Opponency Revisited: Competition and Cooperation Between Dopamine and Serotonin

Y-Lan Boureau[1] and Peter Dayan*,[2]

[1]The Courant Institute of Mathematical Sciences, New York University, New York, NY, USA; [2]Gatsby Computational Neuroscience Unit, London, UK

Affective valence lies on a spectrum ranging from punishment to reward. The coding of such spectra in the brain almost always involves opponency between pairs of systems or structures. There is ample evidence for the role of dopamine in the appetitive half of this spectrum, but little agreement about the existence, nature, or role of putative aversive opponents such as serotonin. In this review, we consider the structure of opponency in terms of previous biases about the nature of the decision problems that animals face, the conflicts that may thus arise between Pavlovian and instrumental responses, and an additional spectrum joining invigoration to inhibition. We use this analysis to shed light on aspects of the role of serotonin and its interactions with dopamine.
Neuropsychopharmacology Reviews advance online publication, 29 September 2010; doi:10.1038/npp.2010.151

INTRODUCTION

The theory of adaptive optimal control concerns learning the actions that maximize rewards and minimize punishments, in both cases over the long run. Optimal control is mathematically straightforward, but suffers from three critical problems: (1) in rich domains, even when given all the information, it can be an extremely hard computation to work out what the optimal action is; (2) it can be expensive to rely on learning in a world in which rewards are scarce and dangers rampant; and (3) some key abstractions in optimal control theory, such as the notion of a single utility function reporting negative and positive subjective values of outcomes, sit ill with the constraints of neural information processing (eg, the firing rates of neurons must be positive). Recent behavioral neuroscience and theoretical studies, combined with a more venerable psychological literature, are providing hints to solutions to these issues. We review these ideas, aiming at the critical impact of dopamine (DA), serotonin (5-HT), and some of their complex interactions.

Optimal control problems are solved through a collection of structurally and functionally different methods (Doya, 1999; Dickinson and Balleine, 2002; Daw *et al*, 2005; Dayan, 2008; Balleine, 2005), each of which realizes a different tradeoff between the difficulty of calculating the optimal action (called computational complexity) and the expense of learning (called sample complexity). Three particularly important methods are: (1) model-based or goal-directed control; (2) model-free, habitual, or cached control; and (3) Pavlovian control. Both goal-directed and cached controls are involved in instrumental conditioning. In the former, a model of the task is constructed and explicitly searched to work out the evolving worth of each action. In the latter, there is a direct mapping from actions to worths, which is learned from experience in a way that obviates acquiring or searching a model. Pavlovian control depends on evolutionarily pre-programmed responses to predictions and occurrences of reinforcers. Pavlovian responses associated with appetitive outcomes include approach (Glickman and Schiff, 1967), engagement, and consumption (Panksepp, 1998). Pavlovian responses to aversive outcomes include a range of species-typical (Bolles, 1970; Schneirla, 1959) defensive and avoidance behaviors such as inhibition and fight/flight/freeze (Gray and McNaughton, 2003) that are sensitive to the proximity of threats (Blanchard and Blanchard, 1989; McNaughton and Corr, 2004).

Our task is to integrate anatomical, pharmacological, physiological, and behavioral neuroscience data to provide a picture of what role(s) DA and 5-HT have in influencing the three types of controls mentioned above. It should be noted that there is a near overwhelming wealth of data; we

had therefore to be highly selective in our discussion, and apologize for the volume we had to omit. Recent reviews such as Haber and Knutson (2010), Jacobs and Fornal (1999), Cooper *et al* (2002), Cools *et al* (2008), Dayan and Huys (2009), Berridge (2007), Everitt *et al* (2008), Tops *et al* (2009), and Iordanova (2009) should collectively be consulted.

Our main argument is that the nature of the roles depends on two critical dimensions influencing control: valence (reward *vs* punishment) and action (invigoration *vs* inhibition) (Ikemoto and Panksepp, 1999). In turn, these dimensions are tied to each other by heuristic biases.

In these terms, one suggested role for DA is initiating appetitively inspired actions such as approach (as in theories of incentive salience; Berridge and Robinson, 1998; Berridge, 2007; Alcaro *et al*, 2007; Ikemoto and Panksepp, 1999). A second role, which is closely related to this, is mediating general appetitive Pavlovian–instrumental transfer (PIT) and vigor (Niv *et al*, 2007; Wyvell and Berridge, 2001; Smith and Dickinson, 1998). This has also been related to the role of DA in overcoming effort costs (Salamone and Correa, 2002), voluntary motivation (Mazzoni *et al*, 2007), and 'seeking' behavior (Panksepp, 1998). A final role for DA is representing the appetitive portion of the temporal difference (TD) prediction error (Sutton, 1988), which is the critical signal in reinforcement learning (RL) for acquiring predictions of long-run future rewards and also for choosing appropriate actions (Montague *et al*, 1996; Schultz *et al*, 1997; Barto, 1995). These three capacities for DA, which are certainly not mutually exclusive (Alcaro *et al*, 2007; McClure *et al*, 2003), all involve rewards. The neuromodulator has a more complex association with punishment, with clear evidence for its release and involvement in some forms in certain aversive paradigms (Beninger *et al*, 1980b; Moutoussis *et al*, 2008; Brischoux *et al*, 2009; Matsumoto and Hikosaka, 2009; Kalivas and Duffy, 1995; Abercrombie *et al*, 1989; Pezze and Feldon, 2004), but also contrary and constraining data (Schultz, 2007; Mirenowicz and Schultz, 1996; Ungless *et al*, 2004).

We recently reviewed computational issues associated with the substantial extra intricacies of 5-HT compared with DA (Dayan and Huys, 2009). Notably, there is much evidence for functional opponency between DA and 5-HT. For instance, one evident association is with behavioral inhibition (opponent to approach) in the face of predictions of aversive outcomes (Soubrié, 1986; Deakin, 1983; Deakin and Graeff, 1991; Graeff *et al*, 1998; Gray and McNaughton, 2003). Equally, just as DA influences vigor in the face of hunger (Niv *et al*, 2007), 5-HT is associated with quiescence in the face of satiety (Gruninger *et al*, 2007; Cools *et al*, 2010). Furthermore, 5-HT neurons are activated (Grahn *et al*, 1999; Takase *et al*, 2004, 2005) and 5-HT is released (Bland *et al*, 2003a) in the face of punishment (Lowry, 2002; Abrams *et al*, 2004).

The notion of appetitive and aversive opponency itself is one of the most venerable ideas for the neural representation of valence (Konorski, 1967; Grossberg, 1984; Solomon and Corbit, 1974; Dickinson and Dearing, 1979; Brodie and Shore, 1957). Indeed, opponency between DA and 5-HT has been considered in detail by Deakin (1983), Deakin and Graeff (1991), and Graeff *et al* (1998). Daw *et al* (2002) suggested a computationally specific version of this idea, for the learning of habitual control. In their model, the phasic activity of DA and 5-HT reported prediction errors for future reward and punishment, and their tonic activity reported long-run average punishment and reward, respectively (ie, the reverse relationship). This account is not tenable in the light of the rich picture of structurally different influences on control, including the Pavlovian controller. It also does not reflect asymmetries in natural environments between reward and punishment, or between safety and danger signaling in terms of their influence on active avoidance. It faces further challenges from recent experiments orthogonalizing valence and activity (Guitart-Masip *et al*, 2010; Crockett *et al*, 2009; Huys *et al*, 2010) that call into question the choice of affective value as the fundamental axis of opponency between the neuromodulators.

In this study we first review optimal control and the different forms of RL with which it is associated. We then consider the critical roles that biases and heuristics have in evading the problems of learning; we argue that these are apparent in the architecture of control as well as in observed behavior. We next consider the implications of these for an updated theory of interaction between 5-HT and DA. Finally, we sum up the claims and main lacunæ of the new view.

The paper in this issue by Cools *et al* (2010) also considers DA and 5-HT interactions, but it started from the important and complementary viewpoint of the issue of vigor and quiescence rather than Pavlovian–instrumental interactions.

OPTIMAL AND SUBOPTIMAL CONTROL

The formal backdrop for the analysis of DA and 5-HT is that of RL and optimal control—how animals (and indeed robots or systems of any sort) can come to choose actions to maximize their long-run rewards and minimize their long-run punishments. This theory stems from operations research and computer science (Sutton and Barto, 1998; Puterman, 2005; Bertsekas and Tsitsiklis, 1996), but has long had rich links with psychology and neuroscience (Klopf, 1982; Barto, 1989, 1995; Schultz, 2002; Montague *et al*, 1996; Gabriel and Moore, 1991; Sutton and Barto, 1981; Daw and Doya, 2006; Niv, 2009). We do not have the space to review all this material here; however, we do need to provide the bones of the issues for the application of RL to understanding the choices of animals (see also the descriptions in Daw *et al*, 2005; Daw and Doya, 2006; Dayan

and Seymour, 2008; Niv, 2009; Schultz, 2007). Central to this section are the differences between goal-directed and habitual control; Pavlovian control will be covered in the next section.

RL, Goal-Directed, and Habitual Control

Consider putting an animal into a maze with distinguishable rooms (in RL these are, more abstractly, called states) in which there can be rewards (eg, food or water) and punishments (shocks or predators). The rooms are connected by a haphazard arrangement of exits and perhaps unidirectional passageways, and hence the animal can choose among a small number of actions in each room, getting to a restricted set of other rooms. We consider its task to be to come to take actions to maximize its expected long-run net utility, that is, benefits minus costs over the whole run through the maze. Of course, these utilities, (particularly the appetitive ones) depend on the animal's motivational state (eg, food being immediately valuable to a hungry, not a sated animal; Niv *et al*, 2006). RL defines a policy to be a systematic set of choices, one exit (or more normally, one probability distribution over exits) for each room in the maze.

RL formalizes the two central and linked concerns for the animal: inference and learning. Take the case that the animal knows the whole layout of the maze. It then still faces the inference problem of choosing which exit to take out of one room. This is hard, because the long-run utility depends not only on the next reward or punishment in the next room, but also on the whole sequence of future rewards and punishments that will unfold from beyond that room. These, in turn, depend on the actions taken in those rooms. This is the same problem that, for instance, chess players face in having to think many steps ahead to work out the benefits of a strategy in terms of winning or losing.

RL crystallizes the problem by noting that, given a whole policy, each room can be endowed with a value. The value of a room is the long-run utility available based on starting in that room, and following the policy. This value incorporates all the rewards and punishments that will be received through following that policy (which incorporates the hard problem mentioned above). Given such values, an optimal choice can be realized by choosing the exit leading to a room that is evaluated most highly. As the values change, the policy that stems from choosing a good exit changes too. If only approximate values are available, choices may, of course, only be suboptimal.

One way to estimate the value of a room is to imagine a fictitious experience (like a form of pre-play; Johnson and Redish, 2007; Foster and Wilson, 2007; Gupta *et al*, 2010), starting from that room, simulating possible whole future paths that follow the policy, and accumulating the estimated utilities of the outcomes that are predicted. This is called model-based RL, as simulating such paths requires knowledge of the domain that amounts to a model of

the passageways between rooms and the available outcomes. Such a model can readily be acquired directly from experience of rooms, transitions, and outcomes. Model-based calculations can also use algorithms that are computationally more efficient than pre-play (Sutton and Barto, 1998; Puterman, 2005; Bertsekas and Tsitsiklis, 1996).

In psychological terms, model-based RL has the property of goal-directed control that changing the utility of an outcome or the contingencies in the world would lead to an immediate change in the choice of actions (Dickinson and Balleine, 2002). This is because the current estimated utilities of all outcomes and the possible transitions are incorporated into the values of rooms through the process of simulation or search in the model.

A different way to perform the estimation is to note that the values of successive rooms should be self-consistent. That is, the value of one room should be the sum of the utility available in that room and the average values of the subsequent rooms to which the policy can lead in a single step. For instance, a room should have a high value if it either offers a large reward itself, or has a passageway to a room that itself has a high value (or both). Eliminating inconsistencies (which are called TD prediction errors; Sutton, 1988) between the estimated values of successive rooms leads to their being correct. Importantly, it is possible to eliminate inconsistencies using just those transitions experienced during learning, without any need for an explicit model of the domain. This method is therefore called model-free (or sometimes cached) RL.

If the values are learned in this way, then, like habits (Dickinson and Balleine, 2002), they will not immediately be sensitive to changes in the utilities of outcomes. This is because explicitly experiencing steps in the maze is necessary to erase inconsistencies.

Model-free and model-based controls differ in the ways they can acquire, generalize, and express information about factors such as the statistical structure or controllability of the environment (Huys and Dayan, 2009), with model-based control likely being able to capture more flexibly, the far finer distinctions.

Controllers in the Brain

There are substantial data on the anatomical substrates of control, reviewed, for instance, in Balleine (2005), Cardinal *et al* (2002), Haber and Knutson (2010), Sesack and Grace (2010), Hollerman *et al* (2000), and Balleine and O'Doherty (2010) (see also Wickens *et al,* 2007; Houk *et al,* 1994; Bezard, 2006). Strikingly, there is evidence that model-based and model-free strategies are simultaneously deployed by partly different neural structures (Dickinson and Balleine, 2002; Killcross and Coutureau, 2003; Balleine, 2005).

Model-free control depends particularly on the dorsolateral striatum (Jog *et al*, 1999; Killcross and Coutureau, 2003;

Yin *et al*, 2004; Tricomi *et al*, 2009). A rather wider set of areas has been implicated in the computationally more complex processes of model-based control, including the ventral (medial) prefrontal cortex (mPFC) and dorsomedial striatum (Balleine and Dickinson, 1998; Killcross and Coutureau, 2003; Yin *et al*, 2005), together with the basolateral amygdala and the orbitofrontal cortex, which are critically implicated in adapting to changes in the motivational value of stimuli (Padoa-Schioppa and Assad, 2006; Schoenbaum *et al*, 2009, 2003; Schultz and Dickinson, 2000; Hollerman *et al*, 2000; Rolls and Grabenhorst, 2008; Mainen and Kepecs, 2009; Fellows, 2007; Cador *et al*, 1989; Killcross *et al*, 1997; Holland and Gallagher, 2003; Corbit and Balleine, 2005; Talmi *et al*, 2008; O'Doherty, 2007; Wallis, 2007; Valentin *et al*, 2007; Baxter *et al*, 2000). Related dependencies include the role of ventral mPFC in extinction (Morgan *et al*, 1993; Morgan and LeDoux, 1995, 1999; Killcross and Coutureau, 2003), the anterior cingulate cortex's putative function in detecting errors and managing response conflicts (Devinsky *et al*, 1995; Pardo *et al*, 1990; Awh and Gehring, 1999; Botvinick *et al*, 2004; Gabriel *et al*, 1991; Bussey *et al*, 1997; Parkinson *et al*, 2000b), and perhaps also the role of hippocampus in non-habitual spatial behavior (White and McDonald, 2002; Doeller and Burgess, 2008).

The nucleus accumbens or ventral striatum is not required for goal-directed (Balleine and Killcross, 1994; Corbit *et al*, 2001) or habitual actions (Reading *et al*, 1991; Robbins *et al*, 1990). However, it is considered an interface between limbic and motor systems and is involved in the expression of Pavlovian responses and the interaction between Pavlovian and instrumental conditioning (Panksepp, 1998; Salamone and Correa, 2002; Balleine and Killcross, 1994; Parkinson *et al*, 2000b; Mogenson *et al*, 1980; Ikemoto and Panksepp, 1999; Killcross *et al*, 1997; Reynolds and Berridge, 2001; Corbit and Balleine, 2003; Hall *et al*, 2001; Talmi *et al*, 2008; Berridge, 2007; Berridge and Robinson, 1998; Sesack and Grace, 2010). These effects also depend on the central nucleus of the amygdala (Hall *et al*, 2001; Killcross *et al*, 1997; Cador *et al*, 1989; Parkinson *et al*, 2000a).

Of particular importance to us, the dopaminergic, serotonergic, and noradrenergic neuromodulatory systems influence, regulate, and plasticize all these other systems in the light of affectively important information. There are extensive, although not complete, interconnections between these structures (see eg, Powell and Leman, 1976; Zahm and Heimer, 1990; Brog *et al*, 1993; Haber *et al*, 2000; Groenewegen *et al*, 1980, 1982, 1999; Joel and Weiner, 2000; Fudge and Haber, 2000; Carr and Sesack, 2000), making for an extremely rich, and as yet incompletely understood, overall network.

As mentioned, goal-directed and habitual control appear to be expressed concurrently (Killcross and Coutureau, 2003). One idea for why this is good is that they represent different tradeoffs between two sorts of uncertainty or inaccuracy: computational noise, which afflicts primarily the model-based controller, and statistically inefficient learning, which afflicts the model-free controller given limited experience (Daw *et al*, 2005). That is, the task for model-based RL of simulating and following deep paths through the maze (or their algorithmic variants; Puterman, 2005) is computationally very taxing, and will therefore lead to inaccurate estimates of values or excessive energy consumption. In comparison, model-free RL does not have to compute its values, but rather has them directly available. However, it learns using inconsistencies between successive values. As these values are all inaccurate at the outset of learning, they do not provide useful error signals. Model-free RL is therefore slower to learn than model-based RL, which absorbs information more optimally. Thus, model-free RL requires more samples to make good choices, and is less adaptive to changes in the world. In sum, model-based control should dominate at the outset, whereas model-free control takes over at the end of learning.

Managing Energy

An additional facet of the optimal decision problem is choosing how vigorously to perform selected actions. Acting more quickly may imply getting rewards more quickly, or being more certain of avoiding being punished; however, it also takes more energy (Niv *et al*, 2007). This tradeoff is easiest to express in a model of optimal control designed for problems that continue for very long epochs (as if the animal is placed back into the maze if it ever reaches an exit), in which they optimize the average utility gained per unit time rather than the sum utility over a path (called average-case RL; Sutton and Barto, 1998; Daw and Touretzky, 2002). In this framework, the passage of time is itself costly, that is, there is an opportunity cost for time that is quantified by this average utility. To take an appetitive case, the greater the average, the more reward the animal should expect to get for each timestep in the maze. Not getting that much reward, for instance, by acting too slothfully is therefore not optimal. Vigor and sloth are thus tied to the overall affective valence of the environment. This average reward can be estimated by learning; but it may also be more directly influenced by innate or acquired previous expectations.

Consideration of vigor points toward a much broader tradeoff between energetically expensive, active reward seeking or threat avoiding, externally directed behaviors, and energy conserving, regenerative, digestive, internally directed behaviors (Handley and McBlane, 1991; Tops *et al*, 2009; Ellison, 1979), a distinction that has been related to that between the sympathetic and the parasympathetic nervous systems (Ellison, 1979). In the terminology of Ellison (1979), these are called, respectively, *ergotropic* (toward work or energy expenditure) and *trophotropic* (toward nourishment). (See Box).

Glossary

Model-based RL

A strategy for calculating the optimal values of states or an optimal policy by building a model of states, transitions, and outcomes, and performing a form of forwards, backwards, or simulation-based tree search using the model, expanding out long sequences of actions. Model-based RL is statistically efficient at learning, but computationally ruinous, and therefore should dominate toward the start of learning when model-free RL is very inaccurate. It has the characteristic of goal-directed behavior of being immediately sensitive to changes in the utilities of outcomes or contingencies. There is evidence that regions of the prefrontal cortex, basolateral nucleus of the amygdala, and the dorsomedial striatum are particularly involved (Killcross and Coutureau, 2003; Balleine, 2005).

