CrossMark

Reward functions of the basal ganglia

Wolfram Schultz[1]

Abstract Besides their fundamental movement function evidenced by Parkinsonian deficits, the basal ganglia are involved in processing closely linked non-motor, cognitive and reward information. This review describes the reward functions of three brain structures that are major components of the basal ganglia or are closely associated with the basal ganglia, namely midbrain dopamine neurons, pedunculopontine nucleus, and striatum (caudate nucleus, putamen, nucleus accumbens). Rewards are involved in learning (positive reinforcement), approach behavior, economic choices and positive emotions. The response of dopamine neurons to rewards consists of an early detection component and a subsequent reward component that reflects a prediction error in economic utility, but is unrelated to movement. Dopamine activations to non-rewarded or aversive stimuli reflect physical impact, but not punishment. Neurons in pedunculopontine nucleus project their axons to dopamine neurons and process sensory stimuli, movements and rewards and reward-predicting stimuli without coding outright reward prediction errors. Neurons in striatum, besides their pronounced movement relationships, process rewards irrespective of sensory and motor aspects, integrate reward information into movement activity, code the reward value of individual actions, change their reward-related activity during learning, and code own reward in social situations depending on whose action produces the reward. These data demonstrate a variety of well-characterized reward processes in specific basal ganglia nuclei consistent with an important function in non-motor aspects of motivated behavior.

Keywords Dopamine · Pedunculopontine nucleus · Striatum

Introduction

The functions of the basal ganglia are closely linked to movements. This view derives from deficits arising after lesions, from anatomical connections with other brain structures with well-understood motor functions, and from the neuronal activity recorded in behaving animals. However, the effects of lesions in nucleus accumbens and of electrical self-stimulation of dopamine neurons point also to non-motor functions, and more specifically to reward and motivation (Kelly et al. 1975; Corbett and Wise 1980; Fibiger et al. 1987). All basal ganglia nuclei show distinct and sophisticated forms of reward processing, often combined with movement-related activity. These motor and non-motor functions are closely linked; for example, large fractions of neurons in the striatum show both movement- and reward-related activity (e.g., Hollerman et al. 1998; Kawagoe et al. 1998; Samejima et al. 2005). Such combined processing may be the hallmark of a neuronal system involved in goal-directed behavior and habit learning (Yin et al. 2004, 2005), which require the processing of reward information and of the actions required to obtain the reward (Dickinson and Balleine 1994).

This review describes electrophysiological recordings of the coding of reward prediction errors by dopamine neurons, the possible contribution of inputs from the pedunculopontine nucleus (PPN) to this signal, and the possible influence of this dopamine signal on action value coding in

✉ Wolfram Schultz
ws234@cam.ac.uk

[1] Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge CB2 3DY, UK

the striatum. However, neuronal signals for various aspects of reward, such as prediction, reception and amount, exist also in all other basal ganglia structures, including the globus pallidus (Gdowski et al. 2001; Tachibana and Hikosaka 2012), subthalamic nucleus (Lardeux et al. 2009) and pars reticulata of substantia nigra (Cohen et al. 2012; Yasuda et al. 2012). The reviewed studies concern neurophysiological recordings from monkeys, rats and mice during performance of controlled behavioral tasks involving learning of new stimuli and choices between known reward-predicting stimuli. Most of the cited studies involve recordings from individual neurons, one at a time, while the animal performs a standard task, often together with specific controls. To this end, monkeys sit, or rodents lie, in specially constructed chairs or chambers in the laboratory where they are fully awake and relaxed and react to stimuli mostly with arm or eye movements to obtain various types and amounts of liquid or food rewards. The stimuli are often presented on computer monitors in front of the animals and pretrained to predict specific rewards. Rewards are precisely quantified drops of fruit juices or water that are quickly delivered at specific time points under computer control, thus eliciting well defined, phasic stimulation of somatosensory receptors at the mouth. The experimental designs are based on constructs from animal learning theory and economic decision theory that conceptualize the functions of rewards in learning and choices. The experiments comprise the learning and updating of behavioral acts, the elicitation of approach behavior and the choice between differently rewarded levers or visual targets. Conceptual and experimental details are found in Schultz (2015).

Midbrain dopamine neurons

Studies of the behavioral deficits in Parkinson's disease and schizophrenia help researchers to develop hypotheses about the functions of dopamine in the brain. The symptoms are probably linked to disorders in slowly changing or tonic dopamine levels. By contrast, electrophysiological recordings from individual dopamine neurons in substantia nigra pars compacta and ventral tegmental area identify a specific, phasic signal reflecting reward prediction error.

Prediction error

Dopamine neurons in monkeys, rats and mice show phasic responses to food and liquid rewards and to stimuli predicting such rewards in a wide variety of Pavlovian and operant tasks (Ljungberg et al. 1991, 1992; Schultz et al. 1993, 1997; Hollerman and Schultz 1998; Satoh et al. 2003; Morris et al. 2004; Pan et al. 2005; Bayer and

Glimcher 2005; Nomoto et al. 2010; Cohen et al. 2012; Kobayashi and Schultz 2014). In all of these diverse tasks, the dopamine signal codes a reward prediction error, namely the difference between received and predicted reward. A reward that is better than predicted at a given moment in time (positive reward prediction error) elicits a phasic activation, a reward that occurs exactly as predicted in value and time (no prediction error) elicits no phasic change in dopamine neurons, and a reward that is worse than predicted at the predicted time (negative prediction error) induces a phasic depression in activity. Reward-predicting stimuli evoke similar prediction error responses, suggesting that dopamine neurons treat rewards and reward-predicting stimuli commonly as events that convey value. These responses occur also in more complex tasks, including delayed response, delayed alternation and delayed matching-to-sample (Ljungberg et al. 1991; Schultz et al. 1993; Takikawa et al. 2004), sequential movements (Satoh et al. 2003; Nakahara et al. 2004; Enomoto et al. 2011), random dot motion discrimination (Nomoto et al. 2010), somatosensory signal detection (de Lafuente and Romo 2011) and visual search (Matsumoto and Takada 2013). The prediction error response involves 70–90 % of dopamine neurons, is very similar in latency across the dopamine neuronal population, and shows only graded rather than categorical differences between medial and lateral neuronal groups (Ljungberg et al. 1992) or between dorsal and ventral groups (Nomoto et al. 2010; Fiorillo et al. 2013a). No other brain structure shows such a global and stereotyped reward signal with similar response latencies and durations across neurons (Schultz 1998). The homogeneous signal leads to locally varied dopamine release that acts on heterogeneous postsynaptic structures and thus results in diverse dopamine functions.

The phasic dopamine reward signal satisfies stringent tests for bidirectional prediction error coding suggested by formal animal learning theory, such as blocking and conditioned inhibition (Waelti et al. 2001; Tobler et al. 2003; Steinberg et al. 2013). With these characteristics, the dopamine prediction error signal implements teaching signals of Rescorla–Wagner and temporal difference (TD) reinforcement learning models (Rescorla and Wagner 1972; Sutton and Barto 1981; Mirenowicz and Schultz 1994; Montague et al. 1996; Enomoto et al. 2011). Via a three-factor synaptic arrangement, the dopamine reinforcing signal would affect coincident synaptic transmission between cortical inputs and postsynaptic striatal, frontal cortex or amygdala neurons (Freund et al. 1984; Goldman-Rakic et al. 1989; Schultz 1998), both immediately and via Hebbian plasticity. Positive dopamine prediction error activation would enhance behavior-related neuronal activity and thus favor behavior that leads to increased reward, whereas negative dopamine prediction error depression

would reduce neuronal activity and thus disfavor behavior resulting diminished reward.

The phasic dopamine signal has all characteristics of effective teaching signals for model-free reinforcement learning. In addition, the signal incorporates predictions from models of the world (acquired by other systems) (Nakahara et al. 2004; Tobler et al. 2005; Bromberg-Martin et al. 2010), thus possibly serving to update predictions using a combination of (model-free) experience and model-based representations. About one-third of dopamine neurons show also a slower, pre-reward activation that varies with reward risk (mixture of variance and skewness) (Fiorillo et al. 2003), which constitutes the first neuronal risk signal ever observed. A risk signal derived purely from variance and distinct from value exists in orbitofrontal cortex (O'Neill and Schultz 2010). The dopamine risk response might be appropriate for mediating the influence of attention on the learning rate of specific learning mechanisms (Pearce and Hall 1980) and thus would support the teaching function of the phasic prediction error signal. The effects of electrical and optogenetic activation further support a teaching function of the phasic dopamine response (Corbett and Wise 1980; Tsai et al. 2009; Adamantidis et al. 2011; Steinberg et al. 2013), thus suggesting a causal influence of phasic dopamine signals on learning.

