OPINION

The short-latency dopamine signal: a role in discovering novel actions?

Peter Redgrave and Kevin Gurney

Abstract | An influential concept in contemporary computational neuroscience is the reward prediction error hypothesis of phasic dopaminergic function. It maintains that midbrain dopaminergic neurons signal the occurrence of unpredicted reward, which is used in appetitive learning to reinforce existing actions that most often lead to reward. However, the availability of limited afferent sensory processing and the precise timing of dopaminergic signals suggest that they might instead have a central role in identifying which aspects of context and behavioural output are crucial in causing unpredicted events.

In his famous experiment on reinforcement learning, Thorndike placed a hungry cat in a cage, the door of which was held closed by a pin[1]. A peddle in the cage was connected to the pin, such that if the cat pressed the peddle, the pin was released and the door fell open. Outside the cage was a piece of fish. Progressively, the cat learned to operate the peddle which opened the door and gave access to the fish. Consequently, Thorndike proposed that "any act which in a given situation produces satisfaction becomes associated with that situation so that when the situation recurs the act is more likely than before to recur also" — the Law of Effect[1]. Had single-unit electrophysiological recording been available to Thorndike, he could have recorded the activity of ventral midbrain dopaminergic (DA) neurons during his experiment. What we now know about the activity of DA neurons suggests that the unpredicted movement or sound of the pin being released would have caused a short-latency, short-duration burst of DA activity, which is referred to as the 'phasic' dopamine response[2]. Evidence is now emerging to suggest that this neural response would have occurred before the cat had turned to see what was happening, long before the door had fallen open, and even longer before the cat had the 'satisfaction' of eating the fish[3–5]. In the light of these and other considerations, we propose that the phasic response of DA neurons provides the learning signal in circuitry that would allow the cat to discover exactly what movements it had to make, and where to make them, to release the pin; in other words, to reinforce the development of an entirely novel action. This suggestion will be contrasted with the currently dominant view that phasic DA responses signal reward prediction errors[2,6–10]. A reward prediction error represents the degree to which a reward cannot be predicted, and is indicated by the difference between the reward obtained by a given action and the reward that was expected to result from that action. In instrumental conditioning paradigms, they are used to reinforce the actions that most frequently lead to satisfaction — that is, presumed pre-existing actions of the cat that led to the door opening and provided access to the fish.

Reward prediction error hypothesis

Given the often overwhelming accumulation of biological information describing the anatomy[11], biochemistry[12,13], physiology[14], pharmacology[15,16] and behaviour[16–18] of central dopamine (DA) systems, it is surprising that there are so few hypotheses concerning the computational task(s) performed by DA neurotransmission (the term 'computational task' in this sense refers to what is being computed and why[19,20]). A notable exception is the reward prediction error hypothesis proposed by Montague *et al.*[6,7] and by Schultz and colleagues[2,8–10]. These investigators suggest that the short-latency, sensory-evoked DA responses signal reward prediction errors, which are used by reinforcement learning mechanisms in the basal ganglia, and elsewhere, to select actions that will maximize the future acquisition of reward. The reward prediction error hypothesis has received much empirical support[21–27] and is now widely accepted by many biological[9,28–30] and computational neuroscientists[7,31–35]. In this article, however, we wish to question this view and make an alternative suggestion. To do this, we first need to outline certain important aspects of phasic DA signalling.

Typically, unexpected biologically significant events including sudden novel stimuli, intense sensory stimuli, primary rewards and arbitrary stimuli classically conditioned by association with primary rewards evoke a stereotypical sensory response from DA neurons in many species[2,36–38]. This response comprises a characteristic short-latency (70–100 ms), short-duration (< 200 ms) burst of activity[2] (FIG. 1b). However, it is the capacity of phasic DA responses to change when experimental conditions are altered that has provoked the most interest[2,9,24–26]. First, the novelty response of DA neurons habituates rapidly when a sensory stimulus is repeated in the absence of behaviourally rewarding consequences[39]. Second, a phasic DA response will emerge following the presentation of a neutral sensory stimulus that predicts a primary reward[39]. Under these conditions the DA responses to the predicted reward gradually diminish[40]. Third, when a predicted reward is omitted, a reliable depression in the spontaneous activity of the DA neurons occurs 70–100 ms after the time of expected reward delivery[41]. It is largely on the basis of these data that the reward prediction error hypothesis was originally formulated[6,8,41].

More recently, additional supporting investigations have established that the phasic DA signal complies with the contiguity, contingency and prediction error tenets

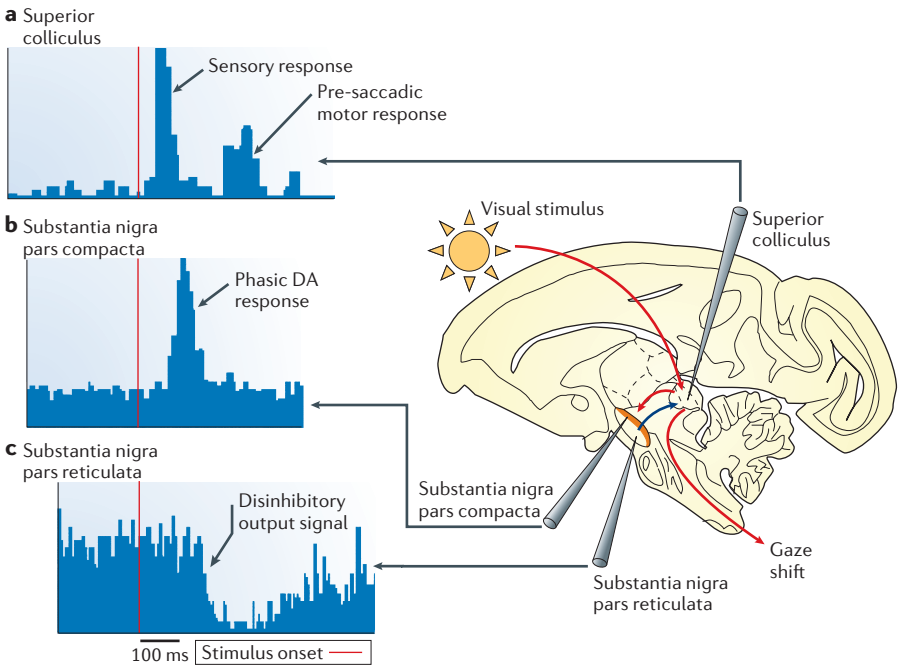

Figure 1 | **A latency constraint associated with visual input to dopaminergic neurons.** Typical examples show the relative timing of responses evoked by unexpected visual stimuli in the superior colliculus, and in the dopaminergic (DA) neurons in the substantia nigra pars compacta and the substantia nigra pars reticulata. Peristimulus histograms showing nerve impulse frequencies from different publications are aligned on stimulus onset. **a** | Activity in the superior colliculus is characterized by an early sensory response (latency ~40 ms) followed by a later motor response (latency ~200 ms). The latter is responsible for driving the orienting gaze shift to bring the stimulus onto the fovea[49]. **b** | The phasic DA response (latency ~70 ms)[2] occurs after the collicular sensory response but prior to the pre-saccadic motor response. **c** | Phasic DA activity also occurs prior to the output signal from the substantia nigra pars reticulata that disinhibits the motor-related activity of target neurons in the superior colliculus[50]. Red arrows, excitatory connections; blue arrows, inhibitory connections. Panel **a** modified, with permission, from REF. 49 © (1987) American Physiological Society. Panel **b** modified, with permission, from REF. 2 © (1998) American Physiological Society. Panel **c** modified, with permission, from REF. 50 © (1983) American Physiological Society.

of contemporary learning theories[9]. A neutral stimulus that is presented contiguously with primary reward acquires the ability to elicit a phasic DA response[42]. The contingency requirement specifies that DA neurons should discriminate between conditioned stimuli that predict reward, predict an absence of reward and neutral stimuli with no predictive value. Under certain conditions (see below) it is clear that DA neurons have this capacity[27]. In the prediction error-defining blocking paradigm (that is, learning is blocked when a stimulus is paired with a fully predicted reward), DA neurons acquire responses to conditioned stimuli only when they are associated with an unpredicted reward[21].