Model-free RL

A strategy for learning the optimal values of states and actions without building a model. This relies on the property of values that they should be sequentially consistent. Inconsistencies amount to a form of prediction error that can be used to train the values. Model-free RL is statistically inefficient as it depends on bootstrapping (ie, learning from its own, initially erroneous, estimates). However, it is computationally trivial, as it depends on stored (cached) values, and thus does not suffer from computational or inferential noise. Thus, as learning proceeds, it produces more accurate values than model-based RL, and hence should come to dominate. Model-free RL cannot incorporate new information about utilities or contingencies without explicit experience from all the states whose values are ultimately affected. It therefore shares the characteristics of habitual behavior. Evidence suggests the involvement of the central nucleus of the amygdala and the dorsolateral striatum, with the impact of model-free choice on behavior also being affected by the infralimbic regions of the medial prefrontal cortex (Killcross and Coutureau, 2003; Balleine, 2005).

Temporal difference (TD) prediction error

The TD prediction error (Sutton, 1988) is the inconsistency between the value estimates at successive states, and has a key role in criticizing and thus teaching model-free values. The appetitive TD prediction error appears to be reported by the phasic activity of at least some dopamine neurons in rats (Pan et al, 2005; Roesch et al, 2007) and monkeys (Schultz et al, 1997; Morris et al, 2006; Bayer et al, 2007; Satoh et al, 2003; Matsumoto and Hikosaka, 2009).

Opportunity cost of time

If animals are seeking to maximize the average rate per unit time of rewards less punishments, then time spent doing nothing implies not getting those outcomes, and hence is either costly (if the rewards that could be available outweigh the punishments) or beneficial (if punishments outweigh rewards). This is quantified as an opportunity cost for the passage of time. It was analyzed as such by Niv et al (2007), and forms a key part of the argument of Cools et al (2010).

BEHAVIORAL PRIORS AND HEURISTICS

As many authors have pointed out, a more pernicious problem than the computational and sample complexities of goal-directed and habitual controls is the potentially calamitous expense for a subject having to engage in learning in the first place. For instance, there is surely an evolutionary disbenefit for organisms that have to learn for themselves to avoid predators by repeated bouts of danger and active escape. Rather, we may expect biases associated with previous expectations about the sorts of decision problems they face. These biases are evident in the choices subjects make, but are also enshrined in the functional architecture of decision-making itself. In this section, we review the biases; in the next section, we draw the relevant conclusions for the interactions between DA and 5-HT.

Layers of Biases

Several types of computational biases can alleviate the need for expensive sampling. The most basic bias is that outcomes, or indeed improvement or worsening in the prospects for future outcomes, have been caused by recent behavior. Thus, as in the law of effect (Thorndike, 1911), whatever actions preceded the delivery of reward should be done more, and whatever preceded punishment should be suppressed. Such a causality bias can be influenced by the recent past, for instance, being overturned by an experience of repeated lack of control. This has been argued to have a particularly important role in phenomena such as learned helplessness (Maier and Watkins, 2005). Here, animals are taught by experiencing inescapable shocks that

they cannot control or influence some aspect of one environment. They generalize this fact to other environments, which they therefore fail to explore or exploit appropriately. There are various ways to formalize this lack of control under which this behavior is actually optimal (Huys and Dayan, 2009).

A second bias is that active engagement is needed to secure possible rewards, and hence prediction of increased availability of rewards should energize behavior.

Other biases have to do with the *type* of response most adapted to a prediction. This is widely evident in Pavlovian responses. In these, predictions are directly tied to actions (ie, requiring value but not action learning), associated, for instance, with species-typical defensive actions (Bolles, 1970; Blanchard and Blanchard, 1971). The potency of Pavlovian conditioning is evident from the ability of Pavlovian responses to compete with instrumental ones, as in omission schedules (Sheffield, 1965; Williams and Williams, 1969), which have been argued as being the tip of an iceberg of more substantial anomalies of human decision making (Dayan et al, 2006; Chen and Bargh, 1999). Biases further arise in 'preparedness' to learn, which suggests that there are constraints as to the stimuli that can support predictions about particular outcomes (McNally, 1987).

Perhaps the most important bias for our argument has to do with the status of emitting (Go) vs withholding (No Go) actions. In principle, there could be an orthogonality between the valence of a possible outcome and the nature of the behavior required to collect or avoid it. Active responding or active or passive non-responding could equally be required for rewards or punishments. However,

when distant, punishments are often avoided by inhibition (Soubrié, 1986; Crockett *et al*, 2009), whereas rewards are gained through approach and engagement (Panksepp, 1998), and it appears that this coupling or non-orthogonality is enshrined in the architecture of control. That is, a fundamental structural principle of the basal ganglia appears to be the intimate coupling of Go, the direct pathway, thalamocortical cortical excitation, and reward, and No Go, the indirect pathway, thalamocortical inhibition, and punishment (Frank, 2005; Gerfen *et al*, 1995; Gerfen, 2000; Brown *et al*, 1999).

Appetitive Responses

We first consider preparatory and consummatory responses to predictions of rewards. These include approach, engagement, and active exploration (Panksepp, 1998). Such actions are consistent with the previous bias that rewards are relatively rare, and typically require active processes to be collected. The predictors are described as having high levels of incentive salience (Berridge, 2007; Berridge and Robinson, 1998; Alcaro *et al*, 2007). An additional appetitive bias is the PIT effect (Estes, 1943; Lovibond, 1983) that Pavlovian predictions of future reward can boost the vigor of instrumental actions, even if instrumental and Pavlovian outcomes are different (the so-called general PIT; Balleine, 2005). One suggestion is that appetitive PIT reflects a bias in the assessment of the overall rate of positive reinforcement, which is linked to the vigor of responding by acting as an opportunity cost for the passage of time (Niv *et al*, 2007).

These appetitively motivated behaviors are associated with the shaded upper right-hand quadrant of the graph in Figure 1. This graph is an adaptation of the so-called affective circumplex (Knutson and Greer, 2008; Larsen and Diener, 1992; Posner *et al*, 2005), replacing arousal, which is the normal ordinate, by action (Crockett *et al*, 2009; Guitart-Masip *et al*, 2010; Huys *et al*, 2010), but leaving the valence axis in its original form. This action axis is intended to include automatic, Pavlovian, responses, partly orchestrated by the nucleus accumbens (Reynolds and Berridge, 2001, 2002, 2008) as well as learned, instrumental, actions associated with the habitual or model-free systems, which are largely associated with the dorsolateral striatum (Yin *et al*, 2004; Everitt *et al*, 2008; Tricomi *et al*, 2009). It is also meant to reflect the level of vigor with which actions are performed. The intimate coupling between invigoration (Ikemoto and Panksepp, 1999) and appetitive valence is a critical part of our argument, and is in line, for instance, with the idea proposed by Wise and Bozarth (1987) that a single common underlying biological structure supports homologous functions in the stimulation of locomotion and positive reinforcement.

How goal-directed actions, and the areas supporting them such as the dorsomedial striatum (Yin *et al*, 2005), fit into this quadrant is unfortunately less clear.
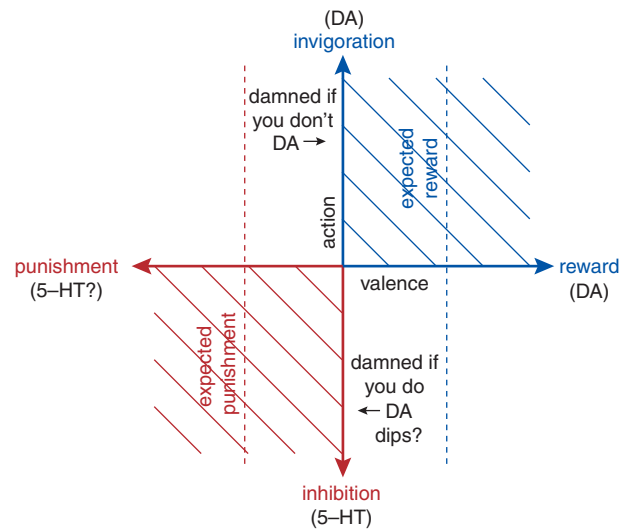

Figure 1. The affect–effect plot. This shows a version of the affective circumplex (Larsen and Diener, 1992; Knutson and Greer, 2008), replacing the arousal axis with action and inhibition. The top-right quadrant is associated with better-than-expected outcomes and their Pavlovian and instrumental effects in learning, approach, and vigor. The bottom-left quadrant is associated with worse-than-expected outcomes, and their effects in terms of inhibition and fear learning. The other quadrants involve more complex interactions between reward and punishment. The top-left arrow shows the case for actions associated with active avoidance of punishments following a movement of the origin left along the valence axis to reflect the prediction of danger; the bottom-right with actions that should be avoided to prevent the loss of otherwise expected rewards, associated with a movement of the origin right along this axis. Dopamine seems particularly associated with the upper right-hand quadrant and, because of its association with active avoidance, effects associated with moving the origin leftward. In the context of this paper, serotonin's prime association is with inhibition, i.e., negative values of the ordinate; the possibility of opponency is that it is responsible for the whole bottom-left quadrant.

Aversive Responses

Predictions of imminent or future punishment have a much more complex effect on behavior. Partly, this reflects the fundamental asymmetry between reward and punishment that successful response learning leads to *repeated* experience of rewards, but *avoided* experience of punishments. Therefore, mechanisms for maintenance and extinction of learnt responses are likely to seem different for rewards and punishments (Mowrer, 1947). However, there is another asymmetry: although it is usually uncomplicatedly safe to approach and engage with rewards, as mentioned above, punishments require a more complex set of species-, threat-, and distance-dependent responses (Blanchard *et al*, 2005; Bolles, 1970; Gray and McNaughton, 2003; McNaughton and Corr, 2004). First, many dangers arise when subjects execute inappropriate actions in dangerous conditions, such as venturing forth into unsafe terrain. A vital generic heuristic in such circumstances is therefore behavioral inhibition, which helps prevent actions of any sort (Gray and McNaughton, 2003). This link is consistent with the shaded, lower-left, quadrant of the graph in Figure 1, which ties inhibition, as an opponent of action,

to aversion, the opponent of reward. Conditioned suppression is another example. This acts like the aversive mirror of appetitive PIT, with a Pavlovian predictor of a shock having the power to suppress or inhibit ongoing, appetitively motivated, instrumental actions.

However, behavioral inhibition is only appropriate in particular circumstances. Often, and particularly in the face of a proximal threat, exactly the opposite is required, namely an active defensive response (Blanchard et al, 2005). The choice of the response depends sensitively on the threat, in ways that animals often cannot afford to have to learn for themselves (Bolles, 1970; McNaughton and Corr, 2004). Indeed, the essential decision as to what to do—fight, freeze, or flee (ultimately controlled by the dorsolateral periacqueductal gray (PAG); the more recuperatory or inhibitory flop being controlled by the ventrolateral PAG; Keay and Bandler, 2001)—depends on a complex risk assessment process. Active, aversively motivated actions can also be elicited by manipulations of activity in the nucleus accumbens, but in the caudal rather than the rostral shell (Reynolds and Berridge, 2001, 2002, 2008). The boundary between appetitively and aversively motivated behaviors depends on contextual factors such as the stressfulness or familiarity of the environment (Reynolds and Berridge, 2008), perhaps consistent with previous expectations about its capacity for potential harm.

The standard way to reconcile the architectural coupling between reward and active, Go, responses and the need for active and instrumental actions in the face of punishments is to introduce the notion of safety, as in forms of two-factor theory (Mowrer, 1947, 1956; Morris, 1975; Kim et al, 2006). This allows the transition from a dangerous state to a neutral or safe one to have the valence and effect of a reward. That is, resetting the baseline expected outcome to be negative (the putative punishment) implies that a neutral outcome will appear positive. It is natural also to extend this explanation to defensive Pavlovian responses. All these responses are occasioned by the opportunity to avoid a threat, and are thus energizing (and hence performed vigorously) and reinforcing (and hence repeated, if possible). They thus fit exactly into the realm of the upper right-hand quadrant of Figure 1 (and illustrated by the rightward-pointing arrow), provided that the origin is shifted leftward, to reflect the inherent danger (Ikemoto and Panksepp, 1999). In terms of Gray and McNaughton (2003), the fight, flight, (active) freezing system, which naturally occupies the upper-left quadrant of the figure, is brought into the upper-right quadrant, which is the territory of the behavioral approach system (BAS). Both are opposed by the behavioral inhibition system, associated with negative values of the ordinate. Coding via safety is parsimonious if the set of actions leading to continuation of punishment or threat is very large, as it is necessary to avoid them all, whereas any action leading to safety is good.

Such a shift of the origin toward negative values is also one way to reconcile the problem posed by Ainslie (2001), in the form of subjects' apparent elective willingness to

experience subjective pain (in the light of mechanisms operating in abnormal circumstances that can suppress it). Pain becomes a predictor that certain protective actions will give rise to future neutral affect (ie, safety), and is thus an appropriate reinforcer.

Safety signaling is thus conceived as the means by which an architecture that couples Go with reward can perform correctly when avoiding punishments requires Go. An obvious question is how an architecture coupling No Go and punishment could perform correctly when rewards require No Go. If the origin of the graph in Figure 1 can be moved rightward, based on the expectation of a reward, then the frustration of losing an expected reward (or of obtaining a smaller one) is endowed with negative valence and processed as punishment. The instrumental difficulties of omission schedules (Williams and Williams, 1969; Sheffield, 1965) or the differential reinforcement of low rates of responding (Staddon, 1965), in which automatic Go actions are penalized by the rules of the experiment, suggest that this shift of the origin may be more complicated than that for the case of punishment.

In sum, the hashed areas in Figure 1 show the seemingly automatic association between predictions of rewards and active engagement and predictions of punishments and behavioral inhibition. In these quadrants, Pavlovian and instrumental responses are generally aligned, and hence learning is easy. The two arrows show safety and frustration signaling that occupy the non-congruent quadrants, and are distinguished by greater difficulties of learning and complexities of the representational structures involved. Table 1 lists some key paradigms that fit into the easy and difficult quadrants of the affect-action graph.

Of course, even if the origin on the valence axis has to move in order to effect these architectural reconciliations, the composite stimulus (danger plus safety, or reward plus frustration) must retain its overall original valence, to guard against masochism or overcautiousness—that is, the aggregate event of punishment followed by its cessation should still be aversive, not appetitive. A way to achieve this would be to ensure that the magnitude of the counterbalancing affective factor be limited by that of the original one. There is some evidence that animals are, in general, surprisingly poor at such reweightings (Pompilio et al, 2006; Clement et al, 2000).

Additional Heuristics

Various other heuristics and priors may also be important for understanding reward and punishment interactions. For instance, one consequence of behavioral inhibition could be pruning of the goal-directed evaluation of states when predicted large punishments arise. This pruning has been suggested as leading to normally overoptimistic evaluation; so that depressive realism (Dayan and Huys, 2008; Alloy and Abramson, 1988; Watson and Clark, 1984) or rumination (Smith and Alloy, 2009) set in when it fails. Over-optimism could equally come from overeager approach and

TABLE 1 Conflicts (top left; bottom right) and congruence (top right; bottom left) between Pavlovian and Instrumental Responses

		Instrumental	
		− (Inhibition)	+ (Action)
Pavlovian	+ (Action)	Omission schedules Delayed responding tasks	Appetitive approach Active responding
	− (Inhibition)	Avoiding threats	Escape from fear (Cain and LeDoux, 2007)

Flight, fight, and active freezing are hard to put in the table because of Pavlovian sensitivity to defensive distance (McNaughton and Corr, 2004).

engagement with apparently rewarding options (Smith *et al*, 2006; Dayan *et al*, 2006). This is a case in which appetitive and aversive Pavlovian responses have similar implications.

Another critical form of pruning is appropriate in the face of substantial threat. In this case, a sensible heuristic is to downweight small rewards, as it is unrealistic to expect sufficient accumulation to overcome the cost of the punishment. One way to implement this would be for the signal that predicts the threat to have a contrast-enhancing effect on the signal reporting possible rewards, such that small potential or actual rewards would have less effect, and large rewards have more. This would make for a magnitude-dependent interaction between reward and punishment. Alternatively, the prospect of punishment might act as a motivational state that differentiates rewards relevant to escape from punishment from other generic rewards. Irrelevant, non-safety-related rewards would then be flattened selectively in a way reminiscent of motivational states such as hunger, thirst, or sodium deprivation (Dickinson and Balleine, 1994).

Talmi *et al* (2008) reported a decrease in appetitive sensitivity in a context in which rewards were financial and the punishments involved electrical shocks. This argues against a generic contrast-enhancing effect of punishment (although it is conceivable that the rewards used all fell into the 'too small' category for one of the controllers). It would be interesting to compare the effects of rewards that were irrelevant (eg, monetary) or relevant (eg, cessation of shocks) to safety to see if punishment creates a specific motivational state, and corresponding reward dissociation. One consequence of the sort of decrease in appetitive sensitivity that Talmi *et al* (2008) observed is enhanced exploration, with the lessened effective difference in the values of available options leading to an increased willingness to try ones that do not appear best. This is the same form of exploration that comes from an increased temperature in a softmax model of choice (Sutton and Barto, 1998), and, among its other properties, is a convenient heuristic for avoiding getting stuck performing suboptimal choices. Naturally, taking persistent advantage of any better action that is found requires an ultimate restoration of the appetitive contrast.

OPPONENCY AND BEYOND

The discussion in previous sections has led us to two coupled spectra associated with habits and Pavlovian influences as shown in Figure 1—from invigoration to inhibition, and from reward to punishment. In this section, we put these spectra into the context of general ideas about affective opponency (Konorski, 1967; Grossberg, 1984; Solomon and Corbit, 1974; Brodie and Shore, 1957) and specific notions about DA–5-HT valence opponency (Deakin, 1983; Deakin and Graeff, 1991; Graeff *et al*, 1998; Kapur and Remington, 1996).

We first briefly introduce the dopaminergic and serotonergic systems, then discuss the properties of these opponencies, consider some additional aspects of the role of 5-HT, and finally consider a third possible form of opponency associated with engagement, but involving norepinephrine (NE) and 5-HT rather than DA and 5-HT. Again, for the sake of space, we had mostly to ignore many factors, including most aspects of anatomical differentiation within DA and 5-HT systems.

The Dopaminergic System

DA cells are located in the midbrain, in the ventral tegmental area (VTA), and in substantia nigra pars compacta (SNc). The dopaminergic system can be divided into three pathways: the nigrostriatal pathway, which projects from the SNc to the dorsal striatum, the mesolimbic pathway, which projects from the VTA to the nucleus accumbens, and the mesocortical pathway, which projects from the VTA to the PFC (Fluxe *et al*, 1974; Swanson, 1982).

Recorded DA neurons fall into one of three categories (Grace and Bunney, 1983; Goto *et al*, 2007): (1) inactive neurons; (2) tonically firing neurons, displaying slow, single-spike firing; and (3) burst-firing neurons, exhibiting phasic firing driven by afferent input. Tonic activity reflects general states of arousal or motivation, whereas phasic activity may be related to the detection and nature of punctate salient events. There is evidence that these modes are separately regulated (Floresco *et al*, 2003), and the functional significance of phasic and tonic firing of DA has

been investigated by several authors (Goto *et al*, 2007; Grace, 1991; Niv *et al*, 2007; Floresco, 2007). As natural DA phasic activity may be obscured by a general tonic increase in DA function (eg, following DA agonist administration), this duality in firing modes complicates the interpretation of many pharmacological studies, as discussed, for instance, by Beninger and Miller (1998).