Unclear movement relationships

Besides serving as reinforcement for learning, stimulation of dopamine neurons elicits immediate behavioral actions, including contralateral rotation, locomotion (Kim et al. 2012), food seeking (Adamantidis et al. 2011) and approach behavior (Hamid et al. 2015); stimulation of striatal neurons expressing specific dopamine receptor subtypes induces differential contralateral or ipsilateral choice preferences (Tai et al. 2012). Although these dopamine effects may be related to the role of dopamine in Parkinson's disease, the phasic dopamine response does not code movement. The dopamine responses to conditioned stimuli occur close to the time of the movement evoked by such stimuli, but dissection of temporal relationships reveals the close association with stimuli rather than movements (Ljungberg et al. 1992). Slower electrophysiological dopamine changes occur with movements (Schultz et al. 1983; Schultz 1986; Romo and Schultz 1990), but are sluggish and inconsistent and often fail to occur in better controlled behavioral tasks (DeLong et al. 1983; Schultz and Romo 1990; Ljungberg et al. 1992; Waelti et al. 2001; Satoh et al. 2003; Cohen et al. 2012; Lak et al. 2014); they seem to reflect general behavioral reactivity rather than specific motor processes. Similar or even slower changes in dopamine release occur with

reward, motivation, stress, punishment, movement and attention (Young et al. 1992; Cheng et al. 2003; Young 2004; Howe et al. 2013). The lack of movement relationships of dopamine neurons is compatible with the fact that dopamine receptor stimulation improves hypokinesia without restoring phasic dopamine activity. Apparently, the Parkinsonian deficits do not reflect phasic movement-related dopamine changes; rather, tonic, ambient dopamine concentrations seem to underlie the movements and other behavioral processes that are deficient in this disorder. Taken together, these dopamine effects demonstrate the different time scales and wide spectrum of dopamine influences (Grace 1991; Grace et al. 2007; Schultz 2007; Robbins and Arnsten 2009) and suggest that the phasic dopamine reward signal does not explain Parkinsonian motor deficits in a simple way.

Two phasic response components

The phasic dopamine signal consists of two distinct components (Mirenowicz and Schultz 1996; Waelti et al. 2001; Tobler et al. 2003; Day et al. 2007; Joshua et al. 2008; Fiorillo et al. 2013b), similar to other, non-dopamine neurons involved in sensory and cognitive processing (Thompson et al. 1996; Kim and Shadlen 1999; Ringach et al. 1997; Shadlen and Newsome 2001; Bredfeldt and Ringach 2002; Roitman and Shadlen 2002; Mogami and Tanaka 2006; Paton et al. 2006; Roelfsema et al. 2007; Ambroggi et al. 2008; Lak et al. 2010; Peck et al. 2013; Stanisor et al. 2013; Pooresmaeili et al. 2014). The first dopamine response component consists of a brief activation that begins with latencies of 60–90 ms and lasts 50–100 ms; it is unselective and arises even with motivationally neutral events, conditioned inhibitors and punishers (Steinfels et al. 1983; Schultz and Romo 1990; Mirenowicz and Schultz 1996; Horvitz et al. 1997; Tobler et al. 2003; Joshua et al. 2008; Kobayashi and Schultz 2014), apparently before the stimuli and their reward values have been properly identified. The component is highly sensitive to sensory intensity (Fiorillo et al. 2013b), reward generalization (due to similarity to unrewarded stimuli) (Mirenowicz and Schultz 1996; Day et al. 2007), reward context (Kobayashi and Schultz 2014), and novelty (Ljungberg et al. 1992); thus it codes distinct forms of salience related to these physical, motivational and novelty aspects. The response is sensitive to prediction (Nomoto et al. 2010) and represents the initial, unselective, salience part of the dopamine prediction error response. The second dopamine response component begins already during the initial component; it codes reward value as prediction error and thus constitutes the specific phasic dopamine reward response. The two components become completely separated in more demanding tasks, such as random dot motion

discrimination, in which the first component stays constant (Fig. 1a, blue), whereas the second dopamine response component begins later, at latencies around 250 ms, and varies with reward value (red) (Nomoto et al. 2010). The transient, initial dopamine response to physically intense stimuli may mislead towards assuming a primary, full attentional dopamine function if rewards are not tested and the second component is not revealed (Steinfels et al. 1983; Horvitz et al. 1997; Redgrave et al. 1999).

After the stimulus identification by the second component, the reward representation stays on in dopamine neurons; this is evidenced by the prediction error response at the time of the reward, which reflects the predicted value at that moment (Fig. 1b) (Tobler et al. 2003; Nomoto et al. 2010). The early onset of the value component before the behavioral action explains why animals usually discriminate well between rewarded and unrewarded stimuli

despite the initial, indiscriminate dopamine response component (Ljungberg et al. 1992; Joshua et al. 2008; Kobayashi and Schultz 2014). Thus, the second, value component contains the principal dopamine reward value message.

Advantage of initial dopamine activation

The initial activation reflects different components of stimulus-driven salience and may be beneficial for neuronal reward processing. The physical and motivational salience components may affect the speed and accuracy of actions (Chelazzi et al. 2014) via similar neuronal mechanisms as the enhancement of sensory processing by stimulus-driven physical salience (Gottlieb et al. 1998; Thompson et al. 2005). The novelty salience component may promote reward learning via the learning rate, as
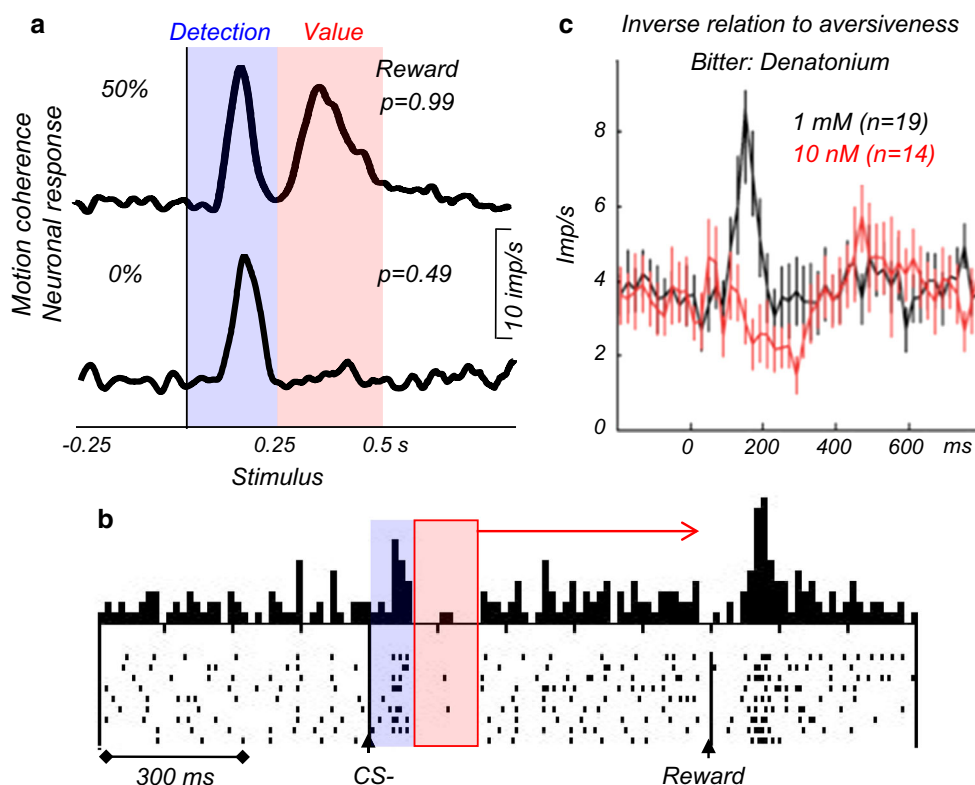
Fig. 1 Basic characteristics of phasic dopamine responses. **a** Two dopamine response components: initial detection response (*blue*), and subsequent value response (*red*) in a dot motion discrimination task. The motion coherence increasing from 0 to 50 % leads to better behavioral dot motion discrimination, which translates into increases of reward probability from $p = 0.49$ to $p = 0.99$ [dopamine neurons process reward probability as value (Fiorillo et al. 2003)]. The first response component is constant (*blue*), whereas the second component grows with reward value derived from probability (reward prediction error). From Nomoto et al. (2010). **b** Accurate value coding at the time of reward despite initial indiscriminate stimulus detection response. *Blue* and *red zones* indicate the initial detection response

and the subsequent value response, respectively. After an unrewarded stimulus (CS-), surprising reward (R) elicits a positive prediction error response, suggesting that the prediction at reward time reflects the lack of value prediction by the CS-. From Waelti et al. (2001). **c** Inverse relationship of dopamine activations to aversiveness of bitter solutions. The activation to the aversive solution (*black*, Denatonium, a strong bitter substance) turns into a depression with increasing aversiveness due to negative value (*red*), suggesting that the activation reflects physical impact rather than punishment. *Imp/s* impulses per second, *n* number of dopamine neurons. Time = 0 indicates onset of liquid delivery. From Fiorillo et al. (2013b)

conceptualized by attentional learning rules (Pearce and Hall 1980). However, the initial dopamine activation is only a transient salience signal, as it is quickly replaced by the subsequent value component that conveys accurate reward value information. In this way, the initial dopamine activation is beneficial for neuronal processing and learning without the cost of unfocusing or misleading the behavior.

The unselective activation may help the animal to gain more rewards. Its high sensitivity to stimulus intensity, reward similarity, reward context and novelty assures the processing of a maximal number of stimuli and avoids missing a reward. Through its short latency, the initial dopamine response detects these stimuli very rapidly, even before having identified their value. As stimulation of dopamine neurons and their postsynaptic striatal neurons induces learning and approach behavior (Tsai et al. 2009; Tai et al. 2012), the fast dopamine response might induce early movement preparation and thus speed up reward acquisition before a competitor arrives, which is particularly precious in times of scarceness. Yet the response is brief enough for canceling movement preparation if the stimulus turns out not to be a reward, and errors and unnecessary energy expenditure can be avoided. Thus, the two-component structure with the early component may facilitate rapid behavioral reactions resulting in more rewards.