This body of evidence provides powerful support for the reward prediction error hypothesis. However, a fundamental aspect of this view is that phasic DA signals result from calculations based, in part, on the capacity of afferent sensory systems to provide an adequate assessment of the reward value of unpredicted events. Despite some seemingly supporting observations[23,27], this is unlikely to be generally the case. Recent evidence from studies that have identified sources of short-latency sensory input to DA neurons[4,5,43–46] indicates that, in real world conditions (and in the example of Thorndike's cat), the reward value of unexpected events (for example, the pin being released) remains to be established at the time of phasic DA signalling. In the following sections, we review this evidence.

Pre-attentive sensory processing

There are three aspects of experimental data concerning phasic DA signalling that suggest it is conducted on the basis of pre-attentive/pre-saccadic sensory processing. Such evidence casts doubt on the general capacity of DA neurons to signal a parameter for which prior determination of the reward value of unpredicted sensory events is essential.

Stimulus diversity. It has already been noted that DA neurons exhibit strong phasic responses to unexpected sensory events that have no obvious appetitive reinforcement consequences[38,47], but are salient by virtue of their novelty, intensity or physical similarity to reward-related stimuli[2]. Studies in which neutral stimuli fail to elicit phasic DA responses[23,27] generally ensure that such stimuli have been previously habituated, that is, they are no longer novel and have been learned previously to have no reward predictive value[48].

Response homogeneity. The latency (70–100 ms following stimulus onset) and duration (100–200 ms) of phasic DA responses (FIG. 1b) are remarkably constant across species and many experimental paradigms, and are largely independent of the modality or perceptual complexity of eliciting events[2]. The stereotypical nature of the DA response creates problems for the reward prediction error hypothesis because it is obvious that the reward value of some stimuli takes longer to establish than others. For example, in Thorndike's experiment the satisfaction of eating the fish, or even the realization that the fish can now be eaten, would probably occur several seconds after the DA response (see next point).

Response latency. FIGURE 1 illustrates how the phasic DA response (latency 70–100 ms)[2] normally precedes the gaze shift (latency 150–200 ms)[49,50] that brings an unpredicted sensory event onto the fovea for analysis by cortical visual systems[51,52]. So far, we know of no examples for which consistent post-saccadic latencies for phasic DA responses (that is, > 200 ms) have been reported. Indeed, in circumstances in which reward prediction errors become apparent shortly after a gaze shift[53], they are notably absent. To the extent that phasic DA responses remain pre-saccadic, they will incorporate only those perceptual characteristics that can be determined on the basis of the pre-attentive afferent sensory processing that typically occurs prior to a foveating gaze shift. It is, therefore, of interest to know where such processing is conducted to determine whether the identified circuitry has the perceptual power required to discriminate the wide range of sensory events in everyday life that signify reward.

Sources of afferent sensory signals

The cell bodies of midbrain DA neurons lie in the densely packed dorsal sector of the substantia nigra (pars compacta) and the

more medially located ventral tegmental area. The principal targets of ascending DA projections include other basal ganglia nuclei (principally the striatum), various limbic structures (for example, the septal area and amygdala) and parts of the frontal cortex[11]. Until recently, and despite the enormous volume of biological data relating to DA systems[11,12,14], little information was available concerning the sources of short-latency sensory inputs to midbrain DA neurons. Because most experiments analysing the sensory properties of DA neurons have used visual stimuli[2,9], from this point we concentrate on probable visual afferents to the ventral midbrain. Note also that our use of the term 'event' refers exclusively to visual stimuli with a phasic onset, as again, to our knowledge, there are no reports indicating that perception of a salient static visual feature can elicit a phasic DA response.

Recent analyses of cortical visual processing (for reviews, see REFS 51,52) indicate that signals related to the identity of objects can be recorded in the inferotemporal cortex ~80–100 ms after stimulus onset. By this time many of the DA neurons have already begun to fire[2], and it is not obvious by which route relevant information could be communicated rapidly from the temporal cortex to the ventral midbrain. Similarly, early visual responses in the striatum[54] and subthalamic nucleus[55] generally occur at about the same time, or after phasic DA signalling. This excludes the possibility that intrinsic basal ganglia processing of reward-related stimuli could provide the requisite short-latency visual input to DA neurons.

By contrast, recent evidence from our laboratory suggests that a subcortical visual structure located in the dorsal midbrain, the superior colliculus, is the most likely source of early visual input to DA neurons[4,5,43]. First, as the superior colliculus receives direct input from retinal ganglion cells, its visual response latencies are always shorter than those of DA neurons[2,4,49] (compare with FIGS 1a,b). Second, a previously unreported direct tectonigral projection connecting the deep layers of the superior colliculus to the substantia nigra pars compacta has been discovered in rats[4] (FIG. 2a), cats[46] and now monkeys[56]. Third, local, visually evoked potentials in the substantia nigra pars compacta can be recorded in the absence of the visual cortex, whereas subsequent removal of the visual layers of the superior colliculus blocks all visually evoked activity in the substantia nigra[4]. Fourth, in urethane anaesthetized rats, neurons in the deep layers of the colliculus, and DA neurons, are unresponsive to visual events. Visual sensitivity can be restored to both collicular[5,57] and DA neurons[5] by a local disinhibitory injection of a GABA (γ-aminobutyric acid) blocker into the superior colliculus (FIG. 2b). Comparable disinhibition of the visual cortex leaves DA neurons unresponsive to visual stimuli[5]. Finally, after application of the anaesthetic, injections of bicuculline into the superior colliculus can also restore a visually evoked phasic release of DA into the striatum[5] (FIG. 2c). For these reasons, we have suggested that the superior colliculus is the primary, if not the exclusive, source of short-latency visual input to ventral midbrain DA neurons[4,5]. If this conclusion is correct, the perceptual properties of early visual processing conducted by the superior colliculus will be an important determinant of the visual information that can be made available to DA neurons.

Visual perception in the superior colliculus

Reviews of visual processing in the mammalian superior colliculus agree that collicular neurons are exquisitely sensitive to spatially localized changes in luminance that signify appearance, disappearance or movement in the visual field[58–61]. They are, however, comparatively insensitive to static contrast, velocity, wavelength and the geometric configuration of visual stimuli[58–61]. Visual events, repeated in the absence of contiguous reward, cause deep layer neurons to habituate rapidly[60,62,63], whereas associating such stimuli with reward can block or reverse habituation and enhance the visual responses of collicular neurons[58,64]. These properties imply that, if early sensory activity is present in the collicular deep layers, the event is likely to be biologically significant, either by virtue of its novelty or because it has been previously associated with reinforcing stimuli (that is, not habituated). So, to the extent that the colliculus has been configured to detect visual transients rather than static features, the short-latency sensitivity of DA neurons to visual stimuli could be similarly constrained.

With such considerations in mind, we should pause to consider how DA neurons seem able to perform the fine perceptual distinctions required to distinguish the complex visual stimuli that have been used to signal different reward magnitudes and probabilities[22,23,27]. Careful reading of procedure indicates that most relevant studies[22–25,27,65] have chosen to present stimuli that predict different levels of reward at different spatial locations. For example, Tobler *et al.*[23] explain that "…to aid discrimination each

