REVIEW

CrossMark

Safety out of control: dopamine and defence

Kevin Lloyd* and Peter Dayan

Abstract

We enjoy a sophisticated understanding of how animals learn to predict appetitive outcomes and direct their behaviour accordingly. This encompasses well-defined learning algorithms and details of how these might be implemented in the brain. Dopamine has played an important part in this unfolding story, appearing to embody a learning signal for predicting rewards and stamping in useful actions, while also being a modulator of behavioural vigour. By contrast, although choosing correct actions and executing them vigorously in the face of adversity is at least as important, our understanding of learning and behaviour in aversive settings is less well developed. We examine aversive processing through the medium of the role of dopamine and targets such as D_2 receptors in the striatum. We consider critical factors such as the degree of control that an animal believes it exerts over key aspects of its environment, the distinction between 'better' and 'good' actual or predicted future states, and the potential requirement for a particular form of opponent to dopamine to ensure proper calibration of state values.

Keywords: Dopamine, Defence, D_2 receptor, Striatum

Background

Our comprehension of appetitive Pavlovian and instrumental conditioning at multiple levels of theory and experiment has progressed dramatically over the last few years. We now enjoy a richly detailed picture, encompassing computational questions about the sorts of prediction and optimization that animals perform, and priors over these; algorithmic issues about the nature of different sorts of learning that get recruited and exploited in various circumstances; and implementational details about the involvement of many structures, including substantial pre-frontal cortical areas, the amygdala, the striatum, and also their respective dopaminergic neuromodulation [1–8]. Along with this evolving understanding of discrete choice, there is evidence that the vigour of engagement in actions is also partly determined through dopaminergic mechanisms associated with the assignment of positive valence, ensuring an alignment of incentive and activity [9–16].

By contrast, the case of aversive Pavlovian and instrumental conditioning is rather less well understood. Perhaps the most venerable puzzle concerns the instrumental case of active avoidance: how could it be that the desired absence of an aversive outcome can influence the choice and motivation of behaviour [17–22]? However, implementational considerations about the architecture of control make for extra problems—if, for instance, vigorous engagement in actions associated with active defence requires recruitment of mechanisms normally thought of as being associated with rewards rather than (potential) punishments [23, 24]. Further, there are alternative passive and active defensive strategies that impose seemingly opposite demands on these systems [25, 26].

In this review, we examine aversion through the medium of dopamine and some of its key targets. Dopamine is by no means the only, or perhaps even the most important, implementational facet of negative valence. For instance, as we will see, complex, species-specific, defensive systems provide an elaborate hard-wired mosaic of responsivity to a panoply of threatening cues [27–29]. Furthermore, cortically-based methods of reasoning that can incorporate and calculate with intricate prior expectations over such things as the degree

*Correspondence: klloyd@gatsby.ucl.ac.uk
Gatsby Computational Neuroscience Unit,
25 Howland Street, London, UK

to which environmental contingencies afford control, play a crucial role in modulating these defences [30–32]. Nevertheless, dopamine is well suited to the purpose of elucidating aversion because of the role it plays in the above enigmas via its influence over learned choice and vigour. Of dopamine's targets, our principal focus here is the striatum, with particular attention to D_2 receptors because of their seemingly special role in passive forms of behavioural inhibition [8, 33].

Almost all the elements of this account have been aired in previous analyses of appetitive and aversive neural reinforcement learning, with the role of dopamine also attracting quite some attention [34–40]. Our main aims are to weave these threads together, using the sophisticated view of appetitive conditioning as a foundation for our treatment of the aversive case, and to highlight issues that remain contentious or understudied. The issue of behavioural control will turn out to be key. We first outline a contemporary view of appetitive conditioning. We then use this to decompose and then recompose the issues concerning innate and learned defence.

Prediction and control of rewards

Reinforcement learning (RL) addresses the following stark problem: learn to choose actions which maximize the sum of a scalar utility or reward signal over the future by interacting with an initially unknown environment. Such environments comprise states or locations, and transitions between these states that may be influenced by actions. What make this problem particularly challenging are both the trial-and-error nature of learning—the effect of actions must be discovered by trying them—and the possibility that actions affect not only immediate but also delayed rewards by changing which states are occupied in the future [41].

Two broad classes of RL algorithms address this computational problem: *model-based* and *model-free* methods [41, 42]. Briefly, model-based methods use experience to construct an internal model of the structure of the environment (i.e. its states and transitions) and the outcomes it affords. Prediction and planning based on the model can then be used to make appropriate choices. Assuming the possibility of constant re-estimation, the flexibility afforded by this class of methods to changes in contingency (i.e. to environmental structure) and motivational state (i.e. to outcome values) has led to the suggestion that it is suitable as a model of goal-directed action [43–46]. Model-based estimates can also encompass comparatively sophisticated 'meta-statistics' of the environment, such as the degree to which rewards and punishments are under the control of the agent [32].

By contrast, model-free methods do not construct an internal model, but rather learn simpler quantities in the service of the same goal. One such is the mean *value* of a state, which summarizes how good it is as judged by the cumulative rewards that are expected to accrue in the future when the subject starts from that state. This is, of course, the quantity that requires optimization. Crucially, the values of successive states satisfy a particular consistency relationship [47], so that states which tend to lead to states of high value will also tend to have high value, and vice-versa for states which tend to lead to low-value states. A broad class of model-free RL methods, known as *temporal difference* (TD) methods, use inconsistencies in the values of sampled successive states—a TD *prediction error* signal—to improve estimates of state values [48].

For selecting appropriate actions, a prominent model-free method is the *actor-critic* [41, 49]. This involves two linked processes. One is the *critic*, which uses the TD error to learn the model-free value of each state. However, future rewards typically depend on the actions chosen, or the behavioural *policy* followed. A policy is a state-response mapping, and is stored in the other component, the *actor*, which determines the relative probabilities of selecting actions. It turns out that the same TD prediction error that can improve the predictions of the critic may also be employed to improve the choices of the actor. There are also other model-free quantities that can be used for action selection. These include the *Q* value [50] of a state-action pair, which reports the expected long-run future reward for taking the particular initial action at the state.

Such model-free methods have the virtue of being able to learn to choose good actions without estimating a world model. However, summarizing experience by simple state values also means that these methods are relatively inflexible in the face of changes in environmental contingencies. Consequently, model-free RL methods have been suggested as a possible model of habitual actions [44, 45].

Both model-based and model-free methods must balance exploration and exploitation. The former is necessary to learn the possibilities associated with a novel domain; the latter then garners the rewards (or avoids the punishments) that the environment has been discovered to afford. This balance depends sensitively on many factors, including prior expectations about the opportunities and threats in the environment, how much control can be exerted over them, and how fast they change [51]. It also requires careful modelling of uncertainty—for instance, it is possible to quantify the value of exploration of unknown options as a function of the expected worth of the exploitation that they could potentially allow in the future [52, 53]. The excess of this over the expected value given current information is sometimes known as

an *exploration bonus*, quantifying optimism in the face of uncertainty [54, 55].

Calculating such bonuses correctly, balancing exploration and exploitation optimally, and even just finding the optimal trajectory of actions in a rich state space, are radically computational intractable; heuristics therefore abound which are differently attuned to different classes of method [51]. Perhaps the most important heuristic is the existence of hard-wired systems that embody pre-specified policies. As we will detail below, these are of particular value in the face of mortal threat—animals will rarely have the luxury of being able to explore to find the best response. However, they are also useful in appetitive cases, obviating learning for actions that are sufficiently evolutionarily stable, such as in food-handling, mating and parenting.

Such hard-wired behaviours may be elicited in the absence of learning by certain stimuli, which are therefore designated *unconditioned stimuli* (USs). Presentation of a US typically inspires what is known as a *consummatory* response, attuned to the particularities of the US. It is through Pavlovian, or classical, conditioning that such innate responses can be attached not only to USs but also to formerly neutral predictors of such outcomes. These predictors are then called *conditioned stimuli* (CSs) since their significance is 'conditioned' by experience. Along with targeted preparation for particular outcomes, CS-elicited *conditioned responses* (CRs) include generic, so-called *preparatory*, actions: typically approach and engagement for appetitive cues, associated with predictions of rewarding outcomes; and inhibition, disengagement and withdrawal for aversive cues, associated with future threats or punishments. The predictions that underpin preparation can be either model-based or model-free [56]. We should note that the long-standing distinction between preparatory and consummatory behaviours [57–59] is not always clear cut; however, it has been usefully invoked—though not always in exactly the same terms—in various related theories of dopamine function [11, 60–66].

The fuller case of RL, in which actions come to be chosen because of their contingent effects rather than being automatically elicited by predictions, corresponds to instrumental conditioning. At least in experimental circumstances such as negative automaintenance [67], automatic, Pavlovian, responses can be placed in direct competition with instrumental choices. Perhaps surprisingly, Pavlovian responses often win [68, 69], leading to inefficient behaviour. A less malign interaction between Pavlovian and instrumental conditioning is called 'Pavlovian-instrumental transfer' (PIT) [45, 70–73]. In this, the vigour of instrumental responding (typically for rewards) is influenced positively or negatively by the presence of Pavlovian CSs associated with appetitive or aversive predictions, respectively.

We start by considering the implications of Pavlovian and instrumental paradigms for the neural realization of control. We use a rather elaborated discussion of appetitive conditioning and rewards as a foundation, since this valence has received more attention and so is better understood. As a preview, we will see that dopamine in the ventral striatum has a special involvement in model-free learning (reporting the TD prediction error). However, dopamine likely also plays an important role in the expression and invigoration of both model-based and model-free behaviour.

Predicting reward: Pavlovian conditioning

Model-free RL, with its TD prediction errors, has played a particularly central role in developing theories of how animals learn state values, the latter interpreted as the predictions of long run rewards that underpin Pavlovian responses [74–77]. There is by now substantial evidence that the phasic activity of midbrain dopamine neurons resembles this TD prediction error in the case of reward [40, 78, 79]. Neural systems in receipt of this dopamine signal are then prime candidates to represent state values. One particularly important such target is the cortical projection to the ventral striatum (or nucleus accumbens; NAc) [78, 80], the plasticity of whose synaptic efficacies may be modulated by dopamine [81–85]. Note that, by contrast, dorsomedial and dorsolateral striatum, which are also targeted by dopamine cells—though by cells in the substantia nigra (SNc) rather than in the ventral tegmental area (VTA)—have been associated respectively with model-based and model-free instrumental behaviour (see below).

Along with its involvement in plasticity, dopamine, particularly in the NAc, has long been implicated in the intensity of the expression of innate, motivated behaviours (i.e., just those behaviours elicited by Pavlovian predictions) in response to both unconditioned and conditioned stimuli [66, 86, 87]. This is a form of Pavlovian vigour [63, 64, 88–91]. Relevant CSs have been described as acquiring 'incentive salience' [65, 92] or 'incentive motivation' [93], possibly via the way that their onset leads to TD errors that reflect state predictions [94]. Perhaps also related to Pavlovian vigour is the observation that the influence of CSs on instrumental responding in PIT paradigms is sensitive to dopamine signalling too [95–97]. It has recently been shown that dopaminergic projections to ventral striatum corelease glutamate [98–100], though see [101], which may modulate these effects.

The influence of dopamine neurons over the expression of behaviour might extend to model-based as well as model-free predictions, based on other afferent

projections to the dopamine system. Model-based values are thought to be stored in, and calculated by, other areas, such as the basolateral amygdala and orbitofrontal cortex [87, 102–109].

Three further details of the ventral striatum and dopamine release in this structure are important. Firstly, anatomically, the NAc is classically subdivided into 'core' (NAcC) and 'shell' (NAcS) subregions [110]. As well as being histochemically distinct, these regions differ in their patterns of connectivity. For example, while NAcC resembles dorsal striatum in projecting extensively to classic basal ganglia output structures, such as the ventral pallidum, NAcS is notable for its projections to subcortical structures outside the basal ganglia, such as lateral hypothalamus and periaqueductal gray (PAG), which are involved in the expression of unlearned behaviours [110–114].

Two related ideas are abroad about the separate roles of these structures. One is that NAcS and NAcC mediate the motivational impact of USs and CSs, respectively [87]. For instance, the projection of NAcS to the lateral hypothalamus is known to play a role in the expression of feeding behaviour [112], requiring intact dopamine signalling within NAcS [115]. Conversely, conditioned approach is impaired by lesions or dopamine depletion of NAcC, but not by lesions of NAcS [116, 117].

The other idea is that NAcS and NAcC are involved in *outcome-specific* and *general* PIT, respectively [118]. The difference concerns whether the Pavlovian prediction is of the same outcome as for the instrumental act (specific PIT), or instead exerts influence according to its valence (general PIT). It has been reported that lesions of NAcS abolished outcome-specific PIT but spared general PIT, while lesions of NAcC abolished general PIT but spared outcome-specific PIT [118].

These ideas are not quite compatible, since both sorts of PIT involve conditioned stimuli. Perhaps, instead, we should think of the NAcC as being more involved in preparatory behaviours, attuned only to the valence (positive or negative) of a predicted outcome but not its particularities, while the NAcS is more involved in consummatory behaviours, which additionally reflect knowledge of the particular expected outcome(s) [118–123]. This is less incompatible with the first idea than it might seem, since outcome-specific PIT presumably relies on representation of the US, even if the US itself is not physically present [56]. This latter interpretation aligns with the distinction between model-free and model-based RL predictions, which would then be associated with NAcC and NAcS, respectively [56].

The second relevant, if somewhat contentious (see below), feature is that, as appears to be the case in the striatum generally, the majority of the principal projection neurons in NAc—medium spiny neurons (MSNs)—may express either D_1 or D_2 receptors, but not both [124, 125]. Briefly, dopamine receptors are currently thought to come in five subtypes, each classified as belonging to one of two families based on their opposing effects on certain intracellular cascades: D_1-like (D_1 and D_5 receptors), and D_2-like (D_2, D_3, and D_4 receptors). D_1 and D_2 receptors are of prime interest here since they are by far the most abundantly expressed dopamine receptors in the striatum and throughout the rest of the brain [126–128]. In the striatum, the majority of D_1 and D_2 receptors are thought to occupy states in which their affinities for dopamine are low and high respectively [129], with the consequence that these receptors are influenced differently by changes in phasic and tonic dopamine release [130]. Furthermore, D_1 and D_2 receptors appear to mediate opposite effects of dopamine on their targets: activation of D_1 receptors tends to excite, and D_2 to inhibit, neurons; this modulation of excitability can then also have consequences for activity-dependent plasticity [131, 132].

In the dorsal striatum, there is substantial evidence for an anatomical segregation between D_1-expressing 'Go' (direct; striatomesencephalic) and D_2-expressing 'NoGo' (indirect; striatopallidal) pathways [131, 133–136]. The effect of these pathways on occurrent and learned choice is consistent with the observations about the activating effect of dopamine [137], as we discuss in more detail below. Equivalent pathways are typically assumed to exist in NAc [138–140] although the segregation here seems more debatable [111, 141–143]. Indeed, D_1-expressing MSNs within NAcC are reported to also project within the striatopallidal ('indirect') pathway [141, 144]; there is evidence for co-expression of D_1 and D_2 receptors, particularly in NAcS [124, 145, 146]; and there are suggestions that D_1 and D_2 receptors can interact to form heteromeric dopamine receptor complexes within the same cell [147, 148], though this appears to be still a matter of question [149]. In functional terms, though, at least in the case of appetitive conditioning, it seems there may be parallel Go and NoGo routes, given evidence that D_1 receptors may be of particular importance in learning Pavlovian contingencies [150–154], while antagonists of either D_1 or D_2 receptors appear to disrupt the expression of such learning [155–159], including the expression of preparatory Pavlovian responses [34, 153]. Unfortunately, given the possible association of core and shell with model-free and model-based systems above, experimental evidence that clearly disentangles the roles of D_1 and D_2 receptors in these respective areas in appetitive conditioning appears to be lacking.

