Volume 23, issue 2   April 2013
ISSN 0959-4388

Current Opinion in
Neurobiology

Cori Bargmann & Edvard Moser, Editors

April 2013
Macrocircuits
Edited by Steve Petersen and Wolf Singer

June 2013 Social and emotional neuroscience
August 2013 Addiction
October 2013 Circadian rhythm and sleep
December 2013 Development of neurons and glia
February 2014 Neurogenetics

Available online at www.sciencedirect.com
SciVerse ScienceDirect

Access *CONB* articles online up to one month before
they appear in your print journal www.sciencedirect.com

CURRENT OPINION
www.current-opinion.com

Available online at www.sciencedirect.com

SciVerse ScienceDirect

Current Opinion in
Neurobiology

Updating dopamine reward signals
Wolfram Schultz

Recent work has advanced our knowledge of phasic dopamine reward prediction error signals. The error signal is bidirectional, reflects well the higher order prediction error described by temporal difference learning models, is compatible with model-free and model-based reinforcement learning, reports the subjective rather than physical reward value during temporal discounting and reflects subjective stimulus perception rather than physical stimulus aspects. Dopamine activations are primarily driven by reward, and to some extent risk, whereas punishment and salience have only limited activating effects when appropriate controls are respected. The signal is homogeneous in terms of time course but heterogeneous in many other aspects. It is essential for synaptic plasticity and a range of behavioural learning situations.

Address
Department of Physiology, Development and Neuroscience, University of Cambridge, Downing Street, Cambridge CB2 3DY, UK

Corresponding author: Schultz, Wolfram (ws234@cam.ac.uk)

Introduction
The smoke signals of the American Indians and the peep-peep from Sputnik demonstrate that information processing systems use signals. The brain is no exception. Its action potentials are instrumental for extracting information from the environment and directing behaviour. The foremost function of the brain is to assure individual survival and gene propagation for ultimate evolutionary fitness. To this end individuals need to acquire specific substances for their bodily functions. The substances come in foods and liquids and require effort to obtain. They are called rewards and support learning, approach behaviour, decision making and positive emotions like pleasure, desire and happiness. Numerous other objects, events and situations have similar reward functions and contribute to evolutionary fitness. To understand reward processing, we need to study neuronal signals for these objects. We can describe reward functions in formal behavioural terms like Pavlovian and operant conditioning, habits and goal directed behaviour, economic

decision making and see how alterations in brain functions influence these processes. But to understand how the brain as an information processing machine mediates reward we need to take these terms apart and identify and understand neuronal signals for their components. This review provides an update on dopamine neurons that provide a reward signal for updating economic decision variables.

Nature of dopamine signal
Most midbrain dopamine neurons (75–80%) show rather stereotyped, phasic activations with latencies of <100 ms and durations of <200 ms following unpredicted food or liquid rewards. This response codes a quantitative prediction error, namely the difference between received and predicted reward value. A reward that is better than predicted elicits an activation (positive prediction error response), a fully predicted reward draws no response, and a reward that is worse than predicted induces a depression (negative error response). Most dopamine neurons (60–75%) respond in similar ways to reward predicting stimuli, coding positive and negative, higher order reward prediction errors at the time of the stimulus relative to the prediction at that moment [1,2,3•,4••].

Power of negative prediction error signal
The positive prediction error response can amount to phasic tripling of impulse rate, whereas the negative response has limited dynamic range owing to the low baseline activity of dopamine neurons (3–5 impulses/s), which might suggest limited negative coding [5,6]. However, synaptic transmission is unlikely to operate on an absolute scale. Completely shutting off a signal is more noticeable than tripling it. Abruptly extinguishing illumination in a dim auditorium has more impact than increasing it threefold. The observation that the pause affects almost all dopamine neurons makes it a powerful negative signal that would phasically stop stimulation of both low affinity D1 and high affinity D2 receptors. Variations of predictions together with proper measurement of negative response intensity, as expressed by duration of depression in extracellular recordings, reveal bidirectional graded reward prediction error coding in dopamine neurons [7]. Thus the negative dopamine error signal is graded and likely to have a major impact on postsynaptic neurons.

Temporal difference neuronal signal
The dopamine response to reward is compatible with the notion of primary reinforcer prediction error according to error driven learning rules [8]. These learning rules are

Figure 1
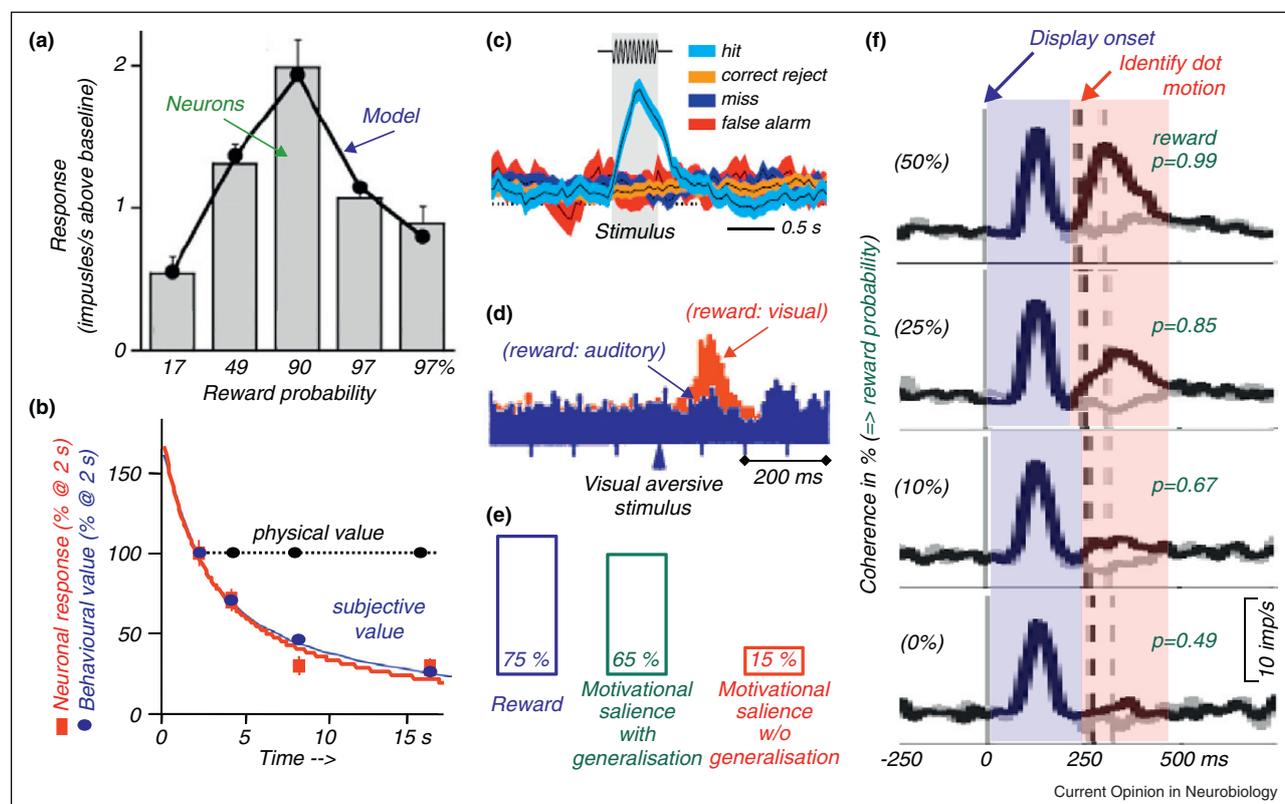
Characteristics of phasic dopamine reward prediction error responses. **(a)** Neuronal coding of reward prediction error closely parallels theoretical prediction error of temporal difference (TD) model ([4••], © National Academy of Sciences USA). **(b)** Temporal discounting of neuronal response to stimulus predicting differently delayed rewards closely parallels behavioural discounting ([15], © Society for Neuroscience). **(c)** Neuronal response depends on subjective stimulus perception ([24••], © National Academy of Sciences USA). **(d)** Stimulus generalisation explains majority of responses to conditioned aversive stimuli. Change in sensory modality of reward predicting stimulus reduces response to unchanged aversive stimulus ([34], © Nature). **(e)** Percentages of dopamine neurons activated by reward (blue, left), motivational salience uncontrolled for stimulus or context generalisation (green) and true motivational salience (red, right). Data from [34]. **(f)** Graded coding of value prediction after initial generalisation coincides with stimulus identification by animal in dot motion task. Percentage of coherently moving dots results in graded percentage of correct performance and reward delivery ([3•], © Society for Neuroscience).