Confounded aversive activations

The initial dopamine response component arising with unrewarded stimuli occurs also with punishers. This activation (Mirenowicz and Schultz 1996) may appear like an aversive signal (Guarraci and Kapp 1999; Joshua et al. 2008) and might suggest a role in motivational salience common to rewards and punishers (Matsumoto and Hikosaka 2009). However, this interpretation fails to take the physical stimulus components of punishers into account, in addition to reward generalization and context. Indeed, independent variations of physical stimulus intensity and aversiveness show positive correlations of dopamine activations with physical intensity, but negative correlations with aversiveness of punishers (Fiorillo et al. 2013a, b). An aversive bitter solution, such as denatonium, induces substantial dopamine activations, whereas its tenfold higher concentration in same-sized drops elicits depressions (Fig. 1c). The activations likely reflect the physical impact of the liquid drops on the monkey's mouth, whereas the depression undercutting the activation may reflect the absence of reward (negative prediction error) or negative punisher value (Fiorillo 2013). The absence of bidirectional prediction error responses with punishers (Joshua et al. 2008; Matsumoto and Hikosaka 2009; Fiorillo 2013)

supports also the physical intensity account. Graded regional differences in responses to aversive stimuli (Matsumoto and Hikosaka 2009; Brischoux et al. 2009) may reflect sensitivity differences of the initial dopamine activation to physical salience, reward generalization and reward context (Mirenowicz and Schultz 1996; Fiorillo et al. 2013b; Kobayashi and Schultz 2014); this might also explain the stronger activations by conditioned compared to unconditioned punishers (Matsumoto and Hikosaka 2009) that defy basic notions of animal learning theory. Correlations with punisher probability (Matsumoto and Hikosaka 2009) may reflect the known sensitivity to salience differences between stimuli (Kobayashi and Schultz 2014). Thus, the dopamine activation to punishers might be explained by other factors than aversiveness, and one may wonder how many aversive dopamine activations remain when all confounds are accounted for.

Subjective value

The value of rewards originates in the organism's requirements for nutritional and other substances. Thus, reward value is subjective and not entirely determined by physical parameters. A good example is satiation, which reduces the value of food rewards, although the food remains physically unchanged. The usual way to assess subjective value involves eliciting behavioral preferences in binary choices between different rewards. The subjective value can then be expressed as measured choice frequencies or as the amount of a reference reward against which an animal is indifferent in binary choices. This measure varies on an objective, physical scale (e.g., ml of juice for animals). Typical for subjective value, preferences differ between individual monkeys (Lak et al. 2014). The choices reveal rank-ordered subjective reward values and satisfy formal transitivity, suggesting meaningful choices by the animals.

The phasic dopamine prediction error signal follows closely the rank-ordered subjective values of different liquid and food rewards (Lak et al. 2014). The dopamine signal reflects also the arithmetic sum of positive and negative subjective values of rewards and punishers (Fiorillo et al. 2013b). Thus, dopamine neurons integrate different outcomes into a subjective value signal.

Subjective reward value is also determined by the risk with which rewards occur. Risk avoiders view lower subjective value, and risk seekers higher value, in risky compared to safe rewards of equal mean physical amount. Correspondingly, dopamine value responses to risk-predicting cues are reduced in monkeys that avoid risk and enhanced when they seek risk (Lak et al. 2014; Stauffer et al. 2014). Voltammetric measurements show similar risk-dependent dopamine changes in rat nucleus

accumbens, which follow the risk attitudes of the individual animals (Sugam et al. 2012). Thus, the reward value signal of dopamine impulses and dopamine release reflects the influence of risk on subjective value.

A further contribution to subjective value is the temporal delay to reward. Temporal discounting reduces reward value even when the physical reward remains unchanged. We usually prefer receiving £100 now than in 3 months. Monkeys show temporal discounting across delays of a few seconds in choices between early and late rewards. The value of the late reward, as assessed by the amount of early reward at choice indifference (point of subjective equivalence), decreases monotonically in a hyperbolic or exponential fashion. Accordingly, phasic dopamine responses to reward-predicting stimuli decrease across these delays (Fiorillo et al. 2008; Kobayashi and Schultz 2008). Voltammetric measurements show corresponding dopamine changes in rat nucleus accumbens reflecting temporal discounting (Day et al. 2010). Taken together, in close association with behavioral preferences, dopamine neurons code subjective value with different rewards, risky rewards and delayed rewards.

Formal economic utility

Utility provides a mathematical characterization of reward preferences (Von Neumann and Morgenstern 1944; Kagel et al. 1995) and constitutes the most theory-constrained measure of subjective reward value. Whereas subjective value estimated from direct preferences or choice indifference points is expressed on an objective, physical scale, formal economic utility provides an internal measure of subjective value (Luce 1959) that is often called utils. Experimental economics tools allow constructing continuous, quantitative, numeric mathematical utility functions from behavioral choices between risky rewards (Von Neumann and Morgenstern 1944; Caraco et al. 1980; Machina 1987). The best-defined test for symmetric, variance risk employs binary equiprobable gambles (Rothschild and Stiglitz 1970). Estimated in this way, utility functions in monkeys are nonlinear and have inflection points between convex, risk seeking and concave, risk avoidance domains (Fig. 2 red) (Stauffer et al. 2014).

In providing the ultimate formal definition of reward value for decision-making, utility should be employed for investigating neuronal reward signals, instead of other, less direct measures of subjective reward value. Dopamine responses to unpredicted, free juice rewards show similar nonlinear increases (Fig. 2 black) (Stauffer et al. 2014). The neuronal responses increase only very slightly with small reward amounts where the behavioral utility function is flat, then linearly with intermediate rewards, and then again more slowly, thus following the nonlinear curvature
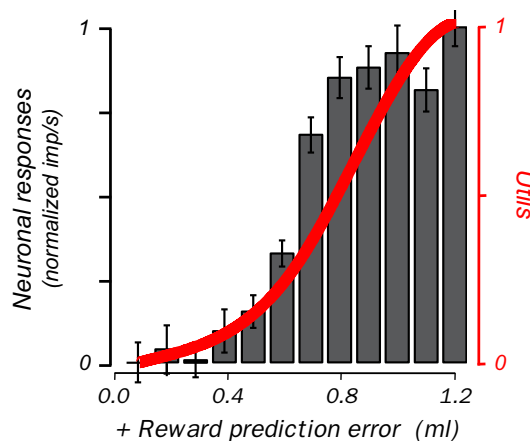
Fig. 2 Utility prediction error signal in monkey dopamine neurons. *Red* utility function derived from behavioral choices using risky gambles. *Black* corresponding, nonlinear increase of population response ($n = 14$ dopamine neurons) in same animal to unpredicted juice. *Norm imp/s* normalized impulses per second. From Stauffer et al. (2014)

of the utility function rather than the linear increase in physical amount. Testing with well-defined risk in binary gambles required by economic theory results in very similar nonlinear changes in prediction error responses (Stauffer et al. 2014). With all factors affecting utility, such as risk, delay and effort cost, held constant, this signal reflects income utility rather than net benefit utility. These data suggest that the dopamine reward prediction error response constitutes a utility prediction error signal and implements the elusive utility in the brain. This neuronal signal reflects an internal metric of subjective value and thus extends well beyond the coding of subjective value derived from choices.

Pedunculopontine nucleus

The PPN projects to midbrain dopamine neurons, pars reticulata of substantia nigra, internal globus pallidus and subthalamic nucleus. Other major inputs to dopamine neurons arise from striatum, subthalamic nucleus and GABAergic neurons of pars reticulata of substantia nigra (Mena-Segovia et al. 2008; Watabe-Uchida et al. 2012). Inputs to PPN derive from cerebral cortex via internal globus pallidus and subthalamic nucleus, and from the thalamus, cerebellum, forebrain, spinal cord, pons and contralateral PPN.

Neurons in the PPN of monkeys, cats and rats show considerable and heterogeneous activity related to a large range of events and behavior. The anatomical and chemical identity of these different neuron types is unknown. Subpopulations of PPN neurons are activated, or

sometimes depressed, by sensory stimuli irrespective of predicting reward (Dormont et al. 1998; Pan and Hyland 2005). Some PPN neurons show spatially tuned activity with saccadic eye movements (Kobayashi et al. 2002; Hong and Hikosaka 2014). In other studies, PPN neurons show differential phasic or sustained activations following reward-predicting stimuli and rewards (Fig. 3) (Kobayashi et al. 2002; Kobayashi and Okada 2007; Okada et al. 2009; Norton et al. 2011; Hong and Hikosaka 2014). Different from dopamine neurons, distinct PPN neurons code reward-predicting stimuli and rewards, rather than both together. Sustained activations following reward-predicting stimuli continue until reward delivery in some PPN neurons; with reward delays, the activation continues until the reward finally occurs (Okada et al. 2009). PPN neurons differentiate between reward amounts. They show usually higher activity to stimuli predicting larger rewards and lower activations or outright depressions to stimuli predicting smaller rewards (Okada et al. 2009; Hong and Hikosaka 2014). Their responses to reward delivery show similarly graded coding, although without displaying depressions. Thus, PPN neurons show various activities during behavioral tasks that are separately related to sensory stimuli, movements, reward-predicting stimuli and reward reception.

