

Serotonin and Dopamine: Unifying Affective, Activational, and Decision Functions

Roshan Cools^{*1}, Kae Nakamura^{2,3} and Nathaniel D Daw⁴

¹Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Centre for Cognitive Neuroimaging, Nijmegen, The Netherlands; ²Department of Physiology, School of Medicine, Kansai Medical University, Moriguchi City, Japan; ³PRESTO, Honcho Kawaguchi, Saitama, Japan; ⁴Center for Neural Science & Department of Psychology, New York University, New York, NY, USA

Serotonin, like dopamine (DA), has long been implicated in adaptive behavior, including decision making and reinforcement learning. However, although the two neuromodulators are tightly related and have a similar degree of functional importance, compared with DA, we have a much less specific understanding about the mechanisms by which serotonin affects behavior. Here, we draw on recent work on computational models of dopaminergic function to suggest a framework by which many of the seemingly diverse functions associated with both DA and serotonin—comprising both affective and activational ones, as well as a number of other functions not overtly related to either—can be seen as consequences of a single root mechanism. *Neuropsychopharmacology Reviews* (2011) **36**, 98–113; doi:10.1038/npp.2010.121; published online 25 August 2010

Keywords: aversion; reward; inhibition; impulsivity; activation; punishment

INTRODUCTION

The ascending monoamine neuromodulatory systems are implicated in healthy and disordered functions so wide ranging and so apparently heterogeneous that characterizing their function more crisply is an important scientific puzzle. In the case of dopamine (DA)—which is involved in cognition, motivation, and movement—notable progress has been made in the last decade using an interdisciplinary and interspecies approach. In particular, computational models of reinforcement learning (RL: trial-and-error learning to obtain rewards) have been used as a framework formally to interpret and connect observations from neurophysiological, brain imaging, and behavioral/pharmacological studies in humans and animals.

In contrast, although the neuromodulator serotonin (5-HT) has functional and clinical importance at least equal to that of DA (eg, it is implicated in impulsivity, depression, and pain), there is no similarly formal and well-developed framework for understanding any of its roles. Here, we take early steps toward such a theoretical framework by reviewing aspects of function that have been prominently associated with 5-HT, namely, aversive processing and

behavioral inhibition, and leveraging the example of DA to suggest how the data supporting these ideas might be interpreted, together with other functions, as manifestations of a common, underlying computational mechanism. In particular, we consider the implications of a recent computational theory of DA (Niv *et al*, 2007) for offering a common explanation for a number of seemingly distinct functional associations of both DA and 5-HT. We discuss the theory informally (omitting equations) and use it as a framework to discuss studies using psychopharmacological manipulations of 5-HT in humans and experimental rodents, as well as single-neuron recording studies in non-human primates. In the first half of the review, we discuss how Niv *et al*'s concept of an opportunity cost of time offers a common explanation for both affective (reward and punishment) and activational (behavioral vigor and withholding) aspects of the neuromodulators' functions. After this, we develop this framework to discuss how a number of additional, seemingly disparate, aspects of decision making that have been associated with these systems, such as time discounting and risk sensitivity, can also be seen as consequences of the same mechanism. Throughout, we stress many caveats, interpretation difficulties, and experimental concerns; our goal here is to articulate a set of important behaviors, computations, and quantities that might guide more definitive experiments. In addition, similar to Boureau and Dayan (2010; this issue) (see also Dayan and Huys, 2008 and Daw, Kakade and Dayan, 2002),

*Correspondence: Dr R Cools, Centre for Cognitive Neuroimaging, Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Kapittelweg 29, Nijmegen 6500HB, The Netherlands, Tel: +31 243 610 656, Fax: +31 243 610 989, E-mail: roshan.cools@gmail.com
Received 12 April 2010; revised 16 July 2010; accepted 16 July 2010

our overall strategy is to push outward from our relatively secure understanding of DA, through what is known about the similarity and differences in DA and 5-HT functions and about how the two neuromodulators interact, to extrapolate a tentative extended understanding encompassing DA and 5-HT collectively in a common framework. Boureau and Dayan take a complementary approach, offering, in particular, a more detailed discussion of the nature of interactions between DA and 5-HT, and between reward and punishment in the context of different components of conditioning.