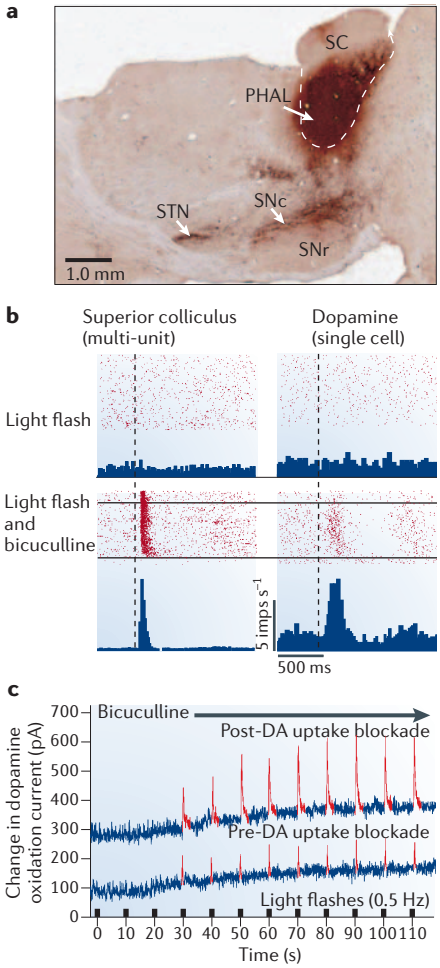

Figure 2 | **Evidence supporting the SC as the primary source of short-latency visual input to DA neurons in the SNc. a** | Anatomy. A direct projection from the superior colliculus (SC) to substantia nigra pars compacta (SNc) was recently discovered[4]. An example of the tectonigral projection in rats revealed by an injection of an anterograde tracer (PHAL) into the rostrolateral deep layers of the superior colliculus is shown. **b** | Electrophysiology. Visual responses of dopaminergic (DA) neurons depend on the visual sensitivity of the superior colliculus[5]. Urethane anaesthesia abolishes sensitivity to a light flash both in the deep layers of the superior colliculus and in an electrophysiologically characterized DA neuron (upper raster displays and peristimulus histograms). Response to the light was restored both to the collicular deep layers and the DA neurons by a local disinhibitory injection of a GABA (γ-aminobutyric acid) antagonist, bicuculline, into the superior colliculus (lower raster displays and peristimulus histograms). **c** | Electrochemistry. After application of the anaesthetic, disinhibition of the superior colliculus by a local injection of bicuculline also restored flash-evoked release of DA into the striatum, measured by fixed-potential amperometry[5]. SNr substantia nigra pars reticulata; STN, subthalamic nucleus. Panels **b** and **c** modified, with permission, from REF. 5 © (2005) American Association for the Advancement of Science.

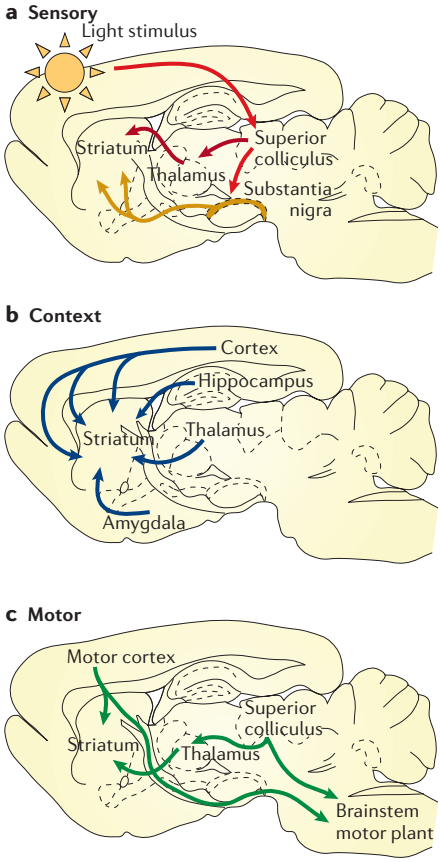

Figure 3 | **Potentially converging inputs to the dorsal striatum. a** | Phasic sensory inputs. Two separate, short-latency representations of unpredicted visual events are likely to converge on striatal circuitry: retino-tecto-thalamo-striatal projections will provide a phasic sensory-related glutamatergic input (red arrows)[68]; and retino-tecto-nigro-striatal projections will provide a phasic dopaminergic input (yellow arrows)[4,5]. **b** | Contextual inputs. Striatal neurons are sensitive to experimental context[26,70–72]. Multidimensional contextual afferents are likely to originate in the cerebral cortex, limbic structures such as the hippocampus and amygdala and the thalamus (blue arrows). **c** | Motor copy inputs. Branched pathways from the motor cortex and subcortical sensorimotor structures (for example, the superior colliculus) reach the striatum directly (cortex) or indirectly via the thalamus (subcortical structures). Motor-related projections are likely to provide the striatum with a running, multidimensional record (motor efference copy) of commands relating to ongoing goals/actions/movements (green arrows)[68,73–77].

stimulus was presented at a unique location on the computer monitor." However, the appearance of visual stimuli at different spatial locations is exactly the parameter that could be readily distinguished in the spatial maps of the superior colliculus[58–61]. Therefore, the predictable association between spatial location and reward value of the stimuli used in these studies is likely to be crucial for DA neurons to signal differences in the reward value of temporally unpredicted events, without having to process information about fine detail.

So, how might we expect DA neurons to behave in the less constrained environments encountered in the natural world? Given that most temporally unexpected transient events in nature are also spatially unpredictable, it should be safe to assume that the predominant phasic activity of DA neurons in natural environments would report the occurrence of events that remain to be identified at the time of DA signalling (that is, prior to the gaze shift that brings the event onto the fovea for detailed analysis by cortical visual systems). In such circumstances, it is unlikely that pre-attentive subcortical visual processing would have the capacity to discriminate the full spectrum of rewarding events, particularly those for which colour and/or high-spatial frequency detail provide the clues to their identity. Perhaps it is time to entertain the possibility that phasic DA signals could be involved in a different computational process — one that has less stringent perceptual requirements.

An alternative functional hypothesis

Essential characteristics of DA signalling. When considering alternative functional possibilities for DA signalling, we shall take into consideration the following two characteristics of the phasic DA response. First, it has striking resemblances to a reinforcement error signal that represents the difference between the anticipated level of future reinforcement predicted immediately prior to an action and the update of that prediction following delivery of a sensory reinforcer[2,6–10,41,66]. Note our use of the term 'reinforcement' rather than 'reward'[67]. Second, its timing is stereotypical and precise (~100 ms latency, ~100 ms duration)[2] (FIGS 1,2b,2c). Together, these characteristics suggest that the DA response is being used in a learning process in which the timing of reinforcement is crucial. A clue to the identity of this process could be obtained by asking what signals are likely to be present in the target regions of the ascending DA projections at the time of the phasic DA response — because it is with these signals that the precisely timed DA release will most readily interact. In view of the comparative availability of relevant information we will, from this point, confine our remarks specifically to afferent projections of the dorsal striatum (caudate/putamen).

Convergent signals. There are likely to be at least three classes of input to the dorsal striatum that would be in a position to interact with phasic DA release (FIG. 3). First, a separate short-latency sensory representation of the same unexpected event that triggered the DA signal, probably relayed via input from the thalamus[68,69] (FIG. 3a). Second, contextual information related to the general sensory, metabolic and cognitive state of the animal[26,70–72] (FIG. 3b). Information related to the animal's current physical location could be particularly important. Third, motor information represented by efference copies or corollary discharges of action decisions and motor commands. Both anatomical and physiological data suggest that copies of motor commands from both cortical and subcortical sensorimotor structures to the brainstem/spinal cord are also directed to the dorsal striatum via branching collaterals[68,73–77]. These efference copy signals are likely to provide the striatum with a running record of current goals, actions and movements (FIG. 3c). It is important to appreciate that, while many of the sensory, contextual and motor signals will arrive via the well-established cortico-basal ganglia-thalamocortical loops[78,79], there seem to be similar loops connecting subcortical sensorimotor structures with the basal ganglia[68]. Within these subcortical loops, sensory and motor input from brainstem structures can access the striatum via relays in the lateral posterior[80], midline and intralaminar nuclei of the thalamus[81–85] (FIG. 3). The latencies of visual activity recorded in the striatum (100–250 ms[54,69]) suggest that short-latency sensory-evoked (glutamatergic[84]) input from the thalamus[86] is likely to be temporally coincident with the phasic DA input from the substantia nigra[5,87,88].