The PPN responses to reward delivery have complex relationships to prediction. Some PPN neurons are activated by both predicted and unpredicted rewards, but their latencies are shorter with predicted rewards, and the responses sometimes anticipate the reward (Kobayashi et al. 2002). Some neurons code reward amount irrespective of the rewards being predicted or not. They are not depressed by omitted or delayed rewards and show an activation to the reward whenever it occurs, irrespective of this being at the predicted or a delayed time (Okada et al. 2009; Norton et al. 2011). Other PPN neurons are depressed by smaller rewards randomly alternating with larger

rewards (Hong and Hikosaka 2014), thus showing relationships to average reward predictions. Thus, reward responses of PPN neurons show some prediction effects, but do not display outright bidirectional reward prediction error responses in the way dopamine neurons do.

Some of the reward responses of PPN neurons may induce components of the dopamine reward prediction error signal after conduction via known PPN projections to the midbrain. Electrical stimulation of PPN under anesthesia induces fast and strong burst activations in 20–40 % of dopamine neurons, in particular in spontaneously bursting dopamine neurons (Scarnati et al. 1984; Lokwan et al. 1999). Non-NMDA receptor and acetylcholine receptor antagonists differentially reduce excitations of dopamine neurons, substantiated as EPSPs or extracellularly recorded action potentials (Scarnati et al. 1986; Di Loreto et al. 1992; Futami et al. 1995), suggesting an involvement of both glutamate and acetylcholine in driving dopamine neurons. Electrical PPN stimulation in behaving monkeys induces activations in monkey midbrain dopamine neurons (Hong and Hikosaka 2014). Correspondingly, inactivation of PPN neurons by local anesthetics in behaving rats reduces dopamine prediction error responses to conditioned, reward-predicting stimuli (Pan and Hyland 2005). PPN neurons differentiating between reward amounts with positive and negative responses project to midbrain targets above the substantia nigra, as shown by their antidromic activation from this region (Hong and Hikosaka 2014). Consistent with conduction of neuronal excitation from PPN to dopamine neurons, latencies of neuronal stimulus responses are slightly shorter in PPN compared to dopamine neurons (Pan and Hyland 2005). Through these synaptic influences, different groups of PPN neurons may separately induce components of the dopamine reward prediction error signal, including responses to reward-predicting stimuli, activations by unpredicted rewards, and depressions by smaller-than-predicted

Fig. 3 Reward processing in monkey pedunculopontine nucleus. **a** Phasic and sustained responses to reward-predicting stimuli. *Imp/s* impulses per second. From Kobayashi and Okada (2007). **b** Magnitude discriminating reward responses. *Norm imp/s* normalized impulses per second. From Okada et al. (2009)
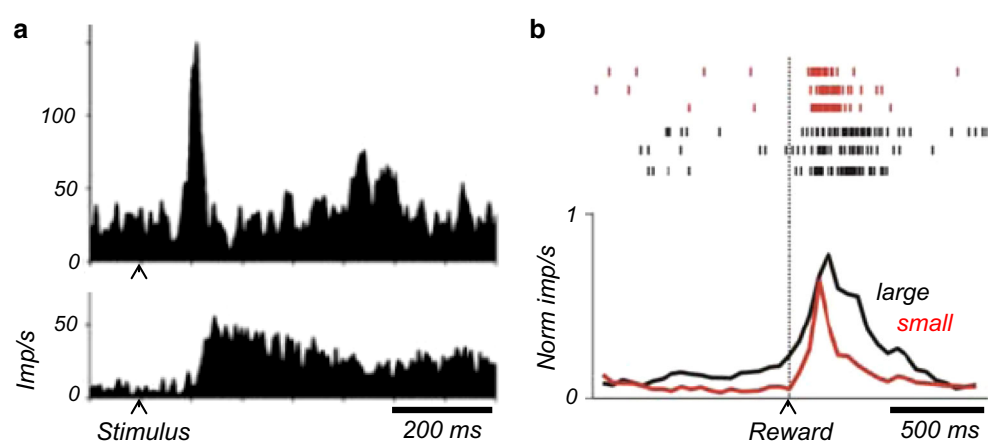
rewards. However, it is unknown whether PPN neurons with response characteristics not seen in dopamine neurons, such as responses to fully predicted rewards, affect dopamine neurons.

Striatum

The deficits arising from Parkinson's disease, Huntington's chorea, dyskinesias, obsessive–compulsive disorder and other movement and cognitive disorders suggest a prominent function of the striatum in motor processes and cognition. Consistent with this functional diversity, the three distinct groups of phasically firing, tonically firing and fast-spiking neurons in the striatum (caudate nucleus, putamen, nucleus accumbens) show a variety of behavioral relationships when sufficiently sophisticated tasks permit their detection. Each of these behavioral relationships engages relatively small fractions of different striatal neurons. However, apart from the small groups of tonically firing interneurons or fast-spiking interneurons, the anatomical and chemical identities of the different functional categories of the large group of medium-spiny striatal neurons are poorly understood. Most of these heterogeneous neurons are influenced by reward information.

Pure reward

All groups of striatal neurons process reward information without reflecting sensory stimulus components or movements. Some of these neurons show selective responses following reward-predicting stimuli or liquid or food rewards (Kimura et al. 1984; Hikosaka et al. 1989; Apicella et al. 1991, 1992; Bowman et al. 1996; Shidara et al. 1998; Ravel et al. 1999; Adler et al. 2013). Striatal responses to reward-predicting stimuli discriminate between different reward types irrespective of predictive stimuli and movements (Fig. 4a) (Hassani et al. 2001). Other groups of striatal neurons show slower, sustained increases of activity for several seconds during the expectation of reward evoked by predictive stimuli (Hikosaka et al. 1989; Apicella et al. 1992). Some of these activations begin and end at specific, predicted time points and reflect the time of reward occurrence (Schultz et al. 1992). Thus, some striatal neurons show passive responses to reward-predicting stimuli and rewards and sustained activities in anticipation of predicted rewards without coding sensory or motor information.

Striatal responses to reward-predicting stimuli vary monotonically with reward amount (Cromwell and Schultz 2003; Báez-Mendoza et al. 2013). Increasing reward probability enhances reward value in a similar way as increasing reward amount. Accordingly, striatal neurons code reward probability (Samejima et al. 2005; Pasquereau et al. 2007; Apicella et al. 2009; Oyama et al. 2010). However, in these tests, objective and subjective reward values are monotonically related to each other; increasing objective value increases also subjective value. The difference in these value measures can be better tested by making different alternative rewards available and thus changing behavioral preferences and subjective reward value without affecting objective value. Correspondingly, some striatal neurons show stronger responses to whichever reward is more preferred by the animal irrespective of its objective value, suggesting subjective rather than objective reward value coding (Cromwell et al. 2005). In a different test for subjective value, striatal reward responses decrease with increasing reward delays, despite unchanged physical reward amount (Roesch et al. 2009; Day et al. 2011). Taken together, groups of striatal neurons signal subjective reward value.

Some striatal neurons respond to surprising rewards and reward prediction errors. Some of them show full, bidirectional coding, being activated by positive prediction errors and depressed by negative errors, although inversely coding neurons exist also (Apicella et al. 2009; Kim et al. 2009; Ding and Gold 2010; Oyama et al. 2010). These responses may affect plasticity during reinforcement learning in a similar manner as dopamine signals, although the anatomically more specific striatal projections onto select groups of postsynaptic neurons would suggest more point-to-point influences on neurons. Other striatal neurons respond either to unpredicted rewards or to reward omission (Joshua et al. 2008; Kim et al. 2009; Asaad and Eskandar 2011). Responses in some of them are stronger in Pavlovian than in operant tasks (Apicella et al. 2011) or occur only after particular behavioral actions (Stalnaker et al. 2012), suggesting selectivity for the behavior that resulted in the error. These unidirectional responses may confer surprise salience or single components of reward prediction errors.