DA, REINFORCEMENT, AND BEHAVIORAL ACTIVATION

The puzzles and controversies of DA have long centered around the question of how to understand its seemingly dual function in both reward and movement (Ungerstedt, 1971; Lyon and Robbins, 1975; Milner, 1977; Evenden and Robbins, 1984; Berridge and Robinson, 1998; Ikemoto and Panksepp, 1999; Schultz, 2007). On the one hand, DA is implicated in motivation and reinforcement, for instance, it is a focus of drugs of abuse and self-stimulation. On the other, it is a facilitator of vigorous action: consider the poverty of movement that accompanies dopaminergic degeneration in Parkinson's disease (PD) or the hyperactivity and stereotypy engendered by psychostimulant drugs that enhance DA, such as methamphetamine (Lyon and Robbins, 1975; Robbins and Sahakian, 1979). In principle, these two axes of behavior might be independent, but they appear instead to be closely coupled through the action of DA.

Thus, one early hypothesis (Mogenson *et al*, 1980) characterized the nucleus accumbens (a key dopaminergic target) as the 'limbic-motor gateway' in which motivational considerations gained access to the control of action. Echoing this idea, more recent RL theories link these aspects by claiming that DA is involved in learning which behaviors are associated with reward. Variants of the reward/action duality also underlie longstanding controversies about what psychological aspects of reward DA might subserve—for instance, hedonics, reinforcement, or motivational and activational (Ikemoto and Panksepp, 1999; Berridge, 2007; Robbins and Everitt, 2007)—and the question whether DA impacts behavior via learning versus performance (Gallistel *et al*, 1974; Berridge, 2007; Niv *et al*, 2007). We focus on this last question here.

Appropriately, given DA's dual nature, theories of its function have grown largely separately on two tracks, rooted in different experimental methodologies and theoretical approaches. The predominant view in computational and systems neuroscience holds that DA serves to promote RL, that is, trial-and-error instrumental learning, to choose rewarding actions (Houk *et al*, 1995; Montague *et al*, 1996; Schultz *et al*, 1997; Samejima *et al*, 2005; Morris *et al*, 2006). This idea is derived from electrophysiological recordings from neurons in the midbrain dopaminergic nuclei of

primates performing simple tasks for reward (Ljungberg *et al*, 1991; Hollerman and Schultz, 1998; Waelti *et al*, 2001), together with the insight that the phasic firing of these neurons quantitatively resembles a 'reward prediction error' signal used in computational algorithms for RL to improve action choice so as to obtain more rewards (Sutton and Barto, 1990; Montague *et al*, 1996; Sutton and Barto, 1998; Montague *et al*, 2004; Bayer and Glimcher, 2005; Frank, 2005). More recently, studies employing temporally precise methods in freely behaving animals, such as electrochemical voltammetric approaches, which enable the measurement of phasic DA release directly (Day *et al*, 2007; Roitman *et al*, 2008), as well as optogenetic approaches, which enable the transient activation of specific DA neurons (Tsai *et al*, 2009), have substantiated these ideas. Furthermore, functional neuroimaging has revealed that similar prediction error signals in humans (McClure *et al*, 2003; O'Doherty *et al*, 2003) might be modulated by DA (Pessiglione *et al*, 2006), whereas microelectrode recordings during deep brain stimulation surgery have demonstrated that such prediction error signals are also encoded by the human midbrain (Zaghloul *et al*, 2009) (see also D'Ardenne *et al*, 2008).