There are two major classes of receptors for DA, D1 and D2. Most critically for us, these may be segregated respectively to the two different direct and indirect pathways through the striatum (Aubert *et al*, 2000; Gerfen, 1992, 2004; but see Aizman *et al*, 2000; Inase *et al*, 1997). To a first approximation, the direct pathway facilitates action (Go) and the indirect pathway suppresses it (No Go). Phasic release of DA increases activity in the Go pathway through stimulation of D1 receptors (Hernández-López *et al*, 1997), which depends on relatively larger concentrations of DA (Goto and Grace, 2005). Conversely, tonic release of DA inhibits activity in the No Go pathway by stimulating D2 receptors (Hernandez-Lopez *et al*, 2000), which are sensitive to much lower concentrations of DA (Creese *et al*, 1983).

We concentrate below on the strong links between DA locomotor activity, RL, and the motivational effect of reinforcers (reviewed, eg, in Berridge and Robinson, 1998; Robbins and Everitt, 1992; Wise, 2008). However, we should also note briefly that DA has also been implicated in many other functions. First, DA modulates several aspects of executive function (as reviewed in Robbins, 2005; Robbins and Roberts, 2007; Robbins and Arnsten, 2009; Di Pietro and Seamans, 2007) and behavioral flexibility, modulating attentional set formation and shifting (but not reversal learning) (Floresco and Magyar, 2006). Its influence over working memory (Williams and Goldman-Rakic, 1995) famously follows an inverted U-shaped function; too little as well as too much DA impairs performance. Such nonmonotonicity has been influential as a more general explanatory schema for apparently paradoxical effects, particularly as different subjects may start from different sides of the peak of such curves. There is evidence for an inverse relationship between prefrontal and nucleus accumbens DA, particularly during aversive stress (Brake *et al*, 2000; Tzschentke, 2001; Del Arco and Mora, 2008; Deutch, 1992, 1993; Pascucci *et al*, 2007; Wilkinson, 1997). DA released in the dorsolateral striatum has also been implicated in cognitive processing (eg, Darvas and Palmiter, 2009; Baunez and Robbins, 1999).

Another important role of DA is that it modulates the interactions between prefrontal and limbic systems at the level of the nucleus accumbens and the amygdala. As reviewed in Grace *et al* (2007) and Sesack and Grace (2010), projections from the hippocampus, amygdala, and PFC converge on single neurons in the nucleus accumbens (Callaway *et al*, 1991; Mulder *et al*, 1998; Finch, 1996; French and Totterdell, 2002, 2003). Their inputs are gated by the hippocampus or by bursts of PFC activity (O'Donnell and Grace, 1995; Gruber and O'Donnell, 2009), and are differentially modulated by accumbal DA (Charara and Grace, 2003; Brady and O'Donnell, 2004; Floresco *et al*, 2001). DA also modulates the control of the basolateral amygdala by the mPFC (Kröner *et al*, 2005; Rosenkranz and Grace, 1999, 2001, 2002; Floresco and Tse, 2007; Grace and Rosenkranz, 2002).

The Serotonergic System and Its Regulation of the Dopaminergic System

5-HT neurons are located in nuclei of the midline of the brain stem. Ascending nuclei projecting to the forebrain mainly comprise the median raphe nucleus (MRN) and dorsal raphe nucleus (DRN); the DRN projects notably to the cortex, amygdala, striatum, thalamus, PAG, and hypothalamus, whereas the MRN innervates the cortex, septal nuclei, hippocampus, and hypothalamus (Azmitia and Segal, 1978; O'Hearn and Molliver, 1984; Geyer *et al*, 1976). The VTA and the substantia nigra pars reticulata, the nucleus accumbens, and the ventromedial aspect of the caudate nucleus receive a dense 5-HT projection, whereas the SNc and the rest of the caudate nucleus are more sparsely innervated (Lavoie and Parent, 1990; Beart and McDonald, 1982; Hervé *et al*, 1987). 5-HT neurons display a slow and 'clock-like' firing pattern (Jacobs and Fornal, 1991, 1999), but can also exhibit phasic activation to pain (Schweimer *et al*, 2008). DRN neurons can even be activated by reward (Nakamura *et al*, 2008; Bromberg-Martin *et al*, 2010; see Kranz *et al*, 2010 for a review of the modulation of reward by 5-HT) and, more generally, respond to a large range of specific sensorimotor information (Ranade and Mainen, 2009).

There are at least 14 different receptor subtypes for 5-HT (Cooper *et al*, 2002; Hoyer *et al*, 2002), making for a highly intricate and complex range of effects. Critical for us is the substantial experimental evidence that 5-HT regulates DA release (Esposito *et al*, 2008; Azmitia and Segal, 1978; Beart and McDonald, 1982; Hervé *et al*, 1987; Parent, 1981; Geyer *et al*, 1976; Egerton *et al*, 2008; De Deurwaerdère *et al*, 2004; Higgins and Fletcher, 2003; Lavoie and Parent, 1990; Spoont, 1992; Harrison *et al*, 1997; Nedergaard *et al*, 1988), and we provide some detail on this to make clear how far there is to go to fit all the interactions together. The precise mechanisms by which this happens appear very diverse, and reflect the complexity of DA regulation itself, for example, reducing the bursting behavior of DA cells (Di Giovanni *et al*, 1999), altering the relative balance between regional DA concentrations (De Deurwaerdère and Spampinato, 1999), or modulating the projections that control DA release (Bortolozzi *et al*, 2005). Regulation can be tonic, or conditional, on DA being activated (Leggio *et al*, 2009b; Lucas *et al*, 2001; Porras *et al*, 2003; De Deurwaerdère *et al*, 2005).

Most importantly, 5-HT displays functional tonic inhibitory control over DA, as lesioning the MRN or DRN increases the metabolism of DA in the nucleus accumbens and either reduces (MRN) or leaves unchanged (DRN) that

in the PFC (Hervé et al, 1979, 1981). Indeed, 5-HT$_{2C}$ receptors in general tonically inhibit DA release. However, many 5-HT receptor types (5-HT$_{1A}$, 5-HT$_{2A}$, 5-HT$_3$, 5-HT$_4$) *stimulate* DA release (Alex and Pehek, 2007; Di Matteo et al, 2008; Higgins and Fletcher, 2003), and 5-HT generally seems to exert an excitatory influence on the VTA (Beart and McDonald, 1982; Van Bockstaele et al, 1994), including enhancing DA release in the nucleus accumbens (Guan and McBride, 1989).

The opposition between 5-HT$_{2A}$ and 5-HT$_{2C}$ seems particularly striking. Both receptor types have been shown to display constitutive activity (Berg et al, 2005; De Deurwaerdère et al, 2004; Navailles et al, 2006), and exert opposite control over the release of DA in the nucleus accumbens and striatum (Porras et al, 2002; Di Giovanni et al, 1999; De Deurwaerdère and Spampinato, 1999) and in the PFC (Millan et al, 1998; Gobert et al, 2000; Pozzi et al, 2002; Alex et al, 2005; Pehek et al, 2006). 5-HT has a critical role in modulating impulsivity, perhaps partly indirectly by modulating DA release (Dalley et al, 2002, 2008; Winstanley et al, 2004, 2006; Millan et al, 2000a). Indeed, these effects may result in the behavioral observation that 5-HT$_{2A}$ receptors are associated with *increased* impulsivity, whereas 5-HT$_{2C}$ activity displays the more general correlation of 5-HT with decreased impulsivity (Robinson et al, 2008; Winstanley et al, 2004; Fletcher et al, 2007). Effects of 5-HT$_{2C}$ appear to be mediated by receptors at the level of the origin (VTA) for PFC DA (Alex et al, 2005; Pozzi et al, 2002), the target (striatum) for dorsolateral striatal DA (Alex et al, 2005), and both for the nucleus accumbens (Navailles et al, 2008).

Unfortunately, it is not even that simple. 5-HT$_{2C}$ activity does not always lead to a decrease in DA; thus, constitutive activity of 5-HT$_{2C}$ receptors in the mPFC contributes to the increase in accumbal DA following morphine, haloperidol, or cocaine (Leggio et al, 2009a, b). This may perhaps be related to the antagonism proposed by certain authors between mPFC DA and accumbal DA (Brake et al, 2000; Tzschentke, 2001; Del Arco and Mora, 2008; Deutch, 1992, 1993). 5-HT$_{2C}$ may exert a tonic inhibitory effect on structures involved in invigoration (De Deurwaerdère et al, 2010) in a way that need not depend on modulating DA.

Furthermore, 5-HT$_{1B}$ has been associated with not only decreased amphetamine-induced enhancement of responding for conditioned reward (Fletcher and Korth, 1999) and satiety (Lee et al, 2002), but also increased DA release (Neumaier et al, 2002; Alex and Pehek, 2007; Yan and Yan, 2001; Millan et al, 2003; Di Matteo et al, 2008) as well as increased amphetamine-induced locomotor hyperactivity (Przegalinski et al, 2001). Also, inhibition of innate escape responses has been linked with 5-HT$_{1A}$ (Deakin and Graeff, 1991; Misane et al, 1998).

Much less data exist as to whether DA exerts regulation over 5-HT release, and it is not clear whether this regulation would be excitatory or inhibitory (Ferré et al, 1994; Matsumoto et al, 1996; Ferré and Artigas, 1993; Thorré

et al, 1998). This could be taken as evidence for a hierarchical arrangement between 5-HT and DA, with the former exerting its effects by manipulating the latter. In this case, such weaker influences might be expected.

Like DA, 5-HT also modulates higher cognitive functions. For instance, PFC 5-HT has been shown to be necessary for reversal learning, but not attentional set formation or shifting, as recently reviewed in Robinson et al (2007), Clarke et al (2007), Robbins (2005), Robbins and Arnsten (2009), and Di Pietro and Seamans (2007).

Complexities of Opponency

When two systems are involved in representing a single spectrum, a number of complexities arise. First, if both systems have baseline activity, then negative values could be expressed by various possible combinations of below-baseline activity of the system representing positive values and above-baseline activity of the system representing negative values. In a precise sense, there is an additional degree of freedom—the net value only constrains the difference in activation of the two systems, leaving the sum of the activations free to be used to represent another quantity.

Second, the combination of Pavlovian and instrumental, and direct and learned, effects associated with these neuromodulators can make it hard to make clear inferences about the effects of manipulations. This point is well made by Bizot and Thiébot (1996) for the case of impulsivity. For instance, as we noted above, Huys and Dayan (2009) argued that 5-HT could have the direct effect of pruning actions associated with potentially negative outcomes, by virtue of a putative role in making aversive predictions. However, this could mean that any effect of reducing 5-HT on eliminating a capacity to make normal aversive predictions, as suggested by Deakin (1983), Deakin and Graeff (1991), and Graeff et al (1998), could be overwhelmed by a concomitant decrease in pruning, and thus increase in actual negative outcomes themselves. The ubiquity of autoreceptor-based negative feedback control over the activity and release of neuromodulatory neurons (see Bonhomme and Esposito, 1998; Millan et al, 2000a) also complicates experimental analyses. The same is true of the non-monotonic inverted U-shaped curves relating release to function (as seen for DA in its modulation of working memory; Williams and Goldman-Rakic, 1995).

Third, we have argued for the case of valence that there may be circumstances under which the origin of the spectrum can be moved to take a positive or negative value. This would imply that the semantic mappings from activity to valence in the individual systems are not fixed. In particular, safety (which implies an aversive context) can be coded in the same way as a truly appetitive outcome (Mowrer, 1947, 1956; Morris, 1975; Kim et al, 2006). This could lead to apparent cooperativity between otherwise competitive opponents. Some of the key paradigms suggesting cooperation and competition are listed in Table 2.

TABLE 2 Some Aspects of Serotonin Function that Suggest Competition or Collaboration with Dopamine or Neither

Opposing	-Tonic inhibition of accumbal DA release by 5-HT (Hervé et al, 1979, 1981)
	-5-HT2c receptors inhibit accumbal and striatal DA release (De Deurwaerdère et al, 2004)
	-DA necessary for active avoidance learning; 5-HT inhibits avoidance learning (Beninger, 1989)
	-DA reduces, 5-HT increases fatigue (Davis et al, 2000, Meeusen et al, 2006)
	-DA involved in appetite (engaging in behavior); 5-HT involved in satiation (ending behavior) (Berridge, 2007, Gruninger et al, 2007)
Collaborating	-Infusion of 5-HT in the nucleus accumbens increases DA levels and enhances responding (Sasaki-Adams and Kelley, 2001, Parsons and Justice, 1993)
	-Increased release of both 5-HT and DA in controllable punishment (Bland et al, 2003b)
	-5-HT2a receptor activity linked to increased hyperactivity and impulsivity (Fletcher et al, 2002, 2007)
	-Antipsychotic effect of combined 5-HT2a and D2 receptors blockade (eg, Schmidt et al, 1995, Meltzer and Huang, 2008)
	-Antidepressant effect of boosting either 5-HT or DA (or NE) (Hirschfeld, 1999, Nutt et al, 2007)
Neither	-Increased 5-HT in uncontrollable punishment (Bland et al, 2003b)
	-Neither DA nor 5-HT necessary for all forms of aversive contingency learning (Beninger, 1989)

Finally, as argued in the section 'Behavioral Priors and Heuristics', the two formally orthogonal spectra of action and valence are anatomically and functionally coupled. This can make it hard to interpret systemic manipulations of one or other system, as factors associated with affect and effect could masquerade as each other.

We organize our discussion around two characteristic types of effect of affective events and predictions: immediate (or proactive), involving modulating subjects' engagement with current and future actions, and retroactive, involving reinforcing or suppressing previous actions.

Immediate Effects

DA, affect, and effect. DA in the nucleus accumbens is known to have a role in Pavlovian responding, incentive salience, appetitive PIT, and vigor (Berridge, 2007; Berridge and Robinson, 1998; Alcaro et al, 2007; Niv et al, 2007;

McClure et al, 2003; Satoh et al, 2003; Dickinson et al, 2000; Lex and Hauber, 2008; Reynolds and Berridge, 2001, 2002, 2008). More generally, it is very well known that DA agonists enhance (whereas DA antagonists reduce) locomotor activity (see, eg, Beninger, 1983). This suggests that DA is responsible for the positive values along the vertical axis in Figure 1, which captures the spectrum from active engagement in choices and actions to inhibition and withdrawal. In line with this, phasic DA stimulates the Go pathway through the basal ganglia, whereas tonic DA inhibits the No Go pathway. Indeed, phasic activation of DA is directly associated with invigoration (Satoh et al, 2003) (putatively via a DA-dependent modulation consistent with appetitive PIT; Murschall and Hauber, 2006; Lex and Hauber, 2008). It has also been observed that animals act more quickly and vigorously when reward rates are higher, in a way that depends positively on (tonic) levels of DA in the striatum (Salamone and Correa, 2002). Such appetitive

aspects are consistent with the model of Niv *et al* (2007) that optimal vigor is tied to the average rates of reward, reported by tonic levels of DA, perhaps also reflecting the integrated phasic signals. (Differences such as this point between the current model and our original account of opponency (Daw *et al*, 2002) are discussed in detail in the Discussion section, and also form a key component of Cools *et al*, 2010.)

However, the involvement of DA in action is not limited to appetitive contexts. Even though many DA neurons are phasically inhibited by aversive outcomes and predictors (Ungless *et al*, 2004; Mirenowicz and Schultz, 1996), the concentration of DA and the phasic activity of other DA neurons (notably in the mesocortical pathway) both increase in the face of aversion (Iordanova, 2009; Sorg and Kalivas, 1991; Guarraci and Kapp, 1999; Kiaytkin, 1988; Abercrombie *et al*, 1989; Louilot *et al*, 1986; Brischoux *et al*, 2009; Lammel *et al*, 2008; Matsumoto and Hikosaka, 2009). There is also evidence that active defensive aversively motivated actions are only elicitable from the nucleus accumbens under normal dopaminergic conditions (Faure *et al*, 2008). Following the interpretation of safety signaling given in the previous section, the benefit of cessation of punishment is a suitable reward, and as such, should also energize behavior.

From a formal viewpoint, only those punishments that are considered controllable should inspire vigor— uncontrollable punishments are not associated with the prospect of safety and should lead to quiescence (Maier and Watkins, 2005; Huys and Dayan, 2009; Cools *et al*, 2010). Somewhat troubling for this interpretation is that DA efflux in the mPFC (though not the nucleus accumbens; Bland *et al*, 2003b) is actually increased during inescapable shock (Bland *et al*, 2003a) (perhaps via the mesocortical pathway; Lammel *et al*, 2008; Brischoux *et al*, 2009). One possibility is that this occurs just during the initial assessment of uncontrollability. Evidence for this is that following uncontrollability, a subsequent challenge with morphine does not boost DA efflux, although it does boost 5-HT efflux (Bland *et al*, 2003a).

The DA signal thus becomes associated more strongly with effect than affect, in that its involvement in active actions remains even when the overall situation is aversive. If the origin of the graph in Figure 1 is moved leftward toward a negative net valence, the former origin now has a positive value—that is, the possibility and means of achieving a neutral state becomes appetitive. This amounts to acknowledging that valence is not absolute (with what is considered reward and punishment being largely dependent on the current baseline), and restores the congruency between action and (apparent) reward.

5-HT and inhibition. The opposite end of the spectrum from invigoration in Figure 1 is inhibition. Consistent with one form of opponency, there are extensive findings on the role of 5-HT in this (Spoont, 1992; Gray and McNaughton, 2003; Soubrié, 1986). This involvement has been especially well documented in aversive situations, where 5-HT inhibits innate responses to

fear in the face of imminent threat as well as responses unrelated to escape in the face of distal threat (Deakin and Graeff, 1991; Graeff, 2004). Via its effects on behavioral inhibition, 5-HT has also been hypothesized to mediate optimistic evaluation (Dayan and Huys, 2008), and thereby be implicated in depressive realism (Alloy and Abramson, 1988; Keller *et al*, 2002) or worse (Carson *et al*, 2010).

Beside their typical slow 'clock-like' firing (Jacobs and Fornal, 1991, 1999), 5-HT neurons can also exhibit phasic activation to pain (Schweimer *et al*, 2008). Although it is not clear if such phasic activation is associated with momentary inhibition in the rather direct way that phasic DA is with invigoration, experiments such as of Shidara and Richmond (2004) and Hikosaka (2007) certainly provide inspiration for some such mechanisms.

There is ample evidence for the role of 5-HT in mediating inaction following uncontrollable punishment (Maier and Watkins, 2005). 5-HT release is increased during stress in the mPFC and amygdala (Kawahara *et al*, 1993; Yoshioka *et al*, 1995; Hashimoto *et al*, 1999), and inescapable shock activates subpopulations of serotonergic neurons in all raphe nuclei in the rat (Takase *et al*, 2004). Uncontrollability potentiates the stress-induced increase in the release of 5-HT in the mPFC and nucleus accumbens shell (Bland *et al*, 2003a, b). Descending connections from the mPFC to the DRN (Peyron *et al*, 1998; Baratta *et al*, 2009), which predominantly synapse onto inhibitory cells, inhibit this uncontrollability response (Amat *et al*, 2005, 2008) when subjects have previous experience with behavioral control over stress. This also blocks the behavioral effects of later uncontrollable stress (Amat *et al*, 2006), allowing normal invigoration.

On the other hand, 5-HT has also been linked with inhibition in non-aversive contexts. Some cases of inhibition, including correct No Go responding in appetitive settings (Fletcher, 1993; Harrison *et al*, 1997, 1999), can be viewed as the mirror images of DA's involvement in safety and active avoidance—with, for instance, the net quiescence needed to avoid impulsive responding mirroring the net vigor needed to reach safety (Cools *et al*, 2010). This preserves the congruence of inhibition with aversion, but does not accommodate other forms of inhibition associated with 5-HT, such as its involvement in the cessation of feeding after satiety (Gruninger *et al*, 2007), or other issues that may have to do with learning, such as the latent inhibition of conditioning to stimuli previously associated with an absence of affective outcome (Weiner, 1990) and extinction (Beninger and Phillips, 1979).