Conjoint reward and action

In contrast to pure reward signals, some reward neurons in the striatum code reward together with specific actions. These neurons differentiate between movement and no-movement reactions (go-nogo) or between spatial target positions during the instruction, preparation and execution of action (Fig. 4b) (Hollerman et al. 1998; Lauwereyns et al. 2002; Hassani et al. 2001; Cromwell and Schultz 2003; Ding and Gold 2010). By processing information about the forthcoming reward during action preparation or execution, these activities may reflect reward representations before and during the action toward the reward, which suggests a relationship to goal-directed behavior
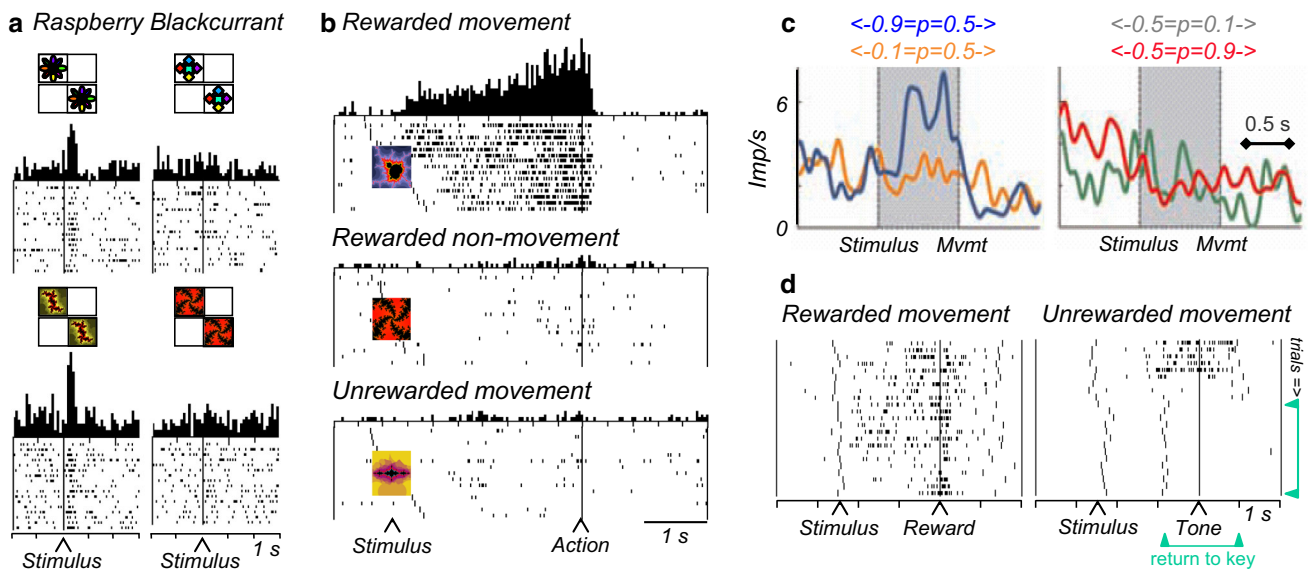
Fig. 4 Reward processing in monkey striatum. **a** Pure reward signal in ventral striatum. The neuron discriminates between raspberry and blackcurrant juice irrespective of movement to *left* or *right* target, and irrespective of the visual image predicting the juice (*top* vs. *bottom*). Trials in rasters are ordered from *top* to *bottom* according to *left* and then *right* stimulus presentation. From Hassani et al. (2001). **b** Conjoint processing of reward (vs. no reward) and movement (vs. no movement) in caudate nucleus (delayed go-nogo-ungo task). The neuronal activities reflect the specific future reward together with the specific action required to obtain that reward. From Hollerman et al. (1998). **c** Action value coding of single striatal neuron. Activity increases with value (probability) for *left* action (*left* panel *blue* vs. *orange*), but is unaffected by value changes for *right* action (*right* panel), indicating *left* action value coding. *Imp/s* impulses per second.

From Samejima et al. (2005). **d** Adaptation of reward expectation activity in ventral striatum during learning. In each learning episode, two new visual stimuli instruct a rewarded and an unrewarded arm movement, respectively, resulting in different reward expectations for the same movement. With rewarded movements (*left*), the animal's hand returns quickly to the resting key after reward delivery (long vertical markers, *right* to reward). With pseudorandomly alternating unrewarded movements, the hand returns quickly after an unrewarded tone to the resting key in initial trials (*top right*), but subsequently returns before the tone (*green arrows*), indicating initial reward expectation that disappears with learning. The reward expectation-related neuronal activity (*short dots*) shows a similar development during learning (from *top* to *bottom*). From Tremblay et al. (1998)

(Dickinson and Balleine 1994). The reward influence on striatal movement activity is so strong that rewards presented at specific positions can alter the spatial preferences of saccade-related activity (Kawagoe et al. 1998). Thus, some striatal neurons integrate reward information into action signals and thus inform about the value of a chosen action. Their activity concerns the integration of reward information into motor processing and thus extends well beyond primary motor functions.

Action value

Specific actions lead to specific rewards with specific values. Action value reflects the reward value (amount, probability, utility) that is obtained by a particular action, thus combining non-motor (reward value) with motor processes (action). If more reward occurs at the left compared to the right, action value is higher for a left than a right movement. Thus, action value is associated with an action, irrespective of this action being chosen. Action value is conceptualized in machine learning as an input variable for competitive decision processes and is updated

by reinforcement processes that are distinct from pure motor learning (Sutton and Barto 1998). The decision process compares the values of the available actions and selects the action that will result in the highest value. Neurons coding action values can serve as suitable inputs for neuronal decision mechanisms if each action is associated with a distinct pool of action value neurons. Thus, action value needs to be coded for each action by separate neurons irrespective of the action being chosen, a crucial characteristic of action value.

Action values are subjective and can be derived from computational models fitted to behavioral choices (Samejima et al. 2005; Lau and Glimcher 2008; Ito and Doya 2009; Seo et al. 2012) or from logistic regressions on the animal's choice frequencies (Kim et al. 2009). Subgroups of neurons code action values in monkey and rat striatum. Their activities reflect the values obtained by specific (left vs. right) arm or eye movements (Fig. 4c) (Samejima et al. 2005; Lau and Glimcher 2008; Ito and Doya 2009; Kim et al. 2009; Seo et al. 2012) and occur irrespective of the animal's choice, thus following the strict definition of action value from machine learning (see Sutton and Barto

1998). Action value neurons are more frequent in monkey striatum than dorsolateral prefrontal cortex (Seo et al. 2012). Thus, the theoretical concept of action value serving as input for competitive decision mechanisms has a biological correlate that is consistent with the important movement function of the striatum.

In a basic economic decision model involving the striatum (Schultz 2015), reinforcement processes would primarily affect neuronal action value signals, as compared to other decision signals. The dopamine prediction error signal from the experienced primary or higher order reward may conform to the formalism of chosen value (Morris et al. 2006; Sugam et al. 2012) and might serve to update synaptic weights on striatal neurons coding action value (Schultz 1998). In a three-factor Hebbian arrangement, the dopamine signal would primarily affect synapses that had been used in the behavior that lead to the reward. As rewards are efficient when occurring after the behavior, rather than before it, these synapses must have been marked by an eligibility trace (Sutton and Barto 1981). The neuronal reinforcement signal would affect only synapses carrying eligibility traces. By contrast, inactive synapses would not carry an eligibility trace and thus remain unchanged or even undergo mild spontaneous decrement. In this way, dopamine activations after a positive prediction error would enhance the synaptic efficacy of active cortical inputs to striatal or other cortical neurons, whereas neuronal depressions after a negative prediction error would reduce the synaptic efficacy. Although this model is simplistic, the reward functions of the basal ganglia could be instrumental in updating economic values in decision processes.

Reward learning

All groups of striatal neurons acquire discriminant responses to visual, auditory and olfactory stimuli predicting liquid or food rewards (Aosaki et al. 1994; Jog et al. 1999; Tremblay et al. 1998; Adler et al. 2013). Neuronal responses to trial outcomes increase and then decrease again during the course of learning, closely following the changes in reward associations measured by the learning rate (Williams and Eskandar 2006). Other striatal neurons respond during initial learning indiscriminately to all novel stimuli and differentiate between rewarded and unrewarded stimuli as learning advances (Tremblay et al. 1998). Striatal neurons with sustained activations preceding reward delivery show expectation-related activations in advance of all outcomes and become selective for reward as the animal learns to distinguish rewarded from unrewarded trials (Fig. 4d) (Tremblay et al. 1998). These changes seem to reflect the adaptation of reward expectation to currently valid predictors. With reversals of reward predictions, striatal neurons switch differential responses rapidly

when a previously rewarded stimulus becomes unrewarded, in close correspondence to behavioral choices (Pasupathy and Miller 2005). Some striatal responses reflect correct or incorrect performance of previous trials (Histed et al. 2009). Neuronal responses in some striatal neurons code reward value using inference from paired associates and exclusion of alternative stimuli, thus reflecting acquired rules (Pan et al. 2014). Taken together, striatal neurons are involved in the formation of reward associations and the adaptation of reward predictions.

Social rewards

Observing reward in others allows social partners to appreciate the outcomes of social interactions, compare other's reward with own reward, and engage in mutually beneficial behavior, such as coordination and cooperation. The basic requirement for these processes involves the distinction between own rewards and the rewards of others. To attribute the reception of a reward to a specific individual requires identifying the agent whose action led to the reward. Once this issue is solved, one can advance to investigating the inequality in reward that different individuals receive for the same labor, and the inequality in labor leading to the same reward between individuals, which has wide ranging social consequences.

In a social reward experiment, two monkeys sit on opposite sides of a horizontally mounted touch-sensitive computer monitor and are presented with visual stimuli indicating who receives reward and who needs to act to produce that reward (Fig. 5a). Phasically active neurons in monkey striatum process mostly own rewards, irrespective of the other animal receiving reward or not (Fig. 5b, red and green) (Báez-Mendoza et al. 2013). Very few striatal neurons signal a reward that is delivered only to another monkey. Thus, striatal reward neurons seem to be primarily interested in own reward. Neurons that signal a reward to another monkey are found more frequently in anterior cingulate cortex (Chang et al. 2013). Striatal social reward neurons show an additional crucial feature. Most of their reward processing depends on the animal whose action produces the reward. It makes a difference when I receive reward because of my own action or because of the action of another individual. Striatal neurons make exactly this distinction. Many of them code own reward only when the animal receiving the reward acts (Fig. 5c), whereas other striatal neurons conversely code own reward only when the conspecific acts (Fig. 5d), thus dissociating actor from reward recipient. These contrasting activities are not due to differences in effort and often disappear when a computer actor replaces the conspecific, thus suggesting a social origin. Taken together, some striatal neurons process reward in a meaningful way during simple social

Fig. 5 Social reward signals in monkey striatum. **a** Behavioral task. Two monkeys sit opposite each other across a horizontally mounted touch-sensitive computer monitor. The acting animal moves from a resting key towards a touch table to give reward to itself, the other animal, both, or none, depending on a specific cue on the monitor (not shown). **b** Neuronal activation when receiving own reward, either only for the actor (*red*) or for both animals (*green*), but no activation with reward only for the other (*violet*) or nobody (*blue*). **c** Activation to the cue predicting own reward (for actor only and for both, *red* and *green*), conditional on own action, and not occurring with conspecific's action. **d** Activation following target touch predicting own reward, conditional on conspecific's action (*dotted lines*), and not occurring with own action. **b–d** *Imp/s* impulses per second, from Báez-Mendoza et al. (2013)

interactions. Such neuronal processes may constitute building blocks of social behavior involving the outcomes of specific actions.