At the same time, more psychological approaches, largely grounded in causal manipulations (eg, drug or lesion) of dopaminergic function, tend to envision DA as being involved less in acquisition and more in the performance of motivated behavior. Indeed, the most pronounced effects of causal DA manipulations tend to be on performance rather than learning, with DA promoting behavioral vigor or activation more generally (Lyon and Robbins, 1975; Ikemoto and Panksepp, 1999; Berridge, 2007; Robbins and Everitt, 2007; Salamone *et al*, 2007). Two current interpretations characterize these effects as arising via dopaminergic mediation of incentive motivation (Berridge, 2007) or cost/benefit tradeoffs (Salamone *et al*, 2007). Other authors writing from a similar tradition have provided a more general activational account, with parallel roles for DA in the dorsal and ventral striatum (Robbins and Everitt, 1982, 1992; Robbins and Everitt, 2007), stressing both a performance-based energetic component to DA and reinforcement-related functions more akin to those posited in the computational RL models, for example, conditioned reinforcement and stamping-in of stimulus-response habits (Wise, 2004). Indeed, early experimental work by Gallistel *et al* (1974) argued for both reinforcing and activational effects of (putatively dopaminergic) brain stimulation reward, distinguished as progressive and immediate effects of contingent versus noncontingent self-stimulation.

MODELING THE DUAL FUNCTION OF DA

One attempt to reconcile these two streams of thought (Niv *et al*, 2007) extended RL accounts, which had traditionally focused on learning which action is most rewarding, into an additional formal analysis of how

vigorously these actions should be performed. The model casts the control of vigor as a problem of trading off the costs (energetic) and benefits (faster reward gathering) of behaving more vigorously, as for a rat pressing a lever for food at a more or less rapid rate. A key outcome of this analysis is that, when all other aspects of a decision are equal, sloth is more costly, and vigor more worthwhile, when rewards are more frequently available. In this case, more reward is foregone, on average, by working more slowly: the opportunity cost of time is higher (Figure 1b). This cost of time can be defined as the amount of reward (or rewards minus punishments) one should expect to receive on an average during some period, that is, the long-term rate at which rewards are received (Figure 1a). In theory this average reward rate is a key variable in determining the rate of responding.

The importance of this hypothesis is that it explicitly relates reward and action vigor, the two axes of DA's function; in particular, it suggests and motivates a mechanism by which a signal carrying average reward information—the opportunity cost—would, causally, influence behavioral vigor. The authors suggest that the hypothesized average reward signal, which (as a prediction about long-term events) should change slowly, would most plausibly be associated with dopaminergic activity at a tonic timescale, rather than a phasic one (Figure 1a). The performance-related effects of dopaminergic manipulations are also, in many cases, seen with treatments such as receptor agonists that are tonic in nature. There are a number of mechanisms by which such tonic DA manipulations may affect behavioral vigor, for instance, by modulating the balance between direct and indirect pathways through the basal