extended by temporal difference (TD) reinforcement learning to higher order reinforcement [9]. Compatible with these notions, dopamine neurons show phasic responses to higher order prediction errors evoked by conditioned, reward predicting stimuli [1]. In a recent multistep task, monkey dopamine neurons coded the TD error quantitatively by reflecting the difference between the sum of multiple future rewards and their prediction [4••]. In the particular task used, the reward probabilities increase towards the end of the multistep sequence, resulting in the highest discounted sum of future reward in the centre from which the lower predictions arising from earlier stimuli are subtracted (Figure 1a). The dopamine responses match this temporal profile of TD error closely, demonstrating the most complete relationship of phasic dopamine responses to the TD model of reinforcement so far reported.

Model-free vs. model-based reinforcement learning

Reinforcement learning establishes reward predictions that are essential for informed decision making [8]. Model-free reinforcement learning occurs through the experienced contiguity between stimuli or actions and reinforcers. However, individuals also learn about states, situations, contexts, rules, sequences, and conditional probabilities, which may be summarily called models of the world. Knowledge derived from models can be useful for improving reward predictions. The processes can include frequentist and Bayesian inferences, logical reasoning, or any method that improves the identification of relevant states. Model-based reinforcement learning involves two separate processes, the acquisition and updating of the overall model, and the influence of the model on current predictions [10,11]. Although it is

currently unknown whether dopamine prediction error responses might be involved in the acquisition or updating of the models, they may incorporate the predictive information from models. As the experienced reward is physical and thus independent of the model, the model's influence on the prediction error would derive from the influence on the reward prediction.

In learning situations governed only by experienced rewards, consecutive unrewarded trials lead to progressively decreasing reward prediction. Correspondingly, positive prediction errors to an unpredicted reward increase, because the repeated absence of reward makes the reward less expected and the delivery of reward becomes more surprising. However, in particular tasks, sequences of unrewarded trials may lead to increasing reward probability (increasing hazard rate) and thus increasing reward prediction. Now the reward becomes progressively less surprising and hence induces a prediction error that decreases over successive trials. Learning in such a task benefits from knowledge about the underlying model of the world. Indeed, dopamine error responses follow precisely this scheme, decreasing with increasing numbers of trials since the last reward [12], suggesting that they process model-based predictions.

Dopamine responses are influenced by previously acquired information about reward probability distributions [13]. The error responses adapt within 2 s to the currently valid expected values and standard deviations despite ten fold differences in reward magnitude, without requiring resampling of the distributions. The adaptations probably reflect scaling of the prediction which is subtracted from the currently experienced reward value to obtain the prediction error.

Repeated reversals of stimulus-reward or position-reward associations lead to acquisition of a rule that helps individuals reverse their reactions within one trial (reversal set). Correspondingly, dopamine responses to reward predicting stimuli reverse after one trial without having experienced the contiguity between the newly valid stimulus position and the reward [14••]. These data suggest that dopamine neurons are influenced by predictions derived from hidden states and Bayesian inference for estimating the reversed reward probabilities.

In each of these examples, the observed dopamine prediction error responses are not fully explained by the actual experience with reward but incorporate information from models of the world. Thus the dopamine prediction error responses occur in both model-free and model-based reinforcement learning situations, without evidence so far for a role in initial model acquisition.

Subjective value coding in temporal discounting

Individuals make choices that maximise their rewards. Although rewards have objective, physical and chemical properties, their value is defined by the needs of the individual decision maker and thus intrinsically subjective. One way to distinguish subjective from objective reward value is temporal discounting, the decrease in subjective reward value with increasing delays despite constant physical size. Psychometric analysis of choices in monkeys with variable early vs. constant 4, 8 and 16 s delayed rewards reveals graded hyperbolic decays in subjective value by 25, 50 and 75%, respectively, compared to reward at 2 s [15]. Exponential decay fit significantly less well.

Dopamine responses to reward predicting stimuli decrease monotonically across reward delays of 2–16 s in monkeys, despite the same physical amount of reward being delivered after each delay [15,16] (Figure 1b). Responses to reception of the constant reward itself show the opposite relationship, namely increasing magnitudes after increasing delays, again being fit generally better by hyperbolic than exponential functions. The increasing response profile may be owing to temporal prediction errors derived from subjective temporal uncertainty that scales with delay or magnitude prediction errors between the full size received reward and the discounted predicted value. Reward magnitude sensitive dopamine neurons in rats show higher sustained activations for several seconds following odours predicting sucrose after a short, fixed delay of 0.5 s compared to long, variable delays of 1–7 s [17].