The third detail, which applies equally to ventral and dorsal striatum, concerns the link between the activity of

dopaminergic cells and the release of dopamine into target areas. While there is little doubt that phasic release of striatal dopamine can be driven by activity in midbrain dopaminergic cells (e.g. [160]), a range of mechanisms local to the striatum is known to play a role in regulating dopamine release, including a host of other neurotransmitters such as glutamate, acetylcholine, and GABA (for recent reviews, see [161, 162]). Indeed, recent evidence suggests that striatal dopamine release can be stimulated axo-axonally by the synchronous activity of cholinergic interneurons, separate from changes in the activity of dopaminergic cells [163]. Furthermore, it has long been suggested that there is at least some independence between fast 'phasic' fluctuations in extracellular dopamine within the ventral striatum and a relatively constant 'tonic' dopamine level; the former are proposed to be spatially restricted signals driven by phasic bursting of dopamine cells, while the latter is thought to be comparatively spatially diffuse and controlled rather by the number of dopamine cells firing in a slower, 'tonic' mode of activity [164–166]. Evidence for co-release of other neurotransmitters alongside dopamine, such as glutamate and GABA, adds further complexity [98, 100, 167, 168].

Controlling reward: instrumental conditioning

In the instrumental, model-free, actor-critic method, the critic is the Pavlovian predictor, associated with the ventral striatum. The actor, by contrast, has been tentatively assigned to the dorsal striatum [78, 80, 169] based on its involvement in instrumental learning and control [170, 171]. The dorsal striatum is also a target of dopamine neurons, albeit from the substantia nigra pars compacta (SNc) rather than the ventral tegmental area (VTA). At a slightly finer grain, habitual behaviour has been particularly associated with dorsolateral striatum [172–175], while goal-directed behaviour has been associated with dorsomedial striatum, as well as ventromedial prefrontal and orbitofrontal cortices (for recent reviews, see [176, 177]). Recent evidence implicates lateral prefrontal cortex and frontopolar cortex in the arbitration between these two different forms of behavioural control in humans [178], and pre- and infra-limbic cortex in rats [179].

As noted above, the classical view of dorsal striatum is that the projections of largely separate populations of D_1-expressing (dMSNs) and D_2-expressing (iMSNs) medium spiny neurons are organised respectively into a direct (striatonigral) pathway, which promotes behaviour, and an indirect (striatopallidal) pathway, which suppresses behaviour [133, 134]. This dichotomous expression of D_1 and D_2 receptors would then allow dopamine to modulate the balance between the two pathways by differentially regulating excitability and plasticity [131]. In particular, activation of D_1 receptors in dMSNs increases their excitability and strengthens the direct pathway via long-term potentiation (LTP) of excitatory synapses. By contrast, activation of D_2 receptors in iMSNs decreases their excitability and weakens the indirect pathway by promoting long-term depression (LTD) of excitatory synapses.

This effect is then the basis of an elegant model-free account of instrumental conditioning [137, 180–182]. The active selection or inhibition of an action is mediated by the balance between direct and indirect pathways. Phasic increases and decreases in dopamine concentration report whether an action results in an outcome that is better or worse than expected, either via direct delivery of reward, or a favourable change in state. An increase consequent on the outcome being better than expected strengthens the direct pathway, making it more likely that the action will be repeated in the future. By contrast, a decrease consequent on the action being worse than expected strengthens the indirect pathway, making a repeat less likely. Much evidence, including recent optogenetic results, appears to support this basic mechanism [181, 183], although it is important to note that recent results suggest a slightly more nuanced view of the simple dichotomy between direct and indirect pathways—for instance, they are reported to be coactive during action initiation [184], consistent with the idea that they form a centre-surround organisation for selecting actions [185–187].

While it is natural to associate a dopamine TD prediction error with model-free prediction and control, there are hints that this signal shows a sophistication which potentially reveals more model-based influences [56, 188–191]. One such influence is exploration: observations of phasic activity of dopamine neurons in response to novel input which is not rewarding in any obvious sense (e.g. a novel auditory stimulus [192]) have been considered as an optimism-based exploration bonus [193]. It is not clear whether such activations depend, as they normatively should, on factors such as reward/punishment controllability that are typically the preserve of model-based calculations. Further, there remains to be a clear analysis of the role dopamine plays in the dorsomedial striatum's known influence over model-based RL [194, 195].

Instrumental vigour

Along with Pavlovian vigour is the possibility of choosing the alacrity or force of an action based on the contingent effects of this choice. Dopamine has also been implicated in this [12], potentially associated with model-based as well as model-free actions [196].

One idea is that there is a coupling between instrumental vigour and relatively tonic levels of dopamine, in the case that the latter report the prevailing average reward rate [9, 197]. This quantity acts as an opportunity cost for sloth, allowing a need for speed to be balanced against the (e.g., energetic) costs of acting quickly. Experiments that directly test this idea have duly supported dopaminergic modulation of vigour in reward-based tasks [13, 14, 16]. Formally, the average rate of TD prediction errors is just the same as the average rate of rewards, suggesting that nothing more complicated would be necessary to implement this effect than averaging phasic fluctuations in dopamine, at least in the model-free case. It could then be that because phasic fluctuations reflect Pavlovian as well as instrumental TD prediction errors, vigour would also be influenced by Pavlovian predictions—something that is contrary to the original instrumental expectation [9] but which is apparent in cases such as PIT [95]. Tonic dopamine has, of course, been suggested to be under somewhat separate control from phasic dopamine [164–166].

The putative involvement of dopamine in both vigour and valence leads to the prediction of a particular sort of hard-wired misbehaviour, or Pavlovian-instrumental conflict, namely that it might be hard to learn to withhold actions in the face of stimuli that predict rewards if inhibition is successful. This is indeed true, for both animals [67] and humans [24].

Defence

The main intent of this review is to understand how the elements of adaptive behaviour that we have just described apply in the aversive case. Coarsely, we need to (i) examine the complexities of consummatory versus preparatory, and active versus passive, defensive choices in the face of unconditioned aversive stimuli and their conditioned predictors; (ii) consider how instrumental avoidance actions can be learned to prevent threats from arising in the first place; and (iii) consider how the vigour of defensive actions is set appropriately.

The reason that we structured this review through the medium of dopamine is that it seems that many of the same dopaminergic mechanisms that we have just described for appetitive conditioning also operate in the aversive case, subject to a few added wrinkles. This makes for puzzles, both for aversion (how one could get vigorous defensive actions when only potential punishments are present and the reward rate is therefore at best negative) and for dopamine (why dopamine would apparently be released in just such purely aversive circumstances).

We argue that it is possible to generalize to these cases an expanded notion of *safety* (cf. [64]), which itself underpins the popular, two-factor solution to instrumental

avoidance [17, 19–22, 198–201]. Amongst other things, this implies subtleties in the semantics of dopamine, and a need to pay attention to the distinctions between reinforcement versus reward, and better versus good. To anticipate, we suggest that evidence for positive phasic and tonic dopamine responses to aversive unconditioned and conditioned stimuli may be explained in terms of a prediction of possible future safety. Furthermore, we suggest that these dopamine responses, and the consequent stimulation of striatal D_2 receptors in particular, play an important role in promoting, or at least licensing, active defensive behaviours.

Aversive unconditioned stimuli

There is some complexity in the consummatory response to an appetitive unconditioned stimulus (US) depending on how it needs to be handled. However, the response elicited by an aversive US—notably fleeing, freezing, or fighting—appears to depend in a richer way on the nature of the perceived threat, and indeed the species of the animal threatened [27]. Different emphases on the nature of the threat, or 'stressor', and the defensive response, or 'coping strategy', have led to subtly different, yet complementary, analyses of defensive behaviour and its neural substrates, which include the amygdala, ventral hippocampus, medial prefrontal cortex (mPFC), ventromedial hypothalamus, and periaqueductal gray (PAG) [25, 26, 28, 202–208] (for a recent review, see [29]).

For our purposes, the most important distinction is between *active* defensive responses, such as fight, flight, or freeze, and *passive* ones, such as quiescence, immobility, or decreased responsiveness. These need to be engaged in different circumstances, subject particularly to whether or not the stressor is perceived as being escapable or controllable [25]. Thus, active responses are adaptive if the stressor is perceived as escapable, since these may cause the stressor to be entirely removed. Conversely, passive responses may be more adaptive in the face of inescapable stress, promoting conservation of resources over the longer term and potential recovery once the stressor is removed. In other words, active responses entail *engagement* with the environment, while passive responses entail a degree of *disengagement* from the environment [25]. Even freezing involves 'attentive immobility', which can be interpreted as a state of high 'internal' engagement in threat monitoring.

The potential link to dopamine here is the proposal, particularly advocated by Cabib and Puglisi-Allegra [209–211] and fleshed out below, that an increased tonic level of dopamine in NAc, and especially the resulting stimulation of dopamine D_2 receptors in this area, promotes active defence, whereas a decreased tonic level of dopamine in NAc, and the resulting decrease in D_2

stimulation, promotes passive defence. This suggestion has clear parallels in the appetitive case. As there, in addition to the canonical direct and indirect pathways, typically associated with dorsal striatum and the expression of instrumental behaviours via disinhibition of cortically-specified actions [133, 137, 185, 212], we should expect accumbens-related Pavlovian defence to involve disinhibition and release of innate behavioural systems organised at the subcortical level, such as in the hypothalamus and PAG [112, 204, 213, 214].

For dopamine release, studies using microdialysis to measure extracellular concentrations of dopamine have reported elevated levels in response to an aversive US in NAc [215, 216], as well as in PFC [217] and amygdala [218, 219]. Using the higher temporal resolution technique of fast-scan cyclic voltammetry (FSCV), it has been reported that an aversive tail pinch US immediately triggers elevated dopamine release in the NAcC which is time-locked to the duration of the stimulus, while in the NAcS dopamine release is predominantly inhibited during the stimulus and either recovers or exceeds baseline levels following US offset [220, 221].

The substrate for this release is less clear. As we noted, many dopamine neurons appear to be activated by unexpectedly appetitive events. Although most studies report that dopamine neurons are inhibited by an aversive US (e.g., an electric shock, tail pinch, or airpuff), there are long-standing reports suggesting that a relatively small proportion may instead be activated [40]. The dopaminergic nature of some such responses appears to have been confirmed more recently via optogenetics [222] and juxtacellular labelling [223]. It has also been suggested that a particular group of ventrally-located dopamine cells in the VTA that projects to mPFC [224, 225] is more uniformly excited by aversive USs [223, 226]. In the SNc, it has recently been reported that dopamine cells projecting to the dorsomedial striatum show immediate suppression of activity, followed by sustained elevation of activity, in response to a brief electrical shock. By contrast, dopamine cells projecting to dorsolateral striatum display an immediate increase in activity before promptly returning to baseline [227].

In relation to defensive behaviour, pharmacological interventions and lesion studies have long suggested that dopamine plays a role (reviews include [12, 34]). More recent evidence supporting a particular role for NAc D_2 receptors in defence comes from a series of experiments exploiting the ability of local disruptions to glutamate signalling in NAcS to elicit motivated behaviours [228, 229]. Thus, Richard and Berridge [230] have shown that expression of certain active defensive behaviours in rats (escape attempts, defensive treading/burying), which can be elicited by local AMPA blockade caudally in medial NAcS, not only requires endogenous dopamine activity [115], but also intact signalling of both D_1 and D_2 receptors. By contrast, (appetitive) feeding behaviour, elicited by glutamate disruption more rostrally in the medial NAcS, only requires intact signalling of D_1 receptors [230]. This result supports a role for D_1 receptors in active defence—as well as particular subregions of NAcS (though see [231] for evidence that the behaviours elicited from these regions is sensitive to context)—but it also seems to indicate an asymmetry in the involvement of D_2 receptors in modulating the expression of innate appetitive versus defensive behaviours.

Other studies also suggest a role for D_2 stimulation in active defence, though do not necessarily trace this to NAcS. For example, the expression of certain defensive behaviours in cats (ear retraction, growling, hissing, and paw striking), elicitable by electrical stimulation in ventromedial hypothalamus, can also be respectively instigated or blocked by direct microinjection into that area of a D_2 agonist or antagonist [232, 233]. Indeed, as mentioned previously, anatomical connections between NAcS and hypothalamus are known to play an important role in controlling motivated behaviours, with NAcS cast in the role of 'sentinel' allowing disinhibition of appropriate behavioural centres located in the hypothalamus [112, 214].

Such lines of evidence are consistent with promotion of active Pavlovian defences via enhanced dopamine release and increased NAc D_2 stimulation. Evidence for the other side of the proposal—promotion of passive Pavlovian defences via a drop in dopamine release and reduced NAc D_2 stimulation—is provided by experiments in which animals are exposed to chronic (i.e. inescapable) aversive stimuli, such as in animal models of depression [234]. Briefly, not only do animals in these settings show diminished expression of active defensive behaviours such as escape attempts over time [235–237], but it has also been observed that an initial increase in NAc tonic dopamine on first exposure to the stressor gradually gives way to reduced, below baseline, dopamine levels [238–241]. Since modifications of the animal's behaviour over time in such cases are presumably driven by experience of the (unsuccessful) outcomes of its escape attempts, and so more naturally fit with an instrumental analysis, we postpone fuller discussion of these results until considering the issue of instrumental behaviour and controllability below. However, we note that these changes in patterns of defence and dopamine release over time potentially yield an interesting case of a model-based influence on dopamine and perhaps model-free behaviours.

Pain research provides a complementary view. Bolles and Fanselow [205] pointed out that efficacious (active) defence requires inhibition of pain-related behaviours oriented towards healing injuries. Thus, it was

hypothesized that activation of a fear motivation system, which promotes defensive behaviours (i.e. fight, flight, or freeze), inhibits—for example, by release of endogenous analgesics—a pain motivation system, which promotes recuperative behaviours (i.e. resting and body-care responses). Similarly, activation of the pain system was hypothesized to inhibit the fear system since (active) defensive behaviours would interfere with recovery via (passive) recuperative behaviours. In this light, it is interesting to note the well-established link between NAc dopamine, and D_2 stimulation in particular, and analgesia [242, 243]. Conversely, reductions in motivation in mouse models of chronic pain—consistent with energy-preserving, recuperative functions—have recently been shown to depend on adaptation of (D_2-expressing) iMSNs in NAc [244], and that this adaptation includes an increase in excitability of iMSNs in medial NAcS [245]. In turn, these results are consistent with previous observations of reduced effortful behaviour caused by blockade of NAc D_2 receptors [246, 247]. Both observations are consistent with reductions of actions involved in active defence being caused by the relative strengthening of a ventral indirect pathway.

While these various lines of evidence point to involvement of accumbens dopamine, and NAc D_2 signalling in particular, in modulating defence, we note some important caveats. As mentioned earlier, the separation of direct and indirect pathways in the accumbens is subject to continuing debate, with evidence that D_1-expressing MSNs in NAc also project within the canonical indirect pathway [141] and that a substantial proportion of NAc MSNs co-express D_1 and D_2 receptors [124]. Furthermore, while D_2 receptors may be more attuned to changes in tonic dopamine levels by virtue of their higher affinity, such changes presumably affect occupancy at both D_1 and D_2 receptors dependent on their affinities [130]. In short, rather than completely separate D_1 and D_2 systems that can be independently switched on and off, the true situation is likely to be more complex. Furthermore, experiments involving dopamine receptor agonists and antagonists can be difficult to interpret, since they may involve certain side-effects—such as the well known extrapyramidal symptoms associated with D_2 antagonists [248]—and placing the system into states not encountered during normal functioning.

From an RL perspective, the roles of dopamine and D_2 receptors raise two salient issues. The first is how to make sense of the apparent asymmetry in the involvement of D_2 receptors in defensive, as opposed to appetitive, behaviours. One possibility starts from the observation that traditional paradigms assessing the interaction of Pavlovian and instrumental conditioning suggest that the Pavlovian defence system is biased towards behavioural

inhibition in the face of threat [249, 250]. This Pavlovian bias may potentially require relatively greater inhibition of the ventral indirect pathway in order to *disinhibit* active defensive responses when required. Of course, this mechanistic speculation merely poses the further question of why the Pavlovian defence system should be biased towards behavioural inhibition in the first place. One dubitable speculation is that this stems from asymmetries in the statistics of rewards and punishments in the environment [251]. However, more work is necessary on this point.