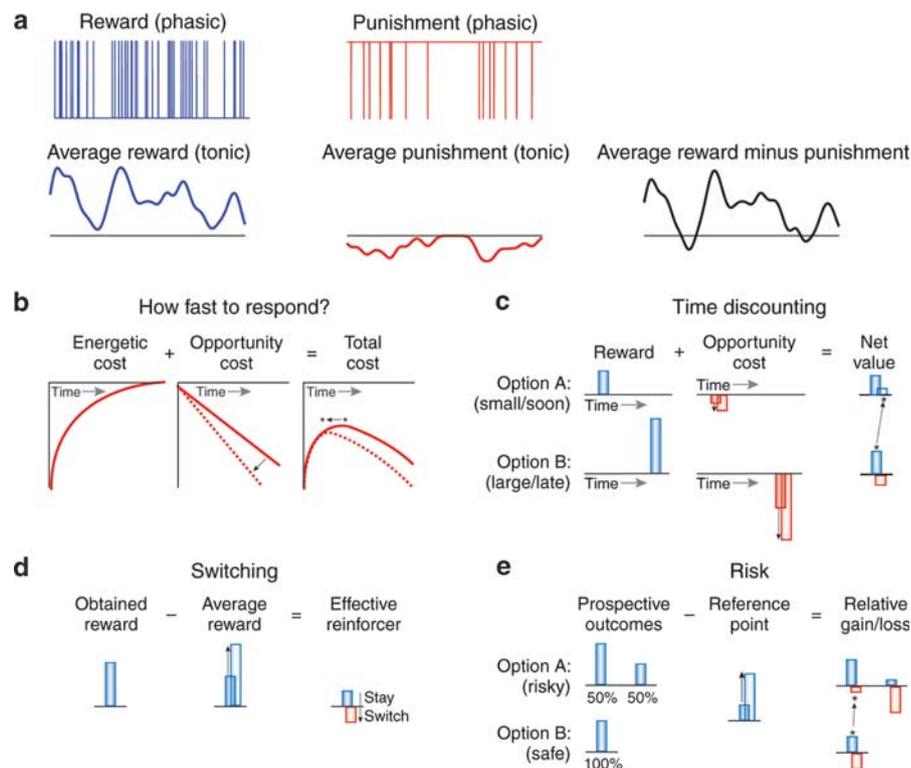


Figure 1. Graphic depiction of the core computational concepts outlined in this article, and their consequences for the functional domains of response vigor, time discounting, switching, and risk sensitivity. (a) Phasic time series of rewards and punishments, together with tonic signals consisting of the slow running average of each. We associate these average reward and punishment signals with tonic dopamine and serotonin, respectively. The difference between them (the overall average outcome), in black, is expected to control (b–e) a number of aspects of decision making. These subfigures illustrate how different decision-related calculations are impacted when the average outcome increases from less rewarding (solid lines and bars) to more rewarding (dashed lines and hollow, offset bars). Black arrows indicate the directions of change, and asterisks indicate preferred options. Rewards are illustrated in blue and costs or punishments in red. (b) The choice of how quickly to perform an instrumental behavior can be guided by trading off the ‘energetic cost’ of performing it more quickly against the ‘opportunity cost’ of the time spent. When the average outcome improves, the opportunity cost grows more quickly with time spent, and the point that minimizes the total cost (the optimal choice of response speed, asterisk) shifts to the left, favoring quicker responding. (c) The choice between a small reward soon (top) or a large reward later (bottom) can be guided by balancing the rewards against the opportunity costs of the delays. When the average outcome improves, the opportunity cost of the longer delay weighs more heavily, shifting the preferred choice from patient to impatient. (d) Learning about the value of an action can be guided by comparing the reward obtained with the average outcome expected. When the average outcome improves, the comparison term can drive the net reinforcement negative, and instead of reinforcing the action (favoring staying) it will be punished, favoring switching. (e) Preference over prospective options involving risk may depend on whether the outcomes are net gains or losses, when measured relative to a reference point. Humans and other animals tend to be risk averse when considering prospects involving (relative) gains and risk prone when considering relative losses. Here, if the average outcome is taken as the reference point, the choice between a sure win (top, safe) and a 50/50 small or large win (top, risky) and a sure win (bottom, safe), shifts from the gain domain to the loss domain when the average outcome improves, leading to a shift in preference from the safe option to the risky one.

ganglia (Mink, 1996), and/or information flow between distinct ventral and dorsal parts of the striatum via spiraling nigro-striatal connections (Nauta, 1979; Nauta, 1982; Haber *et al*, 2000); the suggestion of Niv *et al* (2007) was to interpret these effects teleologically in terms of the action of a hypothetical tonic average reward signal.

