OPINION

Dopamine reward prediction-error signalling: a two-component response

Wolfram Schultz

Abstract | Environmental stimuli and objects, including rewards, are often processed sequentially in the brain. Recent work suggests that the phasic dopamine reward prediction-error response follows a similar sequential pattern. An initial brief, unselective and highly sensitive increase in activity unspecifically detects a wide range of environmental stimuli, then quickly evolves into the main response component, which reflects subjective reward value and utility. This temporal evolution allows the dopamine reward prediction-error signal to optimally combine speed and accuracy.

Rewards induce behaviours that enable animals to obtain necessary objects for survival. Although the term 'reward' is commonly associated with happiness, in scientific terms rewards have three functions. First, they can act as positive reinforcers to induce learning. Second, rewards elicit movements towards the desired object and constitute factors to be considered in economic choices. Their value for the individual decision maker is subjective and can be formalized as economic utility. The third reward function is associated with emotions, such as pleasure and desire. This third function is difficult to test in animals, but the first two and their underlying brain processes can be quantitatively assessed using specific behavioural tasks, and hence are the focus of this article.

Electrophysiological investigations in monkeys, rats and mice have identified individual neurons that signal reward-related information in the midbrain dopamine system (substantia nigra and ventral tegmental area (VTA)), striatum, orbitofrontal cortex, amygdala and associated structures[1]. These reward neurons process specific aspects of rewards — such as their amount, probability, subjective value, formal economic utility and risk — in forms that are suitable for learning and decision-making.

Most dopamine neurons in the substantia nigra and VTA show brief, phasic responses to rewards and reward-predicting stimuli. These responses code a temporal reward prediction error, which reflects the difference in value between a received reward and a predicted reward at each moment in time[2–9]. This fast dopamine signal differs distinctly from the slower dopamine activity increases that reflect reward risk[10] or, more inconsistently, behavioural reactivity[2,11–14]; it differs most from the tonic dopamine level that is necessary to enable neuronal processes underlying a wide range of behaviours (BOX 1).

Despite robust evidence for their involvement in reward coding, research over more than three decades has shown that some dopamine neurons show phasic activity increases in response to non-rewarding and aversive stimuli[2,6,15–27]. This discrepancy does not rule out a role for phasic dopamine release in reward processing; indeed, other components of the brain's reward systems also contain distinct neuronal populations that code non-rewarding events[28–32]. However, the extent to which phasic dopamine responses code reward versus non-reward information has been difficult to resolve because of their sensitivity to experimental conditions[19].

Several recent studies have encouraged a revision of our views on the nature of the phasic dopamine reward response. The studies demonstrate distinct subcomponents of the phasic dopamine response[33], provide an alternative explanation for activations in response to aversive stimuli[25–27] and document strong sensitivity to some unrewarded stimuli[34]. Here, I outline and evaluate the evidence for a more elaborate view of the phasic dopamine reward prediction-error signal, which evolves from an initial response that unselectively detects any potential reward (including stimuli that turn out to be aversive or neutral) to a subsequent main component that codes the by now well-identified reward value. Furthermore, I suggest that the reward prediction-error response should be specifically considered to be a utility prediction error signal[35].

Processing of reward components

Reward components. Rewards consist of distinct sensory and value components (FIG. 1). Their neuronal processing takes time and engages sequential mechanisms, which becomes particularly evident when the rewards consist of more-complex objects and stimuli. Rewards first impinge on the body through their physical sensory impact. They draw attention through their physical salience, which facilitates initial detection. The specific identity of rewards derives from their physical parameters, such as size, form, colour and position, which engage subsequent sensory and cognitive processes. Comparison with known objects determines their novelty, which draws attention through novelty salience and surprise salience. During and after their identification, valuation takes place. Value is the essential feature that distinguishes rewards from other objects and stimuli; it can be estimated from behavioural preferences that are elicited in choices. Value draws attention because it provides motivational salience. The various forms of salience — physical, novelty, surprise and motivational — induce stimulus-driven attention, which selects information and modulates neuronal processing[36–40]. Thus, neuronal reward processing evolves in time from unselective sensory detection to the more demanding and crucial stages of identification and valuation. These processes lead to internal

decisions and overt choices between different options, to actions towards the chosen option and to feedback that updates the neural representation of a reward's value.

Sequential processing in other systems. Research in sensory, cognitive and reward systems has long recognized the component nature of stimuli and objects. Although simple stimuli may be processed too rapidly to reveal their dissociable components, more-sophisticated events take longer to identify, discriminate and valuate. For example, during a visual search task, neurons in the frontal eye fields, lateral intraparietal cortex (LIP) and cortical visual area V4 exhibit an initial unselective response and require a further 50–120 ms to discriminate targets from distractors[41–44] (FIG. 2a). Similarly, during perceptual discriminations between partly coherently moving dots, neuronal activity in the LIP and dorsolateral prefrontal cortex becomes distinctive only 120–200 ms after the initial detection response[45,46]. Even responses in regions of the primary sensory cortex, such as V1 or the mouse somatosensory barrel cortex, show temporal evolutions from gross tuning for stimulus properties to a more finely discriminative response[47–53]. The results of other studies have supported an alternative view in which hierarchically connected neurons evaluate both the available information and prior knowledge about the scene[52]. Although these results demonstrate a sequential neuronal processing flow, typical feature- or category-specific neurons in the inferotemporal cortex process highly specific stimulus properties of visual objects at the same time as they detect and identify them[54]. Despite these exceptions, there is a body of evidence suggesting that many sensory and cognitive neurons process the different components of demanding stimuli in consecutive steps (Supplementary information S1 (table)).

Processing a reward requires an additional valuation step (FIG. 1). Neurons in primary reward structures such as the amygdala show an initial sensory response, followed 60–300 ms later by a separate value component[31,55,56], and V1 and inferotemporal cortex neurons show similar response transitions with differently rewarded stimuli[57,58]. Thus, the neuronal processing of rewards might also involve sequential steps.

Sequential processing in dopamine neurons. Similar to the responses of neurons in other cognition- and reward-related brain regions, phasic dopamine reward prediction error responses show a temporal evolution with sequential components[5,6,9,17,22,25,33,35,59–62] (FIG. 2b–e; Supplementary information S1 (table)). An initial, short latency and short-duration activation of dopamine neurons detects any environmental object before having properly identified and valued it. The subsequently evolving response component properly identifies the object and values it in a finely graded manner. These prediction-error components are often difficult to discern when simple, rapidly distinguishable stimuli are used; however, they can be revealed by specific statistical methods[25] (FIG. 2c) or by using more demanding stimuli that require longer processing[33] (FIG. 2e). Below, I describe in detail the characteristics of these components and the factors influencing them.

The initial component: detection

Effective stimuli. The initial component of the phasic dopamine reward prediction-error response is a brief activation (an increase in the frequency of neuronal impulses) that occurs unselectively in response to a large variety of unpredicted events, such as rewards[9,35], reward-predicting stimuli[9,25,33,59–62] (FIG. 2b–e), unrewarded stimuli[5,16,17,20,34], aversive stimuli[22] and conditioned inhibitors predicting reward omission[6] (FIG. 2b). Owing to their varying sensitivities, some dopamine neurons show little or no initial component. True to the notion that dopamine neurons

Box 1 | Fast, slow and tonic dopamine functions

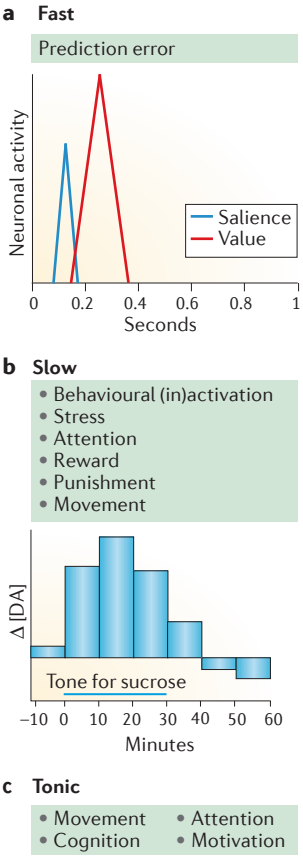
What does dopamine do? How can a single, simple chemical be involved in such diverse processes as movement (and its disruption in Parkinson disease), attention (disrupted in attention deficit hyperactivity disorder (ADHD)), cognition (disrupted in schizophrenia) and motivation[162]? Some answers may lie in the different timescales across which dopamine operates[1].

At the fastest, subsecond timescale (see the figure, part **a**), dopamine neurons show a phasic two-component prediction error response that — I argue in this article — transitions from salience and detection to reward value. This phasic dopamine response can be measured by electrophysiology and voltammetry, and constitutes a highly time specific neuronal signal that is capable of influencing other fast neuronal systems involved in rapid behavioural functions.

At an intermediate timescale (in the seconds to minutes range) a wide variety of behaviours and brain functions are associated with slower changes in dopamine levels that are revealed by dialysis (see the figure, part **b**) and voltammetry. These include behavioural activation or forced inactivation, stress, attention, reward-related behaviour, punishment and movement[156,163–167]. These changes in dopamine levels are unlikely to be driven by subsecond changes in dopamine impulses and thus may be unrelated to reward prediction error. Instead, they may be mediated by slower impulse changes in the seconds range or by presynaptic interactions[157,159,168]. Their function may be to homeostatically adjust the sensitivity of the fast, phasic dopamine reward responses[169].

At the slowest timescale, dopamine exerts an almost tonic influence on postsynaptic structures (see the figure, part **c**). Parkinson disease, ADHD and schizophrenia are associated with deficits in the tonic, finely regulated release of dopamine, which enables the functions of the postsynaptic neurons that mediate movement, cognition, attention and motivation. The effects of tonic dopamine reductions in Parkinson disease are partly remedied by pharmacological dopamine-receptor stimulation, which cannot reinstate phasic dopamine responses but can provide similar receptor occupation to the natural tonic dopamine levels. Thus, the deficits in Parkinson disease are not easily explained by reductions of phasic dopamine changes.

Taken together, dopamine neurotransmission, unlike that mediated by most other neurotransmitters, exerts different influences on neuronal processes and behaviour at different timescales.



Δ [DA], change in dopamine concentration. Part **b** of the figure is adapted with permission from REF. 163, Elsevier.

are involved in temporal prediction-error coding, the initial activation is sensitive to the time of stimulus occurrence and thus codes a temporal-event prediction error[33]. Stimuli of all sensory modalities are effective in eliciting this initial response component; those tested include loud sounds, intense light flashes, rapidly moving visual objects and liquids touching the mouth[2,9,16,19,20,35]. The unselective and multisensory nature of the initial activation thus corresponds to the large range and heterogeneous nature of potentially rewarding stimuli and objects present in the environment.

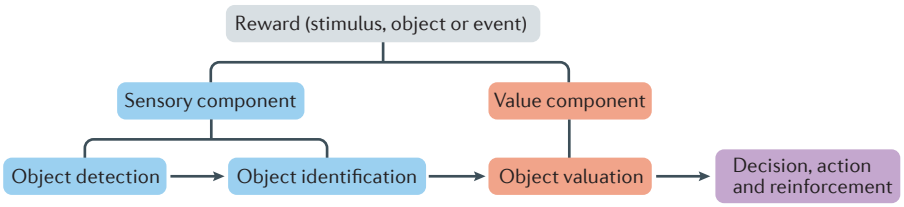


Figure 1 | **Reward components.** A reward is composed of sensory and value-related components. The sensory components have an impact on sensory receptors and neurons, and drive initial sensory processing that detects and subsequently identifies the reward. Reward value, which specifically reflects the positively motivating function of rewards, is processed only after the object has been identified. Value does not primarily reflect physical parameters but rather the brain's subjective assessment of the usefulness of the reward for survival and reproduction. These sequential processes result in decisions and actions, and drive reinforcement as a result of the experienced outcome. For further details, see REF. 114.

Sensitivity to stimulus characteristics. Several factors can enhance the initial dopamine activation. These factors may thus explain why the phasic dopamine response has sometimes been suggested to be involved in aversive processing[21,22,24,63] or to primarily reflect attention[64,65]. Stimuli of sufficient physical intensity elicit the initial dopamine activation in a graded manner (FIG. 3a), irrespective of their positive or negative value associations[25]. Physically weaker stimuli generate little or no initial dopamine activations[6,66]. Such physically weak stimuli will only induce a dopamine activation if they are rewards or are associated with rewards[4,6] (see below).

The context in which a stimulus is presented can also enhance the initial activation. Unrewarded stimuli elicit little dopamine activation when they are well separated from reward; however, the same unrewarded stimuli effectively elicit dopamine activations when presented in the same context in which the animal receives a reward[34] (FIG. 3b). Similarly, increasing the probability that the animal will receive a reward in a given experiment constitutes a more rewarded context and increases the incidence of dopamine activations to unrewarded stimuli[5,6,34]. Neurons might be primed by a rewarded context and initially process every unidentified stimulus as a potential reward until the opposite is proven. These neuronal context sensitivities may involve behavioural pseudoconditioning or higher-order context conditioning[67].

The physical resemblance of a stimulus to other stimuli known to be associated with a reward can enhance the initial dopamine activation through generalization[5,6,17,19,60]. For example, a visual stimulus that is paired with an aversive experience (air puff) leads only to small dopamine activations when it is randomly interspersed between presentations of an auditory reward-predicting stimulus; however, the same conditioned visual aversive stimulus induces substantial activations when it is alternated with a visual reward-predicting stimulus[19] (FIG. 3c). This is analogous to behavioural generalization, in which 'neutral' stimuli elicit similar reactions to physically similar target stimuli[68]. Thus, an otherwise ineffective, unrewarded or even aversive stimulus may drive the initial dopamine neuron activation as a result of generalization with a rewarded stimulus or a reward that requires only superficial identification. The enhancement might also affect the main dopamine response component if more-specific assessment of stimulus similarity requires some identification.

Novel stimuli, whether rewarded or not, can enhance dopamine activations. The response of dopamine neurons to an initial novel stimulus decreases with stimulus repetition[2,4] (FIG. 3d). However, novelty alone is ineffective in activating dopamine neurons: physically weak novel stimuli fail to induce a dopamine response[6]. Similarly to generalization, novelty detection involves comparison with an existing stimulus and thus requires identification, which is proposed to take place in the main dopamine response.

These finding show that the initial dopamine response is sensitive to factors that are related to potential reward availability. Stimuli of high intensity are potential rewards and should be prioritized for processing so as to not miss a reward. Stimuli occurring in reward contexts or resembling known rewards have a fair chance to be rewards themselves. Novel stimuli are potential rewards until their true value has been determined. They are more likely to be rewards than non-novel stimuli whose lack of reward value has already been established. Thus, even the earliest dopamine detection response is already geared towards rewards.

Salience. The factors that enhance the initial dopamine activation are closely associated with different forms of stimulus-driven salience. Stimulus intensity provides physical salience. Stimuli that become effective in rewarding contexts or through response generalization are motivationally salient because of these reward associations. The mechanism by which salience induces the initial dopamine response component may apply primarily to rewarding stimuli, because the negative value of stimuli — including punishers[25], negative reward prediction errors[3,5,69] and conditioned reward inhibitors[6] — is unlikely to induce dopamine activations. Novel or surprising stimuli are salient owing to their rare or unpredicted occurrence. The distinction between these different forms of salience might be important because they are thought to affect different aspects of behaviour, such as the identification of stimuli, valuation of rewards and processing of decisions, actions and reinforcement.

Benefits of initial unselective processing. At first sight, it might be assumed that an unselective response that occurs before value processing constitutes an inaccurate neuronal signal prone to inducing erroneous, disadvantageous behavioural reactions. One may wonder why such unselective responses have survived evolution.