The hypothesis. Our proposal is that the phasic DA signal acts to reinforce the reselection (repetition) of actions/movements that immediately precede an unpredicted biologically salient event (as determined by the presence of short-latency activity in primary subcortical sensory structures such as the superior colliculus). Specifically, in every case in which something done by the animal/agent is the cause of an unexpected sensory event, a crucial conjunction of contextual and motor efference copy inputs to the dorsal striatum will directly precede the simultaneous arrival of the sensory (glutamatergic and DA) representations of the unpredicted event (FIG. 4a). The proposed temporal alignment of these signals could provide a basis for learning, first, whether

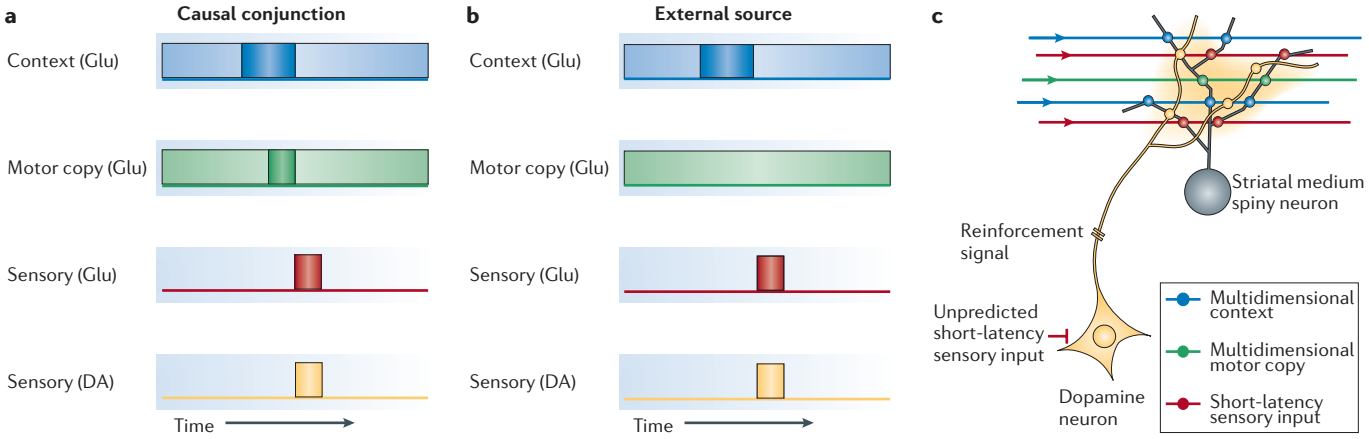

Figure 4 | **The relative timing of proposed inputs to the dorsal striatum could be crucial for determining the source of agency. a** | Event caused by the individual. Whenever the subject is the cause of an unpredicted sensory event, relevant components of the multidimensional contextual (blue) and motor efference copy (green) inputs will directly precede the near-simultaneous short-latency glutamatergic (Glu) sensory input from the thalamus (red) and the phasic dopaminergic (DA) input from the substantia nigra (yellow). **b** | Event caused by an external source. When no relevant motor copy inputs precede the phasic sensory inputs (glutamatergic and DA), the unpredicted event is likely to have been caused by an external source. **c** | Reinforcement identifies causal conjunctions. The proposed function of positive phasic DA signals is to reinforce associations between directly preceding contextual and motor copy signals, thereby promoting the repetition of immediately preceding actions.

any aspect of the animal's current behaviour was the probable cause of the event and, if so, exactly what combination of context, action and movement was crucial. This form of learning could provide the animal with the capacity to distinguish events in the world for which it is responsible from those produced by an external source, and could lead to the development of entirely novel and adaptive responses (BOX 1). We now consider aspects of this hypothesis in more detail.

Biasing action selection. We propose that in cases for which unpredicted sensory events are non-noxious (that is, novel or previously associated with reward), the well-characterized positive DA signal[2] could, through Hebbian-like learning rules[89,90], reinforce the repetition of immediately preceding actions in immediately preceding circumstances (FIG. 4c). Insofar as the basal ganglia is considered to have a central role in action selection[76,91–96], sensory-evoked DA signals could be in a position to promote (reinforce) the reselection (repetition) of recently selected actions/movements.

Action identification. At the outset, the elicited sensory event is entirely unpredicted, so many aspects of the animal's ongoing behaviour are likely to be directed towards entirely different tasks. For example, a confined rat may initially depress an operant lever as part of its attempts to escape from the conditioning chamber. Consequently, at the time of the unpredicted sensory event (arrival of a food pellet caused by the lever press), a possibly large set of immediately preceding, but largely irrelevant, contextual (restraint in the box), motivational (desire to escape) and motor-copy signals (reaching for the edge of the box) are likely to be present in the striatum. Typically, embedded within this large set of inputs, only a small subset of signals (those related to placing a foot on the lever) will be causally related to the unpredicted sensory event (the arrival of food). Discovering precisely which action components, and in which circumstances, are responsible for such events is therefore a computationally difficult problem. So, for the crucial causative component of behaviour to be discovered, DA-evoked repetitions of preceding actions/movements must be sufficiently variable, which is normal[97,98], and must have the component that causes the sensory event occurring sufficiently often. Given these conditions, the proposed DA-driven strengthening of contextual, motivational and motor representations when the sensory event is elicited (long-term potentiation[89,90,99]), coupled with a weakening of representations that are present when the DA signal fails to occur (long-term depression[89,90,99]), could permit successive

Box 1 | **The advantage of knowing who did it**

Our proposal is that sensory-driven dopaminergic (DA) responses provide reinforcement signals that are necessary for the brain, first, to discriminate the unpredicted sensory events for which it is responsible, and second, to discover exactly what new responses are required to make these events happen, irrespective of their immediate reward value; for example, finding out during the day that a particular switch, operated in a particular way, turns on a light could be useful when it gets dark. This simple example highlights some general competencies that would have important adaptive properties. It suggests that the brain should acquire action–outcome routines in circumstances in which the outcome has no immediate benefit. The motivation to learn such associations seems to be intrinsic, that is, done for its own sake; the play exhibited by young animals and children can be viewed in this way. In addition, the acquired action–outcome routine can be stored in the form of a reusable skill that can be deployed in a novel manner, or novel context as circumstances change. Experimental evidence is available to support these ideas. First, it has been shown that stimuli that are normally considered to be neutral have intrinsically reinforcing properties in an instrumental discrimination task[124]. Second, the acquired action of pressing a lever to elicit a neutral light stimulus can be used to effect when the light is subsequently classically conditioned with food in the absence of the lever, and then the lever returned[125]. Finally, the advantage of being able to deploy previously acquired behavioural 'options' in the subsequent learning of goal-directed actions has been demonstrated computationally[117]. It is our contention that the phasic DA response provides a signal, independent of normal goal-directed reward systems (food, drink, temperature, sex, and so on), that reinforces acquisition of the behavioural 'building blocks' necessary for novel sequences of autonomous goal-directed action to be generated.

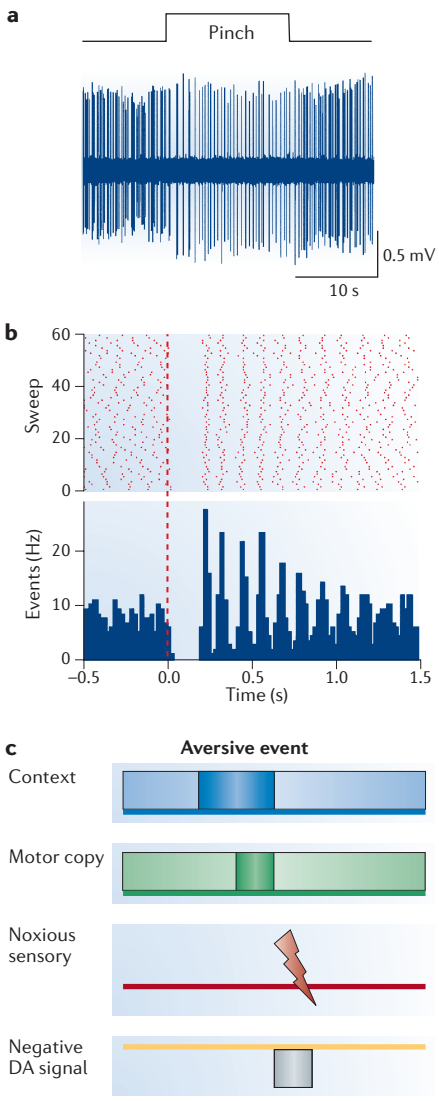

Figure 5 | **Response of dopaminergic neurons to noxious stimuli. a** | Spontaneous activity of an electrophysiologically and histochemically identified dopaminergic (DA) neuron is suppressed for the duration of a noxious foot-pinch[107]. **b** | A peristimulus histogram and raster plot of an electrophysiologically characterized DA neuron showing a similar suppressive response to a noxious footshock (dashed red line)[45]. Note the banding in the histogram and raster plot reflects the regular 7–8 Hz firing of this cell when it begins to fire after the suppression. **c** | A schematic illustrating the probable timing of inputs to the striatum when an action of the subject causes an unpredicted noxious event. Relevant causative components of context and motor copy directly precede the unpredicted noxious event. The observed short-latency negative DA reinforcement signal (panels **a** and **b**) could negatively reinforce future conjunctions of context and motor copy, thereby reducing the tendency to repeat any immediately preceding behaviour. Panel **a** modified, with permission, from REF. 107 © (2004) American Association for the Advancement of Science.

selections by the basal ganglia network to converge on the precise combination of context, motivation and movements responsible for causing the event. Such a combination would represent the emergence of an entirely new action or response on which the traditional mechanisms of reinforcement learning could then operate.