Although the causal effect of tonic DA manipulations is consistent with the effects expected of an average reward signal, there is little evidence as to whether tonic extracellular DA concentrations are sensitive to this variable. One intriguingly simple idea is that, mathematically, the same phasic prediction error signal that RL theories hypothesize is carried by phasic DA responses, also measures the average reward if it is averaged slowly over time. This is simply because when rewards occur more frequently, so equally do reward prediction errors. Temporal averaging of the phasic DA response might, for instance, be realized by synaptic overflow from phasic events followed by slower reuptake. Overflow is indeed measured as extracellular transients in dopaminergic concentrations in many cyclic voltammetry experiments (Garris *et al*, 1997; Phillips *et al*, 2003; Sombers *et al*, 2009). However, regarding filtering this signal by slow reuptake, the large transients from DA bursting are relatively rare and are cleared quickly (Cragg and Rice, 2004); thus, it may be that tonic DA is more influenced by other variables, for example, background levels of dopaminergic spiking or the number of active versus silent DA neurons (Floresco *et al*, 2003; Arbuthnott and Wickens, 2007). This is consistent with the concept of tonic DA as an at least partly independently regulated channel from phasic DA (Grace, 1991), and, in terms of the average reward hypothesis, with a more complex mechanism for computing an average reward signal, drawing on additional sources of information other than the phasic signal (Niv *et al*, 2007).

In summary, the Niv *et al* model argues that the two seemingly separate aspects of dopaminergic action are necessarily and not accidentally related.

SEROTONIN, AVERSIVE PROCESSING, AND BEHAVIORAL INHIBITION

Similar to DA, 5-HT has both affective and activational associations (among many others), although these are less well established empirically, and particular researchers (Soubrié, 1986; Deakin and Graeff, 1991; Deakin, 1998) have argued that one or the other concept may suffice to explain the data. Specifically, some classic accounts of 5-HT propose that the neuromodulator is involved in either of two functions analogous but opposite to those of DA: aversive processing (Deakin, 1983; Deakin and Graeff, 1991) (but see Kranz *et al*, 2010) and behavioral inhibition (Soubrié, 1986). The steps toward reconciliation of the two seemingly disparate functions of DA, discussed above, may point the way toward a similar reconciliation of the analogous aspects of 5-HT function.

Both aversive processing and behavioral inhibition do figure prominently in the data on serotonergic function, although often appearing in tandem rather than separately (for recent reviews see Kranz *et al*, 2010; Cools *et al*, 2008b; Dayan and Huys, 2008; Tops *et al*, 2009; Boureau and Dayan, 2010). Clinically, 5-HT metabolites in cerebrospinal fluid are decreased in impulsive disorders including impulsive aggression, violence, and mania (Linnoila *et al*, 1983; Linnoila and Virkkunen, 1992), which are characterized by both behavioral disinhibition and reduced aversive processing. Increasing 5-HT with selective serotonin reuptake inhibitors (SSRIs) might offer therapeutic benefit for impulse control disorders such as pathological gambling, sexual addiction, and personality disorders (Hollander and Rosen, 2000). These clinical findings are paralleled by observations in the laboratory showing that aversive events activate serotonergic neurons (Takase *et al*, 2004), and depletion of central 5-HT disinhibits responses that are punished by an aversive outcome (Soubrié, 1986). For example, globally reducing forebrain 5-HT through intracerebroventricular infusion of the serotonergic toxin 5,7-dihydroxytryptamine (5,7-DHT) increases premature responding on the five-choice reaction-time task (5CSRTT) (Harrison *et al*, 1997a, 1997b; Harrison *et al*, 1999) (but see Puumala and Sirvio, 1998; Dalley *et al*, 2002); transgenic rats that lack the 5-HT transporter and exhibit enhanced 5-HT transmission display reduced premature responding on the 5CSRTT (Homberg *et al*, 2007), and lowering of the 5-HT precursor tryptophan by means of the dietary acute tryptophan depletion (ATD) procedure in nonhuman primates induces risky decision making on a gambling task in nonhuman primates and rats (Evenden, 1999; Long *et al*, 2009).