The second, and more fundamental, issue is how to interpret variation, particularly enhancement, of NAc dopamine release in response to an aversive US in the first place, given the apparent tie between dopamine, appetitive prediction errors, and reward rates. This is the extended version of the puzzle of active avoidance to which we referred at the beginning. To answer this, we first consider certain similarities and differences between the unexpected arrival of an appetitive or aversive US [252]. This requires us to be more (apparently pedantically) precise about the appetitive case than previously. Here, the unpredicted arrival of the appetitive US (e.g. food) represents an unexpected improvement in the animal's situation. This improvement stems from the fact that the US predicts that an outcome of positive value is immediately attainable. Indeed, all USs can be thought of as predictors, where these predictions are not learned but rather hard-wired. Thus, as previously noted, an appetitive US will engage innate behaviours such as salivation and approach. In turn, these unconditioned responses can be interpreted as reflecting at least an implicit expectation that the predicted reward is attainable/controllable, or at least potentially so, subject to further exploration. Thus, salivation in response to the presence of a food US can be interpreted as reflecting a tacit belief that the food will be consumable (and both require and benefit from ingestion). As reviewed above, the phasic responses of dopamine cells in response to the unexpected presentation of an appetitive US, along with other observations, encourage a TD interpretation in terms of a response to an unexpected predictor of future reward.

Consider now the arrival of an unexpected aversive US (e.g. the sight of a predator). What this event signifies seems more complex. On the one hand, this surprising event presumably indicates that the present situation is *worse* than originally expected, since the animal is now in an undesirable state of danger: i.e., (a) the aversive US is an 'unpredicted predictor of possible future punishment'. As such, we should expect a negative prediction error. Indeed, at least the net value of the prediction error had better be negative to avoid misassignment of positive values to dangerous states and the consequent development

of masochistic tendencies (i.e., the active seeking out of such dangerous states). On the other hand, relative to this new state of danger, the possible prospect of future safety—a positive outcome—comes into play. That is, at the point that the animal would actually manage to eliminate the threat if it can do so, the change in state from danger to safety would lead to an appetitive prediction error—just as with the change in state associated with the unexpected observation of food. Thus, provided the animal has the expectation that it will ultimately be able to achieve safety, i.e., that the situation is controllable, observation of the aversive stimulus should predict this future appetitive outcome, and so (b) lead to an immediate appetitive prediction error. The challenge therefore seems to be that of reconciling (a) and (b), i.e., the role of the aversive US as unpredicted predictor of both danger and possible future safety. To avoid any confusion, note that we discuss learning processes associated with signalling safety below; here, we consider hard-wired assessments of the absence of danger.

One attractive reconciliation comes from appealing to the concept of *opponency* [59, 253, 254]. Here, an aversive process would ensure that the net TD error caused by the unexpected aversive US is negative and that dangerous states are correctly assigned negative value. At the same time, an appetitive process would motivate behaviour towards the comparatively benign state of safety. Indeed, it has previously been proposed that the net prediction error can be decomposed in exactly this way [255], with the phasic activity of dopamine neurons signalling the appetitive component of this signal, while the aversive component is signalled by other means (e.g. by phasic serotonergic activity [23, 249, 252, 256]), such that the net prediction error would actually be negative [252].

A further consideration is the value of exploration. In appetitive contexts, we noted that exploration can be motivated by bonuses associated with the future value of what might be presently discovered. A potential heuristic realization of this was through the phasic activity of dopamine neurons inspired by novel stimuli [192, 193]. Consider the extension of this logic to the unexpected arrival of an aversive US: the animal may have the pragmatic *a priori* belief that safety is controllable, but the unexpected (and therefore 'novel') arrival of an aversive US may nevertheless be attended by uncertainty about how this new situation should be controlled. The issue of how exploration may then be carried out in a benign manner is of course particularly salient here (for a recent view of the issue of safe exploration from the RL perspective, see, e.g. [257]). The idea that a novel stressor elicits exploration in the 'search for effective active coping' has also been suggested by Cabib and Puglisi-Allegra [211]. In their scheme, a novel stressor leads to release of noradrenaline in PFC

and dopamine in NAc; both of these are hypothesized to contribute to an active coping response by encouraging exploration (noradrenaline in PFC) and active removal of the stressor (stimulation of D_2 receptors in NAc). Of particular note is that insufficient exploration can lead to persistent miscalibration [258]. That is, if the subject fails to explore, for instance because it believes the aversive stimulus to be insufficiently controllable, then it would never discover that it actually might be removed. Such a belief could result from a computational-level calculation about generalization from prior experience (as in learned helplessness; [31, 32]). At a different level of explanation, insufficient stimulation of D_2 receptors, leading to a lack of inhibition of passive defensive mechanisms, could readily have the same consequence.

Relevant to the issue of exploration and dopamine's possible involvement is the topic of anxiety. Fear and anxiety can be differentiated both by the behaviours they characteristically involve and their sensitivity to pharmacological challenge [259, 260]. Experimental assays of anxiety typically involve pitting the motivation to approach/explore novel situations against the motivation to avoid potential hazards [261]. According to one influential theory, it is exactly the function of anxiety in such cases of approach-avoidance conflict to move the animal towards potential danger, the better to assess risk [26, 259]. Not only is this thought to involve suppression of incompatible defensive responses, but also stimulation of approach; the associated 'behavioural approach system' is associated with NAc and its modulation by dopamine [259]. It would be interesting to consider a recent Bayesian decision-theoretic view of anxiety [262] that focuses on the opposite aspect, namely behavioural inhibition when there is no information to be gathered, and consider potential anti-correlations with dopaminergic modulation of the NAc.

In addition to evidence that some dopamine cells show phasic excitation in response to an aversive US, we also noted evidence from microdialysis studies for enhanced dopamine release in response to an aversive US over longer periods of time. What is the aversive parallel of the suggestion in the appetitive case that tonic dopamine levels, particularly in NAc, reflect an average reward rate which realizes the opportunity cost for acting slowly [9]? In aversive situations, the average reward rate is never strictly positive but, at least intuitively, time spent not actively engaged in a course of appropriate defensive action could be very costly indeed. For example, if an animal has just detected the presence of a predator, time spent not engaged in a course of defensive action could cost the potential safety that has thereby been missed.

Such considerations indicate the incompleteness of this previous account of tonic dopamine levels. In particular,

dovetailing with our suggestions regarding phasic dopamine above, the suggested mapping of tonic dopamine to the average rate of reward needs to be broadened to include the potentially-achievable rate of safety [252] which, assuming a prior expectation of controllability, will be positive. This provides a possible explanation for why increased tonic dopamine concentrations have been observed in microdialysis studies in response to an aversive US. However, if the aversive US is inescapable or uncontrollable, then the potentially-achievable rate of safety reduces to nothing. Thus, the tonic release of dopamine would also be expected to decrease. This is consistent with evidence already mentioned that the initial increase in tonic NAc dopamine level dissipates over time, giving way to an eventual fall below baseline levels [238–241].

Pavlovian conditioned defence

In relation to conditioning in aversive settings, similar complexities arise due to the fact that learning is likely to result in both aversive (i.e. danger-predicting) and appetitive (i.e. safety-predicting) conditioned stimuli, and may promote passive or active defensive strategies. Again, we use dopamine as a medium through which to view these complexities, with its preferential attachment to single sides of these dichotomies.

Fear conditioning

Conditioning in the aversive case, where animals are exposed to cues predictive of aversive outcomes, is generally known as *fear conditioning* due to the constellation of physiological and behavioural responses that the aversive CS comes to evoke. As in the appetitive case, conditioned and unconditioned responses need not be the same. Take, for instance, the case of conditioning a rat to a footshock US [263]. Here, the predominant response of the rat on exposure to the environment where it has received footshocks in the past, i.e. the CR, is to freeze. By contrast, the immediate response elicited by the shock itself, i.e. the UR, is a vigorous burst of activity. Furthermore, there can be model-based, outcome-specific, predictions allowing tailored responses (e.g., [264]) as well as model-free, outcome-general, predictions leading to generic preparatory responses such as behavioural inhibition.

The intricacies of how CR and UR relate to each other, which are arguably greater in the case of fear conditioning where these may be in conflict, may explain some of the difficulties in explicating dopamine's role in fear conditioning. A role for dopamine in fear conditioning seems to be generally accepted, though there is less consensus on the exact nature of this role (reviews include [265–267]).

Electrophysiological studies report that a substantial fraction (35–65 %) of putative dopamine neurons are activated by an aversive CS which is interleaved with an appetitive CS, a fraction that even exceeds the frequency (<15 %) of activations in response to an aversive US [191]. However, it has been suggested that many, though not all, of these activations may reflect 'false aversive responses', arising principally from generalization from appetitive to aversive CSs of the same sensory modality [191]. Additionally, an aversive CS may allow the animal to reduce the impact of an aversive US or avoid it entirely, and so in effect act as an instrumental 'safety signal', predicting a relatively benign outcome given a suitable defensive strategy. For example, a CS which predicts an aversive airpuff may facilitate a well-timed blink, thereby reducing the airpuff's aversiveness [268]. This fits with the idea, mentioned above, that dopaminergic responses may be instigated by predicted safety, or a relative improvement in expected state of affairs.

Regardless of the interpretation of such activations of dopamine cells by aversive CSs, this activity appears to play a role in fear conditioning. For example, Zweifel et al. [269] have recently shown that disruption of phasic bursting by dopamine neurons via inactivation of their NMDA receptors impairs fear conditioning in mice. These mice apparently develop a 'generalized anxiety-like phenotype', which the authors ascribe to the animals' failure to learn the correct contingencies.

Similar to observations in microdialysis studies of an increase in NAcS dopamine following an aversive US, enhanced NAcS dopamine release is also observed following presentation of an aversive CS [216]. Such enhanced release in NAcS to the onset of an aversive CS is corroborated by a recent FSCV study [270], though the opposite effect—decreased release—was observed in NAcC. Another recent FSCV study suggests that whether an increase or decrease in NAcC dopamine release is observed following an aversive CS depends critically on the animal's ability to avoid the predicted US [271]. Thus, Oleson et al. [271] found that, when trained in a fear conditioning paradigm—where the aversive US (a shock) was necessarily inescapable—presentation of the CS led to a decrease in NAcC dopamine. By contrast, in a conditioned avoidance paradigm—where the animal could potentially avoid the shock—both decreases *and* increases in NAcC dopamine were observed: an *increase* on trials in which animals successfully avoided shock, but a *decrease* on trials in which animals failed to avoid shock.

Dopamine receptor subtypes appear to play distinct roles. There is some consensus that D_1 receptor agonists and antagonists respectively promote or impede learning and expression in fear conditioning paradigms, while

the effect of D_2 manipulations is less clear [265, 267]. One study found that fear-potentiated startle could be restored in dopamine-deficient mice by administration of L-Dopa immediately following fear conditioning, but required intact signalling of D_1 receptors but not of D_2 receptors (although other members of the D_2-like family of receptors were reportedly required; [272]). Consistent with this finding, it has been reported recently that striatal-specific D_1 receptor knock-out mice, but not striatal-specific D_2 receptor knock-out mice, exhibit strongly impaired contextual fear conditioning [273]. Combined with evidence from previous fear conditioning studies [267, 274–277], as well as extensive evidence from the conditioned avoidance literature (see below), it appears that D_2 receptor manipulations affect only the *expression* of conditioned fear, rather than the learning of the association between aversive CS and US. This is consistent with experimental results in appetitive Pavlovian conditioning reviewed above, which suggest that D_1 receptors are particularly important in learning the CS-US contingency, while both D_1 and D_2 receptors are involved in modulating expression of this learning. Further, it has been reported recently that disruption of dopamine signalling in NAcC, but not NAcS, attenuated the ability of an aversive CS to block secondary conditioning of an additional CS, suggesting differential involvement of these areas [278].

Safety conditioning

The situation in which a CS predicts the *absence* of a US is usually known as 'conditioned inhibition' [279, 280]. In the particular case where the predicted absence is of an aversive US, the CS is called a *safety signal* [281, 282]. In considering aversive USs, we previously discussed hard-wired signals for safety—i.e., the absence of danger or threat. By contrast, here we consider previously neutral stimuli whose semantic association with safety is learned.

Such safety signals are capable of inhibiting fear and stress responses, and are known to have rewarding properties. For example, safety signals have been shown to act as conditioned reinforcers of instrumental responses [283]. This is consistent with the proposal of Konorski [59] and subsequent authors [199, 284, 285] that aversive and appetitive motivation systems reciprocally inhibit each other. The idea is that inhibition of the aversive system by a safety signal leads to disinhibition of the appetitive system, and so a safety signal is functionally equivalent to a CS that directly excites the appetitive system.

Neuroscientific study of safety signals is, however, at a relatively early stage (for reviews, see [281, 282]). Studies have identified neural correlates of learned safety in the amygdala [286–288] and striatum [286, 289].

Involvement of dopamine within NAcS in mediating the ability of the safety signal to inhibit fear, and consequently its ability to act as a conditioned reinforcer, is suggested by a recent study [290]. In particular, it was found that both infusion of d-amphetamine, an indirect dopamine agonist, and blockade of D_1/D_2 receptors in NAcS—but not in NAcC—disrupted the fear-inhibiting properties of a safety signal. While this finding implicates a role of NAcS in mediating the impact of the safety signal, why these manipulations had similar, as opposed to contrasting, effects is not clear.

Instrumental defence: learning to avoid

The final form of learning we consider in detail is instrumental avoidance. This is a rich paradigm that involves many of the behaviours and learning processes that we have discussed so far: innate defence mechanisms, fear conditioning, safety conditioning, and instrumental learning (cf. [291]). Furthermore, a role of dopamine in active avoidance, and D_2 receptors in particular, has long been suggested by the fact that dopamine antagonists interfere with avoidance learning [34, 292]. Indeed, such interference led to this paradigm being used to screen dopamine antagonists for antipsychotic activity [10, 12, 248]. Finally, the two-factor theory of active avoidance [17, 200, 201] that we discuss below was actually the genesis of the explanation we have been giving for the ready engagement of dopamine in the case of aversion.

The problem of avoidance and two-factor theory

A typical avoidance learning experiment involves placing an animal (e.g., a rat) in an environment in which a warning signal (e.g., a tone) predicts future experience of an aversive US (e.g., a shock) unless the animal performs a timely instrumental avoidance response (e.g., shuttling to a different location, or pressing a lever). That animals successfully learn to avoid under such conditions posed a problem that concerned early learning theorists [18]: how can the nonoccurrence of an aversive event—a ubiquitous condition—act as a behavioural reinforcer?

A solution to this 'problem of avoidance' has long been suggested in the form of a two-factor theory [17, 59, 200, 293–296]. The name 'two-factor' refers to the hypothesis that two behavioural factors or processes—Pavlovian and instrumental—are involved in the acquisition of conditioned avoidance. Firstly, the warning signal comes to elicit a state of fear through its predictive relationship with the aversive US. Thus, the first factor of the theory refers to the Pavlovian process of fear conditioning. This Pavlovian process then allows the second factor to come into play: if the animal now produces an action leading to the cessation of the warning stimulus, the animal enters a state of relief, or reduced fear, capable of reinforcing the

avoidance response. Thus, the second factor refers to an instrumental process by which the avoidance response is reinforced through fear reduction or relief. Such an account can also include stimuli dependent on the avoidance response and which are anticorrelated with the aversive US, thereby becoming predictive of safety [201]. As discussed, these safety signals (SS) themselves are thought to be capable of inhibiting conditioned fear [280], thereby both preventing Pavlovian fear responses (e.g., freezing) which may interfere with the instrumental avoidance response and reinforcing safety-seeking behaviours in fearful states or environments [294, 296], consistent with theories of opponent motivational processes [59, 285].