Reward neurons in most other primary and associated reward structures of the brain show some degrees of temporal discounting, including orbitofrontal cortex [18], prefrontal cortex [19,20], dorsal and ventral striatum [21,22], premotor cortex [19] and parietal cortex [23]. Thus reward neurons code subjective rather than objective reward value in temporal discounting and thus may provide direct inputs for neuronal mechanisms underlying value based decision processes.

Subjective stimulus coding

If dopamine neurons code reward value subjectively in temporal discounting, would they show other aspects of subjective coding compatible with the notion of an intrinsically subjective survival function of rewards? For instance, would their reward prediction coding reflect the physical presence or the subjective experience of predictive stimuli? Signal detection theory offers well proven tests. Monkeys correctly or incorrectly report the presence of stimuli (hit and false alarm, respectively), and correctly or incorrectly report stimulus absence (correct rejection and miss alarm, respectively) [24••]. Dopamine neurons are activated by an initial task stimulus only

when the animal reports its detection, whereas they are not activated by the same physical stimulus when the animal reports its absence, nor when the stimulus is physically absent irrespective of the animal's report (Figure 1c). The data suggest that both subjective perception and physical presence of a stimulus are necessary to evoke dopamine responses, whereas physical presence alone is insufficient.

The relationship to subjective perception is also seen with dopamine error responses to subsequent cues and reward. Correct and incorrect reports of stimulus presence lead to reward in 76% of trials, whereas correct and incorrect reports of stimulus absence lead to 64% reward [24••]. As dopamine neurons code reward probability as value [25], we calculate the TD prediction errors at a subsequent, noninformative cue as [constant prediction by noninformative cue minus value 76] when the animal reports stimulus presence, and [constant prediction by noninformative cue minus value 64] when reporting stimulus absence. Thus the prediction error is lower when the animal reports detection compared to reporting absence of the same physical cue. Indeed, the dopamine error responses to the cue follow exactly this pattern. At reward reception, the prediction errors and the resulting dopamine responses show exactly the same difference.

Together with the responses to the near threshold stimulus itself and the subjective value coding in temporal discounting, dopamine responses reflect the animal's subjective perception and valuation of the stimuli beyond purely physical reward properties. As decisions are ultimately made according to subjective reward values, the dopamine responses may provide rather direct and parsimonious, and therefore evolutionary beneficial and selected, inputs to neuronal decision processes.

Aversive responses

Electric shocks, painful pinches and airpuffs, as well as conditioned stimuli predicting these events, induce phasic depressions in dopamine impulse activity in anaesthetised monkeys and rodents [26–32,33•] and in awake monkeys, cats and rats [34–37,38•]. The dopaminergic nature of these neurons is verified by optogenetics [39••]. The depression correlates with the length of dopamine dendrites in pars reticulata of substantia nigra [40], suggesting a synaptic drive by GABAergic midbrain neurons which are excited by aversive stimuli and whose stimulation leads to depression in dopamine activity [33•,39••,41•].

A few dopamine neurons (5–15%) are activated by primary aversive stimuli [26,28,34], which is now confirmed [31,33•,36,37,38•,39••]. The overall impact of aversive stimuli on the population of dopamine neurons is difficult to assess in studies not reporting frequencies of responses relative to numbers of tested neurons or testing relatively few neurons (≤30) [32,42]. Midbrain neurons that respond to aversive events but are insensitive to the defining depressant effects of systemic apomorphine [43] are probably not dopaminergic [44,45]. Aversively activated neurons are also activated by rewards and other environmental stimuli that are not directly rewarding, and thus form a separate, limited group of highly sensitive, nonspecifically responding neurons. However, dopamine neurons that are activated by punishers do not seem to show bidirectional punisher prediction error responses [36,37]. The aversively activated dopamine neurons may constitute a distinct ventral tegmental group in rats [33•], although silent dopamine neurons [46] might pick up the juxtacellular staining without being activated.

Aversive stimuli increase striatal dopamine concentrations measured by *in vivo* dialysis over seconds [29,47]. The increase may reflect tonically elevated impulse activity, increased proportions of spontaneously active dopamine neurons [46] and impulse rebound after initial depressions [27,38•,45]. Faster voltammetry reveals striatal dopamine decreases following quinine [48] and increases in patches of striatum and nucleus accumbens core during tail pinch and accumbens shell after termination of tail pinch [49]. The increases probably reflect the limited initial and more frequent rebound dopamine activations by punishers. As termination of an aversive stimulus induces a rewarding rebound according to the opponent process theory of motivation [50], some of the dopamine increases may derive from rebound and reflect reward.

False aversive activations

Substantial fractions of dopamine neurons (35–65%) are activated by conditioned aversive stimuli presented in random or blockwise alternation with reward predicting stimuli [34,36,37], which exceeds the frequency of activations to unpredicted primary aversive stimuli (<15%). These results are paradoxical from a reinforcement perspective because they violate a basic tenet of animal learning theory that postulates stronger motivational and attentional effects of primary, unconditioned rewards or punishers compared to stimuli conditioned to them. Although conditioned aversive stimuli may induce different forms of aversion to which dopamine neurons might be selectively sensitive, such selectivity would be surprising given their common responsiveness to all kinds of rewards. Thus other mechanisms than punishment may play a role. Closer inspections of conditioned aversive activations reveal two response components.

The initial, brief activation is identical for all conditioned reward predicting, unrewarded, differently valued and aversive stimuli, with <100 ms latency and 50–150 ms duration [2,34,51–53]. It is shorter than the full responses to reward predicting stimuli and often curtailed by

subsequent depression. It leads to dopamine release in rats [54]. The responses show correlations with air puff probability across the population, which may arise from a subset of well activated neurons [37]. A proper control procedure maintains the visual aversive stimulus unchanged but changes the modality of the randomly interleaved reward predicting stimulus, for instance from visual to auditory. This manipulation reduces conditioned aversive responses from 65% to <15% of neurons [34] (Figure 1d, e), which approaches the proportion of neurons activated by primary aversive events. Thus physical similarity to rewarding stimuli may produce false aversive responses, possibly through response generalisation to rewarding stimuli. Also, reducing overall reward probability from 75% to 25% between experiments decreases the incidence of false activations by unrewarded stimuli from 63% to 1% [52,53], suggesting an effect of context conditioning (pseudoconditioning). Apparently the initial response occurs before the neuron has properly identified the stimulus.

The second response component constitutes a genuine value response [2] which differentiates well between reward and other events. It codes a graded TD reward value error, consisting of depression following aversive or unrewarded stimuli [36,51–53], and graded activations with different values [17]. The updated prediction in this response is reflected in the error response at reward time [15,52], see Figure 5 in [2].