However, although the initial activation appears unselective, it is (as outlined above) sensitive to modulation by several factors that are related to potential reward availability. Stimuli of high intensity should be prioritized for processing in order to not miss a reward. Stimuli occurring in reward contexts or resembling known rewards have a reasonable chance of being rewards themselves. Novel stimuli are potential

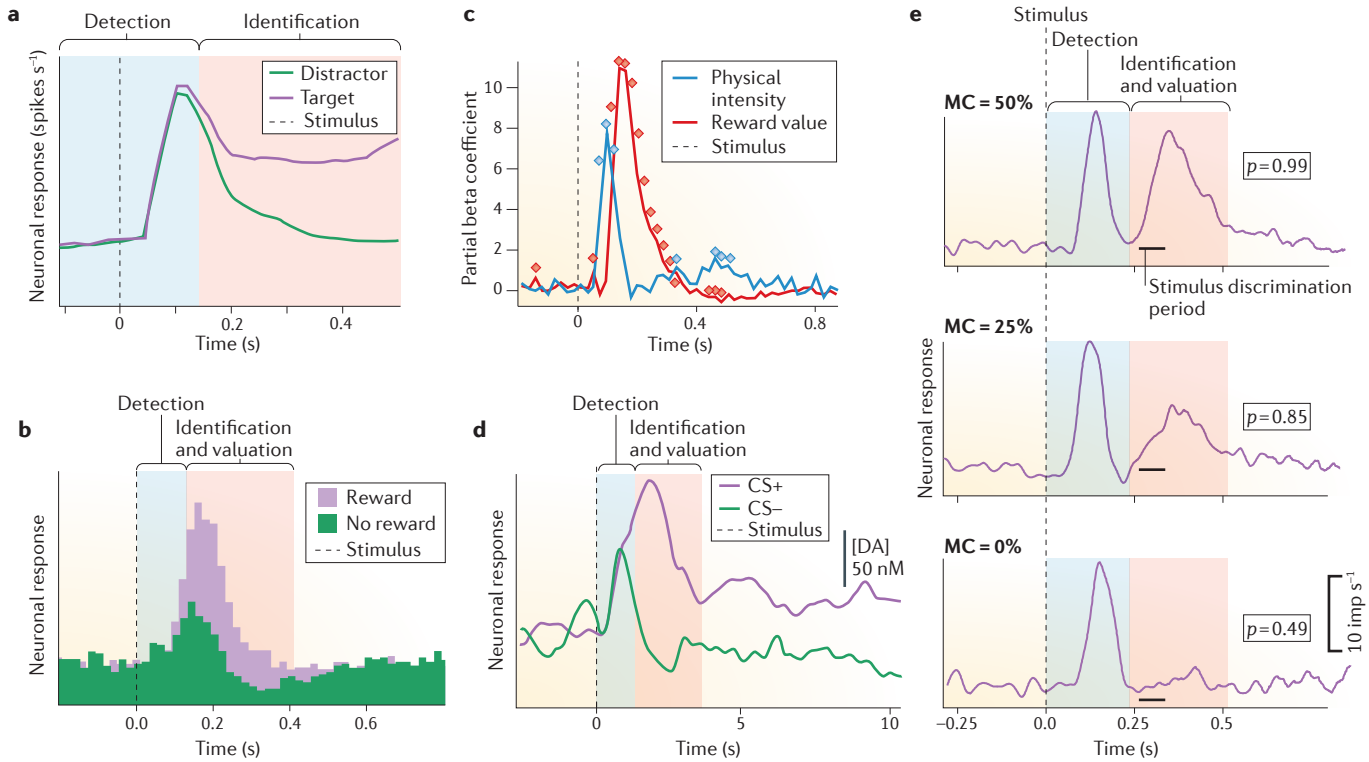


Figure 2 | **Sequential neuronal processing of stimulus and reward components. a** | Sequential processing in a cognitive system. The graph shows the time course of target discrimination in a monkey frontal eye field neuron during attentional selection. The animal's task was to distinguish between a target stimulus and a distractor. The neuron initially detects both stimuli indiscriminately (blue zone); only later does its response differentiate between the stimuli (red zone). **b** | Sequential dopamine processing of reward. The graph shows distinct, sequential components of the dopamine prediction-error response to conditioned stimuli predicting either non-reward or reward delivery. These responses reflect initial transient object detection, which is indiscriminate, and subsequent reward identification and valuation, which distinguishes between reward and no reward prediction. **c** | The components of the dopamine prediction-error response in part **b** that relate to detection and valuation can be distinguished by statistics. The partial beta (slope) coefficients of the double linear regression on physical intensity and reward value show distinct time courses, indicating the dynamic evolution from initial detection to subsequent valuation. **d** | Voltammetric dopamine responses in rat nucleus accumbens distinguish between a reward-predicting conditioned stimulus (CS+) and a non-reward-predicting conditioned stimulus (CS–). Again, the dopamine release comprises an initial indiscriminate detection component and a subsequent identification and value component. **e** | A more demanding random dot motion discrimination task reveals completely separated dopamine response components. Increasing motion coherence (MC) results in more accurate motion discrimination and thus higher reward probability (*p*). The initial, stereotyped, non-differential activation reflects stimulus detection and decreases back to baseline (blue zone); the subsequent separate, graded increase develops when the animal signals stimulus discrimination; it codes reward value (red zone), which in this case derives from reward probability[10,74]. [DA], dopamine concentration. Part **a** is adapted, with permission, from REF. 44, Proceedings of the National Academy of Sciences. Part **b** is adapted from REF. 6, republished with permission of Society for Neuroscience, from Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm, Tobler, P. N., Dickinson, A. & Schultz, W., **23** (32), 2003; permission conveyed through Copyright Clearance Center, Inc. Part **c** is adapted from REF. 25, republished with permission of Society for Neuroscience, from Multiphasic temporal dynamics in responses of midbrain dopamine neurons to appetitive and aversive stimuli, Fiorillo, C. D., Song, M. R. & Yun, S. R., **33** (11), 2013; permission conveyed through Copyright Clearance Center, Inc. Part **d** is from REF. 60, Nature Publishing Group. Part **e** is adapted from REF. 33, republished with permission of Society for Neuroscience, from Temporally extended dopamine responses to perceptually demanding reward-predictive stimuli, Nomoto, K., Schultz, W., Watanabe, T. & Sakagami, M., **30** (32), 2010; permission conveyed through Copyright Clearance Center, Inc.

rewards until their true value has been determined and are thus more likely to be rewards than unrewarded non-novel stimuli whose absent lack of reward value has already been established. Thus, the earliest dopamine detection response is already tuned towards reward. The wide, multisensory sensitivity of the response, on the other hand, would facilitate the detection of a maximal number of potential reward objects that should be attended to and identified quickly to avoid missing a reward.

Through stimulus-driven salience, the early dopamine activation component might serve to transiently enhance the ability of rewards to induce learning and action. This mechanism is formalized in the attentional Pearce–Hall learning rule[70], in which surprise salience derived from reward prediction errors enhances the learning rate, as do physical and motivational salience. Thus, higher salience would induce faster learning, whereas lower salience would result in smaller and more fine-tuned learning steps.

During action generation, stimulus-driven salience and top-down attention are known to enhance the neuronal processing of sensory stimuli and resulting behavioural responses[36–40], and salience processing might also underlie the enhancing effects of reward on the accuracy of spatial target detection[71]. Similarly, by conveying physical, novelty and motivational salience, the initial dopamine response component might boost and sharpen subsequent reward value processing and ultimately increase action accuracy.

This notion mirrors an earlier idea about sequential processing of global and finer stimulus categories in the inferotemporal cortex: that is, "… global information could be used as a 'header' to prepare destination areas for receiving more detailed information" (REF. 48).

Through its rapid detection of potential rewards, the initial dopamine activation might provide a temporal advantage by inducing early preparatory processes that lead to faster behavioural reactions towards important stimuli. As the response occurs more quickly than most behavioural reactions, there would still be time to cancel the behavioural initiation process if the subsequent valuation of a stimulus labels it as worthless or damaging. Thus the lower accuracy of the initial response would not compromise behavioural actions. A temporal gain of several tens of milliseconds, together with attentional response enhancement, might be important in competitive situations that require rapid behavioural reactions. As Darwin said[72], in the long run of evolution, any small edge will ultimately result in an advantage.

I suggest that, through the early detection of salient stimuli, the initial dopamine response component affords a gain in speed and processing without substantially compromising accuracy, thus supporting the function of the phasic dopamine reward signal.

The main component: valuation

The dopamine reward prediction-error response evolves from the initial unselective stimulus detection and gradually sharpens into the increasingly specific identification and valuation of the stimulus[25,33] (FIG. 2c,e). This later component — rather than the initial detection activation described above — defines the function of the phasic dopamine response and reflects the evolving neuronal processing that is required to fully appreciate the value of the stimulus. Higher-than-predicted rewards (generating positive prediction errors) elicit brief dopamine activations, lower-than-predicted rewards (generating negative prediction errors) induce decreases in activity ('depressions'), and accurately predicted rewards do not change the activity. These responses constitute biological implementations of the crucial error term for reinforcement learning according to the Rescorla–Wagner model and temporal difference reinforcement models[73]; such a signal is appropriate for mediating learning and updating of reward predictions for approach behaviour and economic decisions[4].

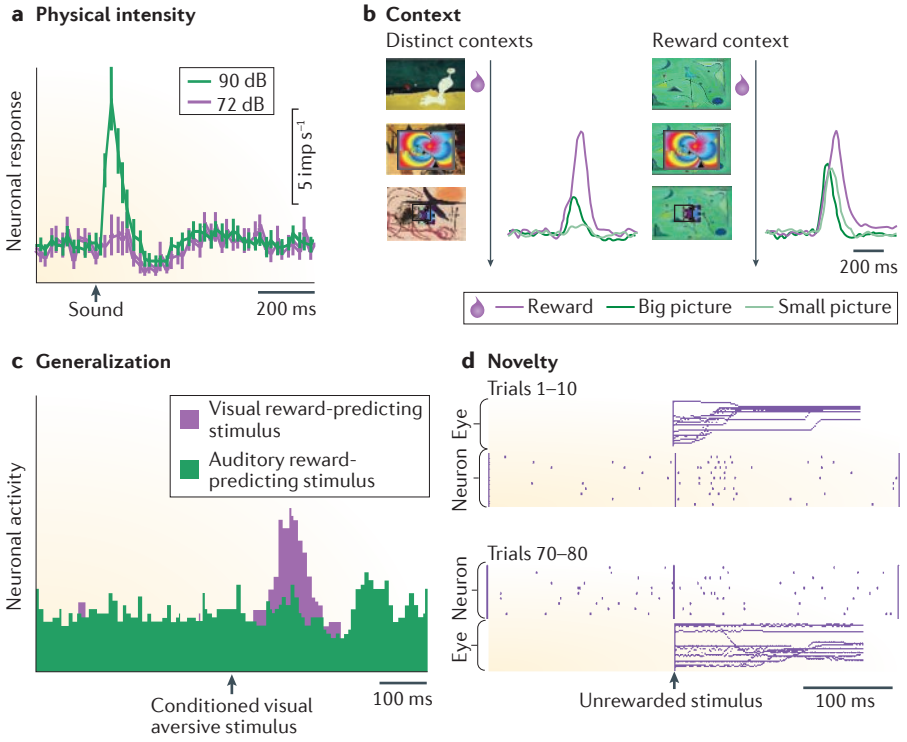


Figure 3 | **Factors influencing the initial dopamine activation. a** | Influence of physical intensity. In the example shown, stronger (yet non-aversive) sounds generate higher initial dopamine responses than weaker sounds. **b** | Influence of reward context. The left graph shows dopamine responses to the presentation of either a reward or one of two unrewarded pictures of different sizes in an experiment in which the contexts in which each is presented are distinct. The right graph shows that the dopamine neuron responses to the pictures are much more substantial when they are presented in the same context as the reward (that is, when the presentations occur in one common trial block with a common background picture and when the liquid spout is always present). **c** | Influence of reward generalization between stimuli that share physical characteristics. In the example shown, a conditioned visual aversive stimulus activates 16% of dopamine neurons when it is pseudorandomly interspersed with an auditory reward-predicting stimulus. However, the same visual aversive stimulus activates 65% of dopamine neurons when the alternating reward-predicting stimulus is also visual. **d** | Influence of stimulus novelty. The graph shows activity in a single dopamine neuron during an experiment in which an animal is repeatedly presented with a novel, unrewarded stimulus (rapid vertical opening of the door of an empty box). Dots representing action potentials are plotted in sequential trials from top down. Activation was substantial in the first ten trials. However, 60 trials later, the same dopamine neuron shows a diminished response. Horizontal eye movement traces displayed above and below the neuronal rasters illustrate the change in behaviour that correlates with decreasing novelty (less eye movement towards the stimulus; *y*-axis represents eye position to the right). Part **a** is adapted from REF. 25, republished with permission of Society for Neuroscience, from Multiphasic temporal dynamics in responses of midbrain dopamine neurons to appetitive and aversive stimuli, Fiorillo, C. D., Song, M. R. & Yun, S. R., **33** (11), 2013; permission conveyed through Copyright Clearance Center, Inc. Part **b** is adapted with permission from REF. 34, Elsevier. Part **c** is from REF. 19, Nature Publishing Group. Part **d** is adapted with permission from REF. 2, American Physiological Society.

Subjective reward value. A reward's value cannot be measured directly but is estimated from observable behavioural choices. Thus, value is a theoretical construct that is used to explain learning and decision-making. In being defined by the individual's needs and behaviour, value is necessarily subjective. The construction of reward value involves brain mechanisms that include those mediated by dopamine neurons in the substantia nigra and VTA.

The material components of rewards are often difficult to assess and, most importantly, do not fully define their subjective value. Although phasic dopamine responses increase with the expected reward value (the summed product of the amounts and probabilities of the rewards, an objective, physical reward measure)[7,10,59,74,75], it is intrinsically unclear whether they code objective or subjective reward value. One way to resolve the issue is to examine

choices between rewards that are objectively equal. When a monkey chooses between identical amounts of blackcurrant juice and orange juice and shows a preference for blackcurrant juice, it can be inferred that the blackcurrant juice has a higher subjective value to the monkey than the orange juice[9]. Similarly, preferences for risky over safe rewards with identical mean volumes suggest increased subjective value due to the risk. Even with a larger range of safe and risky liquid and food rewards, monkeys show well-ranked choice preferences. The animals' preferences satisfy transitivity (when preferring reward A over reward B, and reward B over reward C, they also prefer reward A over reward C), which suggests meaningful rather than chance behaviour[9]. Another way to estimate subjective value is to have an individual choose between the reward in question and a common reference reward; the psychophysically determined amount of the reference reward at which the individual becomes equally likely to select either option (choice indifference) indicates the subjective value of the reward in question. It is measured in physical units of the common reference reward (known as a 'common currency', such as millilitres of blackcurrant juice). Dopamine neurons show higher activations in response to the preferred juice, and their activity also correlates with the indifference amounts in choices between risky and safe rewards, indicating that the neurons consistently code subjective rather than objective reward value[9]. A relationship between dopamine neuron activity and subjective value can also be seen when the reward value is reduced by the addition of an aversive liquid[25] (Supplementary information S2 (box)).

Temporal discounting is another way to dissociate subjective value from objective value. Rewards lose subjective value after delays, even when they remain physically unchanged[76]. Correspondingly, although the initial component of the dopamine response to a stimulus that predicts a delayed reward stays almost constant, the second component of the dopamine response decreases as the delay increases (FIG. 4a) and follows closely the hyperbolic decay of subjective value assessed by measuring behaviour[61,62]. Similar temporal discounting can be observed in dopamine voltammetry[77].