Avoidance, innate defence, and controllability

As mentioned above, the importance of innate defensive behaviours in the avoidance context has long been noted. Bolles [27], highlighting the importance of such 'species-specific defense reactions', argued that if an avoidance behaviour is rapidly acquired, this is because the required avoidance response coincides with the expression of an innate defensive response by the animal, rather than reflecting a learning process; how difficult the animal finds the avoidance task will depend on the extent to which the avoidance response is compatible with its innate defensive repertoire. In turn, which innate behaviour the animal selects will be sensitive to relevant features of the avoidance situation, such as whether there is a visible escape route or not, reminiscent of Tolman's [297] notion of behavioural support stimuli [206, 298].

Just as in the appetitive case, conflict between such Pavlovian behaviours and instrumental contingencies can lead to apparently maladaptive behaviour, albeit in rather unnatural experimental settings. Thus, Seymour et al. [299] highlight experiments in which self-punitive behaviour arises when an animal is (instrumentally) punished for emitting Pavlovian responses in response to that punishment. In one such unfortunate case, squirrel monkeys were apparently unable to decrease the frequency of painful shocks delivered to them by suppressing their shock-induced tendency to pull at a restraining leash attached to their collar; pulling on the restraining leash was exactly the action that hastened arrival of the next shock [300].

Similarly, just as it has been suggested that the animal's appraisal of whether a threat is escapable or not is crucial in determining its defensive strategy in general (e.g., [25]), it was famously shown that the controllability of an aversive US is crucial in determining subsequent avoidance learning performance [301, 302]. In particular, dogs exposed to inescapable shocks in a first environment showed deficits in initiating avoidance or escape responses in a second environment, even though the aversive US was now escapable. This, of course, led to the concept of 'learned helplessness' [237]. Huys and Dayan [32] presented a model-based account of learned helplessness, arguing that the generalization between environments affected the value of exploration, thereby leading to persistent miscalibration.

The issue of model-free versus model-based influences has received rather less attention in aversive than appetitive contexts. However, sensitivity to revaluation of aversive USs in the context of instrumental avoidance has been demonstrated in rats [303, 304] and humans [305, 306], indicating model-based influences under at least some avoidance conditions. Fernando et al. [303] have recently reported that revaluation of a shock US induced by pairing shock with systemic analgesics (morphine or D-amphetamine), leading rats subsequently to decrease their rate of avoidance responding, could also be achieved by pairing the shock with more selective infusions of a mu-opioid agonist into either NAcS or PAG. Involvement of NAcS and related structures in revaluation in this instance is consistent with the idea that the shell is involved in model-based prediction [56].

Dopamine, D_2 receptors, and active avoidance

Two-factor theories of avoidance fit well with the idea that the striatum, in interaction with dopamine, implements an actor-critic algorithm [21, 22, 78, 80, 169, 307]. Thus, an initial period of learning by the critic (in the ventral striatum) of negative state values (i.e., fear conditioning) allows subsequent instrumental training of an avoidance response by the actor (in the dorsal striatum), since actions leading to the unexpected non-delivery of the aversive US are met with a positive prediction error ('better than expected'), as signalled by dopamine neurons in the midbrain.

It was the abilities of certain antipsychotic drugs and neurotoxic lesions to produce active avoidance learning deficits [248, 292] that suggested a critical role for dopamine in the acquisition of conditioned avoidance. Furthermore, localised neurotoxic lesions suggested that dopamine projections to both dorsal and ventral striatum were required for acquisition of active avoidance [308, 309], corroborated by more recent work on selective restoration of dopamine signalling in dopamine-deficient mice [310]. This is consistent with complementary roles of actor (ventral striatum) and critic (dorsal striatum) in adapting behaviour.

Dopamine's action on D_2 receptors appears of particular importance for this. Evidence from active avoidance studies suggests that while blocking D_2 receptors leaves fear conditioning intact, instrumental learning of the avoidance response requires intact D_2 signalling [292, 311, 312]. From the perspective of the actor-critic, one

might conclude that blockade of D_2 receptors therefore does not interfere with the learning of negative state values by the critic but does interfere with the learning of the actor [22]. The finding that D_2 receptor blockade leaves conditioning to aversive stimuli intact in the active avoidance setting is consistent with evidence from fear conditioning studies (see above). Furthermore, that D_2 blockade also disrupts instrumental learning is consistent with dopamine's modulation of direct and indirect pathways in the dorsal striatum, as in Frank's [137] model, since this would be expected to lead to a relative strengthening of the indirect, 'NoGo' pathway and impede acquisition of the appropriate 'Go' response (albeit leaving this model without a means of implementing the preserved fear conditioning). However, this would raise the question of why D_2 receptors within dorsal striatum should be implicated more strongly in learning than are those in ventral striatum. A pertinent observation might be a distinction between the longevity of the effects of tying optogenetic stimulation of (D_1-expressing) dMSNs and (D_2-expressing) iMSNs in the dorsomedial striatum [183] of mice. These authors triggered activation of one or other pathway when the mouse made contact with one of two touch sensors. dMSN stimulation increased preference for its associated lever, whereas iMSN stimulation decreased it. However, whereas the positive preference persisted in extinction throughout a test period, the negative preference rapidly disappeared. Furthermore, it was noted that stimulation of iMSNs elicited brief, immediate freezing followed by an 'escape response', though these behavioural changes were not thought sufficient to explain the bias away from the laser-paired trigger.

Nevertheless, while many findings accord well with an actor-critic account of avoidance learning, there are at least two omissions in such accounts that require correction. Firstly, similar to Bolles' complaints about two-factor theory, actor-critic accounts have largely ignored the role for innate (i.e., Pavlovian) defence mechanisms. Secondly, the key factor of controllability has not been fully integrated with actor-critic models.

Indeed, disruption of innate defensive behaviour by D_2 blockade occurs as well as disruption of instrumental learning of the active avoidance response. There are suggestions that suppression of conditioned avoidance may rely more on disruption of D_2 signalling within ventral, rather than dorsal, striatum [248], consistent with interference with Pavlovian ('critic') rather than instrumental ('actor') processes. For example, post-training injection of a D_2 antagonist into NAcS, but not into dorsolateral striatum, leads to a relatively immediate suppression of a conditioned avoidance response [313]. As we saw above, NAcS, under dopaminergic modulation, is implicated in controlling expression of innate defensive behaviours,

and D_2 activation appears to promote active defensive strategies. Similarly, there is evidence that D_2 blockade leads to enhanced freezing responses—arguably, a more passive form of defence—following footshock, interfering with rats' ability to emit avoidance responses [34, 274], though there remains some doubt about whether fear-induced freezing is an important factor in the disruption of conditioned avoidance [248]. In their review of the role of dopamine in avoidance learning, and defence more generally, Blackburn and colleagues [34] suggest that D_2 blockade does not disrupt defensive behaviour globally but rather 'changes the probability that a given class of defensive response will be selected' ([34], p. 267), in particular increasing the probability of freezing.

In relation to controllability, we have already referred to evidence that exposure to chronic, inescapable stress abolishes stress-induced increases in the concentration of accumbens dopamine [238–241]. Such evidence has led Cabib and Puglisi-Allegra [209–211] to suggest that whether an increase or decrease in accumbens dopamine levels is observed in response to stress depends on whether the stressor is appraised as controllable (increase) or not (decrease). This dissipation of the dopamine response does not appear to be explained by dopamine depletion, since subsequent release from the chronic stressor leads to a large, rapid increase in dopamine concentration [239]. Similarly, Cabib and Puglisi-Allegra [209], using a yoked paradigm in which one of a pair of animals (the 'master') has some control over the amount of shock experienced by means of an escape response while the other ('yoked') animal does not, found evidence consistent with elevated and inhibited NAc dopamine in master and yoked animals, respectively, after an hour of shock exposure.

More recently, Tye et al. [314] used optogenetics to assess the effects of exciting or inhibiting identified VTA dopamine cells in certain rodent models of depression involving inescapable stressors (tail suspension, forced swim, and chronic mild stress paradigms). While optogenetic inhibition of these dopamine cells could induce behaviour that has been related to depression, such as reduced escape attempts, optogenetic activation of the same cells was found to rescue depression-like phenotypes (e.g., promoting escape-related behaviours) induced by chronic stress. Furthermore, it was observed that chronic stress led to a reduction in measures of phasic VTA activity. This latter observation contrasts with studies using repeated social defeat stress, where phasic VTA activity has typically been observed to increase in 'susceptible' animals [315–317]. Apparently contradictory findings regarding stress-induced changes in VTA dopamine activity, and indeed the effects of manipulating this via optogenetic stimulation, might stem from the

subtleties of the different paradigms used, but may also reflect heterogeneity in the properties of different VTA dopamine cells, such as between those projecting to mPFC versus NAc (for a recent discussion of these issues, see [318]).

While there is evidence from microdialysis studies that support a link between controllability, defensive strategy, and tonic NAc dopamine, it should be noted that not all such evidence points in this direction. For example, Bland et al. [319] measured both dopamine and serotonin release in NAcS of rats in the yoked pairs paradigm referred to previously. While they did report a trend for increased dopamine release relative to no-shock controls, this increase was neither significant nor differed between master and yoked animals. By contrast, serotonin levels were found to be significantly increased in yoked animals during and after stress exposure, relative to master and no-shock control animals [319]. Experiments using the same paradigm but taking measurements from mPFC found elevated levels of both dopamine and serotonin in yoked animals compared to master and no-shock controls [320].

These latter studies and others [321–323] highlight that consideration of other neuromodulators, notably serotonin, is crucial for a fuller understanding of defensive behaviour. A role of serotonin has long been suggested both in the particular case of active avoidance [312, 324] and in defence more generally [249, 250]. As mentioned, one suggestion is that the putative opponency between appetitive and aversive motivation systems [59, 254, 285] is at least partly implemented in opponency between dopamine and serotonin, respectively [23, 249–251, 256]. A specific computational model of this idea was suggested by Daw et al. [255], and Dayan [252] has more recently considered such opponency in the particular case of active avoidance. However, a modulatory role of controllability in the active avoidance setting has not yet been fully integrated into RL models.

Conclusions

Here, we have discussed unconditioned/conditioned, Pavlovian/instrumental, and passive/active issues associated with aversion. We used dopamine, and particularly its projection into the striatum and the D_2 system, as a form of canary, since the way that dopamine underpins model-free learning, and model-free and model-based vigour, turns out to be highly revealing for the organization of aversive processing. Our essential explanatory strategy rested on three concepts: safety, opponency, and controllability.

When under threat, safety is a desirable state. We suggested that the prospect of possible future safety underlies positive dopamine responses—both tonic and phasic—in response to aversive stimuli. Indeed, the interpretation of these responses is very similar to the more obviously appetitive case involving rewards, since safety is an appetitive outcome. Thus, phasic activation of dopamine cells in response to an aversive stimulus can be interpreted in TD terms as an 'unpredicted predictor of future safety'. Similarly, increased levels of tonic dopamine in conditions of stress, particularly in NAc, can be interpreted as signalling a potentially-achievable rate of safety.

Of course, what makes safety a more subtle concept is that it is relative; it is defined in opposition to danger. Dangerous states are not, in general, good states, which is why, in opposition to an appetitive process directed at safety in such states, there should be an aversive system which signals the disutility of occupying dangerous states. Therefore, positive dopamine responses which putatively signal the appetitive component of a TD prediction error in such states can only be part of the story—an *opponent* signal is required, marking the value of the path that will (hopefully) not be taken, and providing a new baseline against which to measure outcomes. This results in a form of counterfactual learning signal, a quantity that has also been investigated in purely appetitive contexts, and may have special relevance to the dorsal, rather than the ventral, striatum [325–328].

Unfortunately, while the notion of opponent appetitive and aversive processes is long-standing [59, 253, 254, 285], we still know relatively little about their neural realization. As mentioned, one idea is that this opponency maps to dopamine (appetitive) and serotonin (aversive) signalling [23, 249–251, 256], and specific computational models of this idea have been advanced [252, 255]. Recent attention to electrophysiological recordings from identified serotonergic cells in conditions of reward and punishment is particularly welcome in this regard, albeit offering no comfort to these theoretical ideas [329], and we look forward to further work which leverages advances in neuroscientific techniques to clarify the neural substrate of opponency.

Whether safety is appraised as achievable or not appears to be crucial, hence our appeal to the concept of controllability. We reviewed evidence that tonic levels of dopamine are modulated downwards over time with chronic exposure to aversive stimuli. Further, we reviewed evidence that dopamine, and NAc D_2-receptor stimulation in particular, modulates active versus passive defensive strategies (or perhaps better, defensive versus recuperative behaviours). Modulation of dopamine in this way raises pressing questions about controllability at both more and less abstract levels. Indeed, even formalizing an adequate concept of behavioural control in the first instance is nontrivial [32].

The concept of controllability brings model-based and model-free considerations back into focus since, at least intuitively, this concept seems to imply explicit knowledge of action-outcome statistics in the current environment. In relation to dopamine, this is consistent with evidence that a model-based system could potentially influence model-free learning and performance via the dopaminergic TD signal. However, implementation of heuristics aimed at optimizing the exploration-exploitation trade-off, such as possibly instantiated in a dopaminergic exploration bonus, may provide a model-free proxy for controllability. Thus, further work is required to disentangle the relative contributions of model-based and model-free systems in modulating dopamine signals which, in turn, modulate defensive strategy.

We have focused on dopamine in the accumbens at the expense of other areas—notably the amygdala and mPFC—which are of clear relevance to the themes discussed. For example, intact dopamine signalling in the amygdala, as well as in the striatum, appears to be necessary for acquisition of active avoidance behaviour [310], with the central nucleus particularly implicated in mediating conditioned freezing responses that may interfere with active responding [330]. Indeed, there is evidence that D_2 receptors are particularly prevalent in the central amygdala [265, 331], and a recent review [332] suggests that a key role of D_2 receptors in the central nucleus is to modulate reflex-like defensive behaviours organised in the brain stem. This clearly relates to the proposed importance of D_2 in modulating Pavlovian defence discussed here. Similarly, it is known that stress-induced increases in accumbens dopamine release is constrained by activation of D_1 receptors in mPFC, with both mPFC dopamine depletion or blockade of D_1 receptors leading to enhanced stress-induced accumbens release of dopamine (see [211], and references therein). Furthermore, mPFC is thought to be a key player in the appraisal of whether a stressor is under the animal's control [323].

Throughout the review, we have highlighted various issues that merit experimental and theoretical investigation. Experimentally, the most pressing issue is perhaps heterogeneity in the dopamine system—arriving at a clearer view of the potentially separate roles and activation of different groups of dopamine neurons, and reconciling activation and release. Technical advances have allowed increasingly sophisticated attacks on this issue, though a consensus regarding the degree of heterogeneity, both in terms of activity [333, 334] and connectivity [227, 335, 336], has yet to emerge. To the extent that dopamine neurons with different affective receptive fields project to different targets, there is no need for a shared semantics for their activation [23]. However, if reward and punishment-activated dopamine neurons are interdigitated in the way suggested by some experiments [333], then there is a need for a functional analysis as to how downstream systems might be able to interpret the apparently confusing patterns of dopamine release. One speculation in the former case, e.g. if dopamine cells in the ventral versus dorsal VTA showing different responses to aversive stimuli [223] also differentially project to more ventral versus dorsal regions of striatum, respectively, is that this reflects competing objectives to (a) shape (instrumental) policy *retrospectively*, by assigning (dis)credit to actions that may have led to aversive outcomes, and (b) to promote suitable (Pavlovian) behaviour *prospectively* in light of possible future safety.

Similarly, it would be important to understand the true degree of separation between putative direct and indirect pathways in the core and shell of the accumbens. Heterogeneity in the serotonin system, and its interactions with dopamine in the case of aversion, would also merit investigation. A recent revealing analysis of active avoidance in the zebrafish, showing the critical involvement of a pathway linking the lateral habenula to the median raphe [324] is of importance, particularly since most of the recent studies of optogenetically tagged or manipulated 5-HT neurons have focused instead on the dorsal raphe [329, 337–339]. Integrating the whole array of data on patience, satiety, motor action, behavioural inhibition and aversion associated with 5-HT is a major task.

From a more behavioural viewpoint, it would be interesting to get a clearer view of the scope of model-based aversive conditioning. For instance, take the experiment showing that D_2 blockade does not arrest learning aversive predictions even though it does avoidance responses [311]: it is not clear why model-based predictions would not be capable of generating appropriate avoidance behaviour as soon as the D_2 antagonist is washed out—rather leaving it to be acquired slowly, as if it was purely model-free.