The two response components are better separated in a dot motion task that requires time for stimulus identification [3•] (Figure 1f). The initial dopamine activation before stimulus identification fails to vary with value. However, it decreases when its occurrence becomes more predicted as time goes by (increasing hazard rate), thus coding a temporal prediction error. The second component is well separated and occurs 150–200 ms later when the animal identifies the motion direction that determines reward probability and thus value. It reflects the TD error between the reward value of the specific stimulus and the value predicted by the preceding fixation spot.

Taken together, the more frequent activations to conditioned aversive stimuli compared to primary punishers is probably explained by the nature of the two response components. The initial, stereotyped activation constitutes a pre-identification, higher order temporal TD error response, or possibly response generalisation to rewarding stimuli or contexts. After proper identification of the stimulus, the second component codes a TD error accurately. A fast but inaccurate initial neuronal response may provide survival advantages. It could lead to early initiation of behavioural reactions in the course of which the exact nature of the stimulus becomes identified by the second component. The mechanism would allow the agent to arrive first at the object and thus provide a slight advantage over slower competitors, which would ultimately pan out as higher evolutionary fitness. If the initial response is false and the object unrewarded, behavioural reactions can be cancelled while the second response component identifies the object 200 ms later, which would be a small price to pay.

Whereas interest in dopamine reward responses has overshadowed the long known aversive responses, some current work addresses punishment in identifiable groups of dopamine neurons without assessing their incidence relative to reward responsive neurons [32,42]. Although some dopamine neurons are definitely activated by aversive events, their proportion remains as small as 30 years ago [26,27].

Salience

Salience refers to the capacity of stimuli to elicit arousal, alert and attention which enhance neuronal processing and behavioural reactions. There are, in principle, three forms of salience. Physical salience derives from physically strong stimuli. Novelty or surprise salience originates in novel or surprising events. Motivational salience is produced commonly by rewards and punishers. (Incentive salience refers to motivation for rewards, in contrast to learning, and thus applies to rewards. It elicits the emotion of desire, called 'wanting', rather than valence independent attention [55].)

Would salience explain the higher incidence of responses to conditioned aversive stimuli (35–65%) compared to unpredicted air puffs (<15%) [34,37]? The physical salience of small conditioned visual stimuli is marginal [34,36,37] and higher with air puffs and electric shocks. Thus physical salience would be higher for the unconditioned than conditioned aversive stimuli. Novelty salience should not be a factor with well trained conditioned stimuli, and surprise salience would be equally high with unpredicted conditioned stimuli and unpredicted air puffs. Motivational salience would be higher with unpredicted air puffs than with conditioned stimuli predicting air puffs. All evidence from learning experiments suggests that the highest motivational impact arises from the primary air puff. Even fully conditioned stimuli would not have higher motivational impact, which is further reduced by the delay to air puff (temporal discounting). Thus unpredicted air puff carries the highest total salience of all events, and salience would not explain the higher incidence of activations to conditioned stimuli compared to air puffs. Motivational salience would be restricted to <15% of dopamine activations, and the remaining responses should be owing to pre-identification responses described above, or possibly pseudoconditioning [56,57].

Further arguments limit motivational salience as a major factor determining dopamine responses. First, the change from visual to auditory rewarded stimuli reduces aversive dopamine activations to stimuli without reducing the salience of the unchanged aversive stimulus [34]. Second, conditioned inhibitors for reward are motivationally very salient but fail to activate most dopamine neurons when pre-identification responses are excluded [53]. Third, negative reward prediction errors are motivationally very salient but depress rather than activate dopamine neurons, including those activated by air puffs [36,37].

Whereas 70–80% of dopamine neurons code reward prediction errors, only 10–15% are activated by both rewarding and aversive events and thus motivational salience. Therefore motivational salience induces 5–8 times less dopamine activations than reward, rather than similar proportions as suggested [58]. The depressant responses to the negative value of aversive stimuli are compatible with the predominantly positive value coding nature of dopamine neurons.

Homogenous vs. heterogeneous properties

Together with the previously described risk responses in 30% of dopamine neurons [25], dopamine neurons show several response types with unequal proportions. One large population codes reward value, and smaller populations code reward risk, motivational salience and possibly other event properties. The primarily reward related nature of phasic dopamine signals has been confirmed in the majority of optogenetically identified dopamine neurons [39••] and supported by the restriction of prediction error coding to reward [36,37]. Nonexclusive coding of one particular event type lies within the natural variations of biological systems whose flexibility benefits from minor deviations.

The reward responses are homogeneous in terms of latency, duration, positive monotonicity of value coding and error coding [2] and somewhat heterogeneous owing to risk and salience coding. They occur in a neuronal system with heterogeneous inputs, transmitter colocalisation, axonal projections, and receptor and reuptake transporter expression, which may lead to regionally inhomogeneous dopamine release [49,59]. The net result of optogenetic activation of midbrain dopamine neurons is rewarding, as shown by place preference conditioning and operant nose poking behaviour [60••,61••,62•], although regional stimulations in terminal areas might reveal fine grained heterogeneities. Despite these variations, the high percentage of dopamine neurons coding reward, and their temporal response homogeneity in all sorts of behavioural tasks, is amazing. No other brain structure has such high proportion of reward processing neurons with such low variety of task relationships, including striatum, orbitofrontal cortex and amygdala.

Although complicated tasks may let dopamine responses appear complex, appropriate analysis in terms of TD error coding demonstrates their relatively simplistic and stereotyped character [4].

Effects of dopamine signal

The bidirectional dopamine prediction error response at the time of the reward implements fully the error term of the Rescorla-Wagner learning rule [2]. The dopamine responses to higher order, reward predicting stimuli comply with the notions of efficient TD teaching signals [63]. These dopamine signals may be involved in the two main, objectively measurable reward functions. An error signal would be ideal for mediating value learning and updating for decision making. Considering the brain as a prediction machine, an error signal would also be appropriate for inducing approach behaviour and affecting decision processes.

Postsynaptic neuronal plasticity

A viable neuronal reinforcement mechanism involves three-factor Hebbian plasticity with dopamine and glutamate synapses located at the same dendritic spines of striatal or frontal cortical neurons, the so-called synaptic triad [2,64–66]. Indeed, electrical midbrain stimulation induces dopamine D1 receptor dependent long term potentiation in striatal neurons [66]. Further evidence suggests a crucial role of dopamine in both long-term potentiation (LTP) and depression (LTD) in striatum [67–70], frontal cortex [71,72], hippocampus [73] and amygdala [74]. Iontophoretically applied dopamine puffs induce plasticity in aplysia [75].