These findings indicate that the second dopamine response component codes the subjective value of different types of rewards, risky rewards, composite rewards and delayed rewards.

Utility. Economic utility provides the most constrained, principled and conceptualized definition of the subjective value of rewards. It is the result of 300 years of mathematical and economic theory that incorporates basic ideas about the acquisition and exchange of goods[78–81]. Additions, such as prospect theory[82], provide important extensions but do not question the fundamental role of utility. Probably the most important potential that utility has for neuroscience lies in the assumption that utility provides an internal, private metric for subjective reward value[83]. Utility as an internal value reflects individual choice preferences and constitutes a mathematical function of objective, physical reward amount[84,85]. This function, $u(x)$, is usually, but not necessarily, nonlinear over x. By contrast, the subjective value derived from choice preferences and indifference points described above provides a measure in physical units (such as millilitres, pounds or dollars of a common reference reward) but does not tell us how much a physical unit of the reference reward is privately worth to the individual decision maker. By estimating utility, we could obtain such a private measure.

A private, internal metric for reward value would allow researchers to establish a neuronal value function. This function

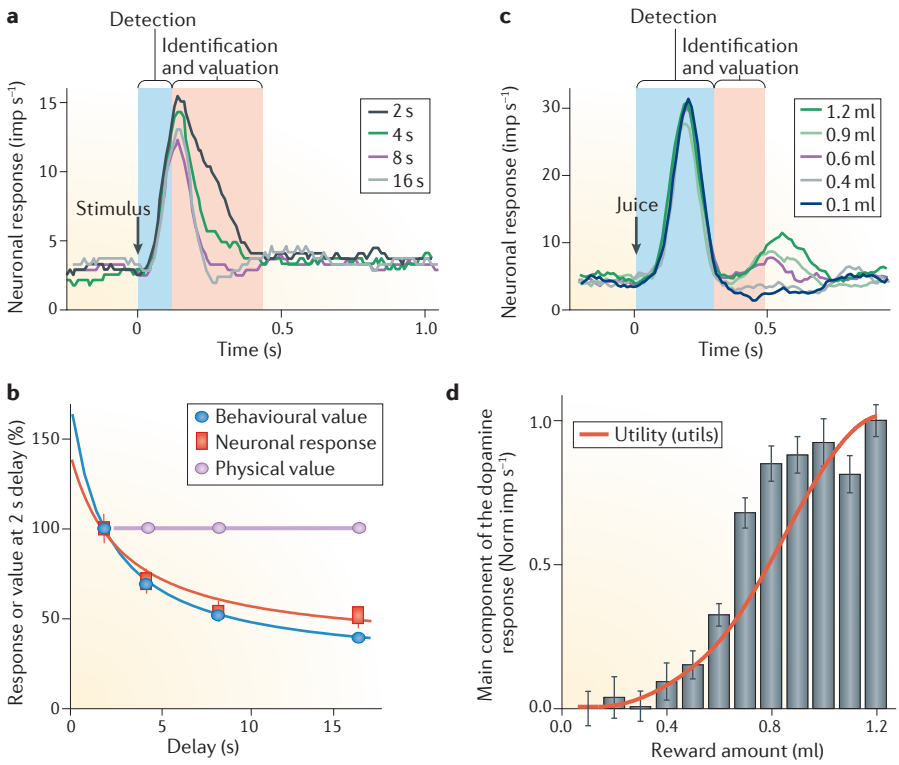


Figure 4 | **Subjective value and utility coding by the main dopamine response component.** **a** | Temporal discounting of a reward's subjective value is reflected in the dopamine response. The figure shows the averaged activity from 54 dopamine neurons responding to 4 stimuli predicting the same amount of a reward that is delivered after delays of 2, 4, 8 or 16 seconds. The stimulus response is reduced as the delay before receiving the reward increases, suggesting that the value coding by these neurons is subjective and decreases with the delay after which the reward is delivered. The value-related response decrement occurs primarily in the later, main response component (red zone). **b** | Corresponding temporal discounting in both the behavioural response and the main dopamine response component reveals subjective value coding. **c** | Monotonic increase of the main dopamine response component (red zone) when an animal is provided with an increasing amount of an unpredicted juice reward. By contrast, the initial response component (blue zone) is unselective and reflects only the detection of juice flow onset. **d** | The dopamine reward prediction-error signal codes formal economic utility. For the experiment described in **c**, the red line shows the utility function as estimated from behavioural choices. The grey bars show a nonlinear increase in the main component of the dopamine response (red zone in **c**) to juice reward, reflecting the positive prediction error generated by its unpredicted delivery. Parts **a** and **b** are adapted from REF. 61, republished with permission of Society for Neuroscience, from Influence of reward delays on responses of dopamine neurons, Kobayashi, S. & Schultz, W., **28** (31), 2008; permission conveyed through Copyright Clearance Center, Inc. Parts **c** and **d** are adapted with permission from REF. 35, Elsevier.

would relate the frequency of action potentials to the internal reward value that matters to the decision maker. The number of action potentials during one well-defined period — for example, for 200 ms after a stimulus — would quantitate how much the reward is valued by the monkey's neurons, and thus how much it is worth privately to the monkey.

Economic theory suggests that numeric estimates of utility can be obtained experimentally in choices involving risky rewards[81,86,87]. The most simple and confound-free form of risk can be tested by using equiprobable gambles in which a small and a large reward occur with equal probability of $p = 0.5$ (REF. 88). To obtain utility functions, we can use specifically structured choices between such gambles and variable safe (riskless) rewards (known as a 'fractile procedure' (REFS 89,90)) and estimate 'certainty equivalents' — the amount of the safe reward that is required for the animal to select this reward as often as the gamble. All certainty equivalents are then used to construct the utility function[89,90]. In such a procedure, a monkey's choices reveal nonlinear utility[35]. When the amount of reward is low, the curvature of the utility function is convex (progressively increasing), which indicates that monkeys tend to be risk-seeking when the stakes are low, as previously observed in other monkeys[91,92] and humans[93–95]. The utility function becomes concave with higher liquid amounts (progressively flattening), which is consistent with the risk avoidance seen in traditional utility functions[85,96]. The convex–concave shape is similar to the inflected utility functions that have been modelled for humans[97,98]. Thus, it is possible to experimentally estimate numeric economic utility functions in monkeys that are suitable for mathematically valid correlations with numeric neuronal reward responses.

The most basic and straightforward method to elicit a dopamine reward prediction-error response involves delivery of an unpredicted reward (juice) at a spout. Reward amount is defined by the duration of juice flow out of the spout. The start of liquid flow thus indicates the onset of reward delivery, but its final amount becomes only appreciable when the liquid flow terminates. Dopamine neurons show an initial, uniform detection response to liquid-flow onset that is unaffected by reward amount, and a second response component that increases monotonically with the final amount and signals value[35] (FIG. 4c). Importantly, the second dopamine response component increases only gradually when juice amounts are small, then more steeply with intermediate juice amounts, and then more gradually again with higher amounts, following the nonlinear curvature of the utility function (FIG. 4d). Thus, the main, fully evolved dopamine reward prediction-error response correlates with numeric, quantitative utility.

To truly determine whether the dopamine signal codes numeric utility, experimental tests should use well-defined gambles that satisfy the conditions for utility[81], rather than unpredicted rewards in which the risk is poorly defined. For example, when using binary, equiprobable gambles with well-defined and identical variance risk, the higher of the two gamble outcomes elicits non-monotonically varying positive dopamine reward prediction-error responses that reflect the nonlinear shape of the utility function[35]. These responses with well-defined gambles match well the responses obtained with free rewards and demonstrate a neuronal signal for formal numeric utility, as stringently defined by economic choice theory.

Because the fully developed main response component codes utility, the phasic dopamine reward prediction-error response can be specified as a utility prediction-error signal. All other factors that might affect utility — including risk, delay and effort cost — were held constant in these experiments; therefore the utility signal reflects income utility rather than net-benefit utility. Although economists consider utility to be a hypothetical construct that explains decision-making but lacks a physical existence, dopamine responses seem to represent a physical correlate for utility.

Downstream influences

Correct behaviour based on late component. As described above, the initial dopamine activation is transient, and it is likely that the accurate value representation of the second dopamine response component can quickly compensate for the initial unselectivity.

Glossary

Behavioural pseudoconditioning
A situation in which the context (environment) is paired, through Pavlovian conditioning, to a reinforcer that is present in this environment. Any stimulus occurring in this context thus reflects the same association, without being explicitly paired with the reinforcer. Pseudoconditioning endows an unpaired stimulus with motivational value.

Context conditioning
An association between a specific stimulus (for example, a reward or punisher) and a context (for example, an environment, including all stimuli except the specific explicit stimulus).

Down states
Neuronal membrane states that are defined by hyperpolarized membrane potentials and very little firing.

Economic utility
A mathematical, usually nonlinear function that derives the internal subjective reward value u from the objective value x. Utility is the fundamental variable that decision-makers maximize in rational economic choices between differently valued options.

Hebbian learning
A cellular mechanism of learning, proposed by Donald Hebb, according to which the connection between a presynaptic and a postsynaptic cell is strengthened if the presynaptic cell is successful in activating a postsynaptic cell.

Motivational salience
The ability of a stimulus to elicit attention due to its positive (reward) or negative (punishment) motivational value. Motivational salience is common to reward and punishment.

Novelty salience
The ability of a stimulus to elicit attention due to its novelty.

Physical salience
The ability of a stimulus to elicit attention by standing out, due to its physical intensity or conspicuousness.

Rescorla–Wagner model
The prime error-driven reinforcement model for Pavlovian conditioning, in which the prediction error (reward or punishment outcome minus current prediction) is multiplied by a learning factor and added to the current prediction to result in an updated prediction.

Surprise salience
The ability of a stimulus to elicit attention due to its unexpectedness.

Temporal difference reinforcement models
A family of non-trial-based reinforcement learning models in which the difference between the expected and actual values of a particular state (prediction error) in a sequence of behaviours is used as a teaching signal to facilitate the acquisition of associative rules or policies to direct future behaviour. Temporal difference learning extends Rescorla–Wagner-type reinforcement models to real time and higher-order reinforcers.

Up states
Neuronal membrane states that are defined by relatively depolarized membrane potentials and lots of action potential firing.

Visual search task
An experimental paradigm in which subjects are asked to detect a 'target' item (for example, a red dot) among an array of distractor items (for example, many green dots).

Voltammetry
An electrochemical measurement of oxidation-reduction currents across a range of imposed voltages, used in neuroscience for assessing concentrations of specific molecules, such as dopamine.

The second response component persists throughout the resulting behaviour until the reward is received, as revealed by the graded positive prediction-error response to reward delivery[5,33]. This prediction-error response is large with intermediate reward probability, which generates an intermediate value prediction[11], and decreases progressively when higher reward-probability predictions lead to less-surprising rewards[33] (FIG. 5a). Similar persistence of the second response component is apparent in temporal discounting experiments, in which longer delays associated with lower values result in higher prediction-error responses to the full reward[61]. Thus, the initial activation lasts only until the subsequent value component conveys the accurate reward value information, which remains present until the reward occurs (FIG. 5b), covering the entire period of planning, initiation, execution and outcome of action. In this way, the initial dopamine signal can be beneficial without unfocusing or misleading the behaviour.

Both the initial, unselective detection information and the specific value information present in the second dopamine response component are propagated to downstream neurons, many of which show similar multi-component responses (FIG. 2a; Supplementary information S1 (table)). Thus, despite the transient, inaccurate, first dopamine response component, the quickly following second response component would allow neurons to distinguish rewards from non-rewards early enough to influence behavioural reactions. This is similar to the multi-component response patterns observed in other sensory, cognitive and reward neurons (FIG. 2a; Supplementary information S1 (table)). Thus, the two-component dopamine signalling mechanism combines speed, efficiency and accuracy in reward processing.

The two-component mechanism operates on a narrow timescale that requires unaltered, precise processing in the 10 ms range. Any changes in the temporal structure of the phasic dopamine response might disturb the valuation and lead to impaired postsynaptic processing of reward information. Thus, stimulant drugs, which are known to prolong increases in dopamine concentration[99], might extend the effects of the initial activation component so that the dopamine surge overlaps with the second, value response and thus generates a false value signal for postsynaptic neurons. This mechanism may contribute to stimulant drug addiction and behavioural alterations in psychiatric disorders (BOX 2).

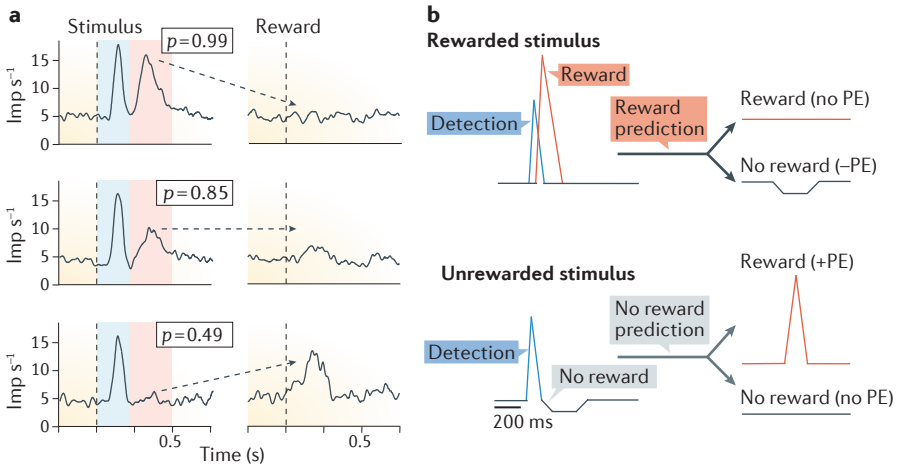


Figure 5 | **Persistent accurate value representation.** The reward prediction error signalled by the main dopamine response component following a conditioned stimulus remains present until the time of reward. **a** | The graphs illustrate persistent reward representation in a random dot motion discrimination task in which distinct dopamine response components can initially be observed (blue and red zones). The reward prediction-error response subsequently decreases in a manner that correlates with increasing reward probability (right), suggesting that a neuronal representation of reward value persists after the onset of the value response component of the dopamine response. **b** | Schematics showing how an accurate reward-value representation may persist until the reward is received. As shown in the top panel, after a rewarded stimulus generates a detection response that develops into a full reward-value activation, reward delivery, which induces no prediction error (no PE), elicits no dopamine response; by contrast, reward omission, generating a negative prediction error (–PE), induces a dopamine depression. However, as shown at in the bottom panel, after an unrewarded stimulus generates a detection response that develops into a dopamine depression, a surprising reward elicits a positive prediction error (+PE) and dopamine activation, whereas no reward fails to generate a prediction error and dopamine response. Thus, the dopamine responses at the time of reward probe the reward prediction that exists at that moment. This proposed mechanism expands on previous suggestions that have not taken into account the two dopamine response components[4,5]. Part **a** is adapted from REF. 33, republished with permission of Society for Neuroscience, from Temporally extended dopamine responses to perceptually demanding reward-predictive stimuli, Nomoto, K., Schultz, W., Watanabe, T. & Sakagami, M. 30 (32), 2010; permission conveyed through Copyright Clearance Center, Inc.