Another important experimental avenue is to try and integrate the processing of costs (and indeed, for humans, outcomes such as financial losses) with that of actual punishment. Costs, which could be either physical or mental [340–342], also exert a negative force on behaviour, and indeed also have a slightly complicated relation to dopamine activation and release [12, 16, 343, 344].

From a more theoretical viewpoint, perhaps the most urgent question concerns pinning down the different facets of controllability, the way that these determine operations such as exploration, and relative model-based and model-free influences. Entropy and reachability of outcomes were considered by [32], but other definitions are possible. Work on learned helplessness suggests a key role for the mPFC in suppressing otherwise exuberant 5-HT activity in animals who have the benefit of

behavioural control—but what exactly mPFC is reporting is unclear.

A further direction is to construct a more comprehensive theory of aversive vigour [252] looked at this in a rather specific set of experimental circumstances. The direct predictions arising even from this have not been thoroughly tested; but a more general theory, also tied to controllability, would be desirable.

Finally, we have noted various structural asymmetries between appetitive and aversive systems, ascribing many of them to asymmetric priors about the structure of rewards and punishments in environments [23]. It would be important to examine these claims in more detail, and indeed look at the effect of changing the statistics of environments to determine the extent of lability.

In conclusion, we have attempted to use our evolving and rich view of the nature and source of learned, appetitive behaviour to examine the case of aversion and defence. Along with substantial commonalities between the two, we have discussed some critical differences—notably in the way that aversive behaviour appears to piggy-back on appetitive processing, leading to various intricate complexities that are incompletely understood. Dopamine plays a number of critical and apparently confounded roles; we therefore used it to lay as bare as possible the extent and limits of our current understanding.

References
1. Schultz W. Neuronal reward and decision signals: from theories to data. Physiol Rev. 2015;95(3):853–951.
2. Kim HF, Hikosaka O. Parallel basal ganglia circuits for voluntary and automatic behaviour to reach rewards. Brain. 2015;138(7):1776–800.
3. Chase HW, Kumar P, Eickhoff SB, Dombrovski AY. Reinforcement learning models and their neural correlates: an activation likelihood estimation meta-analysis. Cogn Affect Behav Neurosci. 2015;15(2):435–59.
4. Ikemoto S, Bonci A. Neurocircuitry of drug reward. Neuropharmacology. 2014;76:329–41.
5. Lee D, Seo H, Jung MW. Neural basis of reinforcement learning and decision making. Annu Rev Neurosci. 2012;35:287–308.
6. Daw ND, Dayan P. The algorithmic anatomy of model-based evaluation. Philos Trans R Soc Lond B Biol Sci. 2014;369(1655):20130478.
7. O'Doherty JP. Contributions of the ventromedial prefrontal cortex to goal-directed action selection. Ann N Y Acad Sci. 2011;1239(1):118–29.
8. Frank MJ, Claus ED. Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. Psychol Rev. 2006;113(2):300–26.
9. Niv Y, Daw ND, Joel D, Dayan P. Tonic dopamine: opportunity costs and the control of response vigor. Psychopharmacology. 2007;191(3):507–20.
10. Salamone JD. The involvement of nucleus accumbens dopamine in appetitive and aversive motivation. Behav Brain Res. 1994;61(2):117–33.
11. Salamone JD, Correa M. Motivational views of reinforcement: implications for understanding the behavioral functions of nucleus accumbens dopamine. Behav Brain Res. 2002;137(1):3–25.
12. Salamone JD, Correa M. The mysterious motivational functions of mesolimbic dopamine. Neuron. 2012;76:470–85.
13. Guitart-Masip M, Beierholm UR, Dolan R, Duzel E, Dayan P. Vigor in the face of fluctuating rates of reward: an experimental examination. J Cogn Neurosci. 2011;23(12):3933–8.
14. Beierholm U, Guitart-Masip M, Economides M, Chowdhury R, Düzel E, Dolan R, Dayan P. Dopamine modulates reward-related vigor. Neuropsychopharmacology. 2013;38:1495–503.
15. Floresco SB. The nucleus accumbens: an interface between cognition, emotion, and action. Annu Rev Psychol. 2015;66:25–52.
16. Hamid AA, Pettibone JR, Mabrouk OS, Hetrick VL, Schmidt R, Vander Weele CM, Kennedy RT, Aragona BJ, Berke JD. Mesolimbic dopamine signals the value of work. Nat Neurosci. 2016;19(1):117–26.
17. Mowrer OH. A stimulus-response analysis of anxiety and its role as a reinforcing agent. Psychol Rev. 1939;46:553–65.
18. Bolles RC. The avoidance learning problem. Psychol Learn Motiv. 1972;6:97–139.
19. Grossberg S. A neural theory of punishment and avoidance, I: qualitative theory. Math Biosci. 1972;15(1):39–67.
20. Johnson JD, Li W, Li J, Klopf AH. A computational model of learned avoidance behavior in a one-way avoidance experiment. Adapt Behav. 2001;9(2):91–104.
21. Maia TV. Two-factor theory, the actor-critic model, and conditioned avoidance. Learn Behav. 2010;38(1):50–67.
22. Moutoussis M, Bentall RP, Williams J, Dayan P. A temporal difference account of avoidance learning. Network. 2008;19(2):137–60.
23. Boureau YL, Dayan P. Opponency revisited: competition and cooperation between dopamine and serotonin. Neuropsychopharmacology. 2010;36(1):74–97.
24. Guitart-Masip M, Duzel E, Dolan R, Dayan P. Action versus valence in decision making. Trends Cogn Sci. 2014;18(4):194–202.
25. Bandler R, Keay KA, Floyd N, Price J. Central circuits mediating patterned autonomic activity during active vs. passive emotional coping. Brain Res Bull. 2000;53(1):95–104.
26. McNaughton N, Corr PJ. A two-dimensional neuropsychology of defense: fear/anxiety and defensive distance. Neurosci Biobehav Rev. 2004;28:285–305.
27. Bolles RC. Species-specific defense reactions and avoidance learning. Psychol Rev. 1970;77(1):32–48.
28. Blanchard RJ, Flannelly KJ, Blanchard DC. Defensive behaviors of laboratory and wild rattus norvegicus. J Comp Psychol. 1986;100(2):101–7.
29. Mobbs D, Kim JJ. Neuroethological studies of fear, anxiety, and risky decision-making in rodents and humans. Curr Opin Behav Sci. 2015;5:8–15.
30. Maier SF, Amal J, Baratta MV, Paul E, Watkins LR. Behavioral control, the medial prefrontal cortex, and resilience. Dialogues Clin Neurosci. 2006;8(4):397–406.
31. Maier SF, Watkins LR. Stressor controllability and learned helplessness: the roles of the dorsal raphe nucleus, serotonin, and corticotropin-releasing factor. Neurosci Biobehav Rev. 2005;29(4):829–41.
32. Huys QJ, Dayan P. A Bayesian formulation of behavioral control. Cognition. 2009;113(3):314–28.
33. Frank MJ, Fossella JA. Neurogenetics and pharmacology of learning, motivation, and cognition. Neuropsychopharmacology. 2011;36(1):133–52.
34. Blackburn JR, Pfaus JG, Phillips AG. Dopamine functions in appetitive and defensive behaviours. Prog Neurobiol. 1992;39:247–79.

35. Brooks AM, Berns GS. Aversive stimuli and loss in the mesocorticolimbic dopamine system. Trends Cogn Sci. 2013;17(6):281–6.
36. Holly EN, Miczek KA. Ventral tegmental area dopamine revisited: effects of acute and repeated stress. Psychopharmacology. 2016;233(2):163–86.
37. Lammel S, Lim BK, Malenka RC. Reward and aversion in a heterogeneous midbrain dopamine system. Neuropharmacology. 2014;76:351–9.
38. McCutcheon JE, Ebner SR, Loriaux AL, Roitman MF. Encoding of aversion by dopamine and the nucleus accumbens. Front Neurosci. 2012;6:137.
39. Pignatelli M, Bonci A. Role of dopamine neurons in reward and aversion: a synaptic plasticity perspective. Neuron. 2015;86(5):1145–57.
40. Schultz W. Dopamine reward prediction-error signalling: a two-component response. Nat Rev Neurosci. 2016;17:183–95.
41. Sutton RS, Barto AG. Reinforcement learning: an introduction. Cambridge: MIT Press; 1998.
42. Kaelbling LP, Littman ML, Moore AW. Reinforcement learning: a survey. J Artif Intell Res. 1996;4:237.
43. Doya K. What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? Neural Netw. 1999;12(7):961–74.
44. Daw ND, Niv Y, Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat Neurosci. 2005;8(12):1704–11.
45. Dickinson A, Balleine BW. The role of learning in motivation. In: Gallistel CR, editor. Steven's handbook of experimental psychology. New York: Wiley; 2002. p. 497–533.
46. Dolan RJ, Dayan P. Goals and habits in the brain. Neuron. 2013;80(2):312–25.
47. Bellman RE. Dynamic programming. Princeton: Princeton University Press; 1957.
48. Sutton RS. Learning to predict by the methods of temporal differences. Mach Learn. 1988;3(1):9–44.
49. Barto AG, Sutton RS, Anderson CW. Neuronlike adaptive elements that can solve difficult learning control problems. IEEE Trans Syst Man Cybern. 1983;13:835–46.
50. Watkins CJCH. Learning from delayed rewards. Ph.D. Thesis, University of Cambridge; 1989.
51. Dayan P. Exploration from generalization mediated by multiple controllers. In: Baldassare G, Mirolli M, editors. Intrinsically motivated learning in natural and artificial systems. Berlin: Springer; 2013. p. 73–91.
52. Howard RA. Information value theory. IEEE Trans Syst Sci Cybern. 1966;2:22–6.
53. Gittins JC. Bandit processes and dynamic allocation indices. J R Stat Soc. 1979;41(2):148–77.
54. Sutton RS. Integrated architecture for learning, planning, and reacting based on approximating dynamic programming. In: Porter BW, Mooney RJ, editors. Proceedings of the seventh international conference on machine learning. Morgan Kaufman Publishers, Inc. 1990. p. 216–24.
55. Dayan P, Sejnowski TJ. Exploration bonuses and dual control. Mach Learn. 1996;25(1):5–22.
56. Dayan P, Berridge KC. Model-based and model-free pavlovian reward learning: revaluation, revision, and revelation. Cogn Affect Behav Neurosci. 2014;14:473–93.
57. Craig W. Appetites and aversions as constituents of instincts. Biol Bull. 1918;34(2):91–107.
58. Sherrington C. The integrative action of the nervous system. New Haven: Yale University Press; 1906.
59. Konorski J. Integrative activity of the brain. Chicago: University of Chicago Press; 1967.
60. Baldo BA, Kelley AE. Discrete neurochemical coding of distinguishable motivational processes: insights from nucleus accumbens control of feeding. Psychopharmacology. 2007;191(3):439–59.
61. Cools R. Role of dopamine in the motivational and cognitive control of behavior. Neuroscientist. 2008;14(4):381–95.
62. Blackburn JR. The role of dopamine in preparatory and consummatory defensive behaviours. Ph.D. Thesis, University of British Columbia; 1989.
63. Nicola SM. The flexible approach hypothesis: unification of effort and cue-responding hypotheses for the role of nucleus accumbens dopamine in the activation of reward-seeking behavior. J Neurosci. 2010;30(49):16585–600.
64. Ikemoto S, Panksepp J. The role of nucleus accumbens dopamine in motivated behavior: a unifying interpretation with special reference to reward-seeking. Brain Res Rev. 1999;31:6–41.
65. Berridge KC, Robinson TE. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? Brain Res Rev. 1998;28(3):309–69.
66. Robbins TW, Everitt BJ. Functions of dopamine in the dorsal and ventral striatum. Semin Neurosci. 1992;4:119–27.
67. Williams DR, Williams H. Auto-maintenance in the pigeon: sustained pecking despite contingent non-reinforcement. J Exp Anal Behav. 1969;12:511–20.
68. Breland K, Breland M. The misbehavior of organisms. Am Psychol. 1961;16:681–4.
69. Dayan P, Niv Y, Seymour B, Daw ND. The misbehavior of value and the discipline of the will. Neural Netw. 2006;19(8):1153–60.
70. Colwill RM, Rescorla RA. Associations between the discriminative stimulus and the reinforcer in instrumental learning. J Exp Psychol Anim Behav Process. 1988;14(2):155–64.
71. Estes WK. Discriminative conditioning. I: a discriminative property of conditioned anticipation. J Exp Psychol. 1943;32:150–5.
72. Holland PC. Relations between pavlovian-instrumental transfer and reinforcer devaluation. J Exp Psychol Anim Behav Process. 2004;30(2):104–17.
73. Lovibond PF. Facilitation of instrumental behavior by a pavlovian appetitive conditioned stimulus. J Exp Psychol Anim Behav Process. 1983;9:225–47.
74. Rescorla RA, Wagner AR. A theory of pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: Black AH, Prokasy WF, editors. Classical conditioning II: current research and theory. New York: Appleton-Century-Crofts Ltd; 1972. p. 64–99.
75. Sutton R, Barto AG. Toward a modern theory of adaptive networks: expectation and prediction. Psychol Rev. 1981;88(2):135–70.
76. Sutton RS, Barto AG. Time-derivative models of pavlovian reinforcement. In: Gabriel M, Moore J, editors. Learning and computational neuroscience: foundations of adaptive networks. Cambridge: MIT Press; 1990. p. 497–537.
77. Dayan P, Kakade S, Montague PR. Learning and selective attention. Nat Neurosci. 2000;3:1218–23.
78. Montague PR, Dayan P, Sejnowski TJ. A framework for mesencephalic dopamine systems based on predictive hebbian learning. J Neurosci. 1996;16(5):1936–47.
79. Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. Science. 1997;275:1593–9.
80. O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. Science. 2004;304:452–4.
81. Calabresi P, Picconi B, Tozzi A, Di Filippo M. Dopamine-mediated regulation of corticostriatal synaptic plasticity. Trends Neurosci. 2007;30(5):211–9.
82. Chen BT, Hopf FW, Bonci A. Synaptic plasticity in the mesolimbic system. Ann N Y Acad Sci. 2010;1187(1):129–39.
83. Reynolds JNJ, Hyland BI, Wickens JR. A cellular mechanism of reward-related learning. Nature. 2001;413(6851):67–70.
84. Reynolds JNJ, Wickens JR. Dopamine-dependent plasticity of corticostriatal synapses. Neural Netw. 2002;15:507–21.
85. Shen W, Flajolet M, Greengard P, Surmeier DJ. Dichotomous dopaminergic control of striatal synaptic plasticity. Science. 2008;321:848–51.
86. Mogenson GJ, Jones DL, Yim CY. From motivation to action: functional interface between the limbic system and the motor system. Prog Neurobiol. 1980;14:69–97.
87. Cardinal RN, Parkinson JA, Hall J, Everitt BJ. Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. Neurosci Biobehav Rev. 2002;26(3):321–52.
88. Di Ciano P, Cardinal RN, Cowell RA, Little SJ, Everitt BJ. Differential involvement of NMDA, AMPA/kainate, and dopamine receptors in the nucleus accumbens core in the acquisition and performance of pavlovian approach behavior. J Neurosci. 2001;21(23):9471–7.
89. Flagel SB, Clark JJ, Robinson TE, Mayo L, Czuj A, Willuhn I, Akers CA, Clinton SM, Phillips PEM, Akil H. A selective role for dopamine in stimulus-reward learning. Nature. 2011;469(7328):53–7.

90. Parkinson JA, Dalley J, Cardinal R, Bamford A, Fehnert B, Lachenal G, Rudarakanchana N, Halkerston K, Robbins T, Everitt B. Nucleus accumbens dopamine depletion impairs both acquisition and performance of appetitive pavlovian approach behaviour: implications for mesoaccumbens dopamine function. Behav Brain Res. 2002;137(1):149–63.