These associations are not perfect. For instance, 5-HT is implicated not only in clinical and laboratory impulsivity but also in depression (Deakin and Graeff, 1991; Cools *et al*, 2008b; Esher and Roiser, 2010). In contrast to impulsivity, depression is characterized by reduced behavioral vigor and enhanced aversive processing, with negative stimuli having a greater impact on behavior and cognition (Clark *et al*, 2009). Yet, like impulsivity, depression has also been associated with low levels of 5-HT, based primarily on the therapeutic efficacy of SSRIs and observations that central 5-HT depletion through dietary manipulation can induce depressive relapse. Indeed, patients with depression show reduced tryptophan levels (Cowen *et al*, 1989), abnormal 5-HT receptor function (Drevets *et al*, 1999), and abnormal 5-HT transporter function (Staley *et al*, 1998). However, the relationship between depression and 5-HT is less clear-cut than that between impulsivity and 5-HT. Thus, although dietary 5-HT depletion can induce negative mood in individuals who have recovered from depression (Delgado *et al*, 1990; Smith *et al*, 1997), these effects seem restricted to those who were previously successfully treated with SSRIs (Booij *et al*, 2003). Moreover, this manipulation has no reliable effects on mood in healthy individuals (Ruhe *et al*, 2007; Robinson and Sahakian, 2009). These observations

have led to a variety of hypotheses that suggest that the link between depression and 5-HT might be indirect and mediated by associative learning (Robinson and Sahakian, 2008) and/or disinhibition of negative thoughts (Dayan and Huys, 2008). In fact, a recent study using direct internal jugular venous blood sampling found brain 5-HT turnover to be elevated in unmedicated patients with major depression and substantially reduced after SSRI treatment (Barton *et al*, 2008). Indeed, although many antidepressants have direct effects on serotonergic neurons, where they inhibit uptake, thus increasing extracellular levels of 5-HT, there is also evidence that the increase in 5-HT produced by (acute) administration of SSRIs might produce a net reduction of activity in the 5-HT system by flooding the somatodendritic inhibitory 5-HT_{1A} autoreceptors.

Thus, the currently dominant hypothesis of 5-HT pertains to a role in counteracting impulsivity, possibly by enhancing aversion and increasing behavioral inhibition, although its precise role in depression is not completely understood. What can we learn from the study of DA when addressing 5-HT's role in these processes?

MODELING THE MULTIPLE FUNCTIONS OF SEROTONIN

As discussed above, the study of DA's function has been strongly influenced by a quantitative computational hypothesis, the prediction error theory. A similarly detailed computational theory has not emerged for 5-HT, in part, perhaps because the extant data (particularly those concerning single neuron responses, discussed below) are less clear. For this reason, one approach has been to attempt to extrapolate from theories of DA to hypotheses for serotonergic function, in part due to empirical evidence for DA-5-HT interactions.

Consistent with the primary behavioral characterization of 5-HT as supporting functions roughly opposite to those of DA, there are also anatomical and neurophysiological reasons to believe that 5-HT serves, at least in some respects, to oppose DA (see Boureau and Dayan, 2010, this issue, for a detailed discussion of these issues). For example, there are direct projections from the 5-HT raphé nuclei to DA neurons in the substantia nigra pars compacta (SNc) and the ventral tegmental area (VTA). Although some of these projections are glutamatergic (Geisler *et al*, 2007), it is unclear whether the release sites for serotonin and glutamate in the VTA are segregated or colocalized (Geisler and Wise, 2008). Electrical stimulation of the raphé inhibits SNc DA neurons, and this effect is mediated by 5-HT (Dray *et al*, 1976; Tsai, 1989; Trent and Tepper, 1991). However, as is the case for the clinical data, this opponency is imperfect; for instance, the effects of 5-HT on DA neurons may depend on their location, with differences between SNc and VTA (Gervais and Rouillard, 2000), and on the receptor type at which it acts (Alex and Pehek, 2007), whereas evidence for

reciprocal effects of DA on 5-HT neurons is less strong than that for serotonergic effects on DA neurons.