Associative learning outside basal ganglia. The proposed mechanisms for DA-driven biasing of action selection probabilities should mean that post-gaze-shift perceptual analysis of the unpredicted event (outcome), plus the motor representations that produced it, will appear more frequently in neural systems external to the basal ganglia. These are likely to include the amygdala[100], hippocampus[101] and limbic cortex[102–105]. It is in the circuitry of these structures that the long-term associations between action and outcome are probably established and stored (BOX 1). We further suggest that, as the behavioural components that elicit the initially unpredicted outcome are gradually identified, they become subject to the normal processes of reinforcement learning. That is, post-gaze-shift representations of the 'economic value' of outcomes[106] can be used to bias future action selections so that actions with high-value outcomes are selected more frequently.

Externally caused events. In cases for which an unpredicted biologically salient event is caused by an external source (for example, when the delivery of a food pellet or onset of a light stimulus is determined by the experimenter rather than by the animal), afferent sensory inputs to the striatum (glutamatergic and DA) would arrive in the absence of any relevant preceding motor efference copy signals (FIG. 4b). Repetition of any 'superstitious' action that happened, by chance, to be present at this time would fail to evoke the sensory event. Presumably, one of the reasons that all short-latency signals associated with non-habituated events, including the phasic DA responses, are relayed to the striatum is to determine whether or not they could have been caused by an action of the agent.

Noxious events. From the perspective of survival, whenever some aspect of an animal's behaviour causes an unpredicted noxious or disadvantageous event, different processes would have to be invoked. In such cases, the evolutionary imperative would be to immediately terminate and then suppress any tendency to repeat immediately preceding actions, and avoid the context(s) in which they occurred. It is therefore significant that recent reports indicate that noxious stimuli elicit a short-latency (< 100 ms) phasic suppression of DA activity that lasts at least for the duration of the noxious event[45,107] (FIG. 5). It is possible that this negative DA signal could act to reduce the likelihood of reselecting the contexts and actions associated with the unpredicted detrimental event. Presumably, the discrimination of noxious events by DA neurons is possible because the somatosensory system contains specialized, high-threshold nociceptors. The output of these nociceptors seems to be wired relatively directly to DA neurons, through relays in the spinal cord and the parabrachial nucleus[44,108], where it has a predominantly inhibitory effect. In the eye, there are no comparable reward detectors. Indeed, central to our argument is that even in the superior colliculus there are no specialized reward discriminators, only discriminators of the different levels of habituation associated with phasic sensory events. Consequently, there is a necessary asymmetry between the comparative inability of pre-attentive visual processing to discriminate reward-related stimuli and specialized nociceptive processing that is designed to detect the occurrence of events that are noxious.

An imperative for short-latency reinforcement. The first part of the current article draws attention to the anomaly of having the brain's principal system for signalling reward prediction errors[6,7,9] reliant on comparatively primitive, pre-attentive sensory processing — that is, processing that seems to be exquisitely sensitive to some stimuli (transient events that appear, disappear or move) and comparatively insensitive to others (static features involving high spatial frequencies and colour)[4,5,43,58,59,61]. However, if rather than directly reinforcing actions that maximize future rewards[7,9], phasic DA responses guide the behavioural selections that can lead to the development of new actions, a possible reason for their stereotypical short latencies and duration becomes evident (FIG. 6). Unpredicted novel, rewarding or aversive (that is, non-habituated) stimuli commonly evoke orienting and/or defensive responses[58–61,109]. Such responses typically comprise variable combinations of eye, head and body movements. Presumably, the efference copy of such movements would be relayed to the striatum as part of the 'running copy' of

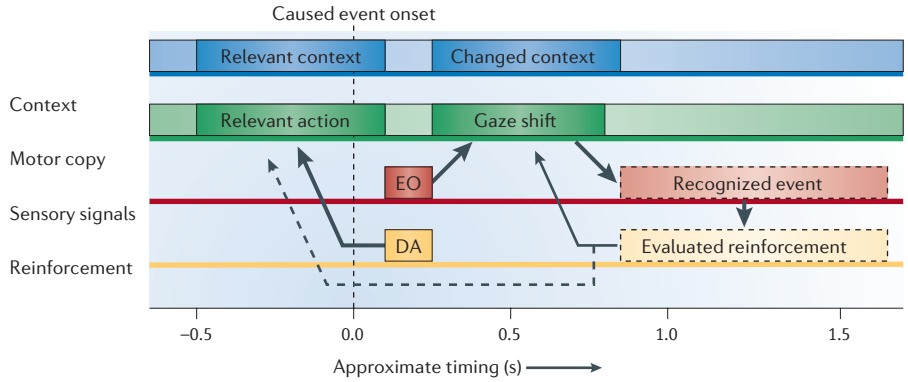

Figure 6 | **A possible explanation for why the phasic dopaminergic reinforcement signal precedes any motor activity elicited by an unpredicted salient sensory event.** For simplicity, only the case for a non-noxious event is illustrated; however, exactly the same rationale applies to negative dopaminergic (DA) responses and defensive reactions elicited by noxious events. The schematic illustrates the approximate timing of hypothesized inputs to the striatum when a particular action (relevant action), occurring in a specific context (relevant context), causes an unpredicted sensory event. Input from the thalamus indicating event onset (EO) and the short-latency phasic DA response occur prior to the orienting gaze shift evoked by the sensory event. The figure illustrates how efference copy signals associated with the gaze shift elicited by the unpredicted event would contaminate the contingency record of potentially causative actions. If the phasic DA response reinforces the repetition of immediately preceding actions/movements, a serious credit assignment problem would result if the reinforcement signal was delayed until after the gaze shift when the reward value of the caused event is fully appreciated — behaviour associated with the gaze shift would receive maximum reinforcement (solid line), rather than the relevant action (dotted line).

ongoing behaviour. The provision of DA reinforcement signals before any movements evoked by the unpredicted event would ensure the reselection or suppression of actions most likely to have caused the unpredicted sensory event. In other words, the maximal positive/negative reinforcing effect of DA would be directed to immediately contiguous motor efference copy (FIG. 6). This analysis would also explain why delaying the sensory event (reinforcement in the case of operant conditioning) by more than a second or so has such a detrimental effect on the rate of learning[9,110,111] — the likelihood of efference copy input to the striatum becoming contaminated with irrelevant actions (that would be reinforced by the sensory-evoked DA response) will increase as a function of the delay.

Implications

Here, we have proposed that the reinforcing function of the phasic DA response has more to do with the discovery of new actions than adjusting the relative probabilities of selecting pre-existing actions to maximize anticipated rewards[2,6–10]. The roots of our idea lie in considerations of basal ganglia circuitry and signal timing. Throughout, we have contrasted the functional implications drawn from this biologically inspired perspective with those originating from computational and behavioural analyses of reinforcement learning. As a different perspective of DA function, the present proposal might offer novel insights into some aspects of the complex relationship between DA neurotransmission and instrumental conditioning paradigms (for contrasting reviews, see REFS 16–18). For example, the reinforcing role of DA in the processes of action identification can be viewed as an essential subcomponent of action–outcome learning, which itself is an essential subcomponent of instrumental conditioning[110]. This analysis is consistent with repeated demonstrations that close contiguity between action and event is a crucial variable in learning action–outcome contingencies[9,110,111] (see above) and in the reliance of instrumental conditioning on intact dopaminergic and basal ganglia functioning[16–18,112].

However, a necessary implication of our current hypothesis is that the reward-related teaching signals (the 'real' reward prediction errors) that drive Law-of-Effect-based instrumental conditioning[1], and are most likely based on post-gaze-shift evaluations of behavioural consequence, must derive from sources other than the pre-saccadic DA response[47,113]. There are plausible alternatives, as longer latency neural responses related to the reward value of sensory stimuli have been detected in several brain regions[102], including the amygdala[100] and limbic prefrontal cortex[114], both of which have strong projections to the basal ganglia[79,115,116].