According to the formal model of opponency, one route by which 5-HT could inspire inhibition is the suppression of activation, for instance, via the suppression of DA. This could be a form of the hierarchical opponency mentioned above, with DA mediating its effects directly, but with 5-HT mediating its affective effects by acting on DA, potentially reassigning behavior away from currently motivated responding, either in appetitive (Sasaki-Adams and Kelley, 2001) or aversive (Archer, 1982) contexts.

Complexities. It has been argued (Frank, 2005; Frank and Claus, 2006; Cohen and Frank, 2009) that dips below baseline in DA activity will exert a particular effect over DA D2 receptors, and thus the indirect or No Go pathway, because D2 receptors have a relatively greater affinity for DA than D1 receptors (Creese *et al*, 1983; Surmeier *et al*, 2007). Such dips have directly been observed in electrophysiological recordings, and constitute some of the strongest evidence that DA reports the TD prediction error (Pan *et al*, 2005; Roesch *et al*, 2007; Schultz *et al*, 1997; Morris *et al*, 2006; Bayer *et al*, 2007). In this model, the indirect pathway suppresses responses that compete with the favored choice. However, it is not clear if this is a general mechanism for inhibition, having particular difficulty, for instance, if it is necessary to wait and inhibit actions in order to get a reward (as in differential reinforcement of low rates of responding). It is also not clear whether the more general sort of response inhibition that appears to be realized by the subthalamic nucleus and that has been associated with competition among different appetitive responses (Frank, 2006; Frank *et al*, 2007b; Bogacz and Gurney, 2007) can be adaptively harnessed to prevent responding altogether.

Along with coarse inhibitory control, such as reduction of vigor by effective inhibition of tonic DA, or, in satiation, decreasing sensitivity to food rewards (Simansky, 1996), the complex array of effects of different 5-HT receptors on DA could allow for a range of quite subtle effects, perhaps mediating forms of previous bias such as that noted in the section 'Behavioral Priors and Heuristics', with the prospect of punishment acting as a motivational state, devaluing apparent rewards not associated with avoidance or escape, or allowing broader exploration by mediating increased temperature in a softmax model of choice (Sutton and Barto, 1998). This would link 5-HT with the reward sensitivity dimension of impulsivity identified by Franken and Muris (2006).

Learned Effects

As mentioned, repetitive pairing of actions such as pressing a lever and outcomes such as rewards or punishments can lead to both stimulus-response (habit) and action-outcome and stimulus-action-outcome (goal-directed) association learning. We examine the roles of DA and 5-HT in both cases.

Appetitively motivated learning. In terms of appetitive learning, substantial experimental evidence suggests that the phasic activation of DA neurons in the VTA and SNc is consistent with its reporting the TD prediction error for reward (Barto, 1995; Montague *et al*, 1996; Schultz *et al*, 1997). This would be an ideal substrate for learning appropriate and inappropriate actions (Wise, 2008; Palmiter, 2008; Frank, 2005; Suri and Schultz, 1999; Montague *et al*, 2004). This proposition is consistent with overwhelming behavioral evidence showing that restricting DA

function often has the same effect as not delivering the reinforcer (in a way that cannot be explained by mere decrease in performance ability, as reviewed in Wise, 2004, 2008), and is further bolstered by the effects on appetitive learning of pharmacological manipulations of DA in normal volunteers (Pessiglione *et al*, 2006), and also in Parkinson's patients on and off DA-boosting medication (Frank *et al*, 2004). One prominent suggestion is that phasic bursts of activity in DA neurons act via D1 receptors and the direct pathway in the striatum to boost actions leading to unexpectedly large rewards (Frank, 2005; Frank and Claus, 2006; Frank *et al*, 2007a). Indeed, recent studies with genetically engineered mice that lack DA (reviewed in Palmiter, 2008) have pinpointed restoration of DA release in the dorsolateral striatum as sufficient for learning reinforced responses.

It is not clear from these results whether learning of actions (as opposed to performing them; Wise, 2004, 2008) by the goal-directed system also has a mandatory dependence on DA. Several paradigms provide elegant approaches to this question by enforcing different dopaminergic conditions during training and testing. Noting the reasonable assumption that aspects of allocentric spatial behavior are associated with non-habitual control (White and McDonald, 2002; Doeller and Burgess, 2008), it is known that the genetically engineered mice that do not produce DA can learn an appetitive T-maze (Robinson *et al*, 2005), as well as conditioned place preference for morphine (Hnasko *et al*, 2005) and cocaine (Hnasko *et al*, 2007) (which, interestingly, appears to be mediated by 5-HT in mutant, but not control, mice), provided DA function is restored during testing. However, in an operant lever pressing task, experimental results are mixed, in that the mice do not appear to have learned the association, but do seem to learn faster than mice that have not been exposed, once DA is restored (Robinson *et al*, 2007). In cases of cognitively more taxing tasks, DA is needed (Darvas and Palmiter, 2009), although it is not yet clear what this indicates about the interaction between goal-directed and habitual control.

Aversively motivated learning. It has also long been known that DA mediates the acquisition of instrumental, active avoidance (Go) responses to aversive stimuli (Beninger, 1989), and DA may be necessary to acquire Pavlovian startle potentiation to aversive stimuli (Fadok *et al*, 2009). As seen above in the case of drive effects, this associates DA more tightly with the action than with the valence dimension. Again, the congruence between reward and action can be restored by viewing safety as a reward (ie, moving the origin of the graph in Figure 1 leftward), which is then again coded by enhanced DA activity. Modeling safety thus leads to correct prediction of observed avoidance behavior (Johnson *et al*, 2001; Moutoussis *et al*, 2008; Maia, 2010).

There is little evidence for preserved avoidance learning by the goal-directed system in the absence of normal DA. Rats under neuroleptics fail to learn an active avoidance

escape response, yet slowly develop the response when tested drug-free in extinction (Beninger *et al*, 1980a; Beninger, 1989). Thus, DA is necessary for learning the active escape, but not for learning about the aversive value of the cue. It thus seems unlikely that learning negative predictions from actual punishments (ie, in the lower left quadrant of Figure 1) is mediated only, if at all, by dips in tonic DA. This raises the question as to the representation of the phasic TD prediction error associated with the delivery of more punishment than expected, that is, leftward excursions from the origin in its original position in Figure 1.

From the perspective of opponency, we would expect this to involve the phasic activity of 5-HT neurons. Indeed, as we have noted, these are activated in aversive contexts. This is clear from cellular imaging data on the activation of selected groups of 5-HT neurons under conditions of shocks and inescapable stress (Lowry, 2002; Takase *et al*, 2004), and direct neurophysiological evidence of the same thing from provably serotonergic neurons (Schweimer *et al*, 2008). Unfortunately, there is far less information about the correlates of phasic 5-HT activity than for phasic DA activity, and indeed newer methods for measuring and manipulating 5-HT in a far more selective manner are needed; we discuss some possibilities later.

Various experimental findings appear to rule out a critical involvement of 5-HT in at least some forms of aversive learning: decreasing 5-HT function seems to facilitate (and increasing 5-HT function to impair) active avoidance learning (eg, Archer, 1982; Archer *et al*, 1982, and see Beninger, 1989). However, the picture is not completely clear as to the effects on phasic signaling of lesions, depletion, or even dietary manipulation of 5-HT. All these manipulations exert primary control over tonic levels instead, leaving the possibility that indirect action via autoreceptors and adaptation of receptor sensitivities could have an opposite effect. Again, it is important to distinguish effects associated with performance from those with learning.

It is known, although, that genetically engineered mice that lack central 5-HT display enhanced contextual fear learning (Dai *et al*, 2008). This form of learning, via the hippocampus, may be more closely associated with the goal-directed than the habitual system. However, it certainly shows that 5-HT is unlikely to have a mandatory role for all types of fear learning.

The characterization of particular groups of dopaminergic or putatively dopaminergic neurons that respond to phasic punishment (Brischoux *et al*, 2009; Lammel *et al*, 2008; Matsumoto and Hikosaka, 2009) raises the intriguing possibility that a separate DA projection may have the role of the phasic prediction error for punishment, in the same way that we discussed above for vigor. Against this, although, is the observation we noted above that DA does not seem to be mandatory for learning predictions of punishment, and only for turning those predictions into appropriate avoidance actions. Of course, the caveat

mentioned above about the need for a better understanding of the effect of experimental manipulations over phasic signaling also applies here.

Also, as mentioned for the case of vigor, it is notable that, in the rat at least, the neurons concerned project to regions likely to be associated with goal-directed control (Brischoux *et al*, 2009; Lammel *et al*, 2008). This might, perhaps, be involved in the assessment of the possibility of safety inherent in the leftward movement of the origin in Figure 1, which is in turn related to the goal-directed notions of controllability that we discussed above (Huys and Dayan, 2009; Maier and Watkins, 2005).

Learning to suppress. There are contexts where the appropriate response is to withhold action, for example, to avoid electrical shocks triggered by lever presses. As we mentioned, the role of phasic dips below baseline of DA activity following non-delivery of expected reward (Pan *et al*, 2005; Roesch *et al*, 2007; Schultz *et al*, 1997; Morris *et al*, 2006; Bayer *et al*, 2007) in response inhibition is not completely obvious. In comparison, the role of these dips in the process of learning not to emit incorrect responses to avoid losing expected rewards is rather clearer. There is direct evidence from Parkinson's disease (Frank *et al*, 2004) and genetic studies (Frank *et al*, 2007a; Frank and Hutchison, 2009) for the involvement in this learning of the D2 receptors believed to be sensitive to these dips (Frank and Claus, 2006; Cohen and Frank, 2009). In terms of Figure 1, expectation of future reward is encoded as a movement rightward of the origin of the graph, so that the formerly neutral zone is now in negative territory; the dips in DA then capture the loss of the expected reward in case of a neutral outcome.

5-HT and Disengagement

DA is not only subject to a potentially hierarchical influence from 5-HT, it is also affected by NE (Millan *et al*, 2000b; Nurse *et al*, 1984; Guiard *et al*, 2008a, b; Villegier *et al*, 2003). This exhibits a similar pattern of net inhibition at the level of the VTA (Guiard *et al*, 2008b) coupled with a potential for targeted excitation (Auclair *et al*, 2004). However, the effect of NE is mostly in the opposite direction, that is, toward ergotropism, and active arousal, seeking, and energy expenditure, rather than trophotropism and replenishment (Ellison and Bresler, 1974; Villegier *et al*, 2003). Remembering the very original suggestions (Brodie and Shore, 1957), and the fact that there are many 5-HT pathways other than those partnering DA, it is intriguing to consider whether there could indeed be opponency between 5-HT and NE (Ellison and Bresler, 1974; Everitt and Robbins, 1991), as well as between 5-HT and DA. Recent experiments showing that NE and 5-HT control DA release, while inhibiting each other (Auclair *et al*, 2004; Tassin, 2008), might add evidence to early suggestions of a competition between NE and 5-HT for behavioral control,

in an opponency paralleling that between sympathetic and parasympathetic nervous systems (Ellison, 1979).

Hints as to the form of this additional opponency come in data associated with the idea that 5-HT is involved in withdrawal or disengagement from the environment, even in the absence of evident threats or reinforcing benefits of No Go (Ellison, 1979; Beninger, 1989; Tops *et al*, 2009). For instance, whereas NE is associated with the arousal, exploration, and the active processing of salient and action-relevant stimuli (Bouret and Sara, 2005; Dayan and Yu, 2006; Aston-Jones and Cohen, 2005), 5-HT is involved in disengagement from sensory stimuli (Beninger, 1989; Handley and McBlane, 1991), for example, because they have been associated with neutral outcomes, as in latent inhibition paradigms (Weiner, 1990), or are no longer associated with affective outcomes, as in extinction (Beninger and Phillips, 1979). Its involvement in fatigue (Newsholme *et al*, 1987; Meeusen *et al*, 2007) and satiation (Gruninger *et al*, 2007) may be related to this too.

DISCUSSION

Fathoming how the processing of reward and punishment are integrated in order to produce appropriate, and approximately appropriate, behavior is critical for understanding healthy and diseased decision making. Perhaps, because punishment and its prospect have such critical roles in sculpting choices, they are embedded deeply, and thus obscurely, in the architectural fabrics concerned. Punishment and threat also enjoy very substantial Pavlovian components, whose logic is only slowly becoming less murky.

It might seem self-evidently clear that reward and punishment are functional opponents, implying that the neural systems involved in processing them should be similarly antagonistic. However, we argued that computational and algorithmic adaptations to expectations about the previous structure of the environment make for substantial complexities in the relationship between the processing of punishment and reward. These collectively give rise to a mix of competitive, cooperative, and interactive associations between these opposing affective facets of the world. We suggested (Figure 1) that it is important to take particular account of an axis associated with invigoration and inhibition along with the one associated with valence, and furthermore that the origin of this graph can move leftward or rightward according to expectations of punishment or reward.

Extending, and sometimes contradicting Deakin (1983), Deakin and Graeff (1991), and Daw *et al* (2002), we considered the role played by DA and 5-HT in the aspects of this that are predominantly associated with Pavlovian and habitual control. DA appears to be responsible for one quadrant of the resulting graph in a rather uncomplicated manner, but dynamic interactions in opponency associated with movement of the origin result in its also being responsible for effects in other quadrants. Most notably,

active avoidance seems to be coded in a locally appetitive manner, as in safety signaling. We considered this as an algorithmic by-product of asymmetries in the effect of reward and punishment, making reinforcement of successful escape actions more parsimonious than inhibition of all actions that do not terminate punishment.

The least complicated association of 5-HT appears to be with behavioral inhibition (Soubrié, 1986), that is, the negative component of the action axis. However, we argued that it could also have a critical role in the analogous affective axis. There are suggestions that dips below baseline in the DA signal can report on the disbenefits of poor actions in the face of expected reward, and 5-HT could also mediate some aspects of action learning by modulating those dips. However, it does not appear crucial for all aspects of avoidance response learning (as opposed to fear learning). We also mentioned the notion that there might be opponency between 5-HT (associated with disengagement; Tops *et al*, 2009) and NE to partner that between 5-HT and DA.

This, and indeed also of Cools *et al* (2010), amount to a significant evolution of our original model (Daw *et al*, 2002) of DA and 5-HT opponency, which focused almost exclusively on valence, ignoring Pavlovian effects, asymmetries coming from previous distributions over environments, and the functional and anatomical association between action and valence. The most obvious contradiction is the role of tonic levels of DA and 5-HT, which we previously suggested as reporting average levels of punishment and reward, respectively. Here, we have adapted Niv *et al* (2007) to suggest that tonic DA, at least, reports the average levels of rewards and controllable punishments as an opportunity cost for the passage of time leading to vigor. Huys and Dayan (2009) and Cools *et al* (2010) suggest, *mutatis mutandis,* that tonic 5-HT reports average levels of punishments as an opportunity benefit for the passage of time, leading to quiescence.

The earlier assignment was based on two notions. First, in the workings of the models that estimate and maximize average rates of reward, phasic reward prediction errors are antagonized by the long-run rate (Sutton and Barto, 1998). We previously suggested that this would therefore be a role for the affective opponent. However, such antagonism could be realized in many other ways, for instance, by a form of adaptation. Second, the assignment of DA to average aversion provided an explanatory schema for the evidence that DA is boosted by punishment as well as reward (Iordanova, 2009; Sorg and Kalivas, 1991; Guarraci and Kapp, 1999; Kiaytkin, 1988; Abercrombie *et al*, 1989; Louilot *et al*, 1986; Brischoux *et al*, 2009; Lammel *et al*, 2008; Matsumoto and Hikosaka, 2009). In the present account, we have considered a finer distinction between controllable and uncontrollable punishment, and the movable origin of the valence axis in Figure 1.

Of course, we have still failed to realize anything like a complete theory of 5-HT function. Indeed, the prospects for this get more fantastic as the complexities within this neuromodulatory system and between this and others

become ever more evident. However, we suggest below that it is concrete enough to suggest some experimental directions.

One particularly critical lacuna has to do with the role the various factors we have discussed have in model-based or goal-directed controls. In the original work on the computational distinction between habitual and goal-directed controls, we suggested that it was in the former that learning was based on neuromodulatory signals such as the dopaminergic report of the appetitive TD prediction error (Daw et al, 2005), with goal-directed learning left to depend on conventional cortical plasticity. However, as we have seen, DA and 5-HT can exert control over processing in the PFC (Williams and Goldman-Rakic, 1995; Robbins, 2005; Robbins and Roberts, 2007; Robbins and Arnsten, 2009), the former via the mesocortical pathway (Fluxe et al, 1974; Swanson, 1982) that is recently understood to include dopaminergic neurons that are excited, rather than inhibited, by punishment (Brischoux et al, 2009; Lammel et al, 2008). DA also seems to influence learning in the hippocampus (see Kumaran and Duzel, 2008), whose involvement in goal-directed control, or perhaps a third form of so-called episodic control (Lengyel and Dayan, 2007), is currently unclear. There are theoretical reasons to do with the impossibility of performing model-based evaluation in complex domains for expecting model-based evaluation to be sometimes grounded in habitual or cached values, and so to expect such interactions. We have also argued that 5-HT can influence model-based evaluation through value-dependent pruning (Dayan and Huys, 2008).

Furthermore, we described model-based and model-free controls in terms of values of states and of performing particular actions at those states. Pavlovian control is also based on (predictions of) the values of states, and hence one might also expect it to reflect both model-based and model-free characteristics under appropriate circumstances. Indeed, there is evidence for this, for instance, in the details of PIT (the so-called outcome-specific PIT; Dickinson and Balleine, 2002; Holland, 2004; Johnson et al, 2009). However, there are many gaps in our knowledge about the nature of and interaction between model-based and model-free Pavlovian evaluation, particularly in aversive contexts, and it is an area that is ripe for study.

We argued that substantial aspects of the relationship between reward and punishment depend on previous expectations about the structure of environmental decision problems. Model-based control may incorporate sophisticated previous information about higher-order aspects of the environment, notably facets such as controllability (Huys and Dayan, 2009), possibly realizing the effects of these on control via neuromodulators such as 5-HT (Maier et al, 1993). However, by itself, habitual control can likely only capture coarser aspects of previous structure, such as overall propensities towards reward or punishment, overall rates of change of contingencies, or overall engagement with the environment.

An alternative set of ideas about the role of 5-HT suggests that it controls the discounting of distant reinforcers compared with proximal ones (Doya, 2000; Schweighofer et al, 2007; Mobini et al, 2000a, b; Schweighofer et al, 2008). These started from the original finding that low 5-HT levels lead to impulsivity in selecting small rewards that are delivered soon, and vice versa (Bizot and Thiébot, 1996; Bizot et al, 1988; Poulos et al, 1996; Wogar et al, 1993), although this is not a ubiquitous finding in humans (Crean et al, 2002; Tanaka et al, 2007) or other animals (Winstanley et al, 2004). Discounting is reviewed in detail by Cardinal (2006) and, in this particular form, can be explained by a large number of different possible factors in RL models (Williams and Dayan, 2005), including a decrease in discounting, and also a bias toward action (eg, from an unwarrantedly large influence of DA favoring Go responses) or too deterministic a choice of actions in the face of insufficient knowledge about the contingencies (a factor we called brittleness; and which could relate to the effect of DA on the gain of recurrently connected cortical neurons; Servan-Schreiber et al, 1990). The explanation associated with Figure 1 would have suggested impaired competition along the action axis for the Pavlovian approach response to the early reward. The discounting view of 5-HT has certainly led to a range of interesting findings about the construction and representation of values at different timescales in the brain (Schweighofer et al, 2007; Tanaka et al, 2004, 2007). This particular theory is hard to integrate with ours—bar the ecumenical possibility is that it involves a different function of a separate group of 5-HT neurons.