Protocols for spike time dependent plasticity (STDP) demonstrate LTP when presynaptic stimulation precedes postsynaptic stimulation by a few tens of milliseconds, whereas LTD occurs with reverse sequence. Intact dopamine D1 receptors are required for both forms of cortically evoked plasticity in striatal neurons involving NMDA receptors [76••]. When dissociating dopamine receptor localisation on striatonigral neurons (D1; 'direct pathway') and striatopallidal neurons (D2; 'indirect pathway'), D1 receptors are involved in LTP in direct pathway neurons and D2 receptors are crucial for LTD in indirect pathway neurons [77••] (Figure 2a, b). Confirming the specificity, LTP in indirect pathway neurons does not depend on dopamine [70], which may explain dopamine independent LTP in striatal neurons not distinguished according to direct vs. indirect pathways [78]. In indirect pathway neurons, LTP is turned into LTD by additional dopamine D2 receptor stimulation [77••]. In hippocampus, stimulation of dopamine D1 receptors enhances STDP LTP and turns LTD into LTP [79]. Neuronal modelling demonstrates the power of a dopamine reward signal occurring a few seconds after a conditioned stimulus typical for STDP learning [80•].

Figure 2
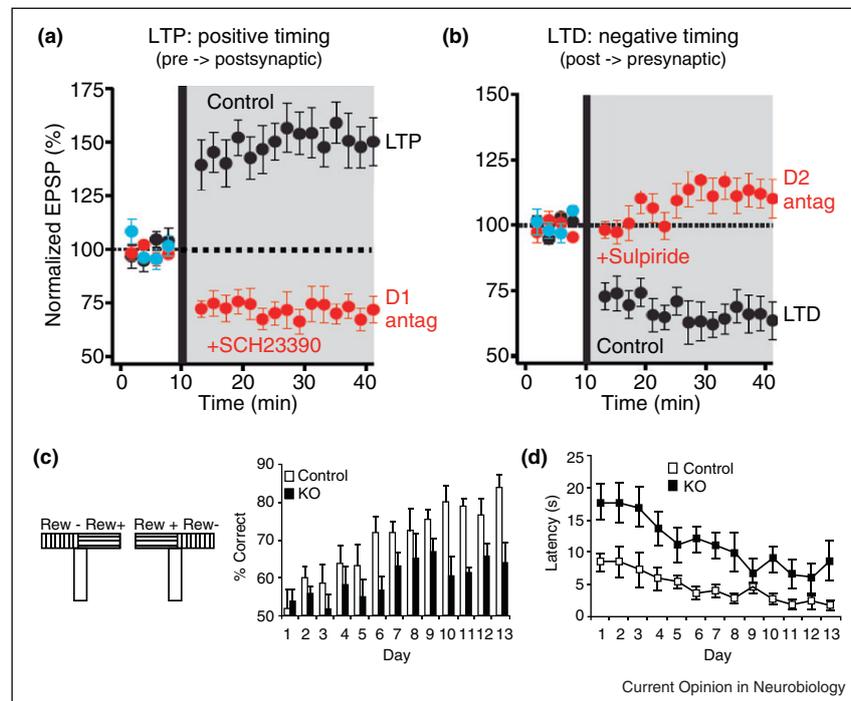
Dopamine dependency of neuronal plasticity and behavioural learning. **(a)** Positive timing in spike time dependent plasticity protocol (STDP) results in long term potentiation (LTP) at synapses from cortical inputs to striato-nigral neurons (direct pathway) (black) and is blocked by dopamine D1 receptor antagonist SCH23390 (red) ([77••], © Science). **(b)** Negative timing in STDP protocol results in long term depression (LTD) at cortical synapses onto striato-pallidal neurons (indirect pathway) (black) and is blocked by dopamine D2 receptor antagonist sulpiride (red) ([77••], © Science). **(c)** T-maze learning deficit in mice with NMDA receptor knock-out in midbrain dopamine neurons impairing dopamine burst firing ([84•], © National Academy of Sciences USA). **(d)** Separate performance deficit in mice tested in (c).

Presynaptic neuronal plasticity

Dopamine neurons show induction of NMDA receptor dependent LTP when a short postsynaptic burst occurs after at least 500 ms of presynaptic stimulation [81•]. A delay of 1 s increases LTP, whereas delays of 0 or 200 ms fail to elicit LTP, and negative delays induce LTD. The delay of 0.5–1 s may reflect the time required to achieve sufficient intracellular IP3 levels, which could constitute a molecular stimulus or eligibility trace [9] tagging synapses for modification by the postsynaptic burst. The protocol follows the requirement for behavioural conditioning in which the unconditioned stimulus (US) follows the conditioned stimulus (CS) by an optimal interval. The postsynaptic burst parallels the dopamine US response, and the NMDA receptor dependent LTP parallels the acquired dopamine CS response. Similar to behavioural and neuronal response extinction, omission of the postsynaptic burst reverses LTP back to baseline [81•]. Synaptic plasticity is not uniform in dopamine neurons but varies according to projection territories [42].

Immediate effects

Dopamine exerts immediate postsynaptic effects during behavioural performance and approach behaviour. At striatal neurons of the direct pathway, dopamine has excitatory effects via the D1 receptor by eliciting or prolonging glutamate inputs and transitions to the up state (depolarisation) of the membrane potential, whereas in indirect pathway neurons D2 receptor activation has inhibitory effects by reducing glutamate release and prolonging membrane down states (hyperpolarisation) [82]. These immediate effects are synergistic with the plasticity function of dopamine in striatal LTP and LTD in direct and indirect pathway neurons [77••].

Behavioural learning

Hundreds of lesioning and psychopharmacological studies using various tasks demonstrate learning deficits with impaired dopamine transmission. At the level of dopamine neurons, knock outs of burst generating NMDA receptors in mice, with resulting decreases in responses to reward predicting stimuli, lead to a range of learning deficits, including conditioned place preference, maze learning and operant conditioning, that are dissociated from performance deficits [83,84•,85] (Figure 2c, d). In confirmation, knock out of GABA-A receptors inhibiting dopamine neurons enhances electrically evoked dopamine release and T-maze and lever press learning [86]. At

the postsynaptic level, intra-accumbens or systemic administration of dopamine D1 receptor blockers impairs simple stimulus-reward learning [87] that engages dopamine neurons, well distinguished from performance [88••]. Learning is also impaired by knock out of NMDA receptors on mouse striatal neurons expressing dopamine D1 receptors [89], closely following the scheme of three-factor Hebbian learning [2,64–66]. Similarly, learning but not performance deficits occur in a visual stimulus-reward association task by administration of a dopamine D1 receptor antagonist into monkey prefrontal cortex [90•].