91. Saunders BT, Robinson TE. The role of dopamine in the accumbens core in the expression of pavlovian-conditioned responses. Eur J Neurosci. 2012;36(4):2521–32.

92. Berridge KC. The debate over dopamine's role in reward: the case for incentive salience. Psychopharmacology. 2007;191:391–431.

93. Wise RA. Dopamine, learning and motivation. Nat Rev Neurosci. 2004;5(6):483–94.

94. McClure SM, Daw ND, Montague PR. A computational substrate for incentive salience. Trends Neurosci. 2003;26(8):423–8.

95. Dickinson A, Smith J, Mirenowicz J. Dissociation of pavlovian and instrumental incentive learning under dopamine antagonists. Behav Neurosci. 2000;114(3):468–83.

96. Hall J, Parkinson JA, Connor TM, Dickinson A, Everitt BJ. Involvement of the central nucleus of the amygdala and nucleus accumbens core in mediating pavlovian influences on instrumental behaviour. Eur J Neurosci. 2001;13(10):1984–92.

97. Lex A, Hauber W. Dopamine D1 and D2 receptors in the nucleus accumbens core and shell mediate Pavlovian-instrumental transfer. Learn Mem. 2008;15:483–91.

98. Stuber GD, Hnasko TS, Britt JP, Edwards RH, Bonci A. Dopaminergic terminals in the nucleus accumbens but not the dorsal striatum corelease glutamate. J Neurosci. 2010;30(24):8229–33.

99. Tecuapetla F, Patel JC, Xenias H, English D, Tadros I, Shah F, Berlin J, Deisseroth K, Rice ME, Tepper JM, et al. Glutamatergic signaling by mesolimbic dopamine neurons in the nucleus accumbens. J Neurosci. 2010;30(20):7105–10.

100. Zhang S, Qi J, Li X, Wang HL, Britt JP, Hoffman AF, Bonci A, Lupica CR, Morales M. Dopaminergic and glutamatergic microdomains in a subset of rodent mesoaccumbens axons. Nat Neurosci. 2015;18(3):386–92.

101. Moss J, Ungless MA, Bolam JP. Dopaminergic axons in different divisions of the adult rat striatal complex do not express vesicular glutamate transporters. Eur J Neurosci. 2011;33(7):1205–11.

102. Gläscher J, Hampton AN, O'Doherty JP. Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. Cereb Cortex. 2009;19(2):483–95.

103. Gottfried JA, O'Doherty J, Dolan RJ. Encoding predictive reward value in human amygdala and orbitofrontal cortex. Science. 2003;301(5636):1104–7.

104. Hatfield T, Han JS, Conley M, Gallagher M, Holland PC. Neurotoxic lesions of basolateral, but not central, amygdala interfere with pavlovian second-order conditioning and reinforcer devaluation effects. J Neurosci. 1996;16(16):5256–65.

105. Holland PC, Gallagher M. Amygdala circuitry in attentional and representational processes. Trends Cogn Sci. 1999;3(2):65–73.

106. Schoenbaum G, Chiba AA, Gallagher M. Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. Nat Neurosci. 1998;1(2):155–9.

107. Schoenbaum G, Chiba AA, Gallagher M. Neural encoding in orbitofrontal cortex and basolateral amygdala during olfactory discrimination learning. J Neurosci. 1999;19(5):1876–84.

108. Valentin VV, Dickinson A, O'Doherty JP. Determining the neural substrates of goal-directed learning in the human brain. J Neurosci. 2007;27(15):4019–26.

109. Dickinson A, Balleine B. Actions and responses: the dual psychology of behaviour. In: Eilan N, McCarthy RA, Brewer B, editors. Spatial representation: problems in philosophy and psychology. Oxford: Blackwell; 1993. p. 277–93.

110. Zahm DS, Brog JS. On the significance of subterritories in the "accumbens" part of the rat ventral striatum. Neuroscience. 1992;50(4):751–67.

111. Humphries MD, Prescott TJ. The ventral basal ganglia, a selection mechanism at the crossroads of space, strategy, and reward. Prog Neurobiol. 2010;90(4):385–417.

112. Kelley AE. Ventral striatal control of appetitive motivation: role in ingestive behavior and reward-related learning. Neurosci Biobehav Rev. 2004;27:765–76.

113. Voorn P, Vanderschuren LJMJ, Groenewegen HJ, Robbins TW, Pennartz CMA. Putting a spin on the dorsal-ventral divide of the striatum. Trends Neurosci. 2004;27(8):468–74.

114. Mogenson G, Swanson L, Wu M. Neural projections from nucleus accumbens to globus pallidus, substantia innominata, and lateral preoptic-lateral hypothalamic area: an anatomical and electrophysiological investigation in the rat. J Neurosci. 1983;3(1):189–202.

115. Faure A, Reynolds SM, Richard JM, Berridge KC. Mesolimbic dopamine in desire and dread: enabling motivation to be generated by localized glutamate disruptions in nucleus accumbens. J Neurosci. 2008;28(28):7184–92.

116. Parkinson JA, Olmstead MC, Burns LH, Robbins TW, Everitt BJ. Dissociation in effects of lesions of the nucleus accumbens core and shell on appetitive pavlovian approach behavior and the potentiation of conditioned reinforcement and locomotor activity by d-amphetamine. J Neurosci. 1999;19(6):2401–11.

117. Parkinson JA, Willoughby PJ, Robbins TW, Everitt BJ. Disconnection of the anterior cingulate cortex and nucleus accumbens core impairs pavlovian approach behavior: further evidence for limbic cortical-ventral striatopallidal systems. Behav Neurosci. 2000;114(1):42–63.

118. Corbit LH, Balleine BW. The general and outcome-specific forms of pavlovian-instrumental transfer are differentially mediated by the nucleus accumbens core and shell. J Neurosci. 2011;31(33):11786–94.

119. Bassareo V, Di Chiara G. Differential responsiveness of dopamine transmission to food-stimuli in nucleus accumbens shell/core compartments. Neuroscience. 1999;89(3):637–41.

120. Loriaux AL, Roitman JD, Roitman MF. Nucleus accumbens shell, but not core, tracks motivational value of salt. J Neurophysiol. 2011;106(3):1537–44.

121. Shiflett MW, Balleine BW. At the limbic-motor interface: disconnection of basolateral amygdala from nucleus accumbens core and shell reveals dissociable components of incentive motivation. Eur J Neurosci. 2010;32(10):1735–43.

122. Saddoris MP, Cacciapaglia F, Wightman RM, Carelli RM. Differential dopamine release dynamics in the nucleus accumbens core and shell reveal complementary signals for error prediction and incentive motivation. J Neurosci. 2015;35(33):11572–82.

123. West EA, Carelli RM. Nucleus accumbens core and shell differentially encode reward-associated cues after reinforcer devaluation. J Neurosci. 2016;36(4):1128–39.

124. Valjent E, Bertran-Gonzalez J, Hervé D, Fisone G, Girault JA. Looking BAC at striatal signalling: cell-specific analysis in new transgenic mice. Trends Neurosci. 2009;32(10):538–47.

125. Tritsch NX, Sabatini BL. Dopaminergic modulation of synaptic transmission in cortex and striatum. Neuron. 2012;76:33–50.

126. Beaulieu JM, Gainetdinov RR. The physiology, signaling, and pharmacology of dopamine receptors. Pharmacol Rev. 2011;63:182–217.

127. Missale C, Nash SR, Robinson SW, Jaber M, Caron MG. Dopamine receptors: from structure to function. Physiol Rev. 1998;78(1):189–225.

128. Vallone D, Picetti R, Borrelli E. Structure and function of dopamine receptors. Neurosci Biobehav Rev. 2000;24:125–32.

129. Richfield EK, Penney JB, Young AB. Anatomical and affinity state comparisons between dopamine D1 and D2 receptors in the rat central nervous system. Neuroscience. 1989;30(3):767–77.

130. Dreyer JK, Herrik KF, Berg RW, Hounsgaard JD. Influence of phasic and tonic dopamine release on receptor activation. J Neurosci. 2010;30(42):14273–83.

131. Gerfen CR, Surmeier DJ. Modulation of striatal projection systems by dopamine. Annu Rev Neurosci. 2011;34:441–66.

132. Surmeier DJ, Ding J, Day M, Wang Z, Shen W. D1 and D2 dopamine-receptor modulation of striatal glutamatergic signaling in striatal medium spiny neurons. Trends Neurosci. 2007;30(5):228–35.

133. Albin RL, Young AB, Penney JB. The functional anatomy of basal ganglia disorders. Trends Neurosci. 1989;12(10):366–75.

134. DeLong MR. Primate models of movement disorders of basal ganglia origin. Trends Neurosci. 1990;13:281–5.

135. Gerfen CR, Engber TM, Mahan LC, Susel Z, Chase TN, Monsma FJ, Sibley DR. D1 and d2 dopamine receptor-regulated gene expression of striatonigral and striatopallidal neurons. Science. 1990;250:1429–32.

136. Kravitz AV, Freeze BS, Parker PRL, Kay K, Thwin MT, Deisseroth K, Kreitzer AC. Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. Nature. 2010;466:622–6.

137. Frank MJ. Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated parkinsonism. J Cogn Neurosci. 2005;17(1):51–72.

138. Carlezone WA Jr, Thomas MJ. Biological substrates of reward and aversion: a nucleus accumbens activity hypothesis. Neuropharmacology. 2009;56:122–32.

139. Grueter BA, Robison AJ, Neve RL, Nestler EJ, Malenka RC. ΔFosB differentially modulates nucleus accumbens direct and indirect pathway function. Proc Natl Acad Sci USA. 2013;110(5):1923–8.

140. Hikida T, Yawata S, Yamaguchi T, Danjo T, Sasaoka T, Wang Y, Nakanishi S. Pathway-specific modulation of nucleus accumbens in reward and aversive behavior via selective transmitter receptors. Proc Natl Acad Sci USA. 2013;110(1):342–7.

141. Kupchik YM, Brown RM, Heinsbroek JA, Lobo MK, Schwartz DJ, Kalivas PW. Coding the direct/indirect pathways by D1 and D2 receptors is not valid for accumbens projections. Nat Neurosci. 2015;18:1230–2.

142. Smith RJ, Lobo MK, Spencer S, Kalivas PW. Cocaine-induced adaptations in D1 and D2 accumbens projection neurons (a dichotomy not necessarily synonymous with direct and indirect pathways). Curr Opin Neurobiol. 2013;23:546–52.

143. Nicola SM, Surmeier DJ, Malenka RC. Dopaminergic modulation of neuronal excitability in the striatum and nucleus accumbens. Annu Rev Neurosci. 2000;23:185–215.

144. Lu XY, Ghasemzadeh MB, Kalivas P. Expression of D1 receptor, D2 receptor, substance P and enkephalin messenger RNAs in the neurons projecting from the nucleus accumbens. Neuroscience. 1997;82(3):767–80.

145. Aizman O, Brismar H, Uhlén P, Zettergren E, Levey AI, Forssberg H, Greengard P, Aperia A. Anatomical and physiological evidence for D_1 and D_2 dopamine receptor colocalization in neostriatal neurons. Nat Neurosci. 2000;3(3):226–30.

146. Bertran-Gonzalez J, Bosch C, Maroteaux M, Matamales M, Hervé D, Valjent E, Girault JA. Opposing patterns of signaling activation in dopamine D_1 and D_2 receptor-expressing striatal neurons in response to cocaine and haloperidol. J Neurosci. 2008;28(22):5671–85.

147. Hasbi A, Fan T, Alijaniaram M, Nguyen T, Perreault ML, O'Dowd BF, George SR. Calcium signaling cascade links dopamine D1–D2 receptor heteromer to striatal BDNF production and neuronal growth. Proc Natl Acad Sci USA. 2009;106(50):21377–82.

148. Rashid AJ, So CH, Kong MM, Furtak T, El-Ghundi M, Cheng R, O'Dowd BF, George SR. D1–D2 dopamine receptor heterooligomers with unique pharmacology are coupled to rapid activation of Gq/11 in the striatum. Proc Natl Acad Sci USA. 2007;104(2):654–9.

149. Frederick A, Yano H, Trifilieff P, Vishwasrao H, Biezonski D, Mészáros J, Urizar E, Sibley D, Kellendonk C, Sonntag K, et al. Evidence against dopamine D1/D2 receptor heteromers. Mol Psychiatry. 2015;20:1373–85.

150. Dalley JW, Lääne K, Theobald DE, Armstrong HC, Corlett PR, Chudasama Y, Robbins TW. Time-limited modulation of appetitive Pavlovian memory by D1 and NMDA receptors in the nucleus accumbens. Proc Natl Acad Sci USA. 2005;102(17):6189–94.

151. Eyny YS, Horvitz JC. Opposing roles of D_1 and D_2 receptors in appetitive conditioning. J Neurosci. 2003;23(5):1584–7.

152. Beninger RJ, Miller R. Dopamine D1-like receptors and reward-related incentive learning. Neurosci Biobehav Rev. 1998;22(2):335–45.

153. Parker JG, Zweifel LS, Clark JJ, Evans SB, Phillips PE, Palmiter RD. Absence of NMDA receptors in dopamine neurons attenuates dopamine release but not conditioned approach during Pavlovian conditioning. Proc Natl Acad Sci USA. 2010;107(30):13491–6.

154. Smith-Roe SL, Kelley AE. Coincident activation of NMDA and dopamine D_1 receptors within the nucleus accumbens core is required for appetitive instrumental learning. J Neurosci. 2000;20(20):7737–42.

155. Bernal SY, Dostova I, Kest A, Abayev Y, Kandova E, Touzani K, Sclafani A, Bodnar RJ. Role of dopamine D1 and D2 receptors in the nucleus accumbens shell on the acquisition and expression of fructose-conditioned flavor-flavor preferences in rats. Behav Brain Res. 2008;190(1):59–66.

156. Fraser KM, Haight JL, Gardner EL, Flagel SB. Examining the role of dopamine D_2 and D_3 receptors in Pavlovian conditioned approach behaviors. Behav Brain Res. 2016;305:87–99.

157. Lopez JC, Karlsson RM, O'Donnell P. Dopamine D2 modulation of sign and goal tracking in rats. Neuropsychopharmacology. 2015;40:2096–102.

158. Ranaldi R, Beninger RJ. Dopamine D1 and D2 antagonists attenuate amphetamine-produced enhancement of responding for conditioned reward in rats. Psychopharmacology. 1993;113(1):110–8.

159. Wolterink G, Phillips G, Cador M, Donselaar-Wolterink I, Robbins T, Everitt B. Relative roles of ventral striatal D1 and D2 dopamine receptors in responding with conditioned reinforcement. Psychopharmacology. 1993;110(3):355–64.

160. Sombers LA, Beyene M, Carelli RM, Wightman RM. Synaptic overflow of dopamine in the nucleus accumbens arises from neuronal activity in the ventral tegmental area. J Neurosci. 2009;29(6):1735–42.

161. Cachope R, Cheer JF. Local control of striatal dopamine release. Front Behav Neurosci. 2014;8:1–7.

162. Rice ME, Patel JC, Cragg SJ. Dopamine release in the basal ganglia. Neuroscience. 2011;198:112–37.

163. Threlfell S, Lalic T, Platt NJ, Jennings KA, Deisseroth K, Cragg SJ. Striatal dopamine release is triggered by synchronized activity in cholinergic interneurons. Neuron. 2012;75:58–64.

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

165. Floresco SB, West AR, Ash B, Moore H, Grace AA. Afferent modulation of dopamine neuron firing differentially regulates tonic and phasic dopamine transmission. Nat Neurosci. 2003;6(9):968–73.

166. Grace AA, Floresco SB, Goto Y, Lodge DJ. Regulation of firing of dopaminergic neurons and control of goal-directed behaviors. Trends Neurosci. 2007;30(5):220–7.

167. Tritsch NX, Ding JB, Sabatini BL. Dopaminergic neurons inhibit striatal output through non-canonical release of GABA. Nature. 2012;490(7419):262–6.

168. Tritsch NX, Granger AJ, Sabatini BL. Mechanisms and functions of GABA co-release. Nat Rev Neurosci. 2016;17:139–45.

169. Suri RE, Schultz W. A neural network model with dopamine-like reinforcement signal that learns a spatial delayed response task. Neuroscience. 1999;91(3):871–90.

