active neurons in the primate's striatum undergo systematic changes during behavioural sensorimotor conditioning. *J. Neurosci.* **14**, 3969–3984 (1994).
67. Apicella, P., Scarnati, E., Ljungberg, T. & Schultz, W. Neuronal activity in monkey striatum related to the expectation of predictable environmental events. *J. Neurophysiol.* **68**, 945–960 (1992).
68. 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).
69. Schoenbaum, G., Chiba, A. A. & Gallagher, M. Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nature Neurosci.* **1**, 155–159 (1998).
70. Hikosaka, K. & Watanabe, M. Delay activity of orbital and lateral prefrontal neurons of the monkey varying with different rewards. *Cerebral Cortex* **10**, 263–271 (2000).
71. Hollerman, J. R., Tremblay, L. & Schultz, W. Influence of reward expectation on behavior-related neuronal activity in primate striatum. *J. Neurophysiol.* **80**, 947–963 (1998).
72. Tremblay, L., Hollerman, J. R. & Schultz, W. Modifications of reward expectation-related neuronal activity during learning in primate striatum. *J. Neurophysiol.* **80**, 964–977 (1998).
73. Tremblay, L. & Schultz, W. Modifications of reward expectation-related neuronal activity during learning in primate orbitofrontal cortex. *J. Neurophysiol.* **83**, 1877–1885 (2000).
74. Dickinson, A. & Balleine, B. Motivational control of goal-directed action. *Animal Learn. Behav.* **22**, 1–18 (1994).
75. Okano, K. & Tanji, J. Neuronal activities in the primate motor fields of the agranular frontal cortex preceding visually triggered and self-paced movement. *Exp. Brain Res.* **66**, 155–166 (1987).
76. Romo, R. & Schultz, W. Neuronal activity preceding self-initiated or externally timed arm movements in area 6 of monkey cortex. *Exp. Brain Res.* **67**, 656–662 (1987).
77. Kurata, K. & Wise, S. P. Premotor and supplementary motor cortex in rhesus monkeys: neuronal activity during externally- and internally-instructed motor tasks. *Exp. Brain Res.* **72**, 237–248 (1988).
78. Schultz, W. & Romo, R. Neuronal activity in the monkey striatum during the initiation of movements. *Exp. Brain Res.* **71**, 431–436 (1988).
79. Watanabe, M. Reward expectancy in primate prefrontal neurons. *Nature* **382**, 629–632 (1996).
    The first demonstration that expected rewards influence the behaviour-related activity of neurons in a manner that is compatible with a goal-directed

account. Neurons in primate dorsolateral prefrontal cortex show different activities depending on the expected reward during a typical prefrontal spatial delayed response task.

80. Leon, M. I. & Shadlen, M. N. Effect of expected reward magnitude on the responses of neurons in the dorsolateral prefrontal cortex of the macaque. *Neuron* **24**, 415–425 (1999).

81. Kawagoe, R., Takikawa, Y. & Hikosaka, O. Expectation of reward modulates cognitive signals in the basal ganglia. *Nature Neurosci.* **1**, 411–416 (1998).

82. Liu, Z. & Richmond, B. J. Response differences in monkey TE and perirhinal cortex: stimulus association related to reward schedules. *J. Neurophysiol.* **83**, 1677–1692 (2000). **The first demonstration of the prominent reward relationships in neurons of temporal cortex. Neuronal responses to task stimuli in the primate perirhinal cortex were profoundly affected by the distance to reward, whereas neuronal responses in the neighbouring TE area predominantly reflected the visual features of the stimuli.**

83. Platt, M. L. & Glimcher, P. W. Neural correlates of decision variables in parietal cortex. *Nature* **400**, 233–238 (1999).
Neurons in primate parietal association cortex were sensitive to two key variables of game theory and decision making: the quantity and the probability of the outcomes. Based on studies of choice behaviour and decision making in human economics and animal psychology, this is the first application of principles of decision theory in primate neurophysiology.