At present, many strands of empirical evidence can be found to support individual components of the proposed network of functionally differentiated inputs to the striatum. However, as with most systems-level hypotheses, much work will be needed to test whether they all work together in the prescribed manner. For example, a crucial evaluation will be to determine whether novel actions fail to develop in the absence of short-latency phasic DA signalling. A second issue will be to determine how converging, functionally designated signals interact at the level of individual striatal neurons[69,89,90]. At a higher level of description, it will also be important to identify neuronal circuits external to the basal ganglia that receive the successive approximations of event-related actions/movements and value-based, post-saccadic perceptual analyses of sensory events[100,102–105]. For it is in these structures that a 'library' of action–outcome routines will most likely be assembled[117] and made available to generate novel sequences of adaptive behaviour (BOX 1).

Finally, the present framework might also provide novel insights about mechanisms that underlie some of the behavioural effects of abnormal DA transmission. For example, high levels of DA activity in animals and humans promote the tendency to repeat chunks of behaviour without apparent purpose — for example, pharmacologically induced behavioural stereotypies[118,119]. With the proposed role of DA to promote the repetition of immediately preceding actions/movements, one might predict that tonically high levels of DA transmission could induce the purposeless repetition of actions/movements that are the cause of or correlate with discrete sensory outcomes. More speculatively, a common feature of schizophrenia is a disturbed 'sense of agency'[120,121]. To the extent that this disease is associated with abnormal DA transmission[122], it is possible that 'sense of agency' disturbances could result from the malfunctioning of processes in the basal ganglia that could identify consequences in the world for which the patient feels responsible. At this time it seems more likely that such disturbances would involve the mesolimbic and mesocortical DA projections from the ventral tegmental area, the targets of which serve a wide range of cognitive functions[123].

Peter Redgrave and Kevin Gurney are at the Neuroscience Research Unit, Department of Psychology, University of Sheffield, Sheffield, S10 2TP, UK.
Correspondence to P.R.
e-mail: P.Redgrave@sheffield.ac.uk