170. Balleine BW, Delgado MR, Hikosaka O. The role of the dorsal striatum in reward and decision-making. J Neurosci. 2007;27(31):8161–5.

171. Packard MG, Knowlton BJ. Learning and memory functions of the basal ganglia. Annu Rev Neurosci. 2002;25(1):563–93.

172. Yin HH, Knowlton BJ, Balleine BW. Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. Eur J Neurosci. 2004;19(1):181–9.

173. Yin HH, Knowlton BJ, Balleine BW. Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. Behav Brain Res. 2006;166(2):189–96.

174. Yin HH, Knowlton BJ. The role of the basal ganglia in habit formation. Nat Rev Neurosci. 2006;7(6):464–76.

175. Tricomi E, Balleine BW, O'Doherty JP. A specific role for posterior dorsolateral striatum in human habit learning. Eur J Neurosci. 2009;29(11):2225–32.

176. Balleine BW, O'Doherty JP. Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. Neuropsychopharmacology. 2010;35:48–69.

177. Doll BB, Simon DA, Daw ND. The ubiquity of model-based reinforcement learning. Curr Opin Neurobiol. 2012;22(6):1075–81.

178. Lee SW, Shimojo S, O'Doherty JP. Neural computations underlying arbitration between model-based and model-free learning. Neuron. 2014;81(3):687–99.

179. Killcross S, Coutureau E. Coordination of actions and habits in the medial prefrontal cortex of rats. Cereb Cortex. 2003;13(4):400–8.

180. Cohen MX, Frank MJ. Neurocomputational models of basal ganglia function in learning, memory and choice. Behav Brain Res. 2009;199:141–56.

181. Collins AGE, Frank MJ. Opponent actor learning (opal): modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. Psychol Rev. 2014;121(3):337–66.

182. Frank MJ, Loughry B, O'Reilly RC. Interactions between frontal cortex and basal ganglia in working memory: a computational model. Cogn Affect Behav Neurosci. 2001;1:137–60.

183. Kravitz AV, Tye LD, Kreitzer AC. Distinct roles for direct and indirect pathway striatal neurons in reinforcement. Nat Neurosci. 2012;15(6):816–9.

184. Cui G, Jun SB, Jin X, Pham MD, Vogel SS, Lovinger DM, Costa RM. Concurrent activation of striatal direct and indirect pathways during action initiation. Nature. 2013;494:238–42.

185. Mink JW. The basal ganglia: focused selection and inhibition of competing motor programs. Prog Neurobiol. 1996;50(4):381–425.

186. Nelson AB, Kreitzer AC. Reassessing models of basal ganglia function and dysfunction. Annu Rev Neurosci. 2014;37:117–35.

187. Calabresi P, Picconi B, Tozzi A, Ghiglieri V, Di Filippo M. Direct and indirect pathways of basal ganglia: a critical reappraisal. Nat Neurosci. 2014;17(8):1022–9.

188. Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ. Model-based influences on humans' choices and striatal prediction errors. Neuron. 2011;69:1204–15.

189. Bromberg-Martin ES, Matsumoto M, Hong S, Hikosaka O. A pallidus-habenula-dopamine pathway signals inferred stimulus values. J Neurophysiol. 2010;104:1068–76.

190. Nakahara H, Itoh H, Kawagoe R, Takikawa Y, Hikosaka O. Dopamine neurons can represent context-dependent prediction error. Neuron. 2004;41(2):269–80.

191. Schultz W. Updating dopamine reward signals. Curr Opin Neurobiol. 2013;23:229–38.

192. Horvitz JC. Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. Neuroscience. 2000;96:651–6.

193. Kakade S, Dayan P. Dopamine: generalization and bonuses. Neural Netw. 2002;15:549–59.

194. Balleine BW. Neural bases of food-seeking: affect, arousal and reward in corticostriatolimbic circuits. Physiol Behav. 2005;86(5):717–30.

195. Yin HH, Ostlund SB, Knowlton BJ, Balleine BW. The role of the dorsomedial striatum in instrumental conditioning. Eur J Neurosci. 2005;22(2):513–23.

196. Cagniard B, Beeler JA, Britt JP, McGehee DS, Marinelli M, Zhuang X. Dopamine scales performance in the absence of new learning. Neuron. 2006;51(5):541–7.

197. Niv Y, Joel D, Dayan P. A normative perspective on motivation. Trends Cogn Sci. 2006;10(8):375–81.

198. Masterson FA, Crawford M. The defense motivation system: a theory of avoidance behavior. Behav Brain Sci. 1982;5(04):661–75.

199. Gray JA. The psychology of fear and stress. Cambridge: Cambridge University Press; 1987.

200. Mowrer OH. On the dual nature of learning: a reinterpretation of "conditioning" and "problem-solving". Harv Educ Rev. 1947;17:102–50.

201. Mowrer OH. Two-factor learning theory reconsidered, with special reference to secondary reinforcement and the concept of habit. Psychol Rev. 1956;63(2):114–28.

202. Canteras NS, Graeff FG. Executive and modulatory neural circuits of defensive reactions: implications for panic disorder. Neurosci Biobehav Rev. 2014;46:352–64.

203. Gross CT, Canteras NS. The many paths to fear. Nat Rev Neurosci. 2012;13(9):651–8.

204. Bandler R, Shipley MT. Columnar organization in the midbrain periaqueductal gray: modules for emotional expression? Trends Neurosci. 1994;17(9):379–89.

205. Bolles RC, Fanselow MS. A perceptual-defensive-recuperative model of fear and pain. Behav Brain Sci. 1980;3:291–323.

206. Fanselow MS. Neural organization of the defensive behavior system responsible for fear. Psychon Bull Rev. 1994;1(4):429–38.

207. Fanselow MS, Lester LS. A functional behavioristic approach to aversive motivated behavior: Predatory imminence as a determinant of the topography of defensive behavior. In: Bolles RC, Beecher MD, editors. Evolution and learning. Hillsdale: Erlbaum; 1988. p. 185–211.

208. Gray JA. The neuropsychology of anxiety: an enquiry into the functions of the septo-hippocampal system. Oxford: Oxford University Press; 1982.

209. Cabib S, Puglisi-Allegra S. Opposite responses of mesolimbic dopamine system to controllable and uncontrollable aversive experiences. J Neurosci. 1994;14(5):3333–40.

210. Cabib S, Puglisi-Allegra S. Stress, depression and the mesolimbic dopamine system. Psychopharmacology. 1996;128:331–42.

211. Cabib S, Puglisi-Allegra S. The mesoaccumbens dopamine in coping with stress. Neurosci Biobehav Rev. 2012;36(1):79–89.

212. Redgrave P, Prescott TJ, Gurney K. The basal ganglia: a vertebrate solution to the selection problem? Neuroscience. 1999;89(4):1009–23.

213. Blanchard DC, Blanchard RJ. Ethoexperimental approaches to the biology of emotion. Annu Rev Psychol. 1988;39:43–68.

214. Swanson LW. Cerebral hemisphere regulation of motivated behavior. Brain Res. 2000;886(1):113–64.

215. Joseph MH, Datla K, Young AMJ. The interpretation of the measurement of nucleus accumbens dopamine by in vivo analysis: the kick, the craving or the cognition? Neurosci Biobehav Rev. 2003;27:527–41.

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

217. Abercrombie ED, Keefe KA, DiFrischia DS, Zigmond MJ. Differential effect of stress on in vivo dopamine release in striatum, nucleus accumbens, and medial frontal cortex. J Neurochem. 1989;52(5):1655–8.

218. Inglis FM, Moghaddam B. Dopaminergic innervation of the amygdala is highly responsive to stress. J Neurochem. 1999;72(3):1088–94.

219. Young AMJ, Rees KR. Dopamine release in the amygdaloid complex of the rat, studied by brain microdialysis. Neurosci Lett. 1998;249(1):49–52.

220. Budygin EA, Park J, Bass CE, Grinevich VP, Bonin KD, Wightman RM. Aversive stimulus differentially triggers subsecond dopamine release in reward regions. Neuroscience. 2012;201:331–7.

221. Park J, Bucher ES, Budygin EA, Wightman RM. Norepinephrine and dopamine transmission in 2 limbic regions differentially respond to acute noxious stimulation. Pain. 2015;156(2):318–27.

222. Cohen JY, Haesler S, Vong L, Lowell BB, Uchida N. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. Nature. 2012;482:85–90.

223. Brischoux F, Chakraborty S, Brierley DI, Ungless MA. Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. Proc Natl Acad Sci USA. 2009;106(12):4894–9.

224. Lammel S, Hetzel A, Häckel O, Jones I, Liss B, Roeper J. Unique properties of mesoprefrontal neurons within a dual mesocorticolimbic dopamine system. Neuron. 2008;57:760–73.

225. Lammel S, Ion DI, Roeper J, Malenka RC. Projection-specific modulation of dopamine neuron synapses by aversive and rewarding stimuli. Neuron. 2011;70(5):855–62.

226. Mantz J, Thierry A, 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. 1989;476(2):377–81.

227. Lerner TN, Shilyansky C, Davidson TJ, Evans KE, Beier KT, Zalocusky KA, Crow AK, Malenka RC, Luo L, Tomer R, et al. Intact-brain analyses reveal distinct information carried by SNc dopamine subcircuits. Cell. 2015;162(3):635–47.

228. Maldonado-Irizarry CS, Swanson CJ, Kelley AE. Glutamate receptors in the nucleus accumbens shell control feeding behavior via the lateral hypothalamus. J Neurosci. 1995;15(10):6779–88.

229. Reynolds SM, Berridge KC. Fear and feeding in the nucleus accumbens shell: rostrocaudal segregation of gaba-elicited defensive behavior versus eating behavior. J Neurosci. 2001;21(9):3261–70.

230. Richard JM, Berridge KC. Nucleus accumbens dopamine/glutamate interaction switches modes to generate desire versus dread: D1 alone for appetitive eating but D1 and D2 together for fear. J Neurosci. 2011;31(36):12866–79.

231. Reynolds SM, Berridge KC. Emotional environments retune the valence of appetitive versus fearful functions in nucleus accumbens. Nat Neurosci. 2008;11:423–5.

232. Sweidan S, Edinger H, Siegel A. The role of D1 and D2 receptors in dopamine agonist-induced modulation of affective defense behavior in the cat. Pharmacol Biochem Behav. 1990;36(3):491–9.

233. Sweidan S, Edinger H, Siegel A. D2 dopamine receptor-mediated mechanisms in the medial preoptic-anterior hypothalamus regulate affective defense behavior in the cat. Brain Res. 1991;549(1):127–37.

234. Willner P. Animal models of depression: an overview. Pharmacol Ther. 1990;45(3):425–55.

235. Steru L, Chermat R, Thierry B, Simon P. The tail suspension test: a new method for screening antidepressants in mice. Psychopharmacology. 1985;85(3):367–70.

236. Porsolt RD, Le Pichon M, Jalfre M. Depression: a new animal model sensitive to antidepressant treatments. Nature. 1977;266(5604):730–2.

237. Maier SF, Seligman ME. Learned helplessness: theory and evidence. J Exp Psychol Gen. 1976;105(1):3–46.

238. Puglisi-Allegra S, Imperato A, Angelucci L, Cabib S. Acute stress induces time-dependent responses in dopamine mesolimbic system. Brain Res. 1991;554:217–22.

239. Imperato A, Angelucci L, Casolini P, Zocchi A, Puglisi-Allegra S. Repeated stressful experiences differently affect limbic dopamine release during and following stress. Brain Res. 1992;577:194–9.

240. Imperato A, Cabib S, Puglisi-Allegra S. Repeated stressful experiences differently affect the time-dependent responses of the mesolimbic dopamine system to the stressor. Brain Res. 1993;601:333–6.

241. Pascucci T, Ventura R, Latagliata EC, Cabib S, Puglisi-Allegra S. The medial prefrontal cortex determines the accumbens dopamine response to stress through the opposing influences of norepinephrine and dopamine. Cereb Cortex. 2007;17(12):2796–804.

242. Leknes S, Tracey I. A common neurobiology for pain and pleasure. Nat Rev Neurosci. 2008;9(4):314–20.

243. Wood PB. Role of central dopamine in pain and analgesia. Expert Rev Neurother. 2008;8(5):781–97.

244. Schwartz N, Temkin P, Jurado S, Lim BK, Heifets BD, Polepalli JS, Malenka RC. Decreased motivation during chronic pain requires long-term depression in the nucleus accumbens. Science. 2014;345(6196):535–42.

245. Ren W, Centeno MV, Berger S, Wu Y, Na X, Liu X, Kondapalli J, Apkarian AV, Martina M, Surmeier DJ. The indirect pathway of the nucleus accumbens shell amplifies neuropathic pain. Nature Neurosci. 2016;19:220–2.

246. Farrar AM, Segovia KN, Randall PA, Nunes EJ, Collins LE, Stopper CM, Port RG, Hockemeyer J, Müller CE, Correa M, Salamone JD. Nucleus accumbens and effort-related functions: behavioral and neural markers of the interactions between adenosine A2A and dopamine D2 receptors. Neuroscience. 2010;166(4):1056–67.

247. Santerre JL, Nunes EJ, Kovner R, Leser CE, Randall PA, Collins-Praino LE, Cruz LL, Correa M, Baqi Y, Müller CE, et al. The novel adenosine A2A antagonist prodrug MSX-4 is effective in animal models related to motivational and motor functions. Pharmacol Biochem Behav. 2012;102(4):477–87.

248. Wadenberg MG, Hicks PB. The conditioned avoidance response test re-evaluated: is it a sensitive test for the detection of potentially atypical antipsychotics? Neurosci Biobehav Rev. 1999;23:851–62.

249. Deakin JFW, Graeff FG. 5-HT and mechanisms of defence. J psychopharmacol. 1991;5:305–15.

250. Graeff FG, Guimarães FS, De Andrade TG, Deakin JF. Role of 5-HT in stress, anxiety, and depression. Pharmacol Biochem Behav. 1996;54(1):129–41.

251. Dayan P, Huys QJM. Serotonin in affective control. Annu Rev Neurosci. 2009;32:95–126.

252. Dayan P. Instrumental vigour in punishment and reward. Eur J Neurosci. 2012;35(7):1152–68.

253. Grossberg S. Some normal and abnormal behavioral syndromes due to transmitter gating of opponent systems. Biol Psychiatry. 1984;19:1075–118.

254. Solomon RL, Corbit JD. An opponent-process theory of motivation: I. temporal dynamics of affect. Psychol Rev. 1974;81(2):119–45.

255. Daw ND, Kakade S, Dayan P. Opponent interactions between serotonin and dopamine. Neural Netw. 2002;15:603–16.

256. Deakin JFW. Roles of serotonergic systems in escape, avoidance and other behaviours. In: Cooper SJ, editor. Theory in psychopharmacology. vol. 2. 2nd edn., New York: Academic Press; 1983. pp. 149–193.

257. García J, Fernández F. A comprehensive survey on safe reinforcement learning. J Mach Learn Res. 2015;16:1437–80.

258. Huys QJ, Daw ND, Dayan P. Depression: a decision-theoretic analysis. Annu Rev Neurosci. 2015;38:1–23.

259. Gray JA, McNaughton N. The neuropsychology of anxiety: an enquiry into the function of the septo-hippocampal system, vol. 33. Oxford: Oxford University Press; 2003.

260. Blanchard RJ, Yudko EB, Rodgers RJ, Blanchard DC. Defense system psychopharmacology: an ethological approach to the pharmacology of fear and anxiety. Behav Brain Res. 1993;58(1):155–65.

261. Lister RG. Ethologically-based animal models of anxiety disorders. Pharmacol Ther. 1990;46(3):321–40.

262. Bach DR. Anxiety-like behavioural inhibition is normative under environmental threat-reward correlations. PLoS Comput Biol. 2015;11(12):1004646.

263. Fanselow MS. The postshock activity burst. Anim Learn Behav. 1982;10(4):448–54.

264. Jenkins H, Moore BR. The form of the auto-shaped response with food or water reinforcers. J Exp Anal Behav. 1973;20(2):163–81.