Conclusions

The motor functions of the basal ganglia known from clinical conditions extend into a more global role in actions that are performed to attain a rewarding goal or occur on a more automatic, habitual basis. The link from movement to reward is represented in neuronal signals in the basal ganglia and the PPN. Of these structures, the current review describes the activity in dopamine neurons, PPN and striatum. Dopamine neurons represent movement and reward with two entirely different time scales and mechanisms. The movement function of dopamine is restricted to ambient levels that are necessary for movements to occur, whereas phasic dopamine changes with movements are not consistently observed. By contrast, the reward function is represented in the phasic dopamine reward utility prediction error signal. PPN neurons show a large variety of phasic movement relationships that are often modulated by reward, or display reward-related activities irrespective of

movements. Thus, PPN neurons show closer neuronal associations between motor and reward processing than dopamine neurons. The striatum shows even more closely integrated motor and reward processing; its neurons process reward conjointly with movement and code very specific, action-related variables suitable for maximizing reward in economic decisions. Although this general framework is likely to be revised in the coming years, the reviewed data suggest non-motor, reward processes as inherent features of basal ganglia function.

References

Adamantidis AR, Tsai H-C, Boutrel B, Zhang F, Stuber GD, Budygin EA, Touriño C, Bonci A, Deisseroth K, de Lecea L (2011) Optogenetic interrogation of dopaminergic modulation of the

multiple phases of reward-seeking behavior. J Neurosci 31:10829–10835

Adler A, Katabi S, Finkes I, Prut Y, Bergman H (2013) Different correlation patterns of cholinergic and GABAergic interneurons with striatal projection neurons. Front Syst Neurosci 7:47

Ambroggi F, Ishikawa A, Fields HL, Nicola SM (2008) Basolateral amygdala neurons facilitate reward-seeking behavior by exciting nucleus accumbens neurons. Neuron 59:648–661

Aosaki T, Tsubokawa H, Ishida A, Watanabe K, Graybiel AM, Kimura M (1994) Responses of tonically active neurons in the primate's striatum undergo systematic changes during behavioral sensorimotor conditioning. J Neurosci 14:3969–3984

Apicella P, Ljungberg T, Scarnati E, Schultz W (1991) Responses to reward in monkey dorsal and ventral striatum. Exp Brain Res 85:491–500

Apicella P, Scarnati E, Ljungberg T, Schultz W (1992) Neuronal activity in monkey striatum related to the expectation of predictable environmental events. J Neurophysiol 68:945–960

Apicella P, Deffains M, Ravel S, Legallet E (2009) Tonically active neurons in the striatum differentiate between delivery and omission of expected reward in a probabilistic task context. Eur J Neurosci 30:515–526

Apicella P, Ravel S, Deffains M, Legallet E (2011) The role of striatal tonically active neurons in reward prediction error signaling during instrumental task performance. J Neurosci 31:1507–1515

Asaad WF, Eskandar EN (2011) Encoding of both positive and negative reward prediction errors by neurons of the primate lateral prefrontal cortex and caudate nucleus. J Neurosci 31:17772–17787

Báez-Mendoza R, Harris C, Schultz W (2013) Activity of striatal neurons reflects social action and own reward. Proc Natl Acad Sci USA 110:16634–16639

Bayer HM, Glimcher PW (2005) Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron 47:129–141

Bowman EM, Aigner TG, Richmond BJ (1996) Neural signals in the monkey ventral striatum related to motivation for juice and cocaine rewards. J Neurophysiol 75:1061–1073

Bredfeldt CE, Ringach DL (2002) Dynamics of spatial frequency tuning in macaque V1. J Neurosci 22:1976–1984

Brischoux F, Chakraborty S, Brierley DI, Ungless MA (2009) Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. Proc Natl Acad Sci USA 106:4894–4899

Bromberg-Martin ES, Matsumoto M, Hon S, Hikosaka O (2010) A pallidus-habenula-dopamine pathway signals inferred stimulus values. J Neurophysiol 104:1068–1076

Caraco T, Martindale S, Whitham TS (1980) An empirical demonstration of risk-sensitive foraging preferences. Anim Behav 28:820–830

Chang SWC, Gariépy J-F, Platt ML (2013) Neuronal reference frames for social decisions in primate frontal cortex. Nat Neurosci 16:243–250

Chelazzi L, Estocinova J, Calletti R, Gerfo EL, Sani I, Della Libera C, Santandrea E (2014) Altering spatial priority maps via reward-based learning. J Neurosci 34:8594–8604

Cheng JJ, de Bruin JPC, Feenstra MGP (2003) Dopamine efflux in nucleus accumbens shell and core in response to appetitive classical conditioning. Eur J Neurosci 18:1306–1314

Cohen JY, Haesler S, Vong L, Lowell BB, Uchida N (2012) Neuron-type-specific signals for reward and punishment in the ventral tegmental area. Nature 482:85–88

Corbett D, Wise RA (1980) Intracranial self-stimulation in relation to the ascending dopaminergic systems of the midbrain: a moveable microelectrode study. Brain Res 185:1–15

Cromwell HC, Schultz W (2003) Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. J Neurophysiol 89:2823–2838

Cromwell HC, Hassani OK, Schultz W (2005) Relative reward processing in primate striatum. Exp Brain Res 162:520–525

Day JJ, Roitman MF, Wightman RM, Carelli RM (2007) Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. Nat Neurosci 10:1020–1028

Day JJ, Jones JL, Wightman RM, Carelli RM (2010) Phasic nucleus accumbens dopamine release encodes effort- and delay-related costs. Biol Psychiat 68:306–309

Day JJ, Jones JL, Carelli RM (2011) Nucleus accumbens neurons encode predicted and ongoing reward costs in rats. Eur J Neurosci 33:308–321

de Lafuente O, Romo R (2011) Dopamine neurons code subjective sensory experience and uncertainty of perceptual decisions. Proc Natl Acad Sci 49:19767–19771

DeLong MR, Crutcher MD, Georgopoulos AP (1983) Relations between movement and single cell discharge in the substantia nigra of the behaving monkey. J Neurosci 3:1599–1606

di Loreto S, Florio T, Scarnati E (1992) Evidence that a non-NMDA receptor is involved in the excitatory pathway from the pedunculopontine region to nigrostriatal dopaminergic neurons. Exp Brain Res 89:79–86

Dickinson A, Balleine B (1994) Motivational control of goal-directed action. Anim Learn Behav 22:1–18

Ding L, Gold JI (2010) Caudate encodes multiple computations for perceptual decisions. J Neurosci 30:15747–15759

Dormont JF, Conde H, Farin D (1998) The role of the pedunculopontine tegmental nucleus in relation to conditioned motor performance in the cat. I. Context-dependent and reinforcement-related single unit activity. Exp Brain Res 121:401–410

Enomoto K, Matsumoto N, Nakai S, Satoh T, Sato TK, Ueda Y, Inokawa H, Haruno M, Kimura M (2011) Dopamine neurons learn to encode the long-term value of multiple future rewards. Proc Natl Acad Sci USA 108:15462–15467

Fibiger HC, LePiane FG, Jakubovic A, Phillips AG (1987) The role of dopamine in intracranial self-stimulation of the ventral tegmental area. J Neurosci 7:3888–3896

Fiorillo CD (2013) Two dimensions of value: dopamine neurons represent reward but not aversiveness. Science 341:546–549

Fiorillo CD, Tobler PN, Schultz W (2003) Discrete coding of reward probability and uncertainty by dopamine neurons. Science 299:1898–1902

Fiorillo CD, Newsome WT, Schultz W (2008) The temporal precision of reward prediction in dopamine neurons. Nat Neurosci 11:966–973

Fiorillo CD, Yun SR, Song MR (2013a) Diversity and homogeneity in responses of midbrain dopamine neurons. J Neurosci 33:4693–4709

Fiorillo CD, Song MR, Yun SR (2013b) Multiphasic temporal dynamics in responses of midbrain dopamine neurons to appetitive and aversive stimuli. J Neurosci 33:4710–4725

Freund TF, Powell JF, Smith AD (1984) Tyrosine hydroxylase-immunoreactive boutons in synaptic contact with identified striatonigral neurons, with particular reference to dendritic spines. Neuroscience 13:1189–1215

Futami T, Takakusaki K, Kitai ST (1995) Glutamatergic and cholinergic inputs from the pedunculopontine tegmental nucleus to dopamine neurons in the substantia nigra pars compacta. Neurosci Res 21:331–342