doi:10.1938/nrn2022
Published online 8 November 2006

1. Thorndike, E. L. *Animal Intelligence* (Macmillan, New York, 1911).
2. Schultz, W. Predictive reward signal of dopamine neurons. *J. Neurophysiol.* **80**, 1–27 (1998).
3. Redgrave, P., Prescott, T. J. & Gurney, K. Is the short latency dopamine response too short to signal reward error? *Trends Neurosci.* **22**, 146–151 (1999).
4. Comoli, E. *et al.* A direct projection from superior colliculus to substantia nigra for detecting salient visual events. *Nature Neurosci.* **6**, 974–980 (2003).
5. Dommett, E. *et al.* How visual stimuli activate dopaminergic neurons at short latency. *Science* **307**, 1476–1479 (2005).
6. 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).
7. Montague, P. R., Hyman, S. E. & Cohen, J. D. Computational roles for dopamine in behavioural control. *Nature* **431**, 760–767 (2004).
8. Schultz, W. Getting formal with dopamine and reward. *Neuron* **36**, 241–263 (2002).
9. Schultz, W. Behavioral theories and the neurophysiology of reward. *Annu. Rev. Psychol.* **57**, 87–115 (2006).
10. Schultz, W. & Dickinson, A. Neuronal coding of prediction errors. *Annu. Rev. Neurosci.* **23**, 473–500 (2000).
11. Gerfen, C. R. & Wilson, C. J. in *Handbook of Chemical Neuroanatomy* Vol. 12 (eds Swanson, L. W., Bjorklund, A. & Hokfelt, T.) Part III, 371–468 (Elsevier, Amsterdam, 1996).
12. Graybiel, A. M. Neurotransmitter and neuromodulators in the basal ganglia. *Trends Neurosci.* **13**, 244–254 (1990).
13. Hiroi, N. *et al.* Molecular dissection of dopamine receptor signaling. *J. Chem. Neuroanat.* **23**, 237–242 (2002).
14. Bergman, H. *et al.* Physiological aspects of information processing in the basal ganglia of normal and Parkinsonian primates. *Trends Neurosci.* **21**, 32–38 (1998).
15. Radad, K., Gille, G. & Rausch, W. D. Short review on dopamine agonists: insight into clinical and research studies relevant to Parkinson's disease. *Pharm. Rep.* **57**, 701–712 (2005).
16. Wise, R. A. Dopamine, learning and motivation. *Nature Rev. Neurosci.* **5**, 483–494 (2004).
17. Berridge, K. C. & Robinson, T. E. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res. Rev.* **28**, 309–369 (1998).
18. Salamone, J. D. & Correa, M. Motivational views of reinforcement: implications for understanding the behavioral functions of nucleus accumbens dopamine. *Behav. Brain Res.* **137**, 3–25 (2002).
19. Marr, D. *Vision: A Computational Approach* (Freeman & Co., San Francisco, 1982).
20. Gurney, K., Prescott, T. J., Wickens, J. R. & Redgrave, P. Computational models of the basal ganglia: from robots to membranes. *Trends Neurosci.* **27**, 453–459 (2004).
21. Waelti, P., Dickinson, A. & Schultz, W. Dopamine responses comply with basic assumptions of formal learning theory. *Nature* **412**, 43–48 (2001).
22. Fiorillo, C. D., Tobler, P. N. & Schultz, W. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* **299**, 1898–1902 (2003).
23. Tobler, P. N., Fiorillo, C. D. & Schultz, W. Adaptive coding of reward value by dopamine neurons. *Science* **307**, 1642–1645 (2005).
24. Bayer, H. M. & Glimcher, P. W. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47**, 129–141 (2005).
25. Satoh, T., Nakai, S., Sato, T. & Kimura, M. Correlated coding of motivation and outcome of decision by dopamine neurons. *J. Neurosci.* **23**, 9913–9923 (2003).
26. Nakahara, H., Itoh, H., Kawagoe, R., Takikawa, Y. & Hikosaka, O. Dopamine neurons can represent context-dependent prediction error. *Neuron* **41**, 269–280 (2004).
27. Tobler, P. N., Dickinson, A. & Schultz, W. Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. *J. Neurosci.* **23**, 10402–10410 (2003).
28. Ungless, M. A. Dopamine: the salient issue. *Trends Neurosci.* **27**, 702–706 (2004).
29. Sugrue, L. P., Corrado, G. S. & Newsome, W. T. Choosing the greater of two goods: neural currencies for valuation and decision making. *Nature Rev. Neurosci.* **6**, 363–375 (2005).
30. Salzman, C. D., Belova, M. A. & Paton, J. J. Beetles, boxes and brain cells: neural mechanisms underlying valuation and learning. *Curr. Opin. Neurobiol.* **15**, 721–729 (2005).
31. Houk, J. C. Agents of the mind. *Biol. Cybern.* **92**, 427–437 (2005).
32. Suri, R. E. TD models of reward predictive responses in dopamine neurons. *Neural Netw.* **15**, 523–533 (2002).
33. Bar-Gad, I. & Bergman, H. Stepping out of the box: information processing in the neural networks of the basal ganglia. *Curr. Opin. Neurobiol.* **11**, 689–695 (2001).
34. Frank, M. J. Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. *J. Cogn. Neurosci.* **17**, 51–72 (2005).
35. Daw, N. D., Niv, Y. & Dayan, P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neurosci.* **8**, 1704–1711 (2005).
36. Freeman, A. S. Firing properties of substantia nigra dopaminergic neurons in freely moving rats. *Life Sci.* **36**, 1983–1994 (1985).
37. 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).
38. 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).
39. Ljungberg, T., Apicella, P. & Schultz, W. Responses of monkey dopamine neurons during learning of behavioural reactions. *J. Neurophysiol.* **67**, 145–163 (1992).
40. Pan, W. X., Schmidt, R., Wickens, J. R. & Hyland, B. I. Dopamine cells respond to predicted events during classical conditioning: evidence for eligibility traces in the reward- learning network. *J. Neurosci.* **25**, 6235–6242 (2005).
41. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
42. Mirenowicz, J. & Schultz, W. Importance of unpredictability for reward responses in primate dopamine neurons. *J. Neurophysiol.* **72**, 1024–1027 (1994).
43. Coizet, V., Comoli, E., Westby, G. W. M. & Redgrave, P. Phasic activation of substantia nigra and the ventral tegmental area by chemical stimulation of the superior colliculus: an electrophysiological investigation in the rat. *Eur. J. Neurosci.* **17**, 28–40 (2003).
44. Overton, P. G., Coizet, V., Dommett, E. J. & Redgrave, P. The parabrachial nucleus is a source of short latency nociceptive input to midbrain dopaminergic neurones in rat. *Soc. Neurosci. Abstr.* 301.5 (2005).
45. Coizet, V., Dommett, E. J., Redgrave, P. & Overton, P. G. Nociceptive responses of midbrain dopaminergic neurones are modulated by the superior colliculus in the rat. *Neuroscience* **139,** 1479–1493 (2006).
46. McHaffie, J. G. *et al.* A direct projection from superior colliculus to substantia nigra pars compacta in the cat. *Neuroscience* **138**, 221–234 (2006).
47. Horvitz, J. C. Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience* **96**, 651–656 (2000).
48. Takikawa, Y., Kawagoe, R. & Hikosaka, O. A possible role of midbrain dopamine neurons in short- and long-term adaptation of saccades to position-reward mapping. *J. Neurophysiol.* **92**, 2520–2529 (2004).
49. Jay, M. F. & Sparks, D. L. Sensorimotor integration in the primate superior colliculus. I. Motor convergence. *J. Neurophysiol.* **57**, 22–34 (1987).
50. Hikosaka, O. & Wurtz, R. H. Visual and oculomotor function of monkey substantia nigra pars reticulata. I. Relation of visual and auditory responses to saccades. *J. Neurophysiol.* **49**, 1230–1253 (1983).
51. Thorpe, S. J. & Fabre-Thorpe, M. Seeking categories in the brain. *Science* **291**, 260–263 (2001).
52. Rousselet, G. A., Thorpe, S. J. & Fabre-Thorpe, M. How parallel is visual processing in the ventral pathway? *Trends Cogn. Sci.* **8**, 363–370 (2004).
53. 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).
54. Hikosaka, O., Sakamoto, M. & Usui, S. Functional properties of monkey caudate neurons. II. Visual and auditory responses. *J. Neurophysiol.* **61**, 799–813 (1989).
55. Matsumura, M., Kojima, J., Gardiner, T. W. & Hikosaka, O. Visual and oculomotor functions of monkey subthalamic nucleus. *J. Neurophysiol.* **67**, 1615–1632 (1992).
56. May, P. J. *et al.* Projections from the superior colliculus to substantia nigra pars compacta in a primate. *Soc. Neurosci. Abstr.* 450.2 (2005).
57. Katsuta, H. & Isa, T. Release from $GABA_A$ receptor-mediated inhibition unmasks interlaminar connection within superior colliculus in anesthetized adult rats. *Neurosci. Res.* **46**, 73–83 (2003).
58. Wurtz, R. H. & Albano, J. E. Visual-motor function of the primate superior colliculus. *Ann. Rev. Neurosci.* **3**, 189–226 (1980).
59. Sparks, D. L. Translation of sensory signals into commands for control of saccadic eye movements: role of the primate superior colliculus. *Physiol. Rev.* **66**, 118–171 (1986).
60. Grantyn, R. in *Neuroanatomy of the Oculomotor System* (ed. Buttner-Ennever, J. A.) 273–333 (Elsevier, Amsterdam, 1988).
61. Stein, B. E. & Meredith, M. A. *The Merging of the Senses* (MIT Press, Cambridge, Massachusetts, 1993).
62. Horn, G. & Hill, R. M. Effect of removing the neocortex on the response to repeated sensory stimulation of neurones in the mid-brain. *Nature* **211**, 754–755 (1966).
63. Sprague, J. M., Marchiafava, P. L. & Rixxolatti, G. Unit responses to visual stimuli in the superior colliculus of the unanesthetized, mid-pontine cat. *Arch. Ital. Biol.* **106**, 169–193 (1968).
64. Ikeda, T. & Hikosaka, O. Reward-dependent gain and bias of visual responses in primate superior colliculus. *Neuron* **39**, 693–700 (2003).
65. Hikosaka, O., Nakamura, K. & Nakahara, H. Basal ganglia orient eyes to reward. *J. Neurophysiol.* **95**, 567–584 (2006).
66. Sutton, R. S. & Barto, A. G. *Reinforcement Learning – an Introduction* (MIT Press, Cambridge, Massachusetts, 1998).
67. White, N. M. Reward or reinforcement: what's the difference? *Neurosci. Biobehav. Rev.* **13**, 181–186 (1989).
68. McHaffie, J. G., Stanford, T. R., Stein, B. E., Coizet, V. & Redgrave, P. Subcortical loops through the basal ganglia. *Trends Neurosci.* **28**, 401–407 (2005).
69. Reynolds, J. N. J., Schulz, J. M. & Wickens, J. R. Visual responsiveness of striatal spiny neurons in anaesthetised rats: an *in vivo* intracellular study. *Proc. Int. Australas. Wint. Conf. Brain Res. Abstr.* 6.4, 39 (2005).
70. Schultz, W., Apicella, P., Romo, R. & Scarnati, E. in *Models of Information Processing in the Basal Ganglia* (eds Houk, J. C., Davis, J. L. & Beiser, D. G.) 11–27 (MIT Press, Cambridge, Massachusetts, 1995).
71. 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).
72. Samejima, K., Ueda, Y., Doya, K. & Kimura, M. Representation of action-specific reward values in the striatum. *Science* **310**, 1337–1340 (2005).
73. Crutcher, M. D. & DeLong, M. R. Single cell studies of the primate putamen. II. Relations to direction of movement and pattern of muscular activity. *Exp. Brain Res.* **53**, 244–258 (1984).
74. Bickford, M. E. & Hall, W. C. Collateral projections of predorsal bundle cells of the superior colliculus in the rat. *J. Comp. Neurol.* **283**, 86–106 (1989).
75. Levesque, M., Charara, A., Gagnon, S., Parent, A. & Deschenes, M. Corticostriatal projections from layer V cells in rat are collaterals of long-range corticofugal axons. *Brain Res.* **709**, 311–315 (1996).
76. Mink, J. W. The basal ganglia: focused selection and inhibition of competing motor programs. *Prog. Neurobiol.* **50**, 381–425 (1996).
77. Reiner, A., Jiao, Y., DelMar, N., Laverghetta, A. V. & Lei, W. L. Differential morphology of pyramidal tract-type and intratelencephalically projecting-type corticostriatal neurons and their intrastriatal terminals in rats. *J. Comp. Neurol.* **457**, 420–440 (2003).
78. Alexander, G. E., DeLong, M. R. & Strick, P. L. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Ann. Rev. Neurosci.* **9**, 357–381 (1986).