By contrast, learning is less impaired by systemic D1 receptor blockade in select tasks with separate stimulus and goal locations that fail to engage dopamine neurons [88••] or when phasic dopamine release remains functional despite NMDA receptor knock-out [91]. Also, neurotoxic dopamine lesions leave reward devaluation learning by taste aversion intact [55]; this learning form may not engage phasic dopamine responses.

Taken together, learning depends on intact dopamine function in simple reward contiguity situations with explicit, easily identifiable rewards and conditioned stimuli that engage phasic dopamine responses. Learning may not depend on dopamine neurons if their phasic responses remain unengaged during learning. Despite this obvious conclusion, learning of tasks normally engaging dopamine responses might even proceed despite lesions when neuronal plasticity allows other learning systems to compensate within the tested time frame. Knock out of specific learning systems may lead to modified ontogenesis with even more opportunity for shifts to other systems. Given the crucial importance of reward for survival, multiple, flexible learning systems are biologically plausible and would enhance evolutionary fitness.

Acknowledgements

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

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

2. Schultz W: **Predictive reward signal of dopamine neurons**. *J Neurophysiol* 1998, **80**:1-27.

3. Nomoto K, Schultz W, Watanabe T, Sakagami M: **Temporally**
• **extended dopamine responses to perceptually demanding reward-predictive stimuli**. *J Neurosci* 2010, **30**:10692-10702.
Dopamine neurons show two well separated error response components to reward predicting stimuli in a random dot motion task.

4. Enomoto K, Matsumoto N, Nakai S, Satoh T, Sato TK, Ueda Y,
•• Inokawa H, Haruno M, Kimura M: **Dopamine neurons learn to encode the long-term value of multiple future rewards**. *Proc Natl Acad Sci USA* 2011, **108**:15462-15467.
The most advanced neurophysiological study to date relating dopamine responses to TD learning.

5. Bayer HM, Glimcher PW: **Midbrain dopamine neurons encode a quantitative reward prediction error signal**. *Neuron* 2005, **47**:129-141.

6. Glimcher PW: **Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis**. *Proc Natl Acad Sci USA* 2011, **108**:15647-15654.

7. Bayer HM, Lau B, Glimcher PW: **Statistics of dopamine neuron spike trains in the awake primate**. *J Neurophysiol* 2007, **98**:1428-1439.

8. Rescorla RA, Wagner AR: **A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement**. In *Classical Conditioning II: Current Research and Theory*. Edited by Black AH, Prokasy WF. Appleton Century Crofts; 1972:64-99.

9. Sutton RS, Barto AG: **Toward a modern theory of adaptive networks: expectation and prediction**. *Psychol Rev* 1981, **88**:135-170.

10. Sutton RS, Barto AG: *Reinforcement Learning*. MIT Press; 1998.

11. Doya K, Samejima K, Katagiri K, Kawato M: **Multiple model-based reinforcement learning**. *Neural Comput* 2002, **14**:1347-1369.

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

13. Tobler PN, Fiorillo CD, Schultz W: **Adaptive coding of reward value by dopamine neurons**. *Science* 2005, **307**:1642-1645.

14. Bromberg-Martin ES, Matsumoto M, Hong S, Hikosaka O: **A**
•• **pallidus-habenula-dopamine pathway signals inferred stimulus values**. *J Neurophysiol* 2010, **104**:1068-1076.
Dopamine responses reflect position-reward reversals before the animal experienced the new position-reward pairings, suggesting an influence of the reversal rule on the prediction error response and possible participation in model-based reinforcement learning.

15. Kobayashi S, Schultz W: **Influence of reward delays on responses of dopamine neurons**. *J Neurosci* 2008, **28**:7837-7846.

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

17. Roesch MR, Calu DJ, Schoenbaum G: **Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards**. *Nat Neurosci* 2007, **10**:1615-1624.

18. Roesch MR, Olson CR: **Neuronal activity in orbitofrontal cortex reflects the value of time**. *J Neurophysiol* 2005, **94**:2457-2471.

19. Roesch MR, Olson CR: **Neuronal activity dependent on anticipated and elapsed delay in macaque prefrontal cortex, frontal and supplementary eye fields, and premotor cortex**. *J Neurophysiol* 2005, **94**:1469-1497.

20. Kim S, Hwang J, Seo H, Lee D: **Valuation of uncertain and delayed rewards in primate prefrontal cortex**. *Neural Networks* 2009, **22**:294-304.

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

22. Cai X, Kim S, Lee D: **Heterogeneous coding of temporally discounted values in the dorsal and ventral striatum during intertemporal choice**. *Neuron* 2011, **69**:170-182.

23. Louie K, Glimcher PW: **Separating value from choice: delay discounting activity in the lateral intraparietal area**. *J Neurosci* 2010, **30**:5498-5507.

24. De Lafuente O, Romo R: **Dopamine neurons code subjective**
•• **sensory experience and uncertainty of perceptual decisions**. *Proc Natl Acad Sci USA* 2011, **49**:19767-19771.
Reward coding reflects the subjective perception rather than physical appearance of reward predicting stimuli in a rigorous signal detection test.

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

26. Tsai CT, Nakamura S, Iwama K: **Inhibition of neuronal activity of the substantia nigra by noxious stimuli and its modification by the caudate nucleus**. *Brain Res* 1980, **195**:299-311.

27. Hommer DW, Bunney BS: **Effect of sensory stimuli on the activity of dopaminergic neurons: involvement of non-dopaminergic nigral neurons and striato-nigral pathways**. *Life Sci* 1980, **27**:377-386.

28. Schultz W, Romo R: **Responses of nigrostriatal dopamine neurons to high intensity somatosensory stimulation in the anesthetized monkey**. *J Neurophysiol* 1987, **57**:201-217.

29. Ungless MA, Magill PJ, Bolam JP: **Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli**. *Science* 2004, **303**:2040-2042.

30. Coizet V, Dommett EJ, Redgrave P, Overton PG: **Nociceptive responses of midbrain dopaminergic neurones are modulated by the superior colliculus in the rat**. *Neuroscience* 2006, **139**:1479-1493.

31. Brown MTC, Henny P, Bolam JP, Magill PJ: **Activity of neurochemically heterogeneous dopaminergic neurons in the substantia nigra during spontaneous and driven changes in brain state**. *J Neurosci* 2009, **29**:2915-2925.