Gdowski MJ, Miller LE, Parrish T, Nenonene EK, Houk JC (2001) Context dependency in the globus pallidus internal segment during targeted arm movements. J Neurophysiol 85:998–1004

Goldman-Rakic PS, Leranth C, Williams MS, Mons N, Geffard M (1989) Dopamine synaptic complex with pyramidal neurons in primate cerebral cortex. Proc Natl Acad Sci USA 86:9015–9019

Gottlieb JP, Kusunoki M, Goldberg ME (1998) The representation of visual salience inmonkey parietal cortex. Nature 391:481–484

Grace AA (1991) Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: a hypothesis for the etiology of schizophrenia. Neuroscience 41:1–24

Grace AA, Floresco SB, Goto Y, Lodge DJ (2007) Regulation of firing of dopaminergic neurons and control of goal-directed behaviors. Trends Neurosci 30:220–227

Guarraci FA, Kapp BS (1999) An electrophysiological characterization of ventral tegmental area dopaminergic neurons during differential pavlovian fear conditioning in the awake rabbit. Behav Brain Res 99:169–179

Hamid AA, Pettibone JR, Mabrouk OS, Hetrick VL, Schmidt R, Vander Weele CM, Kennedy RT, Aragona BJ, Berke JD (2015) Mesolimbic dopamine signals the value of work. Nat Neurosci 19:117–126

Hassani OK, Cromwell HC, Schultz W (2001) Influence of expectation of different rewards on behavior-related neuronal activity in the striatum. J Neurophysiol 85:2477–2489

Hikosaka O, Sakamoto M, Usui S (1989) Functional properties of monkey caudate neurons. III. Activities related to expectation of target and reward. J Neurophysiol 61:814–832

Histed MH, Pasupathy A, Miller EK (2009) Learning substrates in the primate prefrontal cortex and striatum: sustained activity related to successful actions. Neuron 63:244–253

Hollerman JR, Schultz W (1998) Dopamine neurons report an error in the temporal prediction of reward during learning. Nat Neurosci 1:304–309

Hollerman JR, Tremblay L, Schultz W (1998) Influence of reward expectation on behavior-related neuronal activity in primate striatum. J Neurophysiol 80:947–963

Hong S, Hikosaka O (2014) Pedunculopontine tegmental nucleus neurons provide reward, sensorimotor, and alerting signals to midbrain dopamine neurons. Neuroscience 282:139–155

Horvitz JC, Stewart T, Jacobs BL (1997) Burst activity of ventral tegmental dopamine neurons is elicited by sensory stimuli in the awake cat. Brain Res 759:251–258

Howe MW, Tierney PL, Sandberg SG, Phillips PEM, Graybiel AM (2013) Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. Nature 500:575–579

Ito M, Doya K (2009) Validation of decision-making models and analysis of decision variables in the rat basal ganglia. J Neurosci 29:9861–9874

Jog MS, Kubota Y, Connolly CI, Hillegaart V, Graybiel AM (1999) Building neural representations of habits. Science 286:1745–1749

Joshua M, Adler A, Mitelman R, Vaadia E, Bergman H (2008) Midbrain dopaminergic neurons and striatal cholinergic interneurons encode the difference between reward and aversive events at different epochs of probabilistic classical conditioning trials. J Neurosci 28:11673–11684

Kagel JH, Battalio RC, Green L (1995) Economic choice theory: an experimental analysis of animal behavior. Cambridge University Press, Cambridge

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

Kelly PH, Seviour PW, Iversen SD (1975) Amphetamine and apomorphine responses in the rat following 6-OHDA lesions of the nucleus accumbens septi and corpus striatum. Brain Res 94:507–522

Kim JN, Shadlen MN (1999) Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. Nat Neurosci 2:176–185

Kim H, Sul JH, Huh N, Lee D, Jung MW (2009) Role of striatum in updating values of chosen actions. J Neurosci 29:14701–14712

Kim KM, Baratta MV, Yang A, Lee D, Boyden ES, Fiorillo CD (2012) Optogenetic mimicry of the transient activation of

dopamine neurons by natural reward is sufficient for operant reinforcement. PLoS One 7:e33612

Kimura M, Rajkowski J, Evarts E (1984) Tonically discharging putamen neurons exhibit set-dependent responses. Proc Natl Acad Sci USA 81:4998–5001

Kobayashi Y, Okada K-I (2007) Reward prediction error computation in the pedunculopontine tegmental nucleus neurons. Ann NY Acad Sci 1104:310–323

Kobayashi S, Schultz W (2008) Influence of reward delays on responses of dopamine neurons. J Neurosci 28:7837–7846

Kobayashi S, Schultz W (2014) Reward contexts extend dopamine signals to unrewarded stimuli. Curr Biol 24:56–62

Kobayashi Y, Inoue Y, Yamamoto M, Isa T, Aizawa H (2002) Contribution of pedunculopontine tegmental nucleus neurons to performance of visually guided saccade tasks in monkeys. J Neurophysiol 88:715–731

Lak A, Arabzadeh E, Harris JA, Diamond ME (2010) Correlated physiological and perceptual effects of noise in a tactile stimulus. Proc Natl Acad Sci USA 107:7981–7986

Lak A, Stauffer WR, Schultz W (2014) Dopamine prediction error responses integrate subjective value from different reward dimensions. Proc Natl Acad Sci USA 111:2343–2348

Lardeux S, Pernaud R, Paleressompoulle D, Baunez C (2009) Beyond the reward pathway: coding reward magnitude and error in the rat subthalamic nucleus. J Neurophysiol 102:2526–2537

Lau B, Glimcher PW (2008) Value representations in the primate striatum during matching behavior. Neuron 58:451–463

Lauwereyns J, Takikawa Y, Kawagoe R, Kobayashi S, Koizumi M, Coe B, Sakagami M, Hikosaka O (2002) Feature-based anticipation of cues that predict reward in monkey caudate nucleus. Neuron 33:463–473

Ljungberg T, Apicella P, Schultz W (1991) Responses of monkey midbrain dopamine neurons during delayed alternation performance. Brain Res 586:337–341

Ljungberg T, Apicella P, Schultz W (1992) Responses of monkey dopamine neurons during learning of behavioral reactions. J Neurophysiol 67:145–163

Lokwan SJA, Overton PG, Berry MS, Clark S (1999) Stimulation of the pedunculopontine nucleus in the rat produces burst firing in A9 dopaminergic neurons. Neuroscience 92:245–254

Luce RD (1959) Individual choice behavior: a theoretical analysis. Wiley, New York

Machina MJ (1987) Choice under uncertainty: problems solved and unsolved. J Econ Perspect 1:121–154

Matsumoto M, Hikosaka O (2009) Two types of dopamine neuron distinctively convey positive and negative motivational signals. Nature 459:837–841

Matsumoto M, Takada M (2013) Distinct representations of cognitive and motivational signals in midbrain dopamine neurons. Neuron 79:1011–1024

Mena-Segovia J, Winn P, Bolam JP (2008) Cholinergic modulation of midbrain dopaminergic systems. Brtain Res Rev 58:265–271

Mirenowicz J, Schultz W (1994) Importance of unpredictability for reward responses in primate dopamine neurons. J Neurophysiol 72:1024–1027

Mirenowicz J, Schultz W (1996) Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. Nature 379:449–451

Mogami T, Tanaka K (2006) Reward association affects neuronal responses to visual stimuli in macaque TE and perirhinal cortices. J Neurosci 26:6761–6770

Montague PR, Dayan P, Sejnowski TJ (1996) A framework for mesencephalic dopamine systems based on predictive Hebbian learning. J Neurosci 16:1936–1947

Morris G, Arkadir D, Nevet A, Vaadia E, Bergman H (2004) Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. Neuron 43:133–143

Morris G, Nevet A, Arkadir D, Vaadia E, Bergman H (2006) Midbrain dopamine neurons encode decisions for future action. Nat Neurosci 9:1057–1063

Nakahara H, Itoh H, Kawagoe R, Takikawa Y, Hikosaka O (2004) Dopamine neurons can represent context-dependent prediction error. Neuron 41:269–280

Nomoto K, Schultz W, Watanabe T, Sakagami M (2010) Temporally extended dopamine responses to perceptually demanding reward-predictive stimuli. J Neurosci 30:10692–10702

Norton ABW, Jo YS, Clark EW, Taylor CA, Mizumori SJY (2011) Independent neural coding of reward and movement by pedunculopontine tegmental nucleus neurons in freely navigating rats. Eur J Neurosci 33:1885–1896

Okada K-I, Toyama K, Inoue Y, Isa T, Kobayashi Y (2009) Different pedunculopontine tegmental neurons signal predicted and actual task rewards. J Neurosci 29:4858–4870

O'Neill M, Schultz W (2010) Coding of reward risk distinct from reward value by orbitofrontal neurons. Neuron 68:789–800

Oyama K, Hernádi I, Iijima T, Tsutsui K-I (2010) Reward prediction error coding in dorsal striatal neurons. J Neurosci 30:11447–11457

Pan W-X, Hyland BI (2005) Pedunculopontine tegmental nucleus controls conditioned responses of midbrain dopamine neurons in behaving rats. J Neurosci 25:4725–4732

Pan W-X, Schmidt R, Wickens JR, Hyland BI (2005) Dopamine cells respond to predicted events during classical conditioning: evidence for eligibility traces in the reward-learning network. J Neurosci 25:6235–6242