These suggestions of opponency were leveraged in an early attempt (Daw *et al*, 2002) to extend the relatively more detailed computational understanding of DA into a hypothesis about serotonergic function. This model posited that 5-HT might serve as simply a mirror image to the dopaminergic reward prediction error signal, an idea roughly consonant with the aversive processing aspects of 5-HT function (Figure 1a).

If this viewpoint is combined with the Niv *et al* model's insight concerning the relationship between DA's appetitive and activational functions, it immediately suggests a similar resolution of 5-HT's dual roles. Indeed, a straightforward corollary of Niv *et al*'s cost-benefit analysis of rewards and vigor is that when actions are more likely to have aversive outcomes, vigorous action is more costly and sloth preferred: that is, the opportunity cost of delay decreases (Figure 1b). If we hypothesize that 5-HT reports the effects of punishment on the opportunity costs (eg, the average rate of punishment), extending the hypothesized opponency from the phasic reinforcing action to the tonic invigorating action of DA, then this sort of reasoning directly suggests an analogous coupling between aversive and inhibitory functions of 5-HT, as Niv *et al* (2007) suggested for DA. This identification echoes, but reverses, an idea about tonic serotonin from the Daw *et al* (2002) model (see also Boureau and Dayan, 2010); the present review concentrates on many functional consequences of this idea.

Thus, just as for DA, the co-occurrence of these two facets of serotonergic action may be seen as more necessary than accidental.

THE COUPLING BETWEEN INHIBITORY AND AVERSIVE EFFECTS OF SEROTONIN

In considering both DA and 5-HT, it is important to note that Niv *et al*'s formal analysis treated only a particular class of rewards and punishments: those that occur directly as a result of actions and which can, accordingly, be made to arrive earlier or later when the actions are more or less vigorous. This specialization of contingencies is essential to the basic explanation of coupling between motivational and activational variables. Another sort of rewards or punishments is those that arrive in the absence of action. These can add an additional influence on behavioral vigor, which may reverse the couplings so far described. For instance, such events can lead to situations in which vigorous action must be taken to avoid a punishment that would otherwise occur ('active avoidance'), or, conversely, in which a pre-potent action must be inhibited in order to allow a reward to occur. Effectively controlling the activation of behavior in these cases requires additional machinery for taking into account the effect of that behavior on the un-elicited punishments (or rewards) (Dayan and Huys, 2008; Boureau and Dayan, 2010; Maia, 2010). We propose that this

machinery may be separate from a 5-HT system that, by itself, tightly couples aversion and inhibition because it is specialized for the more restricted set of situations, such as passive avoidance, contained in the basic model.

The proposed specialization fits with findings from rodent work showing that performance on passive avoidance tasks is particularly vulnerable to manipulations that lower 5-HT transmission, such as benzodiazepines, *p*-chlorophenylalanine administration, and lesions of the raphé nuclei, while active avoidance is left unaffected (or if anything facilitated) (Lorens, 1978; Soubrié, 1986). Analogous effects are seen on discrimination tasks, in which depleting forebrain 5-HT improves discrimination between two active responses (Ward *et al*, 1999), while impairing discrimination between an active and a passive response (Harrison *et al*, 1999).

Such effects of low 5-HT were originally interpreted to reflect a shift toward active responding, and were emphasized to highlight the observation that effects of 5-HT transmission cannot solely be accounted for by the alleviation of anxiety or aversion (Soubrié, 1986). Indeed, performance on many different tests of impulsivity is affected by 5-HT without necessitating an obvious explanation in terms of aversion, including reversal learning, conditioned suppression, tests of premature responding, and intertemporal choice (Soubrié, 1986; Evenden, 1999; Rogers *et al*, 1999; Leyton *et al*, 2001; Clarke *et al*, 2004) (for recent reviews on the neurochemical modulation of impulsivity see Winstanley *et al*, 2006a; Dalley *et al*, 2008; Pattij and Vanderschuren, 2008).