265. Abraham AD, Neve KA, Lattal KM. Dopamine and extinction: a convergence of theory with fear and reward circuitry. Neurobiol Learn Mem. 2014;108:65–77.

266. Levita L, Dalley JW, Robbins TW. Nucleus accumbens dopamine and learned fear revisited: a review and some new findings. Behav Brain Res. 2002;137:115–27.

267. Pezze MA, Feldon J. Mesolimbic dopaminergic pathways in fear conditioning. Prog Neurobiol. 2004;74:301–20.

268. Frank MJ, Surmeier DJ. Do substantia nigra dopaminergic neurons differentiate between reward and punishment? J Mol Cell Biol. 2009;1:15–6.

269. Zweifel LS, Fadok JP, Argilli E, Garelick MG, Jones GL, Dickerson TMK, Allen JM, Mizumori SJY, Bonci A, Palmiter RD. Activation of dopamine neurons is critical for aversive conditioning and prevention of generalized anxiety. Nat Neurosci. 2011;14(5):620–6.

270. Badrinarayan A, Wescott SA, Vander Weele CM, Saunders BT, Couturier BE, Maren S, Aragona BJ. Aversive stimuli differentially modulate real-time dopamine transmission dynamics within the nucleus accumbens core and shell. J Neurosci. 2012;32(45):15779–90.

271. Oleson EB, Gentry RN, Chioma VC, Cheer JF. Subsecond dopamine release in the nucleus accumbens predicts conditioned punishment and its successful avoidance. J Neurosci. 2012;32(42):14804–8.

272. Fadok JP, Dickerson TMK, Palmiter RD. Dopamine is necessary for cue-dependent fear conditioning. J Neurosci. 2009;29(36):11089–97.

273. Ikegami M, Uemura T, Kishioka A, Sakimura K, Mishina M. Striatal dopamine D1 receptor is essential for contextual fear conditioning. Sci Rep. 2014;4:3976.

274. Blackburn JR, Phillips AG. Enhancement of freezing behaviour by metaclopromide: implications for neuroleptic-induced avoidance deficits. Pharmacol Biochem Behav. 1990;35(3):685–91.

275. de Souza Caetano KA, de Oliveira AR, Brandão ML. Dopamine D$_2$ receptors modulate the expression of contextual conditioned fear: role of the ventral tegmental area and the basolateral amygdala. Behav Pharmacol. 2013;24(4):264–74.

276. Davis M, Falls WA, Campeau S, Kim M. Fear-potentiated startle: a neural and pharmacological analysis. Behav Brain Res. 1993;58:175–98.

277. de Oliveira AR, Reimer AE, Brandão ML. Dopamine D2 receptor mechanisms in the expression of conditioned fear. Pharmacol Biochem Behav. 2006;84(1):102–11.

278. Li SSY, McNally GP. A role of nucleus accumbens dopamine receptors in the nucleus accumbens core, but not shell, in fear prediction error. Behav Neurosci. 2015;129(4):450–6.

279. Pavlov IP. Conditioned reflexes. Oxford: Oxford University Press; 1927.

280. Rescorla RA. Pavlovian conditioned inhibition. Psychol Bull. 1969;72(2):77–94.

281. Christianson JP, Fernando ABP, Kazama AM, Jovanovic T, Ostroff LE, Sangha S. Inhibition of fear by learned safety signals: a mini-symposium review. J Neurosci. 2012;32(41):14118–24.

282. Kong E, Monje FJ, Hirsch J, Pollak DD. Learning not to fear: neural correlates of learned safety. Neuropsychopharmacology. 2014;39:515–27.

283. Fernando ABP, Urcelay GP, Mar AC, Dickinson A, Robbins TW. Safety signals as instrumental reinforcers during free-operant avoidance. Learn Mem. 2014;21:488–97.

284. Dickinson A, Pearce J. Inhibitory interactions between appetitive and aversive stimuli. Psychol Bull. 1977;84:690–711.

285. Dickinson A, Dearing MF. Appetitive-aversive interactions and inhibitory processes. In: Dickinson A, Boakes RA, editors. Mechanisms of learning and motivation. Hillsdale: Erlbaum; 1979. p. 203–31.

286. Rogan MT, Leon KS, Perez DL, Kandel ER. Distinct neural signatures for safety and danger in the amygdala and striatum of the mouse. Neuron. 2005;46:309–20.

287. Genud-Gabai R, Klavir O, Paz R. Safety signals in the primate amygdala. J Neurosci. 2013;33(46):17986–94.

288. Sangha S, Chadick JZ, Janak PH. Safety encoding in the basal amygdala. J Neurosci. 2013;33(9):3744–51.

289. Pollak DD, Rogan MT, Egner T, Perez DL, Yanagihara TK, Hirsch J. A translational bridge between mouse and human models of learned safety. Ann Med. 2010;42(2):127–34.

290. Fernando ABP, Urcelay GP, Mar AC, Dickenson TA, Robbins TW. The role of nucleus accumbens shell in the mediation of the reinforcing properties of a safety signal in free-operant avoidance: dopamine-dependent inhibitory effects of d-amphetamine. Neuropsychopharmacology. 2014;39:1420–30.

291. Bouton ME. Learning and behavior: a contemporary synthesis. Sunderland: Sinauer Associates Inc; 2007.

292. Beninger RJ. The role of dopamine in locomotor activity and learning. Brain Res Rev. 1983;6:173–96.

293. Dinsmoor JA. Punishment: I. the avoidance hypothesis. Psychol Rev. 1954;61:34–46.

294. Dinsmoor JA. Stimuli inevitably generated by behavior that avoids electric shock are inherently reinforcing. J Exp Anal Behav. 2001;75:311–33.

295. Konorski J. Conditioned reflexes and neuron organization. Cambridge: Cambridge University Press; 1948.

296. Miller NE. Studies of fear as an acquirable drive: I. fear as motivation and fear-reduction as reinforcement in the learning of new responses. J Exp Psychol. 1948;38:89–101.

297. Tolman EC. Purposive behavior in animals and men. New York: Century; 1932.

298. Blanchard RJ, Fukunaga KK, Blanchard DC. Environmental control of defensive reactions to footshock. Bull Psychon Soc. 1976;8(2):129–30.

299. Seymour B, Singer T, Dolan R. The neurobiology of punishment. Nat Rev Neurosci. 2007;8(4):300–11.

300. Morse WH, Mead RN, Kelleher RT. Modulation of elicited behavior by a fixed-interval schedule of electric shock presentation. Science. 1967;157(3785):215–7.

301. Overmier JB, Seligman ME. Effects of inescapable shock upon subsequent escape and avoidance responding. J Comp Physiol Psychol. 1967;63(1):28–33.

302. Seligman ME, Maier SF. Failure to escape traumatic shock. J Exp Psychol. 1967;74:1–9.

303. Fernando A, Urcelay G, Mar A, Dickinson A, Robbins T. Free-operant avoidance behavior by rats after reinforcer revaluation using opioid agonists and d-amphetamine. J Neurosci. 2014;34(18):6286–93.

304. Hendersen RW, Graham J. Avoidance of heat by rats: effects of thermal context on rapidity of extinction. Learn Motiv. 1979;10(3):351–63.

305. Declercq M, De Houwer J. On the role of us expectancies in avoidance behavior. Psychon Bull Rev. 2008;15(1):99–102.

306. Gillan CM, Morein-Zamir S, Urcelay GP, Sule A, Voon V, Apergis-Schoute AM, Fineberg NA, Sahakian BJ, Robbins TW. Enhanced avoidance habits in obsessive-compulsive disorder. Biol Psychiatry. 2014;75(8):631–8.

307. Maia TV, Frank MJ. From reinforcement learning models to psychiatric and neurological disorders. Nat Neurosci. 2011;14(2):154–62.

308. Fibiger HC, Phillips AG, Zis AP. Deficits in instrumental responding after 6-hydroxydopamine lesions of the nigro-striatal dopaminergic projection. Pharmacol Biochem Behav. 1974;2:87–96.

309. Koob GF, Simon H, Herman JP, Le Moal M. Neuroleptic-like disruption of the conditioned avoidance response requires destruction of both mesolimbic and nigrostriatal dopamine systems. Brain Res. 1984;303:319–29.

310. Darvas M, Fadok JP, Palmiter RD. Requirement of dopamine signaling in the amygdala and striatum for learning and maintenance of a conditioned avoidance response. Learn Mem. 2011;18(3):136–43.

311. Beninger RJ, Mason ST, Phillips AG, Fibiger HC. The use of extinction to investigate the nature of neuroleptic-induced avoidance deficits. Psychopharmacology. 1980;69:11–8.

312. Beninger RJ. The role of serotonin and dopamine in learning to avoid aversive stimuli. In: Archer T, Nilsson LG, editors. Aversion, avoidance and anxiety: perspectives on aversively motivated behavior. Hillsdale: Lawrence Erlbaum Associates; 1989. p. 265–84.

313. Wadenberg MG, Ericson E, Magnusson O, Ahlenius S. Suppression of conditioned avoidance behavior by the local application of (-)sulpiride into the ventral, but not the dorsal, striatum of the rat. Biol Psychiatry. 1990;28:297–307.

314. Tye KM, Mirzabekov JJ, Warden MR, Ferenczi EA, Tsai HC, Finkelstein J, Kim SY, Adhikari A, Thompson KR, Andalman AS, Gunaydin L, Witten I, Deisseroth K. Dopamine neurons modulate neural encoding and expression of depression-related behaviour. Nature. 2013;493:537–41.

315. Anstrom KK, Woodward DJ. Restraint increases dopaminergic burst firing in awake rats. Neuropsychopharmacology. 2005;10(10):1832–40.

316. Chaudhury D, Walsh JJ, Friedman AK, Juarez B, Ku SM, Koo JW, Ferguson D, Tsai HC, Pomeranz L, Christoffel DJ, et al. Rapid regulation of depression-related behaviours by control of midbrain dopamine neurons. Nature. 2013;493:532–6.

317. Friedman AK, Walsh JJ, Juarez B, Ku SM, Chaudhury D, Wang J, Li X, Dietz DM, Pan N, Vialou VF, et al. Enhancing depression mechanisms in midbrain dopamine neurons achieves homeostatic resilience. Science. 2014;344(6181):313–9.

318. Hollon NG, Burgeno LM, Phillips PE. Stress effects on the neural substrates of motivated behavior. Nat Neurosci. 2015;18(10):1405–12.

319. Bland ST, Twining C, Watkins LR, Maier SF. Stressor controllability modulates stress-induced serotonin but not dopamine efflux in the nucleus accumbens shell. Synapse. 2003;49:206–8.

320. Bland ST, Hargrave D, Pepin JL, Amat J, Watkins LR, Maier SF. Stressor controllability modulates stress-induced dopamine and serotonin efflux and morphine-induced serotonin efflux in the medial prefrontal cortex. Neuropsychopharmacology. 2003;28:1589–96.

321. Amat J, Matus-Amat P, Watkins LR, Maier SF. Escapable and inescapable stress differentially alter extracellular levels of 5-HT in the basolateral amygdala of the rat. Brain Res. 1998;812:113–20.

322. Amat J, Matus-Amat P, Watkins LR, Maier SF. Escapable and inescapable stress differentially and selectively alter extracellular levels of 5-HT in the ventral hippocampus and dorsal periaqueductal gray of the rat. Brain Res. 1998;797:12–22.

323. Amat J, Baratta MV, Paul E, Bland ST, Watkins LR, Maier SF. Medial prefrontal cortex determines how stressor controllability affects behavior and dorsal raphe nucleus. Nat Neurosci. 2005;8:365–71.

324. Amo R, Fredes F, Kinoshita M, Aoki R, Aizawa H, Agetsuma M, Aoki T, Shiraki T, Kakinuma H, Matsuda M, et al. The habenulo-raphe serotonergic circuit encodes an aversive expectation value essential for adaptive active avoidance of danger. Neuron. 2014;84(5):1034–48.

325. Li J, Daw ND. Signals in human striatum are appropriate for policy update rather than value prediction. J Neurosci. 2011;31(14):5504–11.

326. Kishida KT, Saez I, Lohrenz T, Witcher MR, Laxton AW, Tatter SB, White JP, Ellis TL, Phillips PE, Montague PR. Subsecond dopamine fluctuations in human striatum encode superposed error signals about actual and counterfactual reward. Proc Natl Acad Sci USA. 2016;113(1):200–5.

327. D'Ardenne K, Lohrenz T, Bartley KA, Montague PR. Computational heterogeneity in the human mesencephalic dopamine system. Cogn Affect Behav Neurosci. 2013;13(4):747–56.

328. Lohrenz T, McCabe K, Camerer CF, Montague PR. Neural signature of fictive learning signals in a sequential investment task. Proc Natl Acad Sci USA. 2007;104(22):9493–8.

329. Cohen JY, Amoroso MW, Uchida N. Serotonergic neurons signal reward and punishment on multiple timescales. ELife. 2015;4:06346.

330. Choi JS, Cain CK, LeDoux JE. The role of amygdala nuclei in the expression of auditory signaled two-way active avoidance in rats. Learn Mem. 2010;17(3):139–47.

331. Weiner DM, Levey AI, Sunahara RK, Niznik HB, O'Dowd BF, Seeman P, Brann MR. D1 and d2 dopamine receptor mrna in rat brain. Proc Natl Acad Sci USA. 1991;88(5):1859–63.

332. de la Mora MP, Gallegos-Cari A, Arizmendi-García Y, Marcellino D, Fuxe K. Role of dopamine receptor mechanisms in the amygdaloid modulation of fear and anxiety: structural and functional analysis. Prog Neurobiol. 2010;90(2):198–216.

333. Matsumoto M, Hikosaka O. Two types of dopamine neuron distinctly convey positive and negative motivational signals. Nature. 2009;459:837–41.

334. Fiorillo CD. Two dimensions of value: dopamine neurons represent reward but not aversiveness. Science. 2013;341(6145):546–9.

335. Beier KT, Steinberg EE, DeLoach KE, Xie S, Miyamichi K, Schwarz L, Gao XJ, Kremer EJ, Malenka RC, Luo L. Circuit architecture of vta dopamine neurons revealed by systematic input-output mapping. Cell. 2015;162(3):622–34.

336. Menegas W, Bergan JF, Ogawa SK, Isogai Y, Venkataraju KU, Osten P, Uchida N, Watabe-Uchida M. Dopamine neurons projecting to the posterior striatum form an anatomically distinct subclass. ELife. 2015;4:10032.

337. Miyazaki KW, Miyazaki K, Tanaka KF, Yamanaka A, Takahashi A, Tabuchi S, Doya K. Optogenetic activation of dorsal raphe serotonin neurons enhances patience for future rewards. Curr Biol. 2014;24(17):2033–40.

338. Fonseca MS, Murakami M, Mainen ZF. Activation of dorsal raphe serotonergic neurons promotes waiting but is not reinforcing. Curr Biol. 2015;25(3):306–15.

339. Liu Z, Zhou J, Li Y, Hu F, Lu Y, Ma M, Feng Q, Zhang JE, Wang D, Zeng J, et al. Dorsal raphe neurons signal reward through 5-HT and glutamate. Neuron. 2014;81(6):1360–74.

340. Kool W, McGuire JT, Wang GJ, Botvinick MM. Neural and behavioral evidence for an intrinsic cost of self-control. PLoS One. 2013;8(8):72626.

341. McGuire JT, Botvinick MM. Prefrontal cortex, cognitive control, and the registration of decision costs. Proc Natl Acad Sci. 2010;107(17):7922–6.

342. Dayan P. How to set the switches on this thing. Curr Opin Neurobiol. 2012;22(6):1068–74.

343. Gan JO, Walton ME, Phillips PE. Dissociable cost and benefit encoding of future rewards by mesolimbic dopamine. Nat Neurosci. 2010;13(1):25–7.

344. Hollon NG, Arnold MM, Gan JO, Walton ME, Phillips PE. Dopamine-associated cached values are not sufficient as the basis for action selection. Proc Natl Acad Sci. 2014;111(51):18357–62.