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

33. Tan KR, Yvon C, Turiault M, Mirzabekov JJ, Doehner J,
• Labouèbe G, Deisseroth K, Tye KM, Lüscher C: **GABA neurons of the VTA drive conditioned place aversion**. *Neuron* 2012, **73**:1173-1183.
Optogenetic stimulation of midbrain GABA neurons induces depressions of dopamine neurons, as electrical footshock does, and leads to rapid aversion of a chamber in which the animal receives the stimulation.

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

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

36. Joshua M, Adler A, Mitelman R, Vaadia E, Bergman H: **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* 2008, **28**:11673-11684.

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

38. Mileykovskiy B, Morales M: **Duration of inhibition of ventral
• tegmental area dopamine neurons encodes a level of conditioned fear**. *J Neurosci* 2011, **31**:7471-7476.
A heroic effort to investigate the aversive responses of histologically identified dopamine neurons in awake rats, reporting predominantly inhibitory responses.

39. Cohen JY, Haesler S, Vong L, Lowell BB, Uchida N: **Neuron-type-
•• specific signals for reward and punishment in the ventral tegmental area**. *Nature* 2012, **482**:85-88.
The first study to assess reward and punishment responses in optogenetically identified dopamine neurons of behaving mice. The data confirm the earlier reported dopaminergic nature of prediction error responses.

40. Henny P, Brown MT, Northrop A, Faunes M, Ungless MA, Magill PJ, Bolam JP: **Structural correlates of heterogeneous in vivo activity of midbrain dopaminergic neurons**. *Nat Neurosci* 2012, **15**:613-619.

41. van Zessen R, Phillips JL, Budygin EA, Stuber GD: **Activation of
• VTA GABA neurons disrupts reward consumption**. *Neuron* 2012, **73**:1184-1194.
Optogenetic stimulation of midbrain GABA neurons induces depressions of dopamine neurons and reduced behavioural reward consumption.

42. Lammel S, Ion DI, Roeper J, Malenka RC: **Projection-specific modulation of dopamine neuron synapses by aversive and rewarding stimuli**. *Neuron* 2011, **70**:855-862.

43. Wang DV, Tsien JZ: **Convergent processing of both positive and negative motivational signals by the VTA dopamine neuronal populations**. *PLoS ONE* 2011, **6**:e17047.

44. Bunney BS, Aghajanian GK, Roth RH: **Comparison of effects of ʟ-dopa, amphetamine and apomorphine on firing rate of rat dopaminergic neurons**. *Nat New Biol* 1973, **245**:123-125.

45. Freeman AS, Bunney BS: **Activity of A9 and A10 dopaminergic neurons in unrestrained rats: further characterization and effects of cholecystokinin**. *Brain Res* 1987, **405**:46-55.

46. Valenti O, Lodge DJ, Grace AA: **Aversive stimuli alter ventral tegmental area dopamine neuron activity via a common action in the ventral hippocampus**. *J Neurosci* 2011, **31**:4280-4289.

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

48. Roitman MF, Wheeler RA, Wightman RM, Carelli RM: **Real-time chemical responses in the nucleus accumbens differentiate rewarding and aversive stimuli**. *Nat Neurosci* 2008, **11**:1376-1377.

49. Budygin EA, Park J, Bass CE, Grinevich VP, Bonin KD, Wightman RM: **Aversive stimulus differentially triggers subsecond dopamine release in reward regions**. *Neuroscience* 2012, **201**:331-337.

50. SolomonRL, Corbit JD: **An opponent-process theory of motivation**. *Psychol Rev* 1974, **81**:119-145.

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

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

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

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

55. Berridge KC, Robinson TE: **What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience?** *Brain Res Rev* 1998, **28**:309-369.

56. Schultz W: **Dopamine signals for reward value and risk: basic and recent data**. *Behav Brain Funct* 2010, **6**:24.

57. Sheafor PJ: **Pseudoconditioned jaw movements of the rabbit reflect associations conditioned to contextual background cues**. *J Exp Psychol: Anim Behav Proc* 1975, **104**:245-260.

58. Bromberg-Martin ES, Matsumoto M, Hikosaka O: **Dopamine in motivational control: rewarding, aversive, and alerting**. *Neuron* 2010, **68**:815-834.

59. Brown HD, McCutcheon JE, Cone JJ, Ragozzino ME, Roitman MF: **Primary food reward and reward-predictive stimuli evoke different patterns of phasic dopamine signaling throughout the striatum**. *Eur J Neurosci* 2011, **34**:1997-2006.

60. Tsai H-C, Zhang F, Adamantidis A, Stuber GD, Bonci A, de
•• Lecea L, Deisseroth K: **Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning**. *Science* 2009, **324**:1080-1084.
The first study to use optogenetic stimulation of dopamine neurons to induce behavioural responses.

61. Witten IB, Steinberg EE, Lee SY, Davidson TJ, Zalocusky KA,
•• Brodsky M, Yizhar O, Cho SL, Gong S, Ramakrishnan C, Stuber GD, Tye KM, Janak PH, Deisseroth K: **Recombinase-driver rat lines: tools, techniques, and optogenetic application to dopamine-mediated reinforcement**. *Neuron* 2011, **72**:721-733.
The study shows operant behavioural self-stimulation (nose pokes) by optogenetic activation of dopamine neurons and resolves the long disputed issue of the dopaminergic nature of electrical self-stimulation. The paper describes many methodological aspects of optogenetics and shows learning, extinction and reacquisition curves of nose pokes, and correlation of nose pokes with amount of stimulation. Amazing work.

62. Kim KM, Baratta MV, Yang A, Lee D, Boyden ES, Fiorillo CD:
• **Optogenetic mimicry of the transient activation of dopamine neurons by natural reward is sufficient for operant reinforcement**. *PLoS ONE* 2012, **7**:e33612.
The data confirm learning and extinction of nose poke behaviour shown by [61••] and in addition show correlations between channelrhodopsin expression and induced behavioural responses. Together with [61••], the study makes the link to electrical self-stimulation discovered 60 years ago by Olds and Milner.

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

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

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

66. Reynolds JNJ, Hyland BI, Wickens JR: **A cellular mechanism of reward-related learning**. *Nature* 2001, **413**:67-70.

67. Calabresi P, Gubellini P, Centonze D, Picconi B, Bernardi G, Chergui K, Svenningsson P, Fienberg AA, Greengard P: **Dopamine and cAMP-regulated phosphoprotein 32 kDa controls both striatal long-term depression and long-term potentiation, opposing forms of synaptic plasticity**. *J Neurosci* 2000, **20**:8443-8451.

68. Kerr JN, Wickens JR: **Dopamine D-1/D-5 receptor activation is required for long-term potentiation in the rat neostriatum in vitro**. *J Neurophysiol* 2001, **85**:117-124.