Updating predictions and decision variables. The main, utility prediction-error response component (FIG. 4c,d) might provide a suitable reinforcement signal for updating neuronal utility signals[35]. The underlying mechanism may consist of dopamine-dependent plasticity in postsynaptic striatal and cortical neurons[100–106] and involve a three-factor Hebbian learning mechanism with input, output and dopamine signals[4]. A positive prediction error would enhance behaviour-related neuronal activity that resulted in a reward, whereas a negative prediction error would reduce neuronal activity and thus disfavour behaviour associated with a lower reward. Indeed, optogenetic activation of midbrain dopamine neurons induces learning of place preference, nose poking and lever pressing in rodents[107–112]. These learning effects depend on dopamine D1 receptors mediating long-term potentiation (LTP) and long-term depression (LTD) in striatal neurons[103,104]. The teaching effects of phasic dopamine responses may differ between striatal neuron types because optogenetic stimulation of neurons expressing D1 or D2 dopamine receptors induces learning of behavioural preferences and dispreferences, respectively[113].

Dopamine-dependent plasticity may affect neuronal populations whose activity characteristics comply with specific formalisms of competitive decision models, such as object value, action value, chosen value and their derivatives[114]. The dopamine reward prediction-error response conforms to the formalism of chosen value, reflecting the value of the object or action that is chosen by the animal[115]; it might be driven by inputs from striatal and cortical reward neurons[116–119]. The output of the dopamine signal may affect object-value and action-value coding neurons in the striatum and frontal cortex whose particular signals are suitable for competitive decision processes[116,120–124]. Thus, the main, utility prediction-error response component

of dopamine neurons shows appropriate characteristics for inducing learning and value updating in predictions and decision processes, although formal utility coding remains to be established in postsynaptic striatal and cortical neurons.

Immediate influences on behaviour. Phasic dopamine signals may also affect behavioural reactions through an immediate focusing effect on cortico-striatal connections. Through such a mechanism, weak afferent activity might be filtered out, and only information from the most active inputs might be passed on to postsynaptic striatal neurons[125–127]. Optogenetic activation of mouse midbrain dopamine neurons elicits immediate behavioural actions, including contralateral rotation and locomotion[110]. Correspondingly, reduction of dopamine bursting activity through NMDA-receptor knockout prolongs reaction time[128], and dopamine depletion reduces learned neuronal responses in striatum[129]. These behavioural findings might be explained by a neuronal mechanism in which dopamine prolongs transitions to excitatory membrane up states in D1 receptor-expressing striatal direct pathway neurons[130], but reduces membrane up states and prolongs membrane down states in D2 receptor-expressing striatal indirect pathway neurons[131]; both effects conceivably facilitate behavioural reactions. This link to behaviour might be confirmed by the effects of optogenetic stimulation of striatal neurons; stimulation of neurons expressing dopamine D1 receptors increases behavioural choices towards contralateral nose-poke targets, whereas stimulation of D2 receptor-expressing striatal neurons increases ipsilateral choices (or contralateral dispreferences), suggesting that these two populations of neurons have differential effects on the coding of action value[132]. These effects would reflect value influences from the second dopamine response component on neurons in the striatum and frontal cortex.

Addressing open issues

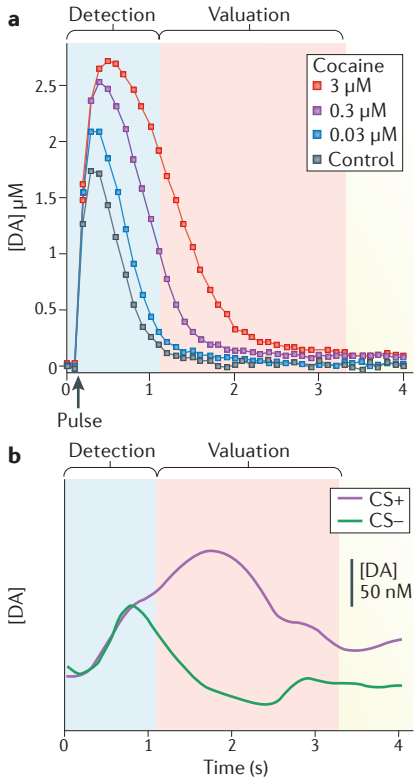
Unlikely aversive activation. Recognition of the two-component dopamine response structure and the influence of physical impact, reward context, reward generalization and novelty on the initial response component may help us to scrutinize the impact of aversive stimuli on dopamine neurons. There is a long history of 'aversive' dopamine activations seen in electrophysiological and voltammetric studies on awake and anaesthetized animals[15,17,19,21–24,63,133]. However, recent reports that distinguished between different stimulus components of punishers suggest that the physical impact of punishers is the major determinant for dopamine activations, and the psychophysically assessed aversive nature of the punishers did not explain the recorded dopamine activations[25–27]. Thus, it is possible that the previously reported activations of dopamine neurons by punishers[15,17,19,21–24,134] may have been due to physical impact of the tested stimuli, rather than true aversiveness. Reward context and reward generalization[19] may have had an additional facilitating influence on these dopamine activations, which would merit further tests.

Some of the dopamine activations observed in response to aversive stimuli might reflect outright reward processes. The end of exposure to a punisher can be rewarding because of the relief it provides[135]. Correspondingly, aversive stimuli induce delayed, post-stimulus dopamine release in the shell of the nucleus accumbens[63,133], the magnitude of which predicts successful punishment avoidance[136]. These responses are distinct from immediate dopamine changes in the nucleus accumbens core[63,133] (which may reflect physical impact, reward generalization or reward context).

Optogenetic stimulation of habenula inputs to dopamine neurons induces behavioural place preference changes[137] that are compatible with both aversive and reward accounts of dopamine function, whereas electrical habenula stimulation affects preferences in a way that is compatible with a

Box 2 | Altered dopamine responses and drug addiction

Recognition of the two-component nature of phasic dopamine responses may allow us to speculate on possible mechanisms of drug addiction. The response structure described here must be precisely transmitted to postsynaptic neurons to preserve the specific information present in each component. Therefore, as is the case for any sophisticated mechanism, it is vulnerable to alterations. Cocaine and other psychostimulants are known to enhance and prolong impulse-dependent changes in dopamine neuron firing (see the figure, part **a**). When a sufficiently strong but motivationally neutral stimulus induces the unselective initial response component, the effects of dopamine uptake inhibition by cocaine prolong the dopamine increase beyond the initial response period (blue zone) and might carry it into the subsequent period of the main response component that signals reward value (red zone). Part **b** of the figure shows the time courses of the changes in dopamine concentration derived from the initial response component to an unrewarded conditioned stimulus (CS−; blue zone) and the continuation of this response into the value component in the case of a reward-predicting conditioned stimulus (CS+; red zone), as shown also in FIG. 2d with a compressed time course. With the prolongation of the first, unselective dopamine response component by cocaine, an unrewarded stimulus would appear to postsynaptic neurons as a reward rather than an undefined stimulus and induce erroneous learning, approach and decision-making. In addition, the cocaine-induced blockade of dopamine uptake may lead to a supranatural dopamine boost that is unfiltered by sensory receptors and is likely to induce exaggerated postsynaptic plasticity effects. Thus, the transformation of the unselective dopamine detection response into a false reward-value response may constitute a possible mechanism contributing to psychostimulant addiction. A similar mechanism may apply to any psychotic or attentional disorder in which reduced neuronal processing precision might compromise critical transitions between the dopamine-response components and induce false value messages and wrong environmental associations.



[DA], dopamine concentration. Figure part **a** is adapted with permission from REF. 99, Jones, S. R., Garris, P. A. & Wightman, R. M. Different effects of cocaine and nomifensine on dopamine uptake in the caudate-putamen and nucleus accumbens. *J. Pharmacol. Exp. Ther.* **274**, 396–403 (1995). Figure part **b** is adapted from REF. 60, Nature Publishing Group.

rewarding dopamine influence[138]. In support of the latter result, habenula stimulation elicits inhibitory electrophysiological dopamine responses[139–141] through an intermediate reticular nucleus[138], rather than the excitatory responses required for a postulated aversive dopamine influence on place preferences[137].

The phasic dopamine activations by aversive stimuli seem to constitute the initial, unselective dopamine response component driven by physical impact[25], and possibly boosted by reward context and reward generalization[19], rather than reflecting a straightforward aversive response; however, the existence of some truly aversive dopamine activations can never be completely ruled out. For a more extensive discussion of the issue, see Supplementary information S2 (box).

Salience. Recognition of the two-component structure of phasic dopamine responses may also resolve earlier controversies, which suggested an attentional rather than rewarding role of the dopamine prediction-error signal[64,65] on the basis of activations following exposure to physically salient stimuli[16,20] and punishers[24]. It might seem as if dopamine neurons were involved in driving attention as a result of physical salience if they were tested in the complete absence of rewards. Without any rewards, reward prediction errors would not occur, and the second, value response component would be absent: an initial, salience response component could then be interpreted as the principal dopamine response. Testing with reward should reveal the complete dopamine signal.

A role for the dopamine prediction-error response in mediating attention derived from motivational salience (which is common to reward and punishers) would be confirmed if dopamine neurons were shown to exhibit the same (activating) response to rewards and punishers[24]. However, the improbability of truly aversive dopamine activations[25] make this interpretation unlikely. This suggests that theories of dopamine function based on motivational salience[142] present an incomplete account of phasic dopamine function.

The incentive-salience hypothesis captures a different form of salience: one that is associated with conditioned stimuli for rewards, not punishers. The hypothesis postulates that dopamine neurons function in approach behaviour rather than in learning[143]. Its experimental basis is the dopamine antagonist-induced deficit in approach behaviour, but not learning, that is seen in so-called sign-tracking rat strains[144]. However, specific learning deficits do occur in mice in which NMDA-receptor knockout in dopamine neurons results in reduced dopamine burst activity, suggesting a connection between phasic dopamine activity and the learning of specific tasks[128]. Thus, the evidence for a strictly differential dopamine role in approach behaviour versus learning is at best inconclusive. The incentive-salience hypothesis and the prediction-error account are difficult to compare and might not be mutually exclusive: incentive salience concerns dopamine's influences on behaviour, whereas prediction-error coding concerns the properties of the dopamine prediction-error signal itself, which can have many functions. Indeed, a prediction-error signal can support both learning and efficient performance[145].

The two-component response structure thus provides a viable account of phasic salience signalling by dopamine neurons. The response to salient stimuli does not represent the full coding potential of dopamine neurons: rather, it constitutes only the initial, undifferentiated component of the dopamine reward prediction-error signal.

Dopamine diversity. The functional interpretation of the phasic dopamine response may shed light on the currently debated diversity of dopamine mechanisms, which focus on differences between phasic dopamine prediction error responses[26] or concern variations of dopamine functions akin to those in other brain systems[146].

The least diverse of dopamine's functions are the phasic electrophysiological dopamine responses, whose latency, duration and type of information coding varies only in a graded (not categorical) manner between neurons. In this respect, dopamine neurons contrast strongly with non-dopamine neurons in the striatum and frontal cortex, which show wide varieties of activations and depressions at different time points before and after different stimuli and behavioural events[114,147–149]. Dopamine neurons in medial and lateral, or dorsal and ventral, midbrain positions do show graded differences in responses to rewarding and aversive stimuli[2,23,24,26,33]. However, because aversive responses may primarily concern the initial response component[25], it is possible that their regional distributions might be explained by varying sensitivities to stimulus intensity, reward context and reward generalization (FIG. 3). Thus, a strong activation in response to the physical impact of an aversive stimulus in particularly sensitive neurons may supersede a depression reflecting the negative value, and this might appear as a categorical difference in aversive and motivational salience coding[15,23,24,134]. Furthermore, the phasic dopamine response has been suggested to process cognitive signals in working memory and visual search tasks[150]. However, on closer inspection, dopamine responses in such elaborate tasks constitute standard reward prediction-error signals[2,33,59,66,69,75,151–153]. Altogether, the phasic dopamine reward signal is remarkably similar across neurons and so far seems to show graded rather than strictly categorical differences.

In contrast to the phasic dopamine reward prediction-error signal, all other aspects of dopamine function are diverse, including dopamine neuron morphology, electrophysiology, neurochemistry, connectivity and contribution to behaviour. For example, various subsets of dopamine neurons (1–44%) show slow, sluggish and diverse changes in the seconds time range that are inconsistently related to various aspects of behavioural task engagement and reactivity with stimulus-triggered or self-initiated arm or mouth movements[11,12,14,18]; these changes fail to occur with more-controlled arm, mouth and eye movements in a considerable range of studies[2,5,9,13,151,152,154] (thus, dopamine impulse activity does not seem to reliably code the movement processes that are deficient in Parkinson disease). In addition, a well-controlled slow activation reflects reward risk during the stimulus–reward interval in about 30% of dopamine neurons[10]. Voltammetric studies in rats show similarly slow striatal dopamine increases with movements towards cocaine levers[155] and with reward proximity and value[156]; the release might derive from the above-mentioned slow impulse activities or from presynaptic influences of non-dopamine terminals[157]. Impulse-dependent dopamine release is also heterogeneous and sensitive to varying degrees of modulation. For example, risky cues induce differential voltammetric dopamine responses in the nucleus accumbens core but non-differential responses in the shell[158]. Further, inhomogeneous neuronal release of acetylcholine, glutamate and substance P affects dopamine release[159–161], which, together with the heterogeneous striosome-matrix compartments, would result in diverse dopamine release. More than with most other neurons, the time course of dopamine neuron function varies vastly[1] (BOX 1). It is unclear whether all dopamine neurons, or only specific subgroups, participate in all of these functions. Furthermore, like most other neurons, dopamine neurons vary in morphology, connections, neurotransmitter

colocalization, receptor location, neurotransmitter sensitivity and membrane channels. Also, the ultimate function of dopamine neurons on behaviour differs depending on the anatomical projection and function of the postsynaptic neurons they are influencing.

The phasic electrophysiological dopamine reward signal is remarkably similar across neurons and shows graded rather than strictly categorical differences. It affects diverse downstream dopamine and non-dopamine mechanisms, which together makes dopamine function as diverse as the function of other neuronal systems. In this way, the homogeneous phasic dopamine signal influences other brain structures with heterogeneous functions and thus exerts differential and specific influences on behaviour.

Conclusions and future directions

Recent research has revealed the interesting and beneficial component structure of the phasic dopamine reward prediction-error signal during its dynamic evolvement. This processing structure is well established in neurons involved in sophisticated, higher-order processes but has long been overlooked for dopamine neurons. It can explain both the salience and punishment accounts of dopamine function. According to this account, salience concerns only the initial and transient part of the dopamine response, whereas punishers activate dopamine neurons through their physical impact rather than their aversiveness. These advances also address the debated issue of dopamine diversity; the phasic reward signal has been shown to be remarkably similar between dopamine neurons and shows only graded variations that are typical for biological phenomena rather than categorical differences; however, all other aspects of dopamine function are as diverse as in other neuronal systems. In moving beyond these issues, we now identify dopamine prediction-error signals for subjective reward value and formal economic utility. Together with recent molecular, cellular and synaptic work, these results will help to better characterize reward signals in other key structures of the brain and construct a neuronal theory of basic learning, utility and economic decision-making.