Future Directions

The most immediately addressable concern associated with the view we have portrayed here involves distinguishing affective value from the requirement for action. That is, it is important to orthogonalize Go, No Go, punishment, and reward, and also the orientation of the action with respect to the cues (to manipulate other aspects of the Pavlovian status of the action), along with the factor controlling whether rewards are related to punishment (eg, money gain vs money loss) or not (eg, money gain vs electric shocks). Crockett et al (2009) is one experiment along these lines; forthcoming behavioral and neuroimaging studies (Guitart-Masip et al, 2010; Huys et al, 2010) should, hopefully, add significantly to our understanding. One facet of these studies is the use of rich RL-based models to fit the behavior; these may help tease apart contributions from different aspects of action and valence.

In terms of theory, we have argued that the architecture of decision making is significantly influenced by aspects of previous expectations about the statistical diet of affective problems. Future work is thus needed to understand this rather understudied aspect of natural environmental statistics, and also how previous biases might be encoded in the different structures involved in choice. It is an intriguing prospect, indeed, deserving substantial study that many apparent anomalies of decision making, including

substantial challenges to the beloved and derided *homo economicus,* arise from interactions between Pavlovian and instrumental systems that harm their collective ability to report veridically on reward and punishment.

However, the key requirement here is a spatially and temporally far finer scale view of the activity of 5-HT (and DA) neurons and release in target structures. New methods based on optogenetics (Tsai *et al*, 2009) and cyclic voltammetry (Hashemi *et al*, 2009), along with technologies such as targeted viral rescue in mice genetically engineered to be deficient in particular neuromodulators (so far, most notably involving DA; Darvas and Palmiter, 2009; Szczypka *et al*, 2001), may collectively finally provide definitive answers as to the roles of genetically defined (sub)populations of DA and 5-HT cells in appetitive and aversive affect and effect.

ACKNOWLEDGEMENTS

DISCLOSURE

The authors declare no conflict of interest.

REFERENCES

Abercrombie ED, Keefe KA, DiFrischia DS, Zigmond MJ (1989). Differential effect of stress on in vivo dopamine release in striatum, nucleus accumbens, and medial frontal cortex. *J Neurochem* **52**: 1655–1658.

Abrams JK, Johnson PL, Hollis JH, Lowry CA (2004). Anatomic and functional topography of the dorsal raphe nucleus. *Ann NY Acad Sci* **1018**: 46–57.

Ainslie G (2001). *Breakdown of Will*. Cambridge University Press: Cambridge, England.

Aizman O, Brismar H, Uhlén P, Zettergren E, Levey AI, Forssberg H *et al* (2000). Anatomical and physiological evidence for D1 and D2 dopamine receptor colocalization in neostriatal neurons. *Nat Neurosci* **3**: 226–230.

Alcaro A, Huber R, Panksepp J (2007). Behavioral functions of the mesolimbic dopaminergic system: an affective neuroethological perspective. *Brain Res Rev* **56**: 283–321.

Alex KD, Pehek EA (2007). Pharmacologic mechanisms of serotonergic regulation of dopamine neurotransmission. *Pharmacol Ther* **113**: 296–320. **Review of the negative and positive influences of serotonin on dopamine, plus the implications of this for disease**.

Alex KD, Yavanian GJ, McFarlane HG, Pluto CP, Pehek EA (2005). Modulation of dopamine release by striatal 5-HT2C receptors. *Synapse* **55**: 242–251.

Alloy L, Abramson L (1988). Depressive realism: four theoretical perspectives. In: Alloy L (ed). *Cognitive Processes in Depression*. Guilford: New York, NY. pp 223–265.

Amat J, Baratta MV, Paul E, Bland ST, Watkins LR, Maier SF (2005). Medial prefrontal cortex determines how stressor controllability affects behavior and dorsal raphe nucleus. *Nat Neurosci* **8**: 365–371.

Amat J, Paul E, Watkins LR, Maier SF (2008). Activation of the ventral medial prefrontal cortex during an uncontrollable stressor reproduces both the immediate and long-term protective effects of behavioral control. *Neuroscience* **154**: 1178–1186.

Amat J, Paul E, Zarza C, Watkins LR, Maier SF (2006). Previous experience with behavioral control over stress blocks the behavioral and dorsal raphe nucleus activating effects of later uncontrollable stress: role of the ventral medial prefrontal cortex. *J Neurosci* **26**: 13264–13272.

Archer T (1982). Serotonin and fear retention in the rat. *J Comp Physiol Psychol* **96**: 491–516.

Archer T, Ogren SO, Ross SB (1982). Serotonin involvement in aversive conditioning: reversal of the fear retention deficit by long-term p-chloroamphetamine but not p-chlorophenylalanine. *Neurosci Lett* **34**: 75–82.

Aston-Jones G, Cohen JD (2005). Adaptive gain and the role of the locus coeruleus-norepinephrine system in optimal performance. *J Comp Neurol* **493**: 99–110.

Aubert I, Ghorayeb I, Normand E, Bloch B (2000). Phenotypical characterization of the neurons expressing the D1 and D2 dopamine receptors in the monkey striatum. *J Comp Neurol* **418**: 22–32.

Auclair A, Drouin C, Cotecchia S, Glowinski J, Tassin J-P (2004). 5-HT2A and alpha1b-adrenergic receptors entirely mediate dopamine release, locomotor response and behavioural sensitization to opiates and psychostimulants. *Eur J Neurosci* **20**: 3073–3084.

Awh E, Gehring WJ (1999). The anterior cingulate cortex lends a hand in response selection. *Nat Neurosci* **2**: 853–854.

Azmitia EC, Segal M (1978). An autoradiographic analysis of the differential ascending projections of the dorsal and median raphe nuclei in the rat. *J Comp Neurol* **179**: 641–667.

Balleine B, Killcross S (1994). Effects of ibotenic acid lesions of the nucleus accumbens on instrumental action. *Behav Brain Res* **65**: 181–193.

Balleine BW (2005). Neural bases of food-seeking: affect, arousal and reward in corticostriatolimbic circuits. *Physiol Behav* **86**: 717–730. **Review of the substantial body of studies on functionally and anatomically different systems involved in appetitive decision-making**.

Balleine BW, Dickinson A (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* **37**: 407–419.

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

Baratta MV, Zarza CM, Gomez DM, Campeau S, Watkins LR, Maier SF (2009). Selective activation of dorsal raphe nucleus-projecting neurons in the ventral medial prefrontal cortex by controllable stress. *Eur J Neurosci* **30**: 1111–1116.

Barto A (1989). From chemotaxis to cooperativity: abstract exercises in neuronal learning strategies. In: Durbin R, Miall C, Mitchison G (eds). *The Computing Neuron*. Addison-Wesley: Wokingham, England. pp 73–98.

Barto A (1995). Adaptive critics and the basal ganglia. In: Houk J, Davis J, Beiser D (eds). *Models of Information Processing in the Basal Ganglia*. MIT Press: Cambridge, MA. pp 215–232.

Baunez C, Robbins TW (1999). Effects of dopamine depletion of the dorsal striatum and further interaction with subthalamic nucleus lesions in an attentional task in the rat. *Neuroscience* **92**: 1343–1356.

Baxter MG, Parker A, Lindner CC, Izquierdo AD, Murray EA (2000). Control of response selection by reinforcer value requires interaction of amygdala and orbital prefrontal cortex. *J Neurosci* **20**: 4311–4319.

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

Beart PM, McDonald D (1982). 5-hydroxytryptamine and 5-hydroxytryptaminergic-dopaminergic interactions in the ventral tegmental area of rat brain. *J Pharm Pharmacol* **34**: 591–593.

Beninger R (1989). The role of serotonin and dopamine in learning to avoid aversive stimuli. In: Archer T, Nilsson L-G (eds). *Aversion, Avoidance, and Anxiety: Perspective on Aversively Motivated Behavior*. Lawrence Erlbaum: Hillsdale, NJ. Ch. 10, pp 265–284 **Review of early experimental findings on one-way active avoidance learning, suggesting that serotonin is involved in tuning out irrelevant stimuli, and dopamine in learning and maintaining avoidance responses**.

Beninger R, Mason S, Phillips A, Fibiger H (1980a). The use of extinction to investigate the nature of neuroleptic-induced avoidance deficits. *Psychopharmacology* **69**: 11–18.

Beninger RJ (1983). The role of dopamine in locomotor activity and learning. *Brain Res* **287**: 173–196.

Beninger RJ, Mason ST, Phillips AG, Fibiger HC (1980b). The use of conditioned suppression to evaluate the nature of neuroleptic-induced avoidance deficits. *J Pharmacol Exp Ther* **213**: 623–627.

Beninger RJ, Miller R (1998). Dopamine D1-like receptors and reward-related incentive learning. *Neurosci Biobehav Rev* **22**: 335–345.

Beninger RJ, Phillips AG (1979). Possible involvement of serotonin in extinction. *Pharmacol Biochem Behav* **10**: 37–41.

Berg KA, Harvey JA, Spampinato U, Clarke WP (2005). Physiological relevance of constitutive activity of 5-HT2A and 5-HT2C receptors. *Trends Pharmacol Sci* **26**: 625–630.

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

Berridge KC, Robinson TE (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res Rev* **28**: 209–269.

Bertsekas DP, Tsitsiklis JN (1996). *Neuro-Dynamic Programming (Optimization and Neural Computation Series, 3)*. Athena Scientific: Nashua, NH.

Bezard E (ed) (2006). *Recent Breakthroughs in Basal Ganglia Research*. Nova Publishers: Hauppauge, NY.

Bizot JC, Thiébot MH (1996). Impulsivity as a confounding factor in certain animal tests of cognitive function. *Brain Res Cogn Brain Res* **3**: 243–250.

Bizot JC, Thiébot MH, Bihan CL, Soubrié P, Simon P (1988). Effects of imipramine-like drugs and serotonin uptake blockers on delay of reward in rats possible implication in the behavioral mechanism of action of antidepressants. *J Pharmacol Exp Ther* **246**: 1144–1151.

Blanchard DC, Blanchard RJ, Griebel G (2005). Defensive responses to predator threat in the rat and mouse. *Current Protocols in Neuroscience*, 8.19.1–8.19.20.

Blanchard R, Blanchard D (1971). Defensive reactions in the albino rat. *Learn Motiv* **2**: 351–362.

Blanchard RJ, Blanchard DC (1989). Antipredator defensive behaviors in a visible burrow system. *J Comp Psychol* **103**: 70–82.

Bland ST, Hargrave D, Pepin JL, Amat J, Watkins LR, Maier SF (2003a). Stressor controllability modulates stress-induced dopamine and serotonin efflux and morphine-induced serotonin efflux in the medial prefrontal cortex. *Neuropsychopharmacology* **28**: 1589–1596.

Bland ST, Twining C, Watkins LR, Maier SF (2003b). Stressor controllability modulates stress-induced serotonin but not dopamine efflux in the nucleus accumbens shell. *Synapse* **49**: 206–208.

Bogacz R, Gurney K (2007). The basal ganglia and cortex implement optimal decision making between alternative actions. *Neural Comput* **19**: 442–477.

Bolles RC (1970). Species-specific defense reactions and avoidance learning. *Psychol Rev* **77**: 32–48. **Underlines the importance of considering innate species-specific defense reactions**.

Bonhomme N, Esposito E (1998). Involvement of serotonin and dopamine in the mechanism of action of novel antidepressant drugs: a review. *J Clin Psychopharmacol* **18**: 447–454.

Bortolozzi A, Díaz-Mataix L, Scorza MC, Celada P, Artigas F (2005). The activation of 5-HT receptors in prefrontal cortex enhances dopaminergic activity. *J Neurochem* **95**: 1597–1607.

Botvinick MM, Cohen JD, Carter CS (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends Cogn Sci* **8**: 539–546.

Bouret S, Sara SJ (2005). Network reset: a simplified overarching theory of locus coeruleus noradrenaline function. *Trends Neurosci* **28**: 574–582.

Brady AM, O'Donnell P (2004). Dopaminergic modulation of prefrontal cortical input to nucleus accumbens neurons in vivo. *J Neurosci* **24**: 1040–1049.

Brake WG, Flores G, Francis D, Meaney MJ, Srivastava LK, Gratton A (2000). Enhanced nucleus accumbens dopamine and plasma corticosterone stress responses in adult rats with neonatal excitotoxic lesions to the medial prefrontal cortex. *Neuroscience* **96**: 687–695.

Brischoux F, Chakraborty S, Brierley DI, Ungless MA (2009). Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. *Proc Natl Acad Sci USA* **106**: 4894–4899. **Uses juxtacellular labeling to show that a selected group of frontally projecting dopamine neurons in the ventral tegmental area are activated rather than suppressed by punishment; thus updating Ungless *et al* (2004)**.

Brodie BB, Shore PA (1957). A concept for a role of serotonin and norepinephrine as chemical mediators in the brain. *Ann NY Acad Sci* **66**: 631–642. **Seminal suggestion about functional opponency between serotonin and norepinephrine, as controlling parasympathetic and sympathetic systems, respectively**.

Brog JS, Salyapongse A, Deutch AY, Zahm DS (1993). The patterns of afferent innervation of the core and shell in the "accumbens" part of the rat ventral striatum: immunohistochemical detection of retrogradely transported fluoro-gold. *J Comp Neurol* **338**: 255–278.

Bromberg-Martin ES, Hikosaka O, Nakamura K (2010). Coding of task reward value in the dorsal raphe nucleus. *J Neurosci* **30**: 6262–6272.

Brown J, Bullock D, Grossberg S (1999). How the basal ganglia use parallel excitatory and inhibitory learning pathways to selectively respond to unexpected rewarding cues. *J Neurosci* **19**: 10502–10511.

Bussey TJ, Muir JL, Everitt BJ, Robbins TW (1997). Triple dissociation of anterior cingulate, posterior cingulate, and medial frontal cortices on visual discrimination tasks using a touchscreen testing procedure for the rat. *Behav Neurosci* **111**: 920–936.

Cador M, Robbins TW, Everitt BJ (1989). Involvement of the amygdala in stimulus-reward associations: interaction with the ventral striatum. *Neuroscience* **30**: 77–86.

Cain CK, LeDoux JE (2007). Escape from fear: a detailed behavioral analysis of two atypical responses reinforced by CS termination. *J Exp Psychol Anim Behav Process* **33**: 451–463.

Callaway CW, Hakan RL, Henriksen SJ (1991). Distribution of amygdala input to the nucleus accumbens septi: an electrophysiological investigation. *J Neural Transm Gen Sect* **83**: 215–225.

Cardinal RN (2006). Neural systems implicated in delayed and probabilistic reinforcement. *Neural Netw* **19**: 1277–1301.

Cardinal RN, Parkinson JA, Hall J, Everitt BJ (2002). Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neurosci Biobehav Rev* **26**: 321–352.

Carr DB, Sesack SR (2000). Projections from the rat prefrontal cortex to the ventral tegmental area: target specificity in the synaptic associations with mesoaccumbens and mesocortical neurons. *J Neurosci* **20**: 3864–3873.

Carson RC, Hollon SD, Shelton RC (2010). Depressive realism and clinical depression. *Behav Res Ther* **48**: 257–265.

Charara A, Grace AA (2003). Dopamine receptor subtypes selectively modulate excitatory afferents from the hippocampus and amygdala to rat nucleus accumbens neurons. *Neuropsychopharmacology* **28**: 1412–1421.

Chen M, Bargh J (1999). Consequences of automatic evaluation: immediate behavioral predispositions to approach or avoid the stimulus. *Pers Soc Psychol Bull* **25**: 215–224.

Clarke HF, Walker SC, Dalley JW, Robbins TW, Roberts AC (2007). Cognitive inflexibility after prefrontal serotonin depletion is behaviorally and neurochemically specific. *Cereb Cortex* **17**: 18–27.

Clement TS, Feltus JR, Kaiser DH, Zentall TR (2000). 'Work ethic' in pigeons: reward value is directly related to the effort or time required to obtain the reward. *Psychon Bull Rev* **7**: 100–106.

Cohen MX, Frank MJ (2009). Neurocomputational models of basal ganglia function in learning, memory and choice. *Behav Brain Res* **199**: 141–156.

Cools R, Nakamura K, Daw N (2010). Serotonin and dopamine: Unifying affective and activational functions. *Neuropsychopharmacology*. (in press)

Cools R, Roberts AC, Robbins TW (2008). Serotonergic regulation of emotional and behavioural control processes. *Trends Cogn Sci* **12**: 31–40.

Cooper J, Bloom F, Roth R (2002). *The Biochemical Basis of Neuropharmacology* 8 edn. OUP: New York, NY.

Corbit LH, Balleine BW (2003). The role of prelimbic cortex in instrumental conditioning. *Behav Brain Res* **146**: 145–157.

Corbit LH, Balleine BW (2005). Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of Pavlovian-instrumental transfer. *J Neurosci* **25**: 962–970.

Corbit LH, Muir JL, Balleine BW (2001). The role of the nucleus accumbens in instrumental conditioning: evidence of a functional dissociation between accumbens core and shell. *J Neurosci* **21**: 3251–3260.

Crean J, Richards JB, de Wit H (2002). Effect of tryptophan depletion on impulsive behavior in men with or without a family history of alcoholism. *Behav Brain Res* **136**: 349–357.

Creese I, Sibley DR, Hamblin MW, Leff SE (1983). The classification of dopamine receptors: relationship to radioligand binding. *Annu Rev Neurosci* **6**: 43–71.

Crockett MJ, Clark L, Robbins TW (2009). Reconciling the role of serotonin in behavioral inhibition and aversion: acute tryptophan depletion abolishes punishment-induced inhibition in humans. *J Neurosci* **29**: 11993–11999.

Dai J-X, Han H-L, Tian M, Cao J, Xiu J-B, Song N-N *et al* (2008). Enhanced contextual fear memory in central serotonin-deficient mice. *Proc Natl Acad Sci USA* **105**: 11981–11986.

Dalley JW, Mar AC, Economidou D, Robbins TW (2008). Neurobehavioral mechanisms of impulsivity: fronto-striatal systems and functional neurochemistry. *Pharmacol Biochem Behav* **90**: 250–260.

Dalley JW, Theobald DE, Eagle DM, Passetti F, Robbins TW (2002). Deficits in impulse control associated with tonically-elevated serotonergic function in rat prefrontal cortex. *Neuropsychopharmacology* **26**: 716–728.

Darvas M, Palmiter RD (2009). Restriction of dopamine signaling to the dorsolateral striatum is sufficient for many cognitive behaviors. *Proc Natl Acad Sci USA* **106**: 14664–14669.

Davis JM, Alderson NL, Welsh RS (2000). Serotonin and central nervous system fatigue: nutritional considerations. *Am J Clin Nutr* **72**(2 Suppl): 573S–578S.

Daw ND, Doya K (2006). The computational neurobiology of learning and reward. *Curr Opin Neurobiol* **16**: 199–204.