79. Haber, S. N. The primate basal ganglia: parallel and integrative networks. *J. Chem. Neuroanat.* **26**, 317–330 (2003).
80. Harting, J. K., Updyke, B. V. & VanLieshout, D. P. The visual-oculomotor striatum of the cat: functional relationship to the superior colliculus. *Exp. Brain Res.* **136**, 138–142 (2001).
81. Krout, K. E., Loewy, A. D., Westby, G. W. M. & Redgrave, P. Superior colliculus projections to midline and intralaminar thalamic nuclei of the rat. *J. Comp. Neurol.* **431**, 198–216 (2001).
82. Krout, K. E., Belzer, R. E. & Loewy, A. D. Brainstem projections to midline and intralaminar thalamic nuclei of the rat. *J. Comp. Neurol.* **448**, 53–101 (2002).
83. Van der Werf, Y. D., Witter, M. P. & Groenewegen, H. J. The intralaminar and midline nuclei of the thalamus. Anatomical and functional evidence for participation in processes of arousal and awareness. *Brain Res. Rev.* **39**, 107–140 (2002).
84. Smith, Y., Raju, D. V., Pare, J. F. & Sidibe, M. The thalamostriatal system: a highly specific network of the basal ganglia circuitry. *Trends Neurosci.* **27**, 520–527 (2004).
85. Parent, M. & Parent, A. Single-axon tracing and three-dimensional reconstruction of centre median-parafascicular thalamic neurons in primates. *J. Comp. Neurol.* **481**, 127–144 (2005).
86. Matsumoto, N., Minamimoto, T., Graybiel, A. M. & Kimura, M. Neurons in the thalamic CM-Pf complex supply striatal neurons with information about behaviorally significant sensory events. *J. Neurophysiol.* **85**, 960–976 (2001).
87. Wightman, R. M. & Robinson, D. L. Transient changes in mesolimbic dopamine and their association with 'reward'. *J. Neurochem.* **82**, 721–735 (2002).
88. Roitman, M. F., Stuber, G. D., Phillips, P. E. M., Wightman, R. M. & Carelli, R. M. Dopamine operates as a subsecond modulator of food seeking. *J. Neurosci.* **24**, 1265–1271 (2004).
89. Centonze, D., Picconi, B., Gubellini, P., Bernardi, G. & Calabresi, P. Dopaminergic control of synaptic plasticity in the dorsal striatum. *Eur. J. Neurosci.* **13**, 1071–1077 (2001).
90. Reynolds, J. N. & Wickens, J. R. Dopamine-dependent plasticity of corticostriatal synapses. *Neural Netw.* **15**, 507–521 (2002).
91. Wickens, J. *A Theory of the Striatum* (Pergamon, Oxford, 1993).
92. Hikosaka, O. in *The Basal ganglia IV: New Ideas and Data on Structure and Function* (eds Percheron, G., McKenzie, J. S. & Feger, J.) 589–596 (Plenum, New York, 1994).
93. Redgrave, P., Prescott, T. & Gurney, K. N. The basal ganglia: a vertebrate solution to the selection problem? *Neuroscience* **89**, 1009–1023 (1999).
94. Gurney, K., Prescott, T. J. & Redgrave, P. A computational model of action selection in the basal ganglia. I. A new functional anatomy. *Biol. Cybern.* **84**, 401–410 (2001).
95. Gurney, K., Prescott, T. J. & Redgrave, P. A computational model of action selection in the basal ganglia. II. Analysis and simulation of behaviour. *Biol. Cybern.* **84**, 411–423 (2001).
96. Prescott, T. J., Gonzalez, F. M. M., Gurney, K., Humphries, M. D. & Redgrave, P. A robot model of the basal ganglia: behavior and intrinsic processing. *Neural Netw.* **19**, 31–61 (2006).
97. Devenport, L. D. & Holloway, F. A. The rat's resistance to superstition: role of the hippocampus. *J. Comp. Physiol. Psychol.* **94**, 691–705 (1980).
98. Roberts, S. & Gharib, A. Variation of bar-press duration: where do new responses come from? *Behav. Processes* **72**, 215–223 (2006).
99. Wickens, J. R., Reynolds, J. N. J. & Hyland, B. I. Neural mechanisms of reward-related motor learning. *Curr. Opin. Neurobiol.* **13**, 685–690 (2003).
100. Paton, J. J., Belova, M. A., Morrison, S. E. & Salzman, C. D. The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature* **439**, 865–870 (2006).
101. Lisman, J. E. & Grace, A. A. The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron* **46**, 703–713 (2005).
102. Schultz, W. Multiple reward signals in the brain. *Nature Rev. Neurosci.* **1**, 199–207 (2000).
103. Schoenbaum, G., Setlow, B., Saddoris, M. P. & Gallagher, M. Encoding predicted outcome and acquired value in orbitofrontal cortex during cue sampling depends upon input from basolateral amygdala. *Neuron* **39**, 855–867 (2003).
104. Corbit, L. H., Ostlund, S. B. & Balleine, B. W. Sensitivity to instrumental contingency degradation is mediated by the entorhinal cortex and its efferents via the dorsal hippocampus. *J. Neurosci.* **22**, 10976–10984 (2002).
105. Corbit, L. H. & Balleine, B. W. The role of prelimbic cortex in instrumental conditioning. *Behav. Brain Res.* **146**, 145–157 (2003).
106. Padoa-Schioppa, C. & Assad, J. A. Neurons in the orbitofrontal cortex encode economic value. *Nature* **441**, 223–226 (2006).
107. Ungless, M. A., Magill, P. J. & Bolam, J. P. Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. *Science* **303**, 2040–2042 (2004).
108. Klop, E. M., Mouton, L. J., Hulsebosch, R., Boers, J. & Holstege, G. In cat four times as many lamina I neurons project to the parabrachial nuclei and twice as many to the periaqueductal gray as to the thalamus. *Neuroscience* **134**, 189–197 (2005).
109. Dean, P., Redgrave, P. & Westby, G. W. M. Event or emergency? Two response systems in the mammalian superior colliculus. *Trends Neurosci.* **12**, 137–147 (1989).
110. Dickinson, A. The 28th Bartlett Memorial Lecture. Causal learning: an associative analysis. *Q. J. Exp. Psychol. B* **54**, 3–25 (2001).
111. Elsner, B. & Hommel, B. Contiguity and contingency in action-effect learning. *Psychol. Res.* **68**, 138–154 (2004).
112. Yin, H. H., Knowlton, B. J. & Balleine, B. W. Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. *Eur. J. Neurosci.* **22**, 505–512 (2005).
113. Burgdorf, J. & Panksepp, J. The neurobiology of positive emotions. *Neurosci. Biobehav. Rev.* **30**, 173–187 (2006).
114. Roesch, M. R. & Olson, C. R. Neuronal activity related to reward value and motivation in primate frontal cortex. *Science* **304**, 307–310 (2004).
115. McDonald, A. J. Topographical organization of amygdaloid projections to the caudatoputamen, nucleus accumbens, and related striatal-like areas of the rat brain. *Neuroscience* **44**, 15–33 (1991).
116. Fudge, J. L., Kunishio, K., Walsh, P., Richard, C. & Haber, S. N. Amygdaloid projections to ventromedial striatal subterritories in the primate. *Neuroscience* **110**, 257–275 (2002).
117. Singh, S., Barto, A. G. & Chentanez, N. in *Advances in Neural Information Processing Systems 17* (eds Saul, L. K., Weiss, H. & Bottou, L.) 1281–1288 (MIT Press, Cambridge, Massachusetts, 2005).
118. Robbins, T. W. & Sahakian, B. J. in *Metabolic Disorders of the Nervous System* (ed. Rose, F. C.) 244–291 (Pitman, London, 1981).
119. Saka, E., Goodrich, C., Harlan, P., Madras, B. K. & Graybiel, A. M. Repetitive behaviors in monkeys are linked to specific striatal activation patterns. *J. Neurosci.* **24**, 7557–7565 (2004).
120. Daprati, E. *et al.* Looking for the agent: an investigation into consciousness of action and self-consciousness in schizophrenic patients. *Cognition* **65**, 71–86 (1997).
121. Spence, S. A. *et al.* A PET study of voluntary movement in schizophrenic patients experiencing passivity phenomena (delusions of alien control). *Brain* **120**, 1997–2011 (1997).
122. Kapur, S., Mizrahi, R. & Li, M. From dopamine to salience to psychosis — linking biology, pharmacology and phenomenology of psychosis. *Schiz. Res.* **79**, 59–68 (2005).
123. Wise, S. P., Murray, E. A. & Gerfen, C. R. The frontal-cortex-basal ganglia system in primates. *Crit. Rev. Neurobiol.* **10**, 317–356 (1996).
124. Reed, P., Mitchell, C. & Nokes, T. Intrinsic reinforcing properties of putatively neutral stimuli in an instrumental two-level discrimination task. *Anim. Learn. Behav.* **24**, 38–45 (1996).
125. St Clair-Smith, R. & MacLaren, D. Response preconditioning effects. *J. Exp. Psychol Anim. Behav. Process.* **9**, 41–48 (1983).

Acknowledgments

This work has been supported by the Wellcome Trust (P.R.) and the Engineering and Physical Sciences Research Council (K.G. and P.R.). For their helpful discussions and/or comments on early drafts of the manuscript the authors would like to acknowledge J. Berke, J. Reynolds, A. Seth, E. Salinas, T. Stanford, J. McHaffie, T. Prescott, P. Overton and T. Dickinson.

Competing interests statement

The authors declare no competing financial interests.

FURTHER INFORMATION

Redgrave's laboratory: http://www.abrg.group.shef.ac.uk/reverb/public

Access to this interactive links box is free online.