69. Tang KC, Low MJ, Grandy DK, Lovinger DM: **Dopamine-dependent synaptic plasticity in striatum during in vivo development**. *Proc Natl Acad Sci USA* 2001, **98**:1255-1260.

70. Kreitzer AC, Malenka RC: **Endocannabinoid-mediated rescue of striatal LTD and motor deficits in Parkinson's disease models**. *Nature* 2007, **445**:643-647.

71. Otani S, Blond O, Desce JM, Crepel F: **Dopamine facilitates long-term depression of glutamatergic transmission in rat prefrontal cortex**. *Neuroscience* 1998, **85**:669-676.

72. Gurden H, Takita M, Jay TM: **Essential role of D1 but not D2 receptors in the NMDA receptor-dependent long-term potentiation at hippocampal-prefrontal cortex synapses in vivo**. *J Neurosci* 2000, **20RC106**:1-5.

73. Otmakhova NA, Lisman JE: **D1/D5 dopamine receptor activation increases the magnitude of early long-term potentiation at CA1 hippocampal synapses**. *J Neurosci* 1996, **16**:7478-7486.

74. Rosenkranz JA, Grace AA: **Dopamine-mediated modulation of odour-evoked amygdala potentials during pavlovian conditioning**. *Nature* 2002, **417**:282-287.

75. Brembs B, Lorenzetti FD, Reyes FD, Baxter DA, Byrne JH: **Operant learning in aplysia: neuronal correlates and mechanisms**. *Science* 2002, **296**:1706-1709.

76. Pawlak V, Kerr JND: **Dopamine receptor activation is required**
•• **for corticostriatal spike-timing-dependent plasticity**. *J Neurosci* 2008, **28**:2435-2446.
Rigorous application of an STDP protocol reveals the necessary role of dopamine receptors in LTP and LTD.

77. Shen W, Flajolet M, Greengard P, Surmeier DJ: **Dichotomous**
•• **dopaminergic control of striatal synaptic plasticity**. *Science* 2008, **321**:848-851.

Classification of D1 receptor carrying 'direct pathway' striatal neurons and D2 receptor carrying 'indirect pathway' neurons allows dissociation of dopamine influences on LTP and LTD. Huge number of tests.

78. Pennartz CMA, Ameerun RF, Groenewegen HJ, Lopes da Silva FH: **Synaptic plasticity in an in vitro slice preparation of the rat nucleus accumbens**. *Europ J Neurosci* 1993, **5**:107-117.

79. Zhang J-C, Lau P-M, Bi G-Q: **Gain in sensitivity and loss in temporal contrast of STDP by dopaminergic modulation at hippocampal synapses**. *Proc Natl Acad Sci USA* 2009, **106**:1328-1333.

80. Izhikevich EM: **Solving the distal reward problem through**
• **linkage of STDP and dopamine signaling**. *Cerebral Cortex* 2007, **17**:2443-2452.
Modelling the dopamine influence on STDP.

81. Harnett MT, Bernier BE, Ahn K-C, Morikawa H: **Burst-timing-**
• **dependent plasticity of NMDA receptor-mediated transmission in midbrain dopamine neurons**. *Neuron* 2009, **62**:826-838.
Plasticity in dopamine neurons follows the same temporal requirement as behavioural Pavlovian conditioning.

82. Hernandez-Lopez S, Bargas J, Surmeier DJ, Reyes A, Galarraga E: **D1 receptor activation enhances evoked discharge in neostriatal medium spiny neurons by modulating an L-type Ca²⁺ conductance**. *J Neurosci* 1997, **17**:3334-3342.

83. Zweifel LS, Argilli E, Bonci A, Palmiter R: **Role of NMDA receptors in dopamine neurons for plasticity and addictive behaviors**. *Neuron* 2008, **59**:486-496.

84. Zweifel LS, Parker JG, Lobb CJ, Rainwater A, Wall VZ, Fadok JP,
• Darvas M, Kim MJ, Mizumori SJ, Paladini CA, Phillips PEM, Palmiter R: **Disruption of NMDAR-dependent burst firing by dopamine neurons provides selective assessment of phasic dopamine-dependent behavior**. *Proc Natl Acad Sci USA* 2009, **106**:7281-7288.
Knock out of NMDA receptors on dopamine neurons reduces burst firing along with behavioural learning in a wide variety of tasks.

85. Wang LP, Li F, Wang D, Xie K, Wang D, Shen X, Tsien JZ: **NMDA receptors in dopaminergic neurons are crucial for habit learning**. *Neuron* 2011, **72**:1055-1066.

86. Parker JG, Wanat MJ, Soden ME, Ahmad K, Zweifel LS, Bamford NS, Palmiter RD: **Attenuating GABAA receptor signaling in dopamine neurons selectively enhances reward learning and alters risk preference in mice**. *J Neurosci* 2011, **31**:17103-17112.

87. Di Ciano P, Cardinal RN, Cowell RA, Little SJ, Everitt BJ: **Differential involvement of NMDA, AMPA/kainate, and dopamine receptors in the nucleus accumbens core in the acquisition and performance of Pavlovian approach behavior**. *J Neurosci* 2001, **21**:9471-9477.

88. Flagel SB, Clark JJ, Robinson TE, Mayo L, Czuj A, Willuhn I,
•• Akers CA, Clinton SM, Phillips PE, Akil H: **A selective role for dopamine in stimulus-reward learning**. *Nature* 2011, **469**:53-57.
Systemically applied dopamine D1 receptor blocker impairs selectively stimulus-reward contiguity ('sign-tracking') learning while leaving more complex learning ('goal-tracking') intact. The differential deficit correlates with differential dopamine responses during sign-tracking but not goal-tracking (without receptor blockade). The learning deficit is well separated from performance deficits.

89. Parker JG, Beutler LR, Palmiter RD: **The contribution of NMDA receptor signaling in the corticobasal ganglia reward network to appetitive Pavlovian learning**. *J Neurosci* 2011, **31**:11362-11369.

90. Puig MV, Miller EK: **The role of prefrontal dopamine D1**
• **receptors in the neural mechanisms of associative learning**. *Neuron* 2012, **74**:874-886.
Local injections of dopamine D1 receptor antagonist into prefrontal cortex of monkeys impair stimulus-reward learning, while leaving performance intact, and confirms crucial involvement of D1 receptors in learning.

91. Parker JG, Zweifel LS, Clark JJ, Evans SB, Phillips PEM, Palmiter RD: **Absence of NMDA receptors in dopamine neurons attenuates dopamine release but not conditioned approach during Pavlovian conditioning**. *Proc Nat Acad Sci USA* 2010, **107**:13491-13496.