Pan X, Fan H, Sawa K, Tsuda I, Tsukada M, Sakagami M (2014) Reward inference by primate prefrontal and striatal neurons. J Neurosci 34:1380–1396

Pasquereau B, Nadjar A, Arkadir D, Bezard E, Goillandeau M, Bioulac B, Gross CE, Boraud T (2007) Shaping of motor responses by incentive values through the basal ganglia. J Neurosci 27:1176–1183

Pasupathy A, Miller EK (2005) Different time courses of learning-related activity in the prefrontal cortex and striaum. Nature 433:873–876

Paton JJ, Belova MA, Morrison SE, Salzman CD (2006) The primate amygdala represents the positive and negative value of visual stimuli during learning. Nature 439:865–870

Pearce JM, Hall G (1980) A model for Pavlovian conditioning: variations in the effectiveness of conditioned but not of unconditioned stimuli. Psychol Rev 87:532–552

Peck CJ, Lau B, Salzman CD (2013) The primate amygdala combines information about space and value. Nat Neurosci 16:340–348

Pooresmaeili A, Poort J, Roelfsema PR (2014) Simultaneous selection by object-based attention in visual and frontal cortex. Proc Natl Acad Sci USA 111:6467–6472

Ravel S, Legallet E, Apicella P (1999) Tonically active neurons in the monkey striatum do not preferentially respond to appetitive stimuli. Exp Brain Res 128:531–534

Redgrave P, Prescott TJ, Gurney K (1999) Is the short-latency dopamine response too short to signal reward? Trends Neurosci 22:146–151

Rescorla RA, Wagner AR (1972) A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: Black AH, Prokasy WF (eds) Classical conditioning II: current research and theory. Appleton Century Crofts, New York, pp 64–99

Ringach DL, Hawken MJ, Shapley R (1997) Dynamics of orientation tuning in macaque primary visual cortex. Nature 387:281–284

Robbins TW, Arnsten AFT (2009) The neuropsychopharmacology of fronto-executive function: monoaminergic modulation. Ann Rev Neurosci 32:267–287

Roelfsema PR, Tolboom M, Khayat PS (2007) Different processing phases for features, figures, and selective attention in the primary visual cortex. Neuron 56:785–792

Roesch MR, Singh T, Brown PL, Mullins SE, Schoenbaum G (2009) Ventral striatal neurons encode the value of the chosen action in rats deciding between differently delayed or sized rewards. J Neurosci 29:13365–13376

Roitman JD, Shadlen MN (2002) Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. J Neurosci 22:9475–9489

Romo R, Schultz W (1990) Dopamine neurons of the monkey midbrain: contingencies of responses to active touch during self-initiated arm movements. J Neurophysiol 63:592–606

Rothschild M, Stiglitz JE (1970) Increasing risk: I. A definition. J Econ Theory 2:225–243

Samejima K, Ueda Y, Doya K, Kimura M (2005) Representation of action-specific reward values in the striatum. Science 310:1337–1340

Satoh T, Nakai S, Sato T, Kimura M (2003) Correlated coding of motivation and outcome of decision by dopamine neurons. J Neurosci 23:9913–9923

Scarnati E, Campana E, Pacitti C (1984) Pedunculopontine-evoked excitation of substantia nigra neurons in the rat. Brain Res 304:351–361

Scarnati E, Proia A, Campana E, Pacitti C (1986) A microiontophoretic study on the nature of the putative synaptic neurotransmitter involved in the pedunculopontine-substantia nigra pars compacta excitatory pathway of the rat. Exp Brain Res 62:470–478

Schultz W (1986) Responses of midbrain dopamine neurons to behavioral trigger stimuli in the monkey. J Neurophysiol 56:1439–1462

Schultz W (1998) Predictive reward signal of dopamine neurons. J Neurophysiol 80:1–27

Schultz W (2007) Multiple dopamine functions at different time courses. Ann Rev Neurosci 30:259–288

Schultz W (2015) Neuronal reward and decision signals: from theories to data. Physiol Rev 95:853–951

Schultz W, Romo R (1990) Dopamine neurons of the monkey midbrain: contingencies of responses to stimuli eliciting immediate behavioral reactions. J Neurophysiol 63:607–624

Schultz W, Ruffieux A, Aebischer P (1983) The activity of pars compacta neurons of the monkey substantia nigra in relation to motor activation. Exp Brain Res 51:377–387

Schultz W, Apicella P, Scarnati E, Ljungberg T (1992) Neuronal activity in monkey ventral striatum related to the expectation of reward. J Neurosci 12:4595–4610

Schultz W, Apicella P, Ljungberg T (1993) Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. J Neurosci 13:900–913

Schultz W, Dayan P, Montague RR (1997) A neural substrate of prediction and reward. Science 275:1593–1599

Seo M, Lee E, Averbeck BB (2012) Action selection and action value in frontal-striatal circuits. Neuron 74:947–960

Shadlen MN, Newsome WT (2001) Neural basis of a perceptual decision in the parietal cortex (Area LIP) of the rhesus monkey. J Neurophysiol 86:1916–1936

Shidara M, Aigner TG, Richmond BJ (1998) Neuronal signals in the monkey ventral striatum related to progress through a predictable series of trials. J Neurosci 18:2613–2625

Stalnaker TA, Calhoon GG, Ogawa M, Roesch MR, Schoenbaum G (2012) Reward prediction error signaling in posterior

dorsomedial striatum is action specific. J Neurosci 32:10296–10305

Stanisor L, van der Togt C, Cyriel MA, Pennartz CMA, Roelfsema PR (2013) A unified selection signal for attention and reward in primary visual cortex. Proc Natl Acad Sci USA 110:9136–9141

Stauffer WR, Lak A, Schultz W (2014) Dopamine reward prediction error responses reflect marginal utility. Curr Biol 24:2491–2500

Steinberg EE, Keiflin R, Boivin JR, Witten IB, Deisseroth K, Janak PH (2013) A causal link between prediction errors, dopamine neurons and learning. Nat Neurosci 16:966–973

Steinfels GF, Heym J, Strecker RE, Jacobs BL (1983) Behavioral correlates of dopaminergic unit activity in freely moving cats. Brain Res 258:217–228

Sugam JA, Day JJ, Wightman RM, Carelli RM (2012) Phasic nucleus accumbens dopamine encodes risk-based decision-making behavior. Biol Psychiat 71:199–2015

Sutton RS, Barto AG (1981) Toward a modern theory of adaptive networks: expectation and prediction. Psychol Rev 88:135–170

Sutton RS, Barto AG (1998) Reinforcement learning. MIT Press, Cambridge

Tachibana Y, Hikosaka O (2012) The primate ventral pallidum encodes expected reward value and regulates motor action. Neuron 76:826–837

Tai L-H, Lee AM, Benavidez N, Bonci A, Wilbrecht L (2012) Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. Nat Neurosci 15:1281–1289

Takikawa Y, Kawagoe R, Hikosaka O (2004) A possible role of midbrain dopamine neurons in short- and long-term adaptation of saccades to position-reward maping. J Neurophysiol 92:2520–2529

Thompson KG, Hanes DP, Bichot NP, Schall JD (1996) Perceptual and motor processing stages identified in the activity of macaque frontal eye field neurons during visual search. J Neurophysiol 76:4040–4055

Thompson KG, Bichot NP, Sato TR (2005) Frontal eye field activity before visual search errors reveals the integration of bottom-up and top-down salience. J Neurophysiol 93:337–351

Tobler PN, Dickinson A, Schultz W (2003) Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. J Neurosci 23:10402–10410

Tobler PN, Fiorillo CD, Schultz W (2005) Adaptive coding of reward value by dopamine neurons. Science 307:1642–1645

Tremblay L, Hollerman JR, Schultz W (1998) Modifications of reward expectation-related neuronal activity during learning in primate striatum. J Neurophysiol 80:964–977

Tsai H-C, Zhang F, Adamantidis A, Stuber GD, Bonci A, de Lecea L, Deisseroth K (2009) Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. Science 324:1080–1084

von Neumann J, Morgenstern O (1944) The theory of games and economic behavior. Princeton University Press, Princeton

Waelti P, Dickinson A, Schultz W (2001) Dopamine responses comply with basic assumptions of formal learning theory. Nature 412:43–48

Watabe-Uchida M, Zhu L, Ogawa SK, Vamanrao A, Uchida N (2012) Whole-brain mapping of direct inputs to midbrain dopamine neurons. Neuron 74:858–873

Williams ZM, Eskandar EN (2006) Selective enhancement of associative learning by microstimulation of the anterior caudate. Nat Neurosci 4:562–568

Yasuda M, Yamamoto S, Hikosaka O (2012) Robust representation of stable object values in the oculomotor basal ganglia. J Neurosci 32:16917–16932

Yin HH, Knowlton BJ, Balleine BW (2004) Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. Eur J Neurosci 19:181–189

Yin HH, Ostlund SB, Knowlton BJ, Balleine BB (2005) The role of the dorsomedial striatum in instrumental conditioning. Eur J Neurosci 23:513–523

Young AMJ (2004) Increased extracellular dopamine in nucleus accumbens in response to unconditioned and conditioned aversive stimuli: studies using 1 min microdialysis in rats. J Neurosci Meth 138:57–63

Young AMJ, Joseph MH, Gray JA (1992) Increased dopamine release in vivo in nucleus accumbens and caudate nucleus of the rat during drinking: a microdialysis study. Neuroscience 48:871–876