Daw ND, Kakade S, Dayan P (2002). Opponent interactions between serotonin and dopamine. *Neural Netw* **15**: 603–616. **Original reinforcement learning theory of tonic and phasic opponency between serotonin and dopamine that the present paper is attempting to refine.**

Daw ND, Niv Y, Dayan P (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci* **8**: 1704–1711. **Computational decomposition of different (model-based and model-free) systems involved in instrumental control, drawing particularly on the results of Dickinson and Balleine (2002) and Killcross and Coutureau (2003).**

Daw ND, Touretzky DS (2002). Long-term reward prediction in TD models of the dopamine system. *Neural Comput* **14**: 2567–2583.

Dayan P (2008). The role of value systems in decision-making. In: Engel C, Singer W (eds). *Better Than Conscious: Decision Making, the Human Mind, and Implications for Institutions*. Ernst Strüngmann Forum, MIT Press: Cambridge, MA. pp 51–70.

Dayan P, Huys QJM (2008). Serotonin, inhibition, and negative mood. *PLoS Comput Biol* **4**: e4. **Theoretical assault on the apparent contradiction between the opponency-based suggestion that serotonin is involved in reporting aversive contingencies, and yet boosting, rather than suppressing; it is a treatment of choice in depression.**

Dayan P, Huys QJM (2009). Serotonin in affective control. *Ann Rev Neurosci* **32**: 95–126.

Dayan P, Niv Y, Seymour B, Daw ND (2006). The misbehavior of value and the discipline of the will. *Neural Netw* **19**: 1153–1160. **Modeling to encompass malign (and benign) Pavlovian effects over instrumental choice, using negative automaintenance as an example.**

Dayan P, Seymour B (2008). Values and actions in aversion. In: Glimcher P, Camerer C, Poldrack R, Fehr E (eds). *Neuroeconomics: Decision Making and the Brain*. Academic Press: New York, NY. pp 175–191.

Dayan P, Yu AJ (2006). Phasic norepinephrine: a neural interrupt signal for unexpected events. *Network* **17**: 335–350.

De Deurwaerdère P, Moine CL, Chesselet M-F (2010). Selective blockade of serotonin 2C receptor enhances Fos expression specifically in the striatum and the subthalamic nucleus within the basal ganglia. *Neurosci Lett* **469**: 251–255.

De Deurwaerdère P, Moison D, Navailles S, Porras G, Spampinato U (2005). Regionally and functionally distinct serotonin3 receptors control in vivo dopamine outflow in the rat nucleus accumbens. *J Neurochem* **94**: 140–149.

De Deurwaerdère P, Navailles S, Berg KA, Clarke WP, Spampinato U (2004). Constitutive activity of the serotonin2C receptor inhibits in vivo dopamine release in the rat striatum and nucleus accumbens. *J Neurosci* **24**: 3235–3241.

De Deurwaerdère P, Spampinato U (1999). Role of serotonin(2A) and serotonin (2B/2C) receptor subtypes in the control of accumbal and striatal serotonin release elicited in vivo by dorsal raphe nucleus electrical stimulation. *J Neurochem* **73**: 1033–1042.

Deakin JFW (1983). Roles of serotonergic systems in escape, avoidance and other behaviours. In: Cooper SJ, (ed.). *Theory in Psychopharmacology* Vol 2, 2nd edn. Academic Press: New York, pp 149–193.

Deakin JFW, Graeff FG (1991). 5-HT and mechanisms of defence. *J Psychopharmacol* **5**: 305–316. **Review of these authors' fecund suggestion about serotonin's role in suppressing and boosting consummatory and preparatory aversive responses, then also associated with dopamine–serotonin opponency.**

Del Arco A, Mora F (2008). Prefrontal cortex-nucleus accumbens interaction: in vivo modulation by dopamine and glutamate in the prefrontal cortex. *Pharmacol Biochem Behav* **90**: 226–235.

Deutch AY (1992). The regulation of subcortical dopamine systems by the prefrontal cortex: interactions of central dopamine systems and the pathogenesis of schizophrenia. *J Neural Transm Suppl* **36**: 61–89.

Deutch AY (1993). Prefrontal cortical dopamine systems and the elaboration of functional corticostriatal circuits: implications for schizophrenia and Parkinson's disease. *J Neural Transm Gen Sect* **91**: 197–221.

Devinsky O, Morrell MJ, Vogt BA (1995). Contributions of anterior cingulate cortex to behaviour. *Brain* **118**(Part 1): 279–306.

Di Giovanni G, De Deurwaerdère P, Di Mascio M, Di Matteo V, Esposito E, Spampinato U (1999). Selective blockade of serotonin-2C/2B receptors enhances mesolimbic and mesostriatal dopaminergic function: a combined in vivo electrophysiological and microdialysis study. *Neuroscience* **91**: 587–597.

Di Matteo V, Di Giovanni G, Pierucci M, Esposito E (2008). Serotonin control of central dopaminergic function: focus on in vivo microdialysis studies. *Prog Brain Res* **172**: 7–44.

Di Pietro NC, Seamans JK (2007). Dopamine and serotonin interactions in the prefrontal cortex: insights on antipsychotic drugs and their mechanism of action. *Pharmacopsychiatry* **40**(Suppl 1): S27–S33.

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

Dickinson A, Balleine B (2002). The role of learning in motivation. In: Gallistel C (ed). *Stevens' Handbook of Experimental Psychology*, Wiley: New York, NY. Vol 3, pp 497–533. **Review of a substantial body of psychological and behavioral neuroscience studies into functional differentiation and interaction between instrumental and Pavlovian learning and response systems.**

Dickinson A, Dearing MF (1979). Appetitive-aversive interactions and inhibitory processes. In: Dickinson A, Boakes RA (eds). *Mechanisms of Learning and Motivation*. Erlbaum, Hillsdale, NJ. pp 203–231.

Dickinson A, Smith J, Mirenowicz J (2000). Dissociation of Pavlovian and instrumental incentive learning under dopamine antagonists. *Behav Neurosci* **114**: 468–483.

Doeller CF, Burgess N (2008). Distinct error-correcting and incidental learning of location relative to landmarks and boundaries. *Proc Natl Acad Sci USA* **105**: 5909–5914.

Doya K (1999). What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Netw* **12**: 961–974.

Doya K (2000). Metalearning, neuromodulation and emotion. In: Hatano G, Okada N, Ta H (eds). *Affective Minds*. Elsevier Science: Amsterdam. pp 101–104.

Egerton A, Ahmad R, Hirani E, Grasby PM (2008). Modulation of striatal dopamine release by 5-HT2A and 5-HT2C receptor antagonists: (11C)raclopride PET studies in the rat. *Psychopharmacology (Berl)* **200**: 487–496.

Ellison G (1979). Chemical systems of the brain and evolution. In: Oakley D, Plotkin H (eds). *Brain, Behaviour and Evolution*. Methuen; London. Ch. 4, pp 78–98.

Ellison G, Bresler D (1974). Tests of emotional behavior in rats following depletion of norepinephrine, of serotonin, or of both. *Psychopharmacology* **34**: 275–288.

Esposito E, Di Matteo V, Di Giovanni G (2008). Serotonin-dopamine interaction: an overview. *Prog Brain Res* **172**: 3–6.

Estes W (1943). Discriminative conditioning. I. a discriminative property of conditioned anticipation. *J Exp Psychol* **32**: 150–155.

Everitt B, Robbins T (1991). Commentary on '5-HT and mechanisms of defence'. *J Psychopharmacol (Oxford)* **5**: 327–329.

Everitt BJ, Belin D, Economidou D, Pelloux Y, Dalley JW, Robbins TW (2008). Neural mechanisms underlying the vulnerability to develop compulsive drug-seeking habits and addiction. *Philos Trans R Soc London B Biol Sci* **363**: 3125–3135 Review.

Fadok JP, Dickerson TMK, Palmiter RD (2009). Dopamine is necessary for cue dependent fear conditioning. *J Neurosci* **29**: 11089–11097.

Faure A, Reynolds SM, Richard JM, Berridge KC (2008). Mesolimbic dopamine in desire and dread: enabling motivation to be generated by localized glutamate disruptions in nucleus accumbens. *J Neurosci* **28**: 7184–7192.

Fellows LK (2007). The role of orbitofrontal cortex in decision making: a component process account. *Ann NY Acad Sci* **1121**: 421–430.

Ferré S, Artigas F (1993). Dopamine D2 receptor-mediated regulation of serotonin extracellular concentration in the dorsal raphe nucleus of freely moving rats. *J Neurochem* **61**: 772–775.

Ferré S, Cortés R, Artigas F (1994). Dopaminergic regulation of the serotonergic raphe-striatal pathway: microdialysis studies in freely moving rats. *J Neurosci* **14**: 4839–4846.

Finch DM (1996). Neurophysiology of converging synaptic inputs from the rat prefrontal cortex, amygdala, midline thalamus, and hippocampal formation onto single neurons of the caudate/putamen and nucleus accumbens. *Hippocampus* **6**: 495–512.

Fletcher PJ (1993). A comparison of the effects of dorsal or median raphe injections of 8-OH-DPAT in three operant tasks measuring response inhibition. *Behav Brain Res* **54**: 187–197. **Evidence implicating median raphe 5-HT neurons in controlling inhibition of behavior induced by non-reward.**

Fletcher PJ, Grottick AJ, Higgins GA (2002). Differential effects of the 5-HT(2A) receptor antagonist M100907 and the 5-HT(2C) receptor antagonist SB242084 on cocaine-induced locomotor activity, cocaine self-administration and cocaine-induced reinstatement of responding. *Neuropsychopharmacology* **27**: 576–586.

Fletcher PJ, Korth KM (1999). Activation of 5-HT1B receptors in the nucleus accumbens reduces amphetamine-induced enhancement of responding for conditioned reward. *Psychopharmacology* **142**: 165–174.

Fletcher PJ, Tampakeras M, Sinyard J, Higgins GA (2007). Opposing effects of 5-HT(2A) and 5-HT(2C) receptor antagonists in the rat and mouse on premature

responding in the five-choice serial reaction time test. *Psychopharmacology (Berl)* **195**: 223–234.

Floresco SB (2007). Dopaminergic regulation of limbic-striatal interplay. *J Psychiatry Neurosci* **32**: 400–411.

Floresco SB, Blaha CD, Yang CR, Phillips AG (2001). Modulation of hippocampal and amygdalar-evoked activity of nucleus accumbens neurons by dopamine: cellular mechanisms of input selection. *J Neurosci* **21**: 2851–2860.

Floresco SB, Magyar O (2006). Mesocortical dopamine modulation of executive functions: beyond working memory. *Psychopharmacology (Berl)* **188**: 567–585.

Floresco SB, Tse MT (2007). Dopaminergic regulation of inhibitory and excitatory transmission in the basolateral amygdala-prefrontal cortical pathway. *J Neurosci* **27**: 2045–2057.

Floresco SB, West AR, Ash B, Moore H, Grace AA (2003). Afferent modulation of dopamine neuron firing differentially regulates tonic and phasic dopamine transmission. *Nat Neurosci* **6**: 968–973.

Fluxe K, Hökfelt T, Johansson O, Jonsson G, Lidbrink P, Ljungdahl A (1974). The origin of the dopamine nerve terminals in limbic and frontal cortex. Evidence for meso-cortico dopamine neurons. *Brain Res* **82**: 349–355.

Foster DJ, Wilson MA (2007). Hippocampal theta sequences. *Hippocampus* **17**: 1093–1099.

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

Frank MJ (2006). Hold your horses: a dynamic computational role for the subthalamic nucleus in decision making. *Neural Netw* **19**: 1120–1136. Neurobiology of decision making.

Frank MJ, Claus ED (2006). Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychol Rev* **113**: 300–326.

Frank MJ, Hutchison K (2009). Genetic contributions to avoidance-based decisions: striatal D2 receptor polymorphisms. *Neuroscience* **164**: 131–140.

Frank MJ, Moustafa AA, Haughey HM, Curran T, Hutchison KE (2007a). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proc Natl Acad Sci USA* **104**: 16311–16316.

Frank MJ, Samanta J, Moustafa AA, Sherman SJ (2007b). Hold your horses: impulsivity, deep brain stimulation, and medication in Parkinsonism. *Science* **318**: 1309–1312.

Frank MJ, Seeberger LC, O'reilly RC (2004). By carrot or by stick: cognitive reinforcement learning in Parkinsonism. *Science* **306**: 1940–1943.

Franken I, Muris P (2006). Gray's impulsivity dimension: a distinction between reward sensitivity versus rash impulsiveness. *Pers Indiv Differ* **40**: 1337–1347.

French SJ, Totterdell S (2002). Hippocampal and prefrontal cortical inputs monosynaptically converge with individual projection neurons of the nucleus accumbens. *J Comp Neurol* **446**: 151–165.

French SJ, Totterdell S (2003). Individual nucleus accumbens-projection neurons receive both basolateral amygdala and ventral subicular afferents in rats. *Neuroscience* **119**: 19–31.

Fudge JL, Haber SN (2000). The central nucleus of the amygdala projection to dopamine subpopulations in primates. *Neuroscience* **97**: 479–494.

Gabriel M, Kubota Y, Sparenborg S, Straube K, Vogt BA (1991). Effects of cingulate cortical lesions on avoidance learning and training-induced unit activity in rabbits. *Exp Brain Res* **86**: 585–600.

Gabriel M, Moore J (eds) (1991). *Learning and Computational Neuroscience: Foundations of Adaptive Networks*. Bradford Books/The MIT Press: Cambridge, MA.

Gerfen C (2004). Basal ganglia. In: Paxinos G (ed). *The Rat Nervous System*. Elsevier Academic Press: San Diego.

Gerfen CR (1992). The neostriatal mosaic: multiple levels of compartmental organization. *Trends Neurosci* **15**: 133–139.

Gerfen CR (2000). Molecular effects of dopamine on striatal-projection pathways. *Trends Neurosci* **23**(10 Suppl): S64–S70.

Gerfen CR, Keefe KA, Gauda EB (1995). D1 and D2 dopamine receptor function in the striatum: coactivation of D1- and D2-dopamine receptors on separate populations of neurons results in potentiated immediate early gene response in D1-containing neurons. *J Neurosci* **15**: 8167–8176.

Geyer MA, Puerto A, Dawsey WJ, Knapp S, Bullard WP, Mandell AJ (1976). Histologic and enzymatic studies of the mesolimbic and mesostriatal serotonergic pathways. *Brain Res* **106**: 241–256.

Glickman SE, Schiff BB (1967). A biological theory of reinforcement. *Psychol Rev* **74**: 81–109.

Gobert A, Rivet JM, Lejeune F, Newman-Tancredi A, Adhumeau-Auclair A, Nicolas JP *et al* (2000). Serotonin(2C) receptors tonically suppress the activity of mesocortical dopaminergic and adrenergic, but not serotonergic, pathways: a combined dialysis and electrophysiological analysis in the rat. *Synapse* **36**: 205–221.

Goto Y, Grace AA (2005). Dopaminergic modulation of limbic and cortical drive of nucleus accumbens in goal-directed behavior. *Nat Neurosci* **8**: 805–812.

Goto Y, Otani S, Grace AA (2007). The Yin and Yang of dopamine release: a new perspective. *Neuropharmacology* **53**: 583–587.

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

Grace AA, Bunney BS (1983). Intracellular and extracellular electrophysiology of nigral dopaminergic neurons-1. Identification and characterization. *Neuroscience* **10**: 301–315.

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

Grace AA, Rosenkranz JA (2002). Regulation of conditioned responses of basolateral amygdala neurons. *Physiol Behav* **77**: 489–493.

Graeff FG (2004). Serotonin, the periaqueductal gray and panic. *Neurosci Biobehav Rev* **28**: 239–259.

Graeff FG, Guimaraes FS, De Andrade TGCS, Deakin JFW (1998). Role of 5HT in stress, anxiety and depression. *Pharm Biochem Behav* **54**: 129–141.

Grahn RE, Will MJ, Hammack SE, Maswood S, McQueen MB, Watkins LR *et al* (1999). Activation of serotonin-immunoreactive cells in the dorsal raphe nucleus in rats exposed to an uncontrollable stressor. *Brain Res* **826**: 35–43.

Gray JA, McNaughton N (2003). *The Neuropsychology of Anxiety*, 2nd edn. OUP: Oxford, England.

Groenewegen HJ, Becker NE, Lohman AH (1980). Subcortical afferents of the nucleus accumbens septi in the cat, studied with retrograde axonal transport of horseradish peroxidase and bisbenzimid. *Neuroscience* **5**: 1903–1916.

Groenewegen HJ, Room P, Witter MP, Lohman AH (1982). Cortical afferents of the nucleus accumbens in the cat, studied with anterograde and retrograde transport techniques. *Neuroscience* **7**: 977–996.

Groenewegen HJ, Wright CI, Beijer AV, Voorn P (1999). Convergence and segregation of ventral striatal inputs and outputs. *Ann NY Acad Sci* **877**: 49–63.

Grossberg S (1984). Some normal and abnormal behavioral syndromes due to transmitter gating of opponent processes. *Biol Psychiatry* **19**: 1075–1118. **Review of one of the more comprehensive and extensive dynamical theories of affective opponency**.

Gruber AJ, O'Donnell P (2009). Bursting activation of prefrontal cortex drives sustained up states in nucleus accumbens spiny neurons in vivo. *Synapse* **63**: 173–180.

Gruninger TR, LeBoeuf B, Liu Y, Garcia LR (2007). Molecular signaling involved in regulating feeding and other motivated behaviors. *Mol Neurobiol* **35**: 1–20.

Guan XM, McBride WJ (1989). Serotonin microinfusion into the ventral tegmental area increases accumbens dopamine release. *Brain Res Bull* **23**: 541–547.

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

Guiard BP, Mansari ME, Blier P (2008a). Cross-talk between dopaminergic and noradrenergic systems in the rat ventral tegmental area, locus ceruleus, and dorsal hippocampus. *Mol Pharmacol* **74**: 1463–1475.

Guiard BP, Mansari ME, Merali Z, Blier P (2008b). Functional interactions between dopamine, serotonin and norepinephrine neurons: an in-vivo electrophysiological study in rats with monoaminergic lesions. *Int J Neuropsychopharmacol* **11**: 625–639.

Guitart-Masip M, Fuentemilla L, Bach D, Huys Q, Dayan P, Dolan R *et al* (2010). Action and valence representations in the human striatum and dopaminergic midbrain. (in submission).

Gupta AS, van der Meer MAA, Touretzky DS, Redish AD (2010). Hippocampal replay is not a simple function of experience. *Neuron* **65**: 695–705.

Haber SN, Fudge JL, McFarland NR (2000). Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *J Neurosci* **20**: 2369–2382.

Haber SN, Knutson B (2010). The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology* **35**: 4–26.

Hall J, Parkinson JA, Connor TM, Dickinson A, Everitt BJ (2001). Involvement of the central nucleus of the amygdala and nucleus accumbens core in mediating Pavlovian influences on instrumental behaviour. *Eur J Neurosci* **13**: 1984–1992.

Handley S, McBlane J (1991). 5-HT the disengaging transmitter? *J Psychopharmacol* **5**: 322–326.

Harrison AA, Everitt BJ, Robbins TW (1997). Central 5-HT depletion enhances impulsive responding without affecting the accuracy of attentional performance: interactions with dopaminergic mechanisms. *Psychopharmacology (Berl)* **133**: 329–342.

Harrison AA, Everitt BJ, Robbins TW (1999). Central serotonin depletion impairs both the acquisition and performance of a symmetrically reinforced go/no-go conditional visual discrimination. *Behav Brain Res* **100**: 99–112.

Hashemi P, Dankoski EC, Petrovic J, Keithley RB, Wightman RM (2009). Voltammetric detection of 5-hydroxytryptamine release in the rat brain. *Anal Chem* **81**: 9462–9471.