Wolfram Schultz is at the Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge CB2 3DY, UK.
ws234@cam.ac.uk

doi:10.1038/nrn.2015.26
Published online 11 Feb 2016

1. Schultz, W. Multiple dopamine functions at different time courses. *Ann. Rev. Neurosci.* **30**, 259–288 (2007).
2. Ljungberg, T., Apicella, P. & Schultz, W. Responses of monkey dopamine neurons during learning of behavioral reactions. *J. Neurophysiol.* **67**, 145–163 (1992).
3. Schultz, W., Dayan, P. & Montague, R. R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
4. Schultz, W. Predictive reward signal of dopamine neurons. *J. Neurophysiol.* **80**, 1–27 (1998).
5. Waelti, P., Dickinson, A. & Schultz, W. Dopamine responses comply with basic assumptions of formal learning theory. *Nature* **412**, 43–48 (2001).
6. Tobler, P. N., Dickinson, A. & Schultz, W. Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. *J. Neurosci.* **23**, 10402–10410 (2003).
7. Bayer, H. M. & Glimcher, P. W. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47**, 129–141 (2005).
8. Pan, W.-X., Schmidt, R., Wickens, J. R. & Hyland, B. I. Dopamine cells respond to predicted events during classical conditioning: evidence for eligibility traces in the reward-learning network. *J. Neurosci.* **25**, 6235–6242 (2005).
9. Lak, A., Stauffer, W. R. & Schultz, W. Dopamine prediction error responses integrate subjective value from different reward dimensions. *Proc. Natl Acad. Sci. USA* **111**, 2343–2348 (2014).
10. Fiorillo, C. D., Tobler, P. N. & Schultz, W. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* **299**, 1898–1902 (2003).
11. Schultz, W., Ruffieux, A. & Aebischer, P. The activity of pars compacta neurons of the monkey substantia nigra in relation to motor activation. *Exp. Brain Res.* **51**, 377–387 (1983).
12. Schultz, W. Responses of midbrain dopamine neurons to behavioral trigger stimuli in the monkey. *J. Neurophysiol.* **56**, 1439–1462 (1986).
13. DeLong, M. R., Crutcher, M. D. & Georgopoulos, A. P. Relations between movement and single cell discharge in the substantia nigra of the behaving monkey. *J. Neurosci.* **3**, 1599–1606 (1983).
14. Romo, R. & Schultz, W. Dopamine neurons of the monkey midbrain: contingencies of responses to active touch during self-initiated arm movements. *J. Neurophysiol.* **63**, 592–606 (1990).
15. Chiodo, L. A., Antelman, S. M., Caggiula, A. R. & Lineberry, C. G. Sensory stimuli alter the discharge rate of dopamine (DA) neurons: evidence for two functional types of DA cells in the substantia nigra. *Brain Res.* **189**, 544–549 (1980).
16. Steinfels, G. F., Heym, J., Strecker, R. E. & Jacobs, B. L. Behavioral correlates of dopaminergic unit activity in freely moving cats. *Brain Res.* **258**, 217–228 (1983).
17. Schultz, W. & Romo, R. Responses of nigrostriatal dopamine neurons to high intensity somatosensory stimulation in the anesthetized monkey. *J. Neurophysiol.* **57**, 201–217 (1987).
18. Schultz, W. & Romo, R. Dopamine neurons of the monkey midbrain: contingencies of responses to stimuli eliciting immediate behavioral reactions. *J. Neurophysiol.* **63**, 607–624 (1990).
19. Mirenowicz, J. & Schultz, W. Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature* **379**, 449–451 (1996).
20. Horvitz, J. C., Stewart, T. & Jacobs, B. L. Burst activity of ventral tegmental dopamine neurons is elicited by sensory stimuli in the awake cat. *Brain Res.* **759**, 251–258 (1997).
21. Guarraci, F. A. & Kapp, B. S. An electrophysiological characterization of ventral tegmental area dopaminergic neurons during differential pavlovian fear conditioning in the awake rabbit. *Behav. Brain Res.* **99**, 169–179 (1999).
22. 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.* **28**, 11673–11684 (2008).
23. Brischoux, F., Chakraborty, S., Brierley, D. I. & Ungless, M. A. Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. *Proc. Natl Acad. Sci. USA* **106**, 4894–4899 (2009).
24. Matsumoto, M. & Hikosaka, O. Two types of dopamine neuron distinctively convey positive and negative motivational signals. *Nature* **459**, 837–841 (2009).
25. Fiorillo, C. D., Song, M. R. & Yun, S. R. Multiphasic temporal dynamics in responses of midbrain dopamine neurons to appetitive and aversive stimuli. *J. Neurosci.* **33**, 4710–4725 (2013).
26. Fiorillo, C. D., Yun, S. R. & Song, M. R. Diversity and homogeneity in responses of midbrain dopamine neurons. *J. Neurosci.* **33**, 4693–4709 (2013).
27. Fiorillo, C. D. Two dimensions of value: dopamine neurons represent reward but not aversiveness. *Science* **341**, 546–549 (2013).
28. Thorpe, S. J., Rolls, E. T. & Maddison, S. The orbitofrontal cortex: neuronal activity in the behaving monkey. *Exp. Brain Res.* **49**, 93–115 (1983).
29. Ravel, S., Legallet, E. & Apicella, P. Responses of tonically active neurons in the monkey striatum discriminate between motivationally opposing stimuli. *J. Neurosci.* **23**, 8489–8497 (2003).
30. Roitman, M. F., Wheeler, R. A. & Carelli, R. M. Nucleus accumbens neurons are innately tuned for rewarding and aversive taste stimuli, encode their predictors, and are linked to motor output. *Neuron* **45**, 587–597 (2005).
31. Paton, J. J., Belova, M. A., Morrison, S. E. & Salzman, C. D. The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature* **439**, 865–870 (2006).
32. Amemori, K.-I. & Graybiel, A. M. Localized microstimulation of primate pregenual cingulate cortex induces negative decision-making. *Nat. Neurosci.* **15**, 776–785 (2012).
33. Nomoto, K., Schultz, W., Watanabe, T. & Sakagami, M. Temporally extended dopamine responses to perceptually demanding reward-predictive stimuli. *J. Neurosci.* **30**, 10692–10702 (2010).
34. Kobayashi, S. & Schultz, W. Reward contexts extend dopamine signals to unrewarded stimuli. *Curr. Biol.* **24**, 56–62 (2014).
35. Stauffer, W. R., Lak, A. & Schultz, W. Dopamine reward prediction error responses reflect marginal utility. *Curr. Biol.* **24**, 2491–2500 (2014).
36. Bushnell, M. C., Goldberg, M. E. & Robinson, D. L. Behavioral enhancement of visual responses in monkey cerebral cortex. I. Modulation in posterior parietal cortex related to selective visual attention. *J. Neurophysiol.* **46**, 755–772 (1981).
37. Treue, S. & Maunsell, J. H. R. Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature* **382**, 539–541 (1996).
38. Womelsdorf, T., Anton-Erxleben, K., Pieper, F. & Treue, S. Dynamic shifts of visual receptive fields in cortical area MT by spatial attention. *Nat. Neurosci.* **9**, 1156–1160 (2006).
39. Nardo, D., Santangelo, V. & Macaluso, E. Stimulus-driven orienting of visuo-spatial attention in complex dynamic environments. *Neuron* **69**, 1015–1028 (2011).
40. Annic, A., Bocquillon, P., Bourriez, J.-L., Derambure, P. & Dujardin, K. Effects of stimulus-driven and goal-directed attention on prepulse inhibition of the cortical responses to an auditory pulse. *Clin. Neurophysiol.* **125**, 1576–1588 (2014).
41. Thompson, K. G., Hanes, D. P., Bichot, N. P. & Schall, J. D. Perceptual and motor processing stages identified in the activity of macaque frontal eye field neurons during visual search. *J. Neurophysiol.* **76**, 4040–4055 (1996).
42. Ipata, A. E., Gee, A. L., Bisley, J. W. & Goldberg, M. E. Neurons in the lateral intraparietal area create a priority map by the combination of disparate signals. *Exp. Brain Res.* **192**, 479–488 (2009).
43. Ipata, A. E., Gee, A. L. & Goldberg, M. E. Feature attention evokes task-specific pattern selectivity in V4 neurons. *Proc. Natl Acad. Sci. USA* **109**, 16778–16785 (2012).
44. Pooresmaeili, A., Poort, J. & Roelfsema, P. R. Simultaneous selection by object-based attention in visual and frontal cortex. *Proc. Natl Acad. Sci. USA* **111**, 6467–6472 (2014).
45. Shadlen, M. N. & Newsome, W. T. Neural basis of a perceptual decision in the parietal cortex (Area LIP) of the rhesus monkey. *J. Neurophysiol.* **86**, 1916–1936 (2001).
46. Roitman, J. D. & Shadlen, M. N. Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci.* **22**, 9475–9489 (2002).
47. Ringach, D. L., Hawken, M. J. & Shapley, R. Dynamics of orientation tuning in macaque primary visual cortex. *Nature* **387**, 281–284 (1997).
48. Sugase, Y., Yamane, S., Ueno, S. & Kawano, K. Global and fine information coded by single neurons in the temporal visual cortex. *Nature* **400**, 869–873 (1999).
49. Bredfeldt, C. E. & Ringach, D. L. Dynamics of spatial frequency tuning in macaque V1. *J. Neurosci.* **22**, 1976–1984 (2002).

50. Hedgé, J. & Van Essen, D. C. Temporal dynamics of shape analysis in macaque visual area V2. *J. Neurophysiol.* **92**, 3030–3042 (2004).
51. Roelfsema, P. R., Tolboom, M. & Khayat, P. S. Different processing phases for features, figures, and selective attention in the primary visual cortex. *Neuron* **56**, 785–792 (2007).
52. Hedgé, J. Time course of visual perception: Coarse-to-fine processing and beyond. *Prog. Neurobiol.* **84**, 405–439 (2008).
53. Lak, A., Arabzadeh, E., Harris, J. A. & Diamond, M. E. Correlated physiological and perceptual effects of noise in a tactile stimulus. *Proc. Natl Acad. Sci. USA* **107**, 7981–7986 (2010).
54. Hung, C. P., Kreiman, G., Poggio, T. & DiCarlo, J. J. Fast readout of object identity from macaque inferior temporal cortex. *Science* **310**, 863–866 (2005).
55. Ambroggi, F., Ishikawa, A., Fields, H. L. & Nicola, S. M. Basolateral amygdala neurons facilitate reward-seeking behavior by exciting nucleus accumbens neurons. *Neuron* **59**, 648–661 (2008).
56. Peck, C. J., Lau, B. & Salzman, C. D. The primate amygdala combines information about space and value. *Nat. Neurosci.* **16**, 340–348 (2013).
57. Mogami, T. & Tanaka, K. Reward association affects neuronal responses to visual stimuli in macaque TE and perirhinal cortices. *J. Neurosci.* **26**, 6761–6770 (2006).
58. Stanisor, L., van der Togt, C., Pennartz, C. M. A. & Roelfsema, P. R. A unified selection signal for attention and reward in primary visual cortex. *Proc. Natl Acad. Sci. USA* **110**, 9136–9141 (2013).
59. Morris, G., Arkadir, D., Nevet, A., Vaadia, E. & Bergman, H. Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. *Neuron* **43**, 133–143 (2004).
60. Day, J. J., Roitman, M. F. & Wightman, R. M. & Carelli, R. M. Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nat. Neurosci.* **10**, 1020–1028 (2007).
61. Kobayashi, S. & Schultz, W. Influence of reward delays on responses of dopamine neurons. *J. Neurosci.* **28**, 7837–7846 (2008).
62. Fiorillo, C. D., Newsome, W. T. & Schultz, W. The temporal precision of reward prediction in dopamine neurons. *Nat. Neurosci.* **11**, 966–973 (2008).
63. Budygin, E. A. *et al.* Aversive stimulus differentially triggers subsecond dopamine release in reward regions. *Neuroscience* **201**, 331–337 (2012).
64. Redgrave, P., Prescott, T. J. & Gurney, K. Is the short-latency dopamine response too short to signal reward? *Trends Neurosci.* **22**, 146–151 (1999).
65. Redgrave, P. & Gurney, K. The short-latency dopamine signal: a role in discovering novel actions? *Nat. Rev. Neurosci.* **7**, 967–975 (2006).
66. Schultz, W., Apicella, P. & Ljungberg, T. Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *J. Neurosci.* **13**, 900–913 (1993).
67. Mitchell, D. S. & Gormezano, I. Effects of water deprivation on classical appetitive conditioning of the rabbit's jaw movement response. *Learn. Motivat.* **1**, 199–206 (1970).
68. Mackintosh, N. J. *The Psychology of Animal Learning* (Academic Press, 1974).
69. Ljungberg, T., Apicella, P. & Schultz, W. Responses of monkey midbrain dopamine neurons during delayed alternation performance. *Brain Res.* **586**, 337–341 (1991).
70. Pearce, J. M. & Hall, G. A model for Pavlovian conditioning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol. Rev.* **87**, 532–552 (1980).
71. Chelazzi, L. *et al.* Altering spatial priority maps via reward-based learning. *J. Neurosci.* **34**, 8594–8604 (2014).
72. Darwin, C. *On the Origin of Species by Natural Selection, or the Preservation of Favoured Races in the Struggle for Life* (John Murray, 1859).
73. Montague, P. R., Dayan, P. & Sejnowski, T. J. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* **16**, 1936–1947 (1996).
74. Tobler, P. N., Fiorillo, C. D. & Schultz, W. Adaptive coding of reward value by dopamine neurons. *Science* **307**, 1642–1645 (2005).
75. Enomoto, K. *et al.* Dopamine neurons learn to encode the long-term value of multiple future rewards. *Proc. Natl Acad. Sci. USA* **108**, 15462–15467 (2011).
76. Richards, J. B., Mitchell, S. H., de Wit, H. & Seiden, L. S. Determination of discount functions in rats with an adjusting-amount procedure. *J. Exp. Anal. Behav.* **67**, 353–366 (1997).
77. Day, J. J., Jones, J. L., Wightman, R. M. & Carelli, R. M. Phasic nucleus accumbens dopamine release encodes effort- and delay-related costs. *Biol. Psychiat.* **68**, 306–309 (2010).
78. Pascal, B. *Pensées (1658–1662)* (Hackett, 2004).
79. Bernoulli, D. *Specimen theoriae novae de mensura sortis. Comentarii Academiae Scientiarum Imperialis Petropolitanae* **5**, 175–192 (1738); English translation available in Exposition of a new theory on the measurement of risk. *Econometrica* **22**, 23–36 (1954).
80. Bentham, J. *An Introduction to the Principle of Morals and Legislations* (Blackwell, 1948).
81. von Neumann, J. & Morgenstern, O. *The Theory of Games and Economic Behavior* (Princeton Univ. Press, 1944).
82. Kahneman, D. & Tversky, A. Prospect theory: an analysis of decision under risk. *Econometrica* **47**, 263–291 (1979).
83. Luce, R. D. *Individual Choice Behavior: A Theoretical Analysis* (Wiley, 1959).
84. Kagel, J. H., Battalio, R. C. & Green, L. *Economic Choice Theory: An Experimental Analysis of Animal Behavior* (Cambridge Univ. Press, 1995).
85. Mas-Colell, A., Whinston, M. & Green, J. *Microeconomic Theory* (Oxford Univ. Press, 1995).
86. Savage, L. J. *The Foundations of Statistics* (Wiley, 1954).
87. Debreu, G. Cardinal utility for even-chance mixtures of pairs of sure prospects. *Rev. Econ. Stud.* **26**, 174–177 (1959).
88. Rothschild, M. & Stiglitz, J. E. Increasing risk: I. A definition. *J. Econ. Theory* **2**, 225–243 (1970).
89. Caraco, T., Martindale, S. & Whitham, T. S. An empirical demonstration of risk-sensitive foraging preferences. *Anim. Behav.* **28**, 820–830 (1980).
90. Machina, M. J. Choice under uncertainty: problems solved and unsolved. *J. Econ. Perspect.* **1**, 121–154 (1987).
91. McCoy, A. N. & Platt, M. L. Risk-sensitive neurons in macaque posterior cingulate cortex. *Nat. Neurosci.* **8**, 1220–1227 (2005).
92. O'Neill, M. & Schultz, W. Coding of reward risk distinct from reward value by orbitofrontal neurons. *Neuron* **68**, 789–800 (2010).
93. Prelec, D. & Loewenstein, G. Decision making over time and under uncertainty: a common approach. *Management Sci.* **37**, 770–786 (1991).
94. Weber, B. J. & Chapman, G. B. Playing for peanuts: why is risk seeking more common for low-stakes gambles? *Organ. Behav. Hum. Decis. Process.* **97**, 31–46 (2005).
95. Fehr-Duda, H., Bruhin, A., Epper, T. & Schubert, R. Rationality on the rise: why relative risk aversion increases with stake size. *J. Risk Uncertain.* **40**, 147–180 (2010).
96. Kreps, D. M. *A Course in Microeconomic Theory* (Pearson Education, 1990).
97. Friedman, M. & Savage, L. J. The utility analysis of choices involving risk. *J. Polit. Econ.* **56**, 279–304 (1948).
98. Markowitz, H. The utility of wealth. *J. Polit. Econ.* **6**, 151–158 (1952).
99. Jones, S. R., Garris, P. A. & Wightman, R. M. Different effects of cocaine and nomifensine on dopamine uptake in the caudate-putamen and nucleus accumbens. *J. Pharmacol. Exp. Ther.* **274**, 396–403 (1995).
100. Calabresi, P. *et al.* Dopamine and cAMP-regulated phosphoprotein 32 kDa controls both striatal long-term depression and long-term potentiation, opposing forms of synaptic plasticity. *J. Neurosci.* **20**, 8443–8451 (2000).
101. Gurden, H., Takita, M. & Jay, T. M. 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.* **20**, RC106 (2000).
102. Reynolds, J. N. J., Hyland, B. I. & Wickens, J. R. A cellular mechanism of reward-related learning. *Nature* **413**, 67–70 (2001).
103. Pawlak, V. & Kerr, J. N. D. Dopamine receptor activation is required for corticostriatal spike-timing-dependent plasticity. *J. Neurosci.* **28**, 2435–2446 (2008).
104. Shen, W., Flajolet, M., Greengard, P. & Surmeier, D. J. Dichotomous dopaminergic control of striatal synaptic plasticity. *Science* **321**, 848–851 (2008).
105. 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* **106**, 1328–1333 (2009).
106. Yagishita, S. *et al.* A critical time window for dopamine actions on the structural plasticity of dendritic spines. *Science* **345**, 1616–1620 (2014).
107. Tsai, H.-C. *et al.* Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science* **324**, 1080–1084 (2009).
108. Witten, I. B. *et al.* Recombinase-driver rat lines: tools, techniques, and optogenetic application to dopamine-mediated reinforcement. *Neuron* **72**, 721–733 (2011).
109. Adamantidis, A. R. *et al.* Optogenetic interrogation of dopaminergic modulation of the multiple phases of reward-seeking behavior. *J. Neurosci.* **31**, 10829–10835 (2011).
110. Kim, K. M. *et al.* Optogenetic mimicry of the transient activation of dopamine neurons by natural reward is sufficient for operant reinforcement. *PLoS ONE* **7**, e33612 (2012).
111. Steinberg, E. E. *et al.* A causal link between prediction errors, dopamine neurons and learning. *Nat. Neurosci.* **16**, 966–973 (2013).
112. Ilango, A. *et al.* Similar roles of substantia nigra and ventral tegmental dopamine neurons in reward and aversion. *J. Neurosci.* **34**, 817–822 (2014).
113. Kravitz, A. V., Tye, L. D. & Kreitzer, A. C. Distinct roles for direct and indirect pathway striatal neurons in reinforcement. *Nat. Neurosci.* **15**, 816–818 (2012).
114. Schultz, W. Neuronal reward and decision signals: from theories to data. *Physiol. Rev.* **95**, 853–951 (2015).
115. Morris, G., Nevet, A., Arkadir, D., Vaadia, E. & Bergman, H. Midbrain dopamine neurons encode decisions for future action. *Nat. Neurosci.* **9**, 1057–1063 (2006).
116. Padoa-Schioppa, C. & Assad, J. A. Neurons in the orbitofrontal cortex encode economic value. *Nature* **441**, 223–226 (2006).
117. Cai, X., Kim, S. & Lee, D. Heterogeneous coding of temporally discounted values in the dorsal and ventral striatum during intertemporal choice. *Neuron* **69**, 170–182 (2011).
118. Kennerley, S. W., Behrens, T. E. J. & Wallis, J. D. Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nat. Neurosci.* **14**, 1581–1589 (2011).
119. So, N.-Y. & Stuphorn, V. Supplementary eye field encodes reward prediction error. *J. Neurosci.* **32**, 2950–2963 (2012).
120. Samejima, K., Ueda, Y., Doya, K. & Kimura, M. Representation of action-specific reward values in the striatum. *Science* **310**, 1337–1340 (2005).
121. Lau, B. & Glimcher, P. W. Value representations in the primate striatum during matching behavior. *Neuron* **58**, 451–463 (2008).
122. Ito, M. & Doya, K. Validation of decision-making models and analysis of decision variables in the rat basal ganglia. *J. Neurosci.* **29**, 9861–9874 (2009).
123. Kim, H., Sul, J. H., Huh, N., Lee, D. & Jung, M. W. Role of striatum in updating values of chosen actions. *J. Neurosci.* **29**, 14701–14712 (2009).
124. Seo, M., Lee, E. & Averbeck, B. B. Action selection and action value in frontal-striatal circuits. *Neuron* **74**, 947–960 (2012).
125. Brown, J. R. & Arbuthnott, G. W. The electrophysiology of dopamine (D2) receptors: a study of the actions of dopamine on corticostriatal transmission. *Neuroscience* **10**, 349–355 (1983).
126. Toan, D. L. & Schultz, W. Responses of rat pallidum cells to cortex stimulation and effects of altered dopaminergic activity. *Neuroscience* **15**, 683–694 (1985).
127. Mink, J. W. The basal ganglia: focused selection and inhibition of competing motor programs. *Prog. Neurobiol.* **50**, 381–425 (1996).
128. Zweifel, L. S. *et al.* Disruption of NMDAR-dependent burst firing by dopamine neurons provides selective assessment of phasic dopamine-dependent behavior. *Proc. Natl Acad. Sci. USA* **106**, 7281–7288 (2009).
129. Aosaki, T., Graybiel, A. M. & Kimura, M. Effect of the nigrostriatal dopamine system on acquired neural responses in the striatum of behaving monkeys. *Science* **265**, 412–415 (1994).
130. Hernández-López, S., Bargas, J., Surmeier, D. J., Reyes, A. & Galarraga, E. D1 receptor activation enhances evoked discharge in neostriatal medium spiny neurons by modulating an L-type Ca^{2+} conductance. *J. Neurosci.* **17**, 3334–3342 (1997).