84. Shima, K. & Tanji, J. Role for cingulate motor area cells in voluntary movement selection based on reward. *Science* **282**, 1335–1338 (1998).
Neurons in the primate cingulate motor area were active when animals switched to a different movement when continuing to perform the current movement would have produced less reward. This study is interesting from the point of view of

movement selection and also for the influence of rewards on behavioural choices.

85. Kelley, A. E. Functional specificity of ventral striatal compartments in appetitive behaviours. *Ann. NY Acad. Sci.* **877**, 71–90 (1999).

86. Carelli, R. M., King, V. C., Hampson, R. E. & Deadwyler, S. A. Firing patterns of nucleus accumbens neurons during cocaine self-administration. *Brain Res.* **626**, 14–22 (1993).
The first demonstration of drug-related changes in neuronal activity in one of the key structures of drug addiction, the nucleus accumbens. Two important neuronal patterns were found in the rat: activity preceding lever presses that led to acquisition of the drug and activity following drug delivery. These findings have been consistently reproduced by other goups.

87. Chang, J. Y., Sawyer, S. F., Lee, R. S. & Woodward, D. J. Electrophysiological and pharmacological evidence for the role of the nucleus accumbens in cocaine self-administration in freely moving rats. *J. Neurosci.* **14**, 1224–1244 (1994).

88. Carelli, R. M. & Deadwyler, S. A. A comparison of nucleus accumbens neuronal firing patterns during cocaine self-administration and water reinforcement in rats. *J. Neurosci.* **14**, 7735–7746 (1994).

89. Chang, J. Y., Janak, P. H. & Woodward, D. J. Comparison of mesocorticolimbic neuronal responses during cocaine and heroin self-administration in freely moving rats. *J. Neurosci.* **18**, 3098–3115 (1998).

90. Carelli, R. M., Ijames, S. G. & Crumling, A. J. Evidence that separate neural circuits in the nucleus accumbens encode cocaine versus 'natural' (water and food) reward. *J. Neurosci.* **20**, 4255–4266 (2000).

91. Chang, J. Y., Paris, J. M., Sawyer, S. F., Kirillov, A. B. & Woodward, D. J. Neuronal spike activity in rat nucleus accumbens during cocaine self-administration under different fixed-ratio schedules. *J. Neurosci.* **74**, 483–497 (1996).

92. Peoples, L. L., Uzwiak, A. J., Gee, F. & West, M. O. Operant behaviour during sessions of intravenous cocaine infusion is

necessary and sufficient for phasic firing of single nucleus accumbens neurons. *Brain Res.* **757**, 280–284 (1997).

93. West, M. O., Peoples, L. L., Michael, A. J., Chapin, J. K. & Woodward, D. J. Low-dose amphetamine elevates movement-related firing of rat striatal neurons. *Brain Res.* **745**, 331–335 (1997).

94. Peoples, L. L. & West, M. O. Phasic firing of single neurons in the rat nucleus accumbens correlated with the timing of intravenous cocaine self-administration. *J. Neurosci.* **16**, 3459–3473 (1996).

95. Peoples, L. L., Gee, F., Bibi, R. & West, M. O. Phasic firing time locked to cocaine self-infusion and locomotion: dissociable firing patterns of single nucleus accumbens neurons in the rat. *J. Neurosci.* **18**, 7588–7598 (1998).

96. 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).

97. Otmakhova, N. A. & Lisman, J. E. D1/D5 dopamine receptor activation increases the magnitude of early long-term potentiation at CA1 hippocampal synapses. *J. Neurosci.* **16**, 7478–7486 (1996).

98. 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).

99. Calabresi, P. *et al.* Abnormal synaptic plasticity in the striatum of mice lacking dopamine D2 receptors. *J. Neurosci.* **17**, 4536–4544 (1997).