Hashimoto S, Inoue T, Koyama T (1999). Effects of conditioned fear stress on serotonin neurotransmission and freezing behavior in rats. *Eur J Pharmacol* **378**: 23–30.

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

Hernandez-Lopez S, Tkatch T, Perez-Garci E, Galarraga E, Bargas J, Hamm H et al (2000). D2 dopamine receptors in striatal medium spiny neurons reduce L-type Ca2+ currents and excitability via a novel PLC(beta)1-IP3-calcineurin-signaling cascade. *J Neurosci* **20**: 8987–8995.

Hervé D, Pickel VM, Joh TH, Beaudet A (1987). Serotonin axon terminals in the ventral tegmental area of the rat: fine structure and synaptic input to dopaminergic neurons. *Brain Res* **435**: 71–83.

Hervé D, Simon H, Blanc G, Lemoal M, Glowinski J, Tassin JP (1981). Opposite changes in dopamine utilization in the nucleus accumbens and the frontal cortex after electrolytic lesion of the median raphe in the rat. *Brain Res* **216**: 422–428.

Hervé D, Simon H, Blanc G, Lisoprawski A, Moal ML, Glowinski J et al (1979). Increased utilization of dopamine in the nucleus accumbens but not in the cerebral cortex after dorsal raphe lesion in the rat. *Neurosci Lett* **15**: 127–133.

Higgins GA, Fletcher PJ (2003). Serotonin and drug reward: focus on 5-HT2C receptors. *Eur J Pharmacol* **480**: 151–162.

Hikosaka O (2007). Basal ganglia mechanisms of reward-oriented eye movement. *Ann NY Acad Sci* **1104**: 229–249.

Hirschfeld RM (1999). Efficacy of SSRIs and newer antidepressants in severe depression: comparison with TCAs. *J Clin Psychiatry* **60**: 326–335.

Hnasko TS, Sotak BN, Palmiter RD (2005). Morphine reward in dopamine-deficient mice. *Nature* **438**: 854–857.

Hnasko TS, Sotak BN, Palmiter RD (2007). Cocaine-conditioned place preference by dopamine-deficient mice is mediated by serotonin. *J Neurosci* **27**: 12484–12488.

Holland PC (2004). Relations between Pavlovian-instrumental transfer and reinforcer devaluation. *J Exp Psychol Anim Behav Process* **30**: 104–117.

Holland PC, Gallagher M (2003). Double dissociation of the effects of lesions of basolateral and central amygdala on conditioned stimulus-potentiated feeding and Pavlovian-instrumental transfer. *Eur J Neurosci* **17**: 1680–1694.

Hollerman JR, Tremblay L, Schultz W (2000). Involvement of basal ganglia and orbitofrontal cortex in goal-directed behavior. *Prog Brain Res* **126**: 193–215.

Houk JC, Davis JL, Beiser DG (eds) (1994). *Models of Information Processing in the Basal Ganglia*. MIT Press: Cambridge, MA.

Hoyer D, Hannon JP, Martin GR (2002). Molecular, pharmacological and functional diversity of 5-HT receptors. *Pharmacol Biochem Behav* **71**: 533–554.

Huys Q, Cools R, Gölzer M, Fiedel E, Heinz A, Dolan R et al (2010). Approaching avoidance: instrumental and Pavlovian asymmetries in the processing of rewards and punishments. (in submission).

Huys QJM, Dayan P (2009). A Bayesian formulation of behavioral control. *Cognition* **113**: 314–328.

Ikemoto S, Panksepp J (1999). The role of nucleus accumbens dopamine in motivated behavior: a unifying interpretation with special reference to reward-seeking. *Brain Res Brain Res Rev* **31**: 6–41.

Inase M, Li BM, Tanji J (1997). Dopaminergic modulation of neuronal activity in the monkey putamen through D1 and D2 receptors during a delayed Go/Nogo task. *Exp Brain Res* **117**: 207–218.

Iordanova MD (2009). Dopaminergic modulation of appetitive and aversive predictive learning. *Rev Neurosci* **20**: 383–404.

Jacobs BL, Fornal CA (1991). Activity of brain serotonergic neurons in the behaving animal. *Pharmacol Rev* **43**: 563–578.

Jacobs BL, Fornal CA (1999). Activity of serotonergic neurons in behaving animals. *Neuropsychopharmacology* **21**(2 Suppl): 9S–15S.

Joel D, Weiner I (2000). The connections of the dopaminergic system with the striatum in rats and primates: an analysis with respect to the functional and compartmental organization of the striatum. *Neuroscience* **96**: 451–474.

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

Johnson A, Redish AD (2007). Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *J Neurosci* **27**: 12176–12189.

Johnson AW, Gallagher M, Holland PC (2009). The basolateral amygdala is critical to the expression of Pavlovian and instrumental outcome-specific reinforcer devaluation effects. *J Neurosci* **29**: 696–704.

Johnson J, Li W, Li J, Klopf A (2001). A computational model of learned avoidance behavior in a one-way avoidance experiment. *Adapt Behav* **9**: 91.

Kalivas PW, Duffy P (1995). Selective activation of dopamine transmission in the shell of the nucleus accumbens by stress. *Brain Res* **675**: 325–328.

Kapur S, Remington G (1996). Serotonin-dopamine interaction and its relevance to schizophrenia. *Am J Psychiatry* **153**: 466–476.

Kawahara H, Yoshida M, Yokoo H, Nishi M, Tanaka M (1993). Psychological stress increases serotonin release in the rat amygdala and prefrontal cortex assessed by in vivo microdialysis. *Neurosci Lett* **162**: 81–84.

Keay KA, Bandler R (2001). Parallel circuits mediating distinct emotional coping reactions to different types of stress. *Neurosci Biobehav Rev* **25**: 669–678.

Keller P, Lipkus I, Rimer B (2002). Depressive realism and health risk accuracy: the negative consequences of positive mood. *J Consum Res* **29**: 57–69.

Kiaytkin E (1988). Functional properties of presumed doparmine-containing and other ventral tegmental area neurons in conscious rats. *Int J Neurosci* **42**: 21–43.

Killcross S, Coutureau E (2003). Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb Cortex* **13**: 400–408. **Evidence for parallel, rather than serial, computation of model-based (goal-directed) and model-free (habitual) choices; these putatively depend on neuromodulators in different ways**.

Killcross S, Robbins TW, Everitt BJ (1997). Different types of fear-conditioned behaviour mediated by separate nuclei within amygdala. *Nature* **388**: 377–380.

Kim H, Shimojo S, O'Doherty J (2006). Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. *PLoS Biol* **4**: e233.

Klopf A (1982). *The Hedonistic Neuron: A Theory of Memory, Learning and Intelligence*. Hemisphere: Washington, London.

Knutson B, Greer SM (2008). Anticipatory affect: neural correlates and consequences for choice. *Philos Trans R Soc London B Biol Sci* **363**: 3771–3786.

Konorski J (1967). *Integrative Activity of the Brain: An Interdisciplinary Approach*. University of Chicago Press: Chicago, IL.

Kranz GS, Kasper S, Lanzenberger R (2010). Reward and the serotonergic system. *Neuroscience* **166**: 1023–1035.

Kröner S, Rosenkranz JA, Grace AA, Barrionuevo G (2005). Dopamine modulates excitability of basolateral amygdala neurons in vitro. *J Neurophysiol* **93**: 1598–1610.

Kumaran D, Duzel E (2008). The hippocampus and dopaminergic midbrain: old couple, new insights. *Neuron* **60**: 197–200.

Lammel S, Hetzel A, Häckel O, Jones I, Liss B, Roeper J (2008). Unique properties of mesoprefrontal neurons within a dual mesocorticolimbic dopamine system. *Neuron* **57**: 760–773.

Larsen R, Diener E (1992). Promises and problems with the circumplex model of emotion. In: Clark M (ed). *Review of Personality and Social Psychology: Emotion*. Sage: Newbury Park, CA. Vol. 13, pp 25–29.

Lavoie B, Parent A (1990). Immunohistochemical study of the serotoninergic innervation of the basal ganglia in the squirrel monkey. *J Comp Neurol* **299**: 1–16.

Lee MD, Kennett GA, Dourish CT, Clifton PG (2002). 5-HT1B receptors modulate components of satiety in the rat: behavioural and pharmacological analyses of the selective serotonin1B agonist CP-94,253. *Psychopharmacology (Berl)* **164**: 49–60.

Leggio GM, Cathala A, Moison D, Cunningham KA, Piazza PV, Spampinato U (2009a). Serotonin2C receptors in the medial prefrontal cortex facilitate cocaine-induced dopamine release in the rat nucleus accumbens. *Neuropharmacology* **56**: 507–513.

Leggio GM, Cathala A, Neny M, Rouge-Pont F, Drago F, Piazza PV et al (2009b). In vivo evidence that constitutive activity of serotonin2C receptors in the medial prefrontal cortex participates in the control of dopamine release in the rat nucleus accumbens: differential effects of inverse agonist versus antagonist. *J Neurochem* **111**: 614–623.

Lengyel M, Dayan P (2007). Hippocampal contributions to control: The third way. In: Platt J, Koller D, Singer Y and Roweis S (eds). *Advances in Neural Information Processing Systems* Vol 20. MIT Press, pp 889–896.

Lex A, Hauber W (2008). Dopamine D1 and D2 receptors in the nucleus accumbens core and shell mediate Pavlovian-instrumental transfer. *Learn Mem* **15**: 483–491.

Louilot A, Moal ML, Simon H (1986). Differential reactivity of dopaminergic neurons in the nucleus accumbens in response to different behavioral situations. an in vivo voltammetric study in free moving rats. *Brain Res* **397**: 395–400.

Lovibond PF (1983). Facilitation of instrumental behavior by a Pavlovian appetitive conditioned stimulus. *J Exp Psychol Anim Behav Process* **9**: 225–247.

Lowry CA (2002). Functional subsets of serotonergic neurones: implications for control of the hypothalamic-pituitary-adrenal axis. *J Neuroendocrinol* **14**: 911–923. **Evidence for extensive functional segregation between different subsets of serotonin neurons, significantly complicating the task for electrophysiologists**.

Lucas G, Matteo VD, De Deurwaerdère P, Porras G, Martín-Ruiz R, Artigas F et al (2001). Neurochemical and electrophysiological evidence that 5-HT4 receptors

exert a state-dependent facilitatory control in vivo on nigrostriatal, but not mesoaccumbal, dopaminergic function. *Eur J Neurosci* **13**: 889–898.

Maia TV (2010). Two-factor theory, the actor-critic model, and conditioned avoidance. *Learn Behav* **38**: 50–67.

Maier SF, Grahn RE, Kalman BA, Sutton LC, Wiertelak EP, Watkins LR (1993). Role of amygdala and dorsal raphe nucleus in mediating the behavioural consequences of inescapable shock. *Behav Neurosci* **107**: 377–388.

Maier SF, Watkins LR (2005). Stressor controllability and learned helplessness: the roles of the dorsal raphe nucleus, serotonin, and corticotropin-releasing factor. *Neurosci Biobehav Rev* **29**: 829–841. **Review of learned helplessness, a key animal model for depression, in which 5-HT has a central role in realizing the consequences of uncontrollability.**

Mainen ZF, Kepecs A (2009). Neural representation of behavioral outcomes in the orbitofrontal cortex. *Curr Opin Neurobiol* **19**: 84–91.

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

Matsumoto M, Yoshioka M, Togashi H, Ikeda T, Saito H (1996). Functional regulation by dopamine receptors of serotonin release from the rat hippocampus: in vivo microdialysis study. *Naunyn Schmiedebergs Arch Pharmacol* **353**: 621–629.

Mazzoni P, Hristova A, Krakauer JW (2007). Why don't we move faster? Parkinson's disease, movement vigor, and implicit motivation. *J Neurosci* **27**: 7105–7116.

McClure SM, Daw ND, Montague PR (2003). A computational substrate for incentive salience. *Trends Neurosci* **26**: 423–428.

McNally RJ (1987). Preparedness and phobias: a review. *Psychol Bull* **101**: 283–303.

McNaughton N, Corr PJ (2004). A two-dimensional neuropsychology of defense: fear/anxiety and defensive distance. *Neurosci Biobehav Rev* **28**: 285–305. **Discusses the sophistication of Pavlovian defensive reactions (Bolles, 1970), in terms of their adaptivity to the nature, extent, and direction of putative threats.**

Meeusen R, Watson P, Hasegawa H, Roelands B, Piacentini MF (2006). Central fatigue: the serotonin hypothesis and beyond. *Sports Med* **36**: 881–909.

Meeusen R, Watson P, Hasegawa H, Roelands B, Piacentini MF (2007). Brain neurotransmitters in fatigue and overtraining. *Appl Physiol Nutr Metab* **32**: 857–864.

Meltzer HY, Huang M (2008). In vivo actions of atypical antipsychotic drug on serotonergic and dopaminergic systems. *Prog Brain Res* **172**: 177–197.

Millan MJ, Dekeyne A, Gobert A (1998). Serotonin (5-HT)2C receptors tonically inhibit dopamine (DA) and noradrenaline (NA), but not 5-HT, release in the frontal cortex in vivo. *Neuropharmacology* **37**: 953–955.

Millan MJ, Lejeune F, Gobert A (2000a). Reciprocal autoreceptor and heteroreceptor control of serotonergic, dopaminergic and noradrenergic transmission in the frontal cortex: relevance to the actions of antidepressant agents. *J Psychopharmacol* **14**: 114–138.

Millan MJ, Lejeune F, Gobert A (2000b). Reciprocal autoreceptor and heteroreceptor control of serotonergic, dopaminergic and noradrenergic transmission in the frontal cortex: relevance to the actions of antidepressant agents. *J Psychopharmacol* **14**: 114–138.

Millan MJ, Veiga S, Girardon S, Brocco M (2003). Blockade of serotonin 5-HT1B and 5-HT2A receptors suppresses the induction of locomotor activity by 5-HT reuptake inhibitors, citalopram and fluvoxamine, in NMRI mice exposed to a novel environment: a comparison to other 5-HT receptor subtypes. *Psychopharmacology (Berl)* **168**: 397–409.

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

Misane I, Johansson C, Ogren SO (1998). Analysis of the 5-HT1A receptor involvement in passive avoidance in the rat. *Br J Pharmacol* **125**: 499–509.

Mobini S, Chiang TJ, Al-Ruwaitea AS, Ho MY, Bradshaw CM, Szabadi E (2000a). Effect of central 5-hydroxytryptamine depletion on inter-temporal choice: a quantitative analysis. *Psychopharmacology (Berl)* **149**: 313–318.

Mobini S, Chiang TJ, Ho MY, Bradshaw CM, Szabadi E (2000b). Effects of central 5-hydroxytryptamine depletion on sensitivity to delayed and probabilistic reinforcement. *Psychopharmacology (Berl)* **152**: 390–397.

Mogenson GJ, Jones DL, Yim CY (1980). From motivation to action: functional interface between the limbic system and the motor system. *Prog Neurobiol* **14**: 69–97.

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

Montague PR, Hyman SE, Cohen JD (2004). Computational roles for dopamine in behavioural control. *Nature* **431**: 760–767.

Morgan MA, LeDoux JE (1995). Differential contribution of dorsal and ventral medial prefrontal cortex to the acquisition and extinction of conditioned fear in rats. *Behav Neurosci* **109**: 681–688.

Morgan MA, LeDoux JE (1999). Contribution of ventrolateral prefrontal cortex to the acquisition and extinction of conditioned fear in rats. *Neurobiol Learn Mem* **72**: 244–251.

Morgan MA, Romanski LM, LeDoux JE (1993). Extinction of emotional learning: contribution of medial prefrontal cortex. *Neurosci Lett* **163**: 109–113.

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

Morris R (1975). Preconditioning of reinforcing properties to an exteroceptive feedback stimulus. *Learn Motiv* **6**: 289–298.

Moutoussis M, Bentall RP, Williams J, Dayan P (2008). A temporal difference account of avoidance learning. *Network* **19**: 137–160.

Mowrer O (1947). On the dual nature of learning: a reinterpretation of conditioning and problem-solving. *Harv Educ Rev* **17**: 102–150.

Mowrer O (1956). Two-factor learning theory reconsidered, with special reference to secondary reinforcement and the concept of habit. *Psychol Rev* **63**: 114–128.

Mulder AB, Hodenpijl MG, da Silva FHL (1998). Electrophysiology of the hippocampal and amygdaloid projections to the nucleus accumbens of the rat: convergence, segregation, and interaction of inputs. *J Neurosci* **18**: 5095–5102.

Murschall A, Hauber W (2006). Inactivation of the ventral tegmental area abolished the general excitatory influence of Pavlovian cues on instrumental performance. *Learn Mem* **13**: 123–126.

Nakamura K, Matsumoto M, Hikosaka O (2008). Reward-dependent modulation of neuronal activity in the primate dorsal raphe nucleus. *J Neurosci* **28**: 5331–5343.

Navailles S, Moison D, Cunningham KA, Spampinato U (2008). Differential regulation of the mesoaccumbens dopamine circuit by serotonin2C receptors in the ventral tegmental area and the nucleus accumbens: an in vivo microdialysis study with cocaine. *Neuropsychopharmacology* **33**: 237–246.

Navailles S, Moison D, Ryczko D, Spampinato U (2006). Region-dependent regulation of mesoaccumbens dopamine neurons in vivo by the constitutive activity of central serotonin2C receptors. *J Neurochem* **99**: 1311–1319.

Nedergaard S, Bolam JP, Greenfield SA (1988). Facilitation of a dendritic calcium conductance by 5-hydroxytryptamine in the substantia nigra. *Nature* **333**: 174–177.

Neumaier JF, Vincow ES, Arvanitogiannis A, Wise RA, Carlezon WA (2002). Elevated expression of 5-HT1B receptors in nucleus accumbens efferents sensitizes animals to cocaine. *J Neurosci* **22**: 10856–10863.

Newsholme E, Acworth I, Blomstrand E (1987). Amino acids, brain neurotransmitters and a functional link between muscle and brain that is important in sustained exercise. In: Benzi G (ed). *Advances in Myochemistry.* John Libbey Eurotext: London. pp 127–133.

Niv Y (2009). Reinforcement learning in the brain. *J Math Psychol* **53**: 139–154.

Niv Y, Daw ND, Joel D, Dayan P (2007). Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl)* **191**: 507–520.

Niv Y, Joel D, Dayan P (2006). A normative perspective on motivation. *Trends Cogn Sci* **10**: 375–381.

Nurse B, Russell VA, Taljaard JJ (1984). Alpha 2 and beta-adrenoceptor agonists modulate (3H)dopamine release from rat nucleus accumbens slices: implications for research into depression. *Neurochem Res* **9**: 1231–1238.

Nutt D, Demyttenaere K, Janka Z, Aarre T, Bourin M, Canonico PL et al (2007). The other face of depression, reduced positive affect: the role of catecholamines in causation and cure. *J Psychopharmacol* **21**: 461–471.

O'Doherty JP (2007). Lights, camembert, action! The role of human orbitofrontal cortex in encoding stimuli, rewards, and choices. *Ann NY Acad Sci* **1121**: 254–272.

O'Donnell P, Grace AA (1995). Synaptic interactions among excitatory afferents to nucleus accumbens neurons: hippocampal gating of prefrontal cortical input. *J Neurosci* **15**(5 Part 1): 3622–3639.

O'Hearn E, Molliver ME (1984). Organization of raphe-cortical projections in rat: a quantitative retrograde study. *Brain Res Bull* **13**: 709–726.