131. Hernández-López, S. *et al.* D2 dopamine receptors in striatal medium spiny neurons reduce L-type Ca^{2+} currents and excitability via a novel PLCβ1-IP_3-calcineurin-signaling cascade. *J. Neurosci.* **20**, 8987–8995 (2000).
132. Tai, L.-H. & Lee, A. M., Benavidez, N., Bonci, A. & Wilbrecht, L. Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. *Nat. Neurosci.* **15**, 1281–1289 (2012).
133. Badrinarayan, A. *et al.* Aversive stimuli differentially modulate real-time dopamine transmission dynamics within the nucleus accumbens core and shell. *J. Neurosci.* **32**, 15779–15790 (2012).
134. Lerner, T. *et al.* Intact-brain analyses reveal distinct information carried by SNc dopamine subcircuits. *Cell* **162**, 635–647 (2015).
135. Solomon, R. L. & Corbit, J. D. An opponent-process theory of motivation. *Psychol. Rev.* **81**, 119–145 (1974).
136. Oleson, E. B., Gentry, R. N., Chioma, V. C. & Cheer, J. F. Subsecond dopamine release in the nucleus accumbens predicts conditioned punishment and its successful avoidance. *J. Neurosci.* **32**, 14804–14808 (2012).
137. Lammel, S. *et al.* Input-specific control of reward and aversion in the ventral tegmental area. *Nature* **491**, 212–217 (2012).
138. Stopper, C. M., Tse, M. T. L., Montes, D. R., Wiedman, C. R. & Floresco, S. B. Overriding phasic dopamine signals redirects action selection during risk/reward decision making. *Neuron* **84**, 177–189 (2014).
139. Christoph, G. R., Leonzio, R. J. & Wilcox, K. S. Stimulation of the lateral habenula inhibits dopamine-containing neurons in the substantia nigra and ventral tegmental area of the rat. *J. Neurosci.* **6**, 613–619 (1986).
140. Ji, H. & Shepard, P. D. Lateral habenula stimulation inhibits rat midbrain dopamine neurons through a $GABA_A$ receptor-mediated mechanism. *J. Neurosci.* **27**, 6923–6930 (2007).
141. Matsumoto, M. & Hikosaka, O. Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature* **447**, 1111–1115 (2007).
142. Kapur, S. Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *Am. J. Psychiatry* **160**, 13–23 (2003).
143. Robinson, T. E. & Berridge, K. C. The neural basis for drug craving: an incentive-sensitization theory of addiction. *Brain Res. Rev.* **18**, 247–291 (1993).
144. Saunders, B. T. & Robinson, T. E. The role of dopamine in the accumbens core in the expression of Pavlovian-conditioned responses. *Eur. J. Neurosci.* **36**, 2521–2532 (2012).
145. Rao, R. P. N. & Ballard, D. H. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* **2**, 79–87 (1999).
146. Roeper, J. Dissecting the diversity of midbrain dopamine neurons. *Trends Neurosci.* **36**, 336–342 (2013).
147. Fuster, J. M. Unit activity of prefrontal cortex during delayed-response performance: neuronal correlates of transient memory. *J. Neurophysiol.* **36**, 61–78 (1973).
148. Apicella, P., Scarnati, E., Ljungberg, T. & Schultz, W. Neuronal activity in monkey striatum related to the expectation of predictable environmental events. *J. Neurophysiol.* **68**, 945–960 (1992).
149. Hollerman, J. R., Tremblay, L. & Schultz, W. Influence of reward expectation on behavior-related neuronal activity in primate striatum. *J. Neurophysiol.* **80**, 947–963 (1998).
150. Matsumoto, M. & Takada, M. Distinct representations of cognitive and motivational signals in midbrain dopamine neurons. *Neuron* **79**, 1011–1024 (2013).
151. Satoh, T., Nakai, S., Sato, T. & Kimura, M. Correlated coding of motivation and outcome of decision by dopamine neurons. *J. Neurosci.* **23**, 9913–9923 (2003).
152. Bromberg-Martin, E. S. & Hikosaka, O. Lateral habenula neurons signal errors in the prediction of reward information. *Nature Neurosci.* **14**, 1209–1216 (2011).
153. de Lafuente, O. & Romo, R. Dopamine neurons code subjective sensory experience and uncertainty of perceptual decisions. *Proc. Natl Acad. Sci. USA* **49**, 19767–19771 (2011).
154. Cohen, J. Y., Haesler, S., Vong, L., Lowell, B. B. & Uchida, N. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature* **482**, 85–88 (2012).
155. Stuber, G. D., Wightman, R. M. & Carelli, R. M. Extinction of cocaine self-administration reveals functionally and temporally distinct dopaminergic signals in the nucleus accumbens. *Neuron* **46**, 661–669 (2005).
156. Howe, M. W., Tierney, P. L., Sandberg, S. G., Phillips, P. E. M. & Graybiel, A. M. Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature* **500**, 575–579 (2013).
157. Chesselet, M. F. Presynaptic regulation of neurotransmitter release in the brain: facts and hypothesis. *Neuroscience* **12**, 347–375 (1984).
158. Sugam, J. A., Day, J. J., Wightman, R. M. & Carelli, R. M. Phasic nucleus accumbens dopamine encodes risk-based decision-making behavior. *Biol. Psychiat.* **71**, 199–205 (2012).
159. Threlfell, S. *et al.* Striatal dopamine release is triggered by synchronized activity in cholinergic interneurons. *Neuron* **75**, 58–64 (2012).
160. Chuhma, N., Mingote, S., Moore, H. & Rayport, S. Dopamine neurons control striatal cholinergic neurons via regionally heterogeneous dopamine and glutamate signaling. *Neuron* **81**, 901–912 (2014).
161. Brimblecombe, K. R. & Cragg, S. J. Substance P weights striatal dopamine transmission differently within the striosome-matrix axis. *J. Neurosci.* **35**, 9017–9023 (2015).
162. Robbins, T. W. & Arnsten, A. F. T. The neuropsychopharmacology of fronto-executive function: monoaminergic modulation. *Ann. Rev. Neurosci.* **32**, 267–287 (2009).
163. Young, A. M. J., Joseph, M. H. & Gray, J. A. Increased dopamine release *in vivo* in nucleus accumbens and caudate nucleus of the rat during drinking: a microdialysis study. *Neuroscience* **48**, 871–876 (1992).
164. Grace, A. A. Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: a hypothesis for the etiology of schizophrenia. *Neuroscience* **41**, 1–24 (1991).
165. Datla, K. P., Ahier, R. G., Young, A. M. J., Gray, J. A. & Joseph, M. H. Conditioned appetitive stimulus increases extracellular dopamine in the nucleus accumbens of the rat. *Eur. J. Neurosci.* **16**, 1987–1993 (2002).
166. Cheng, J. J., de Bruin, J. P. C. & Feenstra, M. G. P. Dopamine efflux in nucleus accumbens shell and core in response to appetitive classical conditioning. *Eur. J. Neurosci.* **18**, 1306–1314 (2003).
167. Young, A. M. J. 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 (2004).
168. Anzalone, A. *et al.* Dual control of dopamine synthesis and release by presynaptic and postsynaptic dopamine D2 receptors. *J. Neurosci.* **32**, 9023–9034 (2012).
169. Grace, A. A., Floresco, S. B., Goto, Y. & Lodge, D. J. Regulation of firing of dopaminergic neurons and control of goal-directed behaviors. *Trends Neurosci.* **30**, 220–227 (2007).

Acknowledgements

The author thanks A. Dickinson, P. Bossaerts, C. R. Plott and C. Harris for discussions about animal learning theory and experimental economics; his collaborators on the cited studies for their ingenuity, work and patience; and three anonymous referees for comments. The author is also indebted to K. Nomoto, M. Sakagami and C. D. Fiorillo, whose recent experiments encouraged the ideas proposed in this article. The author acknowledges grant support from the Wellcome Trust (Principal Research Fellowship, Programme and Project Grants: 058365, 093270 and 095495), the European Research Council (ERC Advanced Grant 293549) and the US National Institutes of Health Caltech Conte Center (P50MH094258).

Competing interests statement

The author declares no competing interests.

SUPPLEMENTARY INFORMATION

See online article: S1 (table) | S2 (box)

ALL LINKS ARE ACTIVE IN THE ONLINE PDF

In format provided by Schultz *et al.* (March 2016)

SUPPLEMENTARY INFORMATION 1 (TABLE)

Multi-component neuronal responses

Reference	Task	Brain structure	Initial latency	Main latency	Figure number
Thompson et al. 1996[1]	Visual search	FEF	45 ms	100 ms	9a
Ringach et al. 1997[2]	(Orientation tuning)	V1	45 ms	55 ms	3
Kim & Shadlen 1999[3]	Motion discrimination	DLPFC	~50 ms	200 ms	5, 7
Sugase et al. 1999[4]	(Face categorisation)	IT	40 ms	120 ms	2a
Shadlen & Newsome 2001[5]	Motion discrimination	LIP	175 ms	~300 ms	8, 10
Bredfeldt & Ringach 2002[6]	(Spatial frequency)	V1	36-42 ms	66 ms	3
Roitman & Shadlen 2002[7]	Motion discrimination	LIP	100 ms	200 ms	7, 9
Hedgé & Van Essen 2004[8]	(Shape discrimination)	V2	40-80 ms	120-180 ms	5
Mogami & Tanaka 2006[9]	Visual go-nogo reward	IT	~110 ms	~210 ms	7
Paton et al. 2006[10]	Pavlovian CS+/CS-	Amygdala	?	?	3a, 3b
Roelfsema et al. 2007[11]	Visual discrimination	V1	48 ms	137 ms	3e
Ambroggi et al. 2008[12]	Pavlovian CS+/CS-	Amygdala	20 ms	120 ms	3b, 3c, 3g, 4a
Ipata et al. 2009[13]	Visual search	LIP	25 ms	90-100 ms	3, 4
Lak et al. 2010[14]	Frequency discr	Barrel cx	140 ms	200 ms	4b
Ipata et al. 2012[15]	Visual search	V4	50 ms	100-125 ms	3, 4
Peck et al. 2013[16]	Pavlovian CS+/CS-	Amygdala	90 ms	170 ms	3a
Stanisor et al. 2013[17]	Operant/Pavlovian CSs	V1	50 ms	120 ms	2b, 2c, 3d
Pooresmaeili et al. 2014[18]	Visual search	V1	50 ms	120-190 ms	2b
Pooresmaeili et al. 2014[18]	Visual search	FEF	70 ms	110-190 ms	2a
Lorteije et al. 2015[19]	Decision tree	V1, V4	~40 ms	140-180 ms	4
Schultz & Romo 1990[20]	go-nogo reaching	DA	85 ms	135 ms	11
Waelti et al. 2001[21]	Pavlovian CS+/CS-	DA	90 ms	150 ms	4
Tobler et al. 2003[22]	Pavlovian CS+/CS-	DA	80 ms	180 ms	3c
Morris et al. 2004[23]	Operant probability	DA	125 ms	210 ms	3b
Day et al. 2007[24]	Pavlovian CS+/CS-	DAvolt	?	?	4b
Kobayashi & Schultz 2008[25]	Pavlovian discounting	DA	75 ms	145 ms	5a, 5c
Fiorillo et al. 2008[26]	Pavlovian discounting	DA	105 ms	165 ms	2a
Joshua et al. 2008[27]	Pavlovian probability	DA	80 ms	180 ms	6a
Nomoto et al. 2010[28]	Random dot motion	DA	90 ms	250 ms	3
Fiorillo et al. 2013[29]	Pavlovian CSs, USs	DA	40-120 ms	150-250 ms	7a
Kobayashi & Schultz 2014[30]	Pavlovian CS1-CS3	DA	85 ms	160 ms	3
Lak et al. 2014[31]	Pavlovian CSs, USs	DA	75 ms	120 ms	2d, S6a, S9c
Stauffer et al. 2014[32]	Pavlovian CSs, USs	DA	75 ms	130 ms	5e, 5f, 6b

Abbreviations: FEF: frontal eye fields, V1: primary visual cortex, DLPFC: dorsolateral prefrontal cortex, IT: inferotemporal cortex, LIP: lateral intraparietal cortex, DA: midbrain dopamine neurons, DAvolt: voltammetrically assessed striatal dopamine concentration change. ?: not indicated, CS: conditioned stimulus, US: unconditioned stimulus. Parentheses indicate passive stimulation.

REFERENCES

1. Thompson, K. G., Hanes, D. P., Bichot, N. P. & Schall, J.D. Perceptual and motor processing stages identified in the activity of macaque frontal eye field neurons during visual search. *J. Neurophysiol*. **76**, 4040-4055 (1996).
2. Ringach, D. L., Hawken, M. J. & Shapley, R. Dynamics of orientation tuning in macaque primary visual cortex. *Nature* **387**, 281–284 (1997).
3. Kim, J. N., Shadlen, M. N. Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nat. Neurosci*. **2**, 176-185 (1999).
4. Sugase, Y., Yamane, S., Ueno, S. & Kawano, K. Global and fine information coded by single neurons in the temporal visual cortex. *Nature* **400**, 869-873 (1999).
5. Shadlen, M. N. & Newsome, W. T. Neural basis of a perceptual decision in the parietal cortex (Area LIP) of the rhesus monkey. *J. Neurophysiol*. **86**, 1916-1936 (2001).
6. Bredfeldt, C. E. & Ringach, D. L. Dynamics of Spatial Frequency Tuning in Macaque V1. *J. Neurosci*. **22**, 1976-1984 (2002).
7. Roitman, J. D. & Shadlen, M. N. Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci*. **22**, 9475–9489 (2002).
8. Hedgé, J. & Van Essen D. C. Temporal dynamics of shape analysis in macaque visual area V2. *J. Neurophysiol*. **92**, 3030-3042 (2004).
9. Mogami, T. & Tanaka, K. Reward association affects neuronal responses to visual stimuli in macaque TE and perirhinal cortices. *J. Neurosci*. **26**, 6761-6770 (2006).
10. Paton, J. J., Belova, M. A., Morrison, S. E.& Salzman, C. D. The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature* **439**, 865-870 (2006).
11. Roelfsema, P. R., Tolboom, M. & Khayat, P. S. Different processing phases for features, figures, and selective attention in the primary visual cortex. *Neuron* 56:785–792 (2007).
12. Ambroggi, F., Ishikawa, A., Fields, H. L. & Nicola, S. M. Basolateral amygdala neurons facilitate reward-seeking behavior by exciting nucleus accumbens neurons. *Neuron* **59**, 648-661 (2008).
13. Ipata A. E., Gee, A. L., Bisley J. W. & Goldberg, M. E. Neurons in the lateral intraparietal area create a priority map by the combination of disparate signals. Exp. Brain Res. 192, 479–488 (2009).
14. Lak, A., Arabzadeh, E., Harris, J. A. & Diamond, M. E. Correlated physiological and perceptual effects of noise in a tactile stimulus. *Proc. Natl. Acad. Sci. (USA)* **107**, 7981-7986 (2010).
15. Ipata A. E., Gee, A. L. & Goldberg, M. E. Feature attention evokes task-specific pattern selectivity in V4 neurons. Proc. Natl. Acad. Sci. (USA) **109**, 16778–16785 (2012).
16. Peck, C. J., Lau, B. & Salzman, C. D. The primate amygdala combines information about space and value. *Nat. Neurosci*. **16**, 340-348 (2013).
17. Stanisor, L., van der Togt, C., Pennartz, C. M. A. & Roelfsema, P. R. A unified selection signal for attention and reward in primary visual cortex. *Proc. Natl. Acad. Sci. (USA)* **110**, 9136-9141 (2013).

18. Pooresmaeili, A., Poort, J. & Roelfsema, P. R. Simultaneous selection by object-based attention in visual and frontal cortex. *Proc. Natl. Acad. Sci. (USA)* **111**, 6467–6472 (2014).
19. Lorteije, J. A. M., Zylberberg, A., Ouellette, B. G., De Zeeuw, C. I., Sigman, M. & Roelfsema, P. R. The Formation of Hierarchical Decisions in the Visual Cortex. *Neuron* **87**, 1344–1356 (2015).
20. Schultz, W. & Romo, R. Dopamine neurons of the monkey midbrain: Contingencies of responses to stimuli eliciting immediate behavioral reactions. *J. Neurophysiol.* **63**, 607-624 (1990).
21. Waelti, P., Dickinson, A. & Schultz, W. Dopamine responses comply with basic assumptions of formal learning theory. *Nature* **412**, 43-48 (2001).
22. Tobler, P. N., Dickinson, A. & Schultz, W. Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. *J. Neurosci.* **23**, 10402-10410 (2003).
23. Morris, G., Arkadir, D., Nevet, A., Vaadia, E. & Bergman, H. Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. *Neuron* **43**, 133-143 (2004).
24. Day, J. J., Roitman, M. F., Wightman, R. M. & Carelli. R. M. Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nat. Neurosci.* **10**, 1020-1028 (2007).
25. Kobayashi, S. & Schultz, W. Influence of reward delays on responses of dopamine neurons. *J. Neurosci.* **28**, 7837-7846 (2008).
26. Fiorillo, C. D., Newsome, W. T. & Schultz, W. The temporal precision of reward prediction in dopamine neurons. *Nat. Neurosci.* **11**, 966-973 (2008).
27. 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.* **28**, 11673–11684 (2008).
28. Nomoto, K., Schultz, W., Watanabe, T. & Sakagami, M. Temporally extended dopamine responses to perceptually demanding reward-predictive stimuli. *J. Neurosci.* **30**, 10692–10702 (2010).
29. Fiorillo, C. D., Song, M. R & Yun, S. R. Multiphasic temporal dynamics in responses of midbrain dopamine neurons to appetitive and aversive stimuli. *J. Neurosci.* **33**, 4710–4725 (2013).
30. Kobayashi, S. & Schultz, W. Reward contexts extend dopamine signals to unrewarded stimuli. *Curr. Biol.* **24**, 56-62 (2014).
31. Lak, A., Stauffer, W. R. & Schultz, W. Dopamine prediction error responses integrate subjective value from different reward dimensions. *Proc. Natl. Acad. Sci. (USA)* **111**, 2343-2348 (2014).
32. Stauffer, W. R., Lak, A. & Schultz, W. Dopamine reward prediction error responses reflect marginal utility. *Curr. Biol.* **24**, 2491-2500 (2014).

Supplementary Information 2 (Box)

Aversive dopamine activations?

Electrophysiological studies have reported for more than 30 years that aversive stimuli induce phasic activations in subpopulations of dopamine neurons[1-5]; this result was replicated recently[6-8]. Voltammetric studies describe analogous dopamine increases in nucleus accumbens in response to aversive stimuli[9,10]. Similar to rewards, punishers comprise several stimulus components, including physical stimulus impact and negative value, and exert their motivational effects based on their subjective value rather than physical stimulus properties. In line with the theme of this article, recognition of these factors has led to a reassessment of aversive dopamine activations.

Absence of 'aversive' dopamine activations in behaviourally controlled studies. Only one set of experiments has addressed the issue of the subjective value of aversive stimuli[11-13]. The study dissociated the physical impact of an aversive bitter liquid from its subjective value (left part of the figure, which shows the psychophysical assessment of subjective aversive value in choices between two options: a variable singular juice reward (red drop), and a combination of the same but fixed juice reward and an aversive liquid). The subjective aversive value increased as the aversive solution became more concentrated, while the physical impact of the same-amount of liquid drops remained constant. Dopamine neurons showed lower activation with increasing bitterness, suggesting an inverse relationship with its aversiveness (centre part of the figure, reprinted with permission[12]). These results substantiate the notion that the dopamine activation induced by punishers reflects their physical impact rather than their aversiveness[12].

The dopamine responses can be analysed by using a standard economic model, which sums the subjective values from different outcomes into one final net value. According to this model, the negative value of the punisher can be estimated as the amount of the variable singular juice reward at which both options are chosen equally often (choice indifference). At this point, both options have equal subjective value and the subjective value of the punisher is the difference between the amount of singular juice and the amount of juice given together with the aversive liquid. Thus, the subjective value is measured in physical units of juice amount ('common currency'). The indifference point for 1mM denatonium plus 180 microlitres of apple juice was at 80 microlitres of apple juice alone, suggesting a 1mM denatonium value of -100 microlitres apple juice; in analogy, the indifference point for 10 mM denatonium plus 180 microlitres apple juice was at 0 microlitres and suggests a value of -180 microlitres for the 10 mM denatonium solution (denatonium is one of the most bitter substances known to humans). By contrast, the physical impact of the denatonium drops, which have different concentrations but the same volume, remains constant, effectively dissociating aversiveness from physical impact. With this analysis, the reduced dopamine activation with the more concentrated bitter solution would reflect either the lower summed subjective value from the reward juice combined with the bitter liquid or a strong dopamine depression induced by the negative punisher value that cancels the activation from physical impact; the latter mechanism would contradict the suggestion that punishers induce dopamine depressions as negative reward prediction error due to reward absence rather than through their negative value[13]. Interestingly, the quantitative behavioural tests[11-13] reveal that many of the tested 'aversive' stimuli, including airpuffs, tones and chemical solutions, have no appreciable negative subjective value for the animals, although they activate some dopamine neurons in a similar way as the lower denatonium solution of the figure. Also, studies on anaesthetized animals, in which subjective values are intrinsically difficult to ascertain, show 'aversive' activations in some dopamine neurons[1-3,8]. Thus, it is unclear how no or negative subjective value of 'punishers' can explain dopamine activations.

Nature of 'aversive' activations. If negative subjective value fails to explain 'aversive' dopamine activations, what is it that drives these activations? The experimental dissociation of sensory and value stimulus components reveals that physical intensity, rather than negative subjective values of various punishers, correlates with dopamine activations [11-13]. Thus, the 'aversive' dopamine activation may belong primarily to the initial, unselective dopamine response component that reflects stimulus detection. This would also explain why airpuffs, tones and chemical solutions without appreciable negative subjective value activate dopamine neurons[12]. The genuine response to aversive value in dopamine neurons seems to be a depression that reflects either the negative punisher value, or a negative reward prediction error arising from the absence of reward[13]. This response would belong to the second, value dopamine

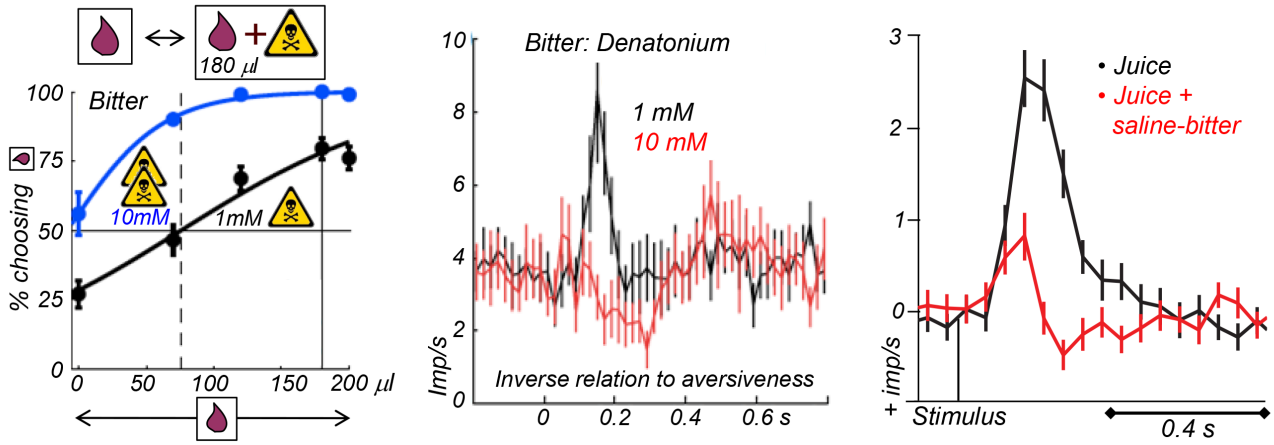


Figure 1. The left and central panels of the figure are adapted with permission from Fiorillo, C. D., Song, M. R. & Yun, S. R. Multiphasic temporal dynamics in responses of midbrain dopamine neurons to appetitive and aversive stimuli. *J. Neurosci.* **33**, 4710–25 (2013). The right panel of the figure is adapted from Fiorillo, C. D. Two dimensions of value: dopamine neurons represent reward but not aversiveness. *Science* **341**, 546–549 (2013). Reprinted with permission from AAAS.

response component. The right part of the figure supports this notion in showing that adding a saline-bitter solution to apple juice results in a lower activation than that evoked by the juice alone. In following the standard economic model of summed positive and negative values, the positively valued apple juice elicits an activation (black), which is reduced by the added negatively valued saline-bitter solution (red). Thus, the frequently reported activation by aversive stimuli may constitute the first dopamine response component and is driven by physical salience, whereas the dopamine depression induced by the negative subjective value of punishers, or by the absence of reward[13], would belong to the second, value response component. As the phasic dopamine activation does not appear to reflect punishment, the main, value response component of dopamine neurons seems to process subjective value in a positive monotonic manner across rewards and punishers.