Padoa-Schioppa C, Assad JA (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature* **441**: 223–226.

Palmiter RD (2008). Dopamine signaling in the dorsal striatum is essential for motivated behaviors: lessons from dopamine-deficient mice. *Ann NY Acad Sci* **1129**: 35–46.

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

Panksepp J (1998). *Affective Neuroscience.* OUP: New York, NY.

Pardo JV, Pardo PJ, Janer KW, Raichle ME (1990). The anterior cingulate cortex mediates processing selection in the Stroop attentional conflict paradigm. *Proc Natl Acad Sci USA* **87**: 256–259.

Parent A (1981). Comparative anatomy of the serotoninergic systems. *J Physiol (Paris)* **77**: 147–156.

Parkinson JA, Robbins TW, Everitt BJ (2000a). Dissociable roles of the central and basolateral amygdala in appetitive emotional learning. *Eur J Neurosci* **12**: 405–413.

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

Parsons LH, Justice JB (1993). Perfusate serotonin increases extracellular dopamine in the nucleus accumbens as measured by in vivo microdialysis. *Brain Res* **606**: 195–199.

Pascucci T, Ventura R, Latagliata EC, Cabib S, Puglisi-Allegra S (2007). The medial prefrontal cortex determines the accumbens dopamine response to stress through the opposing influences of norepinephrine and dopamine. *Cereb Cortex* **17**: 2796–2804.

Pehek EA, Nocjar C, Roth BL, Byrd TA, Mabrouk OS (2006). Evidence for the preferential involvement of 5-HT2A serotonin receptors in stress- and drug-induced dopamine release in the rat medial prefrontal cortex. *Neuropsychopharmacology* **31**: 265–277.

Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* **442**: 1042–1045.

Peyron C, Petit JM, Rampon C, Jouvet M, Luppi PH (1998). Forebrain afferents to the rat dorsal raphe nucleus demonstrated by retrograde and anterograde tracing methods. *Neuroscience* **82**: 443–468.

Pezze MA, Feldon J (2004). Mesolimbic dopaminergic pathways in fear conditioning. *Prog Neurobiol* **74**: 301–320.

Pompilio L, Kacelnik A, Behmer ST (2006). State-dependent learned valuation drives choice in an invertebrate. *Science* **311**: 1613–1615.

Porras G, De Deurwaerdère P, Moison D, Spampinato U (2003). Conditional involvement of striatal serotonin3 receptors in the control of in vivo dopamine outflow in the rat striatum. *Eur J Neurosci* **17**: 771–781.

Porras G, Di Matteo V, Fracasso C, Lucas G, De Deurwaerdère P, Caccia S et al (2002). 5-HT2A and 5-HT2C/2B receptor subtypes modulate dopamine release induced in vivo by amphetamine and morphine in both the rat nucleus accumbens and striatum. *Neuropsychopharmacology* **26**: 311–324.

Posner J, Russell J, Peterson B (2005). The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. *Dev Psychopathol* **17**: 715–734.

Poulos CX, Parker J, Le A (1996). Dexfenfluramine and 8-OH-DPAT modulate impulsivity in a delay-of-reward paradigm: implications for a correspondence with alcohol consumption. *Behav Pharmacol* **7**: 395–399.

Powell EW, Leman RB (1976). Connections of the nucleus accumbens. *Brain Res* **105**: 389–403.

Pozzi L, Acconcia S, Ceglia I, Invernizzi RW, Samanin R (2002). Stimulation of 5-hydroxytryptamine (5-HT(2C)) receptors in the ventrotegmental area inhibits stress-induced but not basal dopamine release in the rat prefrontal cortex. *J Neurochem* **82**: 93–100.

Przegalinski E, Siwanowicz J, Nowak E, Papla I, Filip M (2001). Role of 5-HT(1B) receptors in the sensitization to amphetamine in mice. *Eur J Pharmacol* **422**: 91–99.

Puterman ML (2005). *Markov Decision Processes: Discrete Stochastic Dynamic Programming (Wiley Series in Probability and Statistics)*. Wiley-Interscience: Hoboken, NJ.

Ranade SP, Mainen ZF (2009). Transient firing of dorsal raphe neurons encodes diverse and specific sensory, motor, and reward events. *J Neurophysiol* **102**: 3026–3037.

Reading PJ, Dunnett SB, Robbins TW (1991). Dissociable roles of the ventral, medial and lateral striatum on the acquisition and performance of a complex visual stimulus-response habit. *Behav Brain Res* **45**: 147–161.

Reynolds SM, Berridge KC (2001). Fear and feeding in the nucleus accumbens shell: rostrocaudal segregation of GABA-elicited defensive behavior versus eating behavior. *J Neurosci* **21**: 3261–3270.

Reynolds SM, Berridge KC (2002). Positive and negative motivation in nucleus accumbens shell: bivalent rostrocaudal gradients for GABA-elicited eating, taste 'liking'/'disliking' reactions, place preference/avoidance, and fear. *J Neurosci* **22**: 7308–7320.

Reynolds SM, Berridge KC (2008). Emotional environments retune the valence of appetitive versus fearful functions in nucleus accumbens. *Nat Neurosci* **11**: 423–425. **Latest in a series of papers (including (Reynolds and Berridge, 2001, Reynolds and Berridge, 2002)) showing the orderly arrangement across the nucleus accumbens of chemically stimulable approach and avoidance responses; this paper showing that the boundary between appetitive and aversive responses is sensitive to the stressfulness of the environment.**

Robbins T, Everitt B (1992). Functions of dopamine in the dorsal and ventral striatum. *SeminNeurosci* **4**: 119–127.

Robbins TW (2005). Chemistry of the mind: neurochemical modulation of prefrontal cortical function. *J Comp Neurol* **493**: 140–146.

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

Robbins TW, Giardini V, Jones GH, Reading P, Sahakian BJ (1990). Effects of dopamine depletion from the caudate-putamen and nucleus accumbens septi on the acquisition and performance of a conditional discrimination task. *Behav Brain Res* **38**: 243–261.

Robbins TW, Roberts AC (2007). Differential regulation of fronto-executive function by the monoamines and acetylcholine. *Cereb Cortex* **17**(Suppl 1): i151–i160.

Robinson ESJ, Dalley JW, Theobald DEH, Glennon JC, Pezze MA, Murphy ER et al (2008). Opposing roles for 5-HT2A and 5-HT2C receptors in the nucleus accumbens on inhibitory response control in the 5-choice serial reaction time task. *Neuropsychopharmacology* **33**: 2398–2406.

Robinson S, Rainwater AJ, Hnasko TS, Palmiter RD (2007). Viral restoration of dopamine signaling to the dorsal striatum restores instrumental conditioning to dopamine-deficient mice. *Psychopharmacology (Berl)* **191**: 567–578.

Robinson S, Sandstrom SM, Denenberg VH, Palmiter RD (2005). Distinguishing whether dopamine regulates liking, wanting, and/or learning about rewards. *Behav Neurosci* **119**: 5–15.

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

Rolls ET, Grabenhorst F (2008). The orbitofrontal cortex and beyond: from affect to decision-making. *Prog Neurobiol* **86**: 216–244.

Rosenkranz JA, Grace AA (1999). Modulation of basolateral amygdala neuronal firing and afferent drive by dopamine receptor activation in vivo. *J Neurosci* **19**: 11027–11039.

Rosenkranz JA, Grace AA (2001). Dopamine attenuates prefrontal cortical suppression of sensory inputs to the basolateral amygdala of rats. *J Neurosci* **21**: 4090–4103.

Rosenkranz JA, Grace AA (2002). Cellular mechanisms of infralimbic and prelimbic prefrontal cortical inhibition and dopaminergic modulation of basolateral amygdala neurons in vivo. *J Neurosci* **22**: 324–337.

Salamone JD, Correa M (2002). Motivational views of reinforcement: implications for understanding the behavioral functions of nucleus accumbens dopamine. *Behav Brain Res* **137**: 3–25.

Sasaki-Adams DM, Kelley AE (2001). Serotonin-dopamine interactions in the control of conditioned reinforcement and motor behavior. *Neuropsychopharmacology* **25**: 440–452.

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

Schmidt CJ, Sorensen SM, Kehne JH, Carr AA, Palfreyman MG (1995). The role of 5-HT2A receptors in antipsychotic activity. *Life Sci* **56**: 2209–2222.

Schneirla T (1959). An evolutionary and developmental theory of biphasic processes underlying approach and withdrawal. In: Jones M (ed). *Nebraska Symposium on Motivation*. University of Nebraska Press: Lincoln, NE. pp 1–42.

Schoenbaum G, Roesch MR, Stalnaker TA, Takahashi YK (2009). A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. *Nat Rev Neurosci* **10**: 885–892.

Schoenbaum G, Setlow B, Saddoris MP, Gallagher M (2003). Encoding predicted outcome and acquired value in orbitofrontal cortex during cue sampling depends upon input from basolateral amygdala. *Neuron* **39**: 855–867.

Schultz W (2002). Getting formal with dopamine and reward. *Neuron* **36**: 241–263.

Schultz W (2007). Behavioral dopamine signals. *Trends Neurosci* **30**: 203–210.

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

Schultz W, Dickinson A (2000). Neuronal coding of prediction errors. *Annu Rev Neurosci* **23**: 473–500.

Schweighofer N, Bertin M, Shishida K, Okamoto Y, Tanaka SC, Yamawaki S et al (2008). Low-serotonin levels increase delayed reward discounting in humans. *J Neurosci* **28**: 4528–4532.

Schweighofer N, Tanaka SC, Doya K (2007). Serotonin and the evaluation of future rewards: theory, experiments, and possible neural mechanisms. *Ann NY Acad Sci* **1104**: 289–300.

Schweimer J, Brierley D, Ungless M (2008). Phasic nociceptive responses in dorsal raphe serotonin neurons. *Fundam Clin Pharmacol* **22**: 119.

Servan-Schreiber D, Printz H, Cohen JD (1990). A network model of catecholamine effects: gain, signal-to-noise ratio, and behavior. *Science* **249**: 892–895.

Sesack SR, Grace AA (2010). Cortico-basal ganglia reward network: microcircuitry. *Neuropsychopharmacology* **35**: 27–47.

Sheffield F (1965). Relation between classical conditioning and instrumental learning. In: Prokasy W (ed). *Classical Conditioning*. Appelton-Century-Crofts: New York, NY. pp 302–322.

Shidara M, Richmond BJ (2004). Differential encoding of information about progress through multi-trial reward schedules by three groups of ventral striatal neurons. *Neurosci Res* **49**: 307–314.

Simansky KJ (1996). Serotonergic control of the organization of feeding and satiety. *Behav Brain Res* **73**: 37–42.

Smith A, Li M, Becker S, Kapur S (2006). Dopamine, prediction error and associative learning: a model-based account. *Network* **17**: 61–84.

Smith J, Dickinson A (1998). The dopamine antagonist, pimozide, abolishes Pavlovian-instrumental transfer. *J Psychopharmacol* **12**: A6.

Smith JM, Alloy LB (2009). A roadmap to rumination: a review of the definition, assessment, and conceptualization of this multifaceted construct. *Clin Psychol Rev* **29**: 116–128.

Solomon RL, Corbit JD (1974). An opponent-process theory of motivation. I. Temporal dynamics of affect. *Psychol Rev* **81**: 119–145.

Sorg BA, Kalivas PW (1991). Effects of cocaine and footshock stress on extracellular dopamine levels in the ventral striatum. *Brain Res* **559**: 29–36.

Soubrié P (1986). Reconciling the role of central serotonin neurons in human and animal behaviour. *Behav Brain Sci* **9**: 319–364. **Notably influential review providing the foundations for the behavioral inhibition account of 5-HT**.

Spoont MR (1992). Modulatory role of serotonin in neural information processing: implications for human psychopathology. *Psychol Bull* **112**: 330–350.

Staddon JE (1965). Some properties of spaced responding in pigeons. *J Exp Anal Behav* **8**: 19–27.

Suri RE, Schultz W (1999). A neural network model with dopamine-like reinforcement signal that learns a spatial delayed response task. *Neuroscience* **91**: 871–890.

Surmeier DJ, Ding J, Day M, Wang Z, Shen W (2007). D1 and D2 dopamine-receptor modulation of striatal glutamatergic signaling in striatal medium spiny neurons. *Trends Neurosci* **30**: 228–235.

Sutton R (1988). Learning to predict by the methods of temporal differences. *Mach Learn* **3**: 9–44.

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

Sutton RS, Barto AG (1998). *Reinforcement Learning: An Introduction (Adaptive Computation and Machine Learning*. The MIT Press: Cambridge, MA. **Key treatise on reinforcement learning, providing a tutorial introduction to the basic methods used throughout the affective decision-making literature to model behavior and neural responses**.

Swanson LW (1982). The projections of the ventral tegmental area and adjacent regions: a combined fluorescent retrograde tracer and immunofluorescence study in the rat. *Brain Res Bull* **9**: 321–353.

Szczypka MS, Kwok K, Brot MD, Marck BT, Matsumoto AM, Donahue BA *et al* (2001). Dopamine production in the caudate putamen restores feeding in dopamine-deficient mice. *Neuron* **30**: 819–828.

Takase LF, Nogueira MI, Baratta M, Bland ST, Watkins LR, Maier SF *et al* (2004). Inescapable shock activates serotonergic neurons in all raphe nuclei of rat. *Behav Brain Res* **153**: 233–239.

Takase LF, Nogueira MI, Bland ST, Baratta M, Watkins LR, Maier SF *et al* (2005). Effect of number of tailshocks on learned helplessness and activation of serotonergic and noradrenergic neurons in the rat. *Behav Brain Res* **162**: 299–306.

Talmi D, Seymour B, Dayan P, Dolan RJ (2008). Human Pavlovian-instrumental transfer. *J Neurosci* **28**: 360–368.

Tanaka SC, Doya K, Okada G, Ueda K, Okamoto Y, Yamawaki S (2004). Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nat Neurosci* **7**: 887–893.

Tanaka SC, Schweighofer N, Asahi S, Shishida K, Okamoto Y, Yamawaki S *et al* (2007). Serotonin differentially regulates short- and long-term prediction of rewards in the ventral and dorsal striatum. *PLoS One* **2**: e1333.

Tassin J-P (2008). Uncoupling between noradrenergic and serotonergic neurons as a molecular basis of stable changes in behavior induced by repeated drugs of abuse. *Biochem Pharmacol* **75**: 85–97.

Thorndike E (1911). *Animal Intelligence*. MacMillan: New York, NY.

Thorré K, Sarre S, Smolders I, Ebinger G, Michotte Y (1998). Dopaminergic regulation of serotonin release in the substantia nigra of the freely moving rat using microdialysis. *Brain Res* **796**: 107–116.

Tops M, Russo S, Boksem MAS, Tucker DM (2009). Serotonin: modulator of a drive to withdraw. *Brain Cogn* **71**: 427–436. **Recent review of the data suggesting that 5-HT is involved as much in sensory disengagement as behavioral inhibition, an idea we tentatively related to opponency with norepinephrine**.

Tricomi E, Balleine BW, O'Doherty JP (2009). A specific role for posterior dorsolateral striatum in human habit learning. *Eur J Neurosci* **29**: 2225–2232.

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

Tzschentke TM (2001). Pharmacology and behavioral pharmacology of the mesocortical dopamine system. *Prog Neurobiol* **63**: 241–320.

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

Valentin VV, Dickinson A, O'Doherty JP (2007). Determining the neural substrates of goal-directed learning in the human brain. *J Neurosci* **27**: 4019–4026.

Van Bockstaele EJ, Cestari DM, Pickel VM (1994). Synaptic structure and connectivity of serotonin terminals in the ventral tegmental area: potential sites for modulation of mesolimbic dopamine neurons. *Brain Res* **647**: 307–322.

Villegier A-S, Drouin C, Bizot J-C, Marien M, Glowinski J, Colpaërt F *et al* (2003). Stimulation of postsynaptic alpha1b- and alpha2-adrenergic receptors amplifies dopamine-mediated locomotor activity in both rats and mice. *Synapse* **50**: 277–284.

Wallis JD (2007). Orbitofrontal cortex and its contribution to decision-making. *Annu Rev Neurosci* **30**: 31–56.

Watson D, Clark LA (1984). Negative affectivity: the disposition to experience aversive emotional states. *Psychol Bull* **96**: 465–490.

Weiner I (1990). Neural substrates of latent inhibition: the switching model. *Psychol Bull* **108**: 442–461.

White NM, McDonald RJ (2002). Multiple parallel memory systems in the brain of the rat. *Neurobiol Learn Mem* **77**: 125–184.

Wickens JR, Horvitz JC, Costa RM, Killcross S (2007). Dopaminergic mechanisms in actions and habits. *J Neurosci* **27**: 8181–8183.

Wilkinson LS (1997). The nature of interactions involving prefrontal and striatal dopamine systems. *J Psychopharmacol* **11**: 143–150.

Williams DR, Williams H (1969). Auto-maintenance in the pigeon: sustained pecking despite contingent non-reinforcement. *J Exp Anal Behav* **12**: 511–520.

Williams GV, Goldman-Rakic PS (1995). Modulation of memory fields by dopamine D1 receptors in prefrontal cortex. *Nature* **376**: 572–575.

Williams J, Dayan P (2005). Dopamine, learning, and impulsivity: a biological account of attention-deficit/hyperactivity disorder. *J Child Adolesc Psychopharmacol* **15**: 160–179;discussion 157–159.

Winstanley CA, Theobald DEH, Dalley JW, Cardinal RN, Robbins TW (2006). Double dissociation between serotonergic and dopaminergic modulation of medial prefrontal and orbitofrontal cortex during a test of impulsive choice. *Cereb Cortex* **16**: 106–114.

Winstanley CA, Theobald DEH, Dalley JW, Glennon JC, Robbins TW (2004). 5-HT2A and 5-HT2C receptor antagonists have opposing effects on a measure of impulsivity: interactions with global 5-HT depletion. *Psychopharmacology (Berl)* **176**: 376–385.

Wise RA (2004). Dopamine, learning and motivation. *Nat Rev Neurosci* **5**: 483–494.

Wise RA (2008). Dopamine and reward: the anhedonia hypothesis 30 years on. *Neurotox Res* **14**: 169–183.

Wise RA, Bozarth MA (1987). A psychomotor stimulant theory of addiction. *Psychol Rev* **94**: 469–492.

Wogar MA, Bradshaw CM, Szabadi E (1993). Effect of lesions of the ascending 5-hydroxytryptaminergic pathways on choice between delayed reinforcers. *Psychopharmacology (Berl)* **111**: 239–243.

Wyvell CL, Berridge KC (2001). Incentive sensitization by previous amphetamine exposure: increased cue-triggered 'wanting' for sucrose reward. *J Neurosci* **21**: 7831–7840.

Yan QS, Yan SE (2001). Activation of 5-HT(1B/1D) receptors in the mesolimbic dopamine system increases dopamine release from the nucleus accumbens: a microdialysis study. *Eur J Pharmacol* **418**: 55–64.

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

Yin HH, Ostlund SB, Knowlton BJ, Balleine BW (2005). The role of the dorsomedial striatum in instrumental conditioning. *Eur J Neurosci* **22**: 513–523.

Yoshioka M, Matsumoto M, Togashi H, Saito H (1995). Effects of conditioned fear stress on 5-HT release in the rat prefrontal cortex. *Pharmacol Biochem Behav* **51**: 515–519.

Zahm DS, Heimer L (1990). Two transpallidal pathways originating in the rat nucleus accumbens. *J Comp Neurol* **302**: 437–446.