Reassessment of previous 'aversive' activations. Several factors may explain the previously reported 'aversive' dopamine activations.

First, as argued above, 'aversive' dopamine activations may reflect physical stimulus intensity. Hence, the different fractions of 'aversively' activated dopamine neurons along the dorsal-ventral or medio-lateral extent of the ventral midbrain may reflect gradients of neuronal sensitivities to physical stimulus components[11], rather than categorically distinct phasic dopamine responses[1,7,8,14]. Voltammetric dopamine increases in nucleus accumbens core with tail pinch may similarly reflect physical impact[9].

Second, the activations may reflect punishment relief, which is rewarding[15,16]. As an aversive stimulus not only predicts punishment but also relief, relief prediction might also be rewarding. A system that is activated by reward but not punishment might show activations to rewarding relief; these would be indistinguishable from true aversive activations. Indeed, termination, rather than onset, of tail pinch and footshock elicits voltammetric dopamine responses in nucleus accumbens shell[9,10], which predicts successful punishment avoidance[17]. Relief mechanisms might also explain the observed aversive dopamine activations[4,6,7]. If physical impact and relief combine to elicit 'aversive' dopamine activations, the immediate voltammetric dopamine increase to tail pinch, which is seen in nucleus accumbens core[9], may reflect high physical sensitivity and low relief sensitivity, whereas activations seen with punishment termination[9,10,17] would suggest the opposite.

Third, aversive stimuli that activate dopamine neurons fail to induce full, bidirectional dopamine prediction error responses of punishment in the same way that rewards do[6,7,13]. This result fits with the notion that the 'aversive' activations belong to the first, salience response component but not the second, value component.

Fourth, dopamine activations increase with punisher probability that is predicted by different conditioned stimuli[7]. These conditioned stimuli likely differed in terms of their physical parameters, which were not measured. A physical intensity explanation may also hold for the more frequent dopamine activations to conditioned stimuli compared to primary punishers in the same experiments, which contradict fundamental notions of animal learning theory that postulates lower associative strength for conditioned than unconditioned stimuli[18].

Fifth, the more frequent dopamine activations to conditioned than unconditioned punishers[18] might also be explained by reward context, which, possibly together with reward generalisation, exists in a laboratory that provides monkeys with rewards.

Sixth, some 'aversive responses' may not have come from dopamine neurons. A recent study on five VTA neurons that were activated by footshock determined their dopaminergic nature through juxtacellular labeling[8]. However, 30-45% of dopamine neurons are known to have no appreciable spontaneous activity[19-22]. If these silent dopamine neurons also fail to be activated by aversive stimuli, and if no other physically strong or rewarding stimuli are tested, these neurons may go undetected, despite picking up the label. Instead, neighbouring, non-dopamine neurons, some of which which are well known to be activated by footshocks[23,24], may have responded. Although this may not apply to all dopamine neurons in that study, it might provide a partial explanation whose relevance is difficult to assess in light of the small number of dopamine neurons reported (five)[8]. In another study reporting activations to aversive stimuli[7], many neurons were located above the lateral substantia nigra, where dopamine neurons are rare.

Seventh, habenula neurons, which provide disynaptic inhibitory inputs to dopamine neurons, show similar excitations to punishers as to negative reward prediction errors[25], which would likely result in depressant dopamine responses to punishers, rather than activations. Habenula depressions in response to punishers, which would result in dopamine activations, have not been reported.

Taken together, the two-component account of dopamine responses, and the exquisite sensitivity of the unselective first response component to widely reward-related factors, may explain many of the 'aversive' dopamine activations. When all confounds are considered, the phasic dopamine activation does not seem to reflect punishment. It would be interesting to see how many truly aversively driven dopamine neurons remain when all confounds are accounted for.

Effects of habenula inputs. Results from stimulation of lateral habenula inputs to dopamine neurons contribute diverging arguments for aversive dopamine activations. Optogenetic stimulation of the axons of habenula neurons in their projection site in the medial VTA leads to c-fos activation and excitatory postsynaptic currents (EPSCs) in presumptive dopamine neurons, and induces behavioural place avoidance[26]. However, previous experiments showed inhibitory rather than excitatory dopamine action potential responses that are elicited by habenula stimulation[25,27-29], presumably via the adjacent inhibitory rostromedial reticular GABAergic neurons[30,31]. The frequent rebound excitation in these inhibited dopamine neurons[25,27-29] might explain the (relatively time-insensitive) c-fos activation. Also, EPSCs do not necessarily lead to action potentials that are crucial for dopamine release. Further, the dopaminergic nature of the c-fos activated neurons and those in which EPSCs were recorded is unclear, as these neurons were investigated in a strain of TH:Cre mice that show also labeling of neighbouring non-dopamine GABA neurons[32]. Finally, the employed conditioned place preference task allows only relative value comparisons and thus cannot distinguish between place avoidance induced by aversive dopamine activation and place dispreference evoked by negative reward value from dopamine inhibition. Due to the relative, rather than absolute, valuation in such choices, both mechanisms would result in higher choice frequency of the more valuable place compartment. Indeed, place dispreference is elicited by direct optogenetic dopamine inhibition[24,33] and by optogenetic activation of GABA neurons inhibiting dopamine neurons[34]. These arguments put doubts on an interpretation of habenula fibre stimulation effects in terms of aversive dopamine activation. Indeed, another study using electrical stimulation of the habenula or of rostromedial reticular GABA neurons reports an opposite, inhibitory effect on dopamine neurons and behavioural dispreference compatible with dopamine inhibition[29]. Although electrical stimulation is neuron-unselective, these results do not require the assumption of aversive dopamine activations and agree with the known inhibitory habenula influence on dopamine[25,27-29] and the rewarding effects of direct dopamine stimulation[33,35-39]. These data suggest that habenula stimulation most likely inhibits dopamine neurons and induces place dispreference via a negative dopamine reward signal, rather than supporting an aversive function of phasic dopamine activations.

References

1. Chiodo, L. A., Antelman, S. M., Caggiula, A. R. & Lineberry, C. G. Sensory stimuli alter the discharge rate of dopamine (DA) neurons: Evidence for two functional types of DA cells in the substantia nigra. *Brain Res.* **189**, 544-549 (1980).
2. Schultz, W. & Romo, R. Responses of nigrostriatal dopamine neurons to high intensity somatosensory stimulation in the anesthetized monkey. *J. Neurophysiol* **57**, 201-217 (1987).
3. Mantz. J., Thierry, A. M. & Glowinski, J. Effect of noxious tail pinch on the discharge rate of mesocortical and mesolimbic dopamine neurons: selective activation of the mesocortical system. *Brain Res.* **476**, 377-81 (1989).
4. Mirenowicz, J. & Schultz, W. Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature* **379**, 449-451 (1996).
5. Guarraci, F. A. & Kapp, B. S. An electrophysiological characterization of ventral tegmental area dopaminergic neurons during differential pavlovian fear conditioning in the awake rabbit. Behav. *Brain Res.* **99**, 169-179 (1999).
6. 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.* **28**, 11673–11684 (2008).
7. Matsumoto, M. & Hikosaka, O. Two types of dopamine neuron distinctively convey positive and negative motivational signals. *Nature* **459**, 837-841 (2009).
8. Brischoux, F., Chakraborty, S., Brierley, D.I. & Ungless, M.A. Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. *Proc. Natl. Acad. Sci. (USA)* **106**, 4894–4899 (2009).
9. Budygin, E.A., Park, J., Bass, C.E., Grinevich, V. P., Bonin, K. D. & Wightman, R.M. Aversive stimulus differentially triggers subsecond dopamine release in reward regions. *Neuroscience* **201**, 331-337 (2012).
10. Badrinarayan, A., Wescott, S. A., Vander Weele, C. M., Saunders, B. T., Brenann E. Couturier, B. E., Maren, S. & Aragona, B. J. Aversive stimuli differentially modulate real-time dopamine transmission dynamics within the nucleus accumbens core and shell. *J. Neurosci*. **32**, 15779 –15790 (2012).
11. Fiorillo, C. D., Yun, S. R. & Song, M. R. Diversity and homogeneity in responses of midbrain dopamine neurons. *J. Neurosci.* **33**, 4693–4709 (2013).
12. Fiorillo, C. D., Song, M. R & Yun, S. R. Multiphasic temporal dynamics in responses of midbrain dopamine neurons to appetitive and aversive stimuli. *J. Neurosci.* **33**, 4710–4725 (2013).
13. Fiorillo, C. D. Two dimensions of value: Dopamine neurons represent reward but not aversiveness. *Science* **341**, 546-549 (2013).
14. Lerner, T., N., Shilyansky, C., Davidson, T., J., Luo, L., Tomer, R. & Deisseroth, K. Intact-brain analyses reveal distinct information carried by SNc dopamine subcircuits. *Cell* **162**, 635–647 (2015).
15. Solomon, R.L. & Corbit, J.D. An opponent-process theory of motivation. *Psychol. Rev.* **81**, 119-145 (1974).
16. Gerber, B., Yarali, A., Diegelmann, S., Wotjak, C.T., Pauli, P., Fendt, M. Pain-relief learning in flies, rats, and man: basic research and applied perspectives. *Learn. Mem*. **21**, 232-252 (2014).
17. Oleson, E.B., Gentry, R.N., Chioma, V.C, & Cheer, J.F. Subsecond dopamine release in the nucleus accumbens predicts conditioned punishment and its successful avoidance. *J. Neurosci.* **32**, 14804-14808 (2012).
18. Mackintosh, N. J. *The Psychology of Animal Learning*. (Academic Press, London, 1974).
19. Bunney, B. S. & Grace, A.A. Acute and chonic haloperidol treatment: Comparison of effects on nigral dopaminergic cell activity. *Life Sci.* **23**, 1715-1728 (1978).
20. Henry, D. J., Greene, M. A. & White, F. J. Electrophysiological effects of cocaine in the mesoaccumbens dopamine system: Repeated administration. *J. Pharm. Exp. Ther.* **251**, 833-839 (1989).
21. Gao, W.-Y., Lee, T. H., King, G. R. & Ellinwood, E. H. Alterations in baseline activity and quinpirole sensitivity in putative dopamine neurons in the substantia nigra and ventral tegmental area after withdrawal from cocaine pretreatment. *Neuropsychopharmacology* **18**, 222-232 (1998).
22. Floresco, S. B., West, A. R., Ash, B., Moore, H. & Grace, A. A. Afferent modulation of dopamine neuron firing differentially regulates tonic and phasic dopamine transmission. *Nat. Neurosci.* **6**, 968-973 (2003).
23. Cohen, J. Y., Haesler, S., Vong, L., Lowell, B. B. & Uchida, N. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature* **482**, 85-88 (2012).
24. Tan, K. R., Yvon, C., Turiault, M., Mirzabekov, J. J., Doehner, J., Labouèbe, G., Deisseroth, K., Tye, K. M. & Lüscher, C. GABA neurons of the VTA drive conditioned place aversion. *Neuron* **73**, 1173-1183 (2012).
25. Bromberg-Martin, E. S. & Hikosaka, O. Lateral habenula neurons signal errors in the prediction of reward information. *Nat. Neurosci.* **14**, 1209-1216 (2011).
26. Lammel, S., Lim, B. K., Ran, C., Huang, K. W., Betley, M. J., Tye, K. M., Deisseroth, K. & Malenka, R. C. Input-specific control of reward and aversion in the ventral tegmental area. *Nature* **491**, 212-217 (2012).
27. Christoph, G. R, Leonzio, R. J. & Wilcox, K. S. Stimulation of the lateral habenula inhibits dopamine-containing neurons in the substantia nigra and ventral tegmental area of the rat. *J. Neurosci.* **6,** 613-619 (1986).
28. Ji, H. & Shepard, P. D. Lateral habenula stimulation inhibits rat midbrain dopamine neurons through a GABAA receptor-mediated mechanism. *J. Neurosci.* **27**, 6923-6930 (2007).
29. Stopper, C. M., Tse, M. T. L., Montes, D. R., Wiedman, C. R. & Floresco, S. B. Overriding phasic dopamine signals redirects action selection during risk/reward decision making. *Neuron* **84**, 177-189 (2014).
30. Jhou, T. C., Fields, H. L., Baxter, M. B., Saper, C. B. & Holland, P. C. The rostromedial tegmental nucleus (RMTg), a GABAergic afferent to midbrain dopamine neurons, encodes aversive stimuli and inhibits motor responses. *Neuron* **61**, 786–800 (2009).
31. Hong, S., Jhou, T. C., Smith, M., Saleem, K. S. & Hikosaka, O. Negative reward signals from the lateral habenula to dopamine neurons are mediated by rostromedial tegmental nucleus in primates. *J. Neurosci.* **31**, 11457-11471 (2011).
32. Lammel, S., Steinberg, E., Földy, C., Wall, N. R., Beier, K., Luo, L. & Malenka, R. C. Diversity of transgenic mouse models for selective targeting of midbrain dopamine neurons. *Neuron* **85**, 429-438 (2015).
33. Ilango, A., Kesner, A. J., Keller, K. L., Stuber, G. D., Bonci, A. & Ikemoto, S. Similar roles of substantia nigra and ventral tegmental dopamine neurons in reward and aversion. *J. Neurosci.* **34**, 817-822 (2014).
34. van Zessen, R., Phillips, J. L., Budygin, E. A. & Stuber, G. D. Activation of VTA GABA neurons disrupts reward consumption. *Neuron* **73**, 1184–1194 (2012).
35. Tsai, H.-C., Zhang, F., Adamantidis, A., Stuber, G. D., Bonci, A., de Lecea, L. & Deisseroth, K. Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science* **324**, 1080-1084 (2009).
36. Witten, I. B., Steinberg, E. E., Lee, S. Y., Davidson, T. J., Zalocusky, K. A., Brodsky, M., Yizhar, O., Cho, S. L., Gong, S., Ramakrishnan, C., Stuber, G. D., Tye, K. M., Janak, P. H. & Deisseroth K. Recombinase-driver rat lines: tools, techniques, and optogenetic application to dopamine-mediated reinforcement. *Neuron* **72**, 721–733 (2011).
37. Adamantidis, A. R., Tsai, H.-C., Boutrel, B, Zhang, F., Stuber, G. D., Budygin, E.A., Touriño, C., Bonci, A., Deisseroth, K. & de Lecea, L. Optogenetic interrogation of dopaminergic modulation of the multiple phases of reward-seeking behavior. *J. Neurosci.* **31**, 10829-10835 (2011).
38. Kim, K. M., Baratta, M. V., Yang, A., Lee, D., Boyden, E. S. & Fiorillo, C. D. Optogenetic mimicry of the transient activation of dopamine neurons by natural reward is sufficient for operant reinforcement. *PLoS ONE* **7,** e33612, 2012.
39. Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K. & Janak, P. H. A causal link between prediction errors, dopamine neurons and learning. *Nat. Neurosci.* **16**, 966-973 (2013).