However, purely inhibitory accounts have difficulties similar to those faced by the pure anxiety accounts, with explaining effects of 5-HT manipulations on other tasks. Thus, studies in rats and humans have shown that manipulating 5-HT does not affect performance on tasks of inhibition that have no clear affective component, such as the stop-signal reaction-time task (Clark *et al*, 2005; Cools *et al*, 2005; Chamberlain *et al*, 2006; Bari *et al*, 2009; Eagle *et al*, 2009), the self-ordered spatial working memory task (Walker *et al*, 2009), and the go–nogo task (Rubia *et al*, 2005; Evers *et al*, 2006) (but see LeMarquand *et al*, 1999).

Thus, as is the case for DA, the two seemingly separate aspects of 5-HT appear to be intertwined. More specific empirical evidence for this theoretical idea comes from a recent study by Crockett *et al* (2009), who tested both activational (go–nogo) and affective (reward *vs* punishment) factors in the context of the dietary ATD procedure in healthy human volunteers (Figure 2a). This procedure is well known to reduce central 5-HT levels, although to a modest extent. Consistent with the current hypothesis, they revealed that the 5-HT manipulation affected the factors in an interactive way rather than separately. Specifically, ATD abolished punishment-related slowing of responding in a go–nogo task, in which go- and nogo-responding were differentially rewarded or punished. Although ATD did not affect response biases toward or away from ‘nogo’, it did abolish the slowing of responding seen on correct go

reaction time periods in punished relative to rewarded conditions, with this effect on performance correlating with the effect of ATD on plasma tryptophan levels.

Further evidence for a role for 5-HT in the vigor of responding in an affective context comes from another ATD study in healthy volunteers (Cools *et al*, 2005). In this study, the effect of motivationally relevant affective signals on response vigor was measured in a reaction-time task, while the stop-signal reaction-time task was used to measure response inhibition in an affectively more neutral context. In the affective task, cues predictive of high reinforcement likelihood (high reward probability for fast, correct responding, and high punishment probability for slow or incorrect responding) induced faster, but less accurate responses compared with cues predictive of low reinforcement certainty. Depletion of central 5-HT modulated this coupling between motivation and action, so that response speed and accuracy no longer varied as a function of cued incentive certainty. Specifically, response latencies were much faster on the low reinforcement trials after ATD than after placebo, possibly reflecting disinhibition of responding in the context of a negative reward signal (Figure 2b). In contrast, ATD left the ability to inhibit prepotent responses in the stop-signal reaction-time test in the same subjects unaltered, consistent with the general set of findings (mentioned above) that 5-HT does not affect response inhibition outside an affective context.

AFFECTIVE AND ACTIVATIONAL FACTORS IN UNIT RECORDINGS FROM SEROTONERGIC NUCLEI

As is the case for DA, unit recordings from the serotonergic raphé nuclei do not entirely track the suggestions from the more causal manipulations discussed above. In addition, unlike DA, they have so far not revealed a signal with a specific computational interpretation. However, recordings do at least broadly suggest roles in both affective/motivational and activational processes, and the example of DA offers some suggestions how this work might be refined in future.

In early studies, activity of single neurons in the raphé nuclei was associated with changes in muscle tone during sleep, as well as responses mediated by central pattern generators such as chewing, locomotion, and respiration, leading to the notion that one general function of the brain serotonergic system is to facilitate motor output (Jacobs and Fornal, 1993).

On the other hand, more specific transient event-locked responses of neurons in the dorsal raphé nucleus (DRN) were recently found to depend on motivational factors. For example, Ranade and Mainen (2009) have found that such transient responses of rodent DRN neurons sometimes correlated with reward parameters, including the omission of reward, but also encoded specific sensorimotor events, suggesting that the DRN does not encode a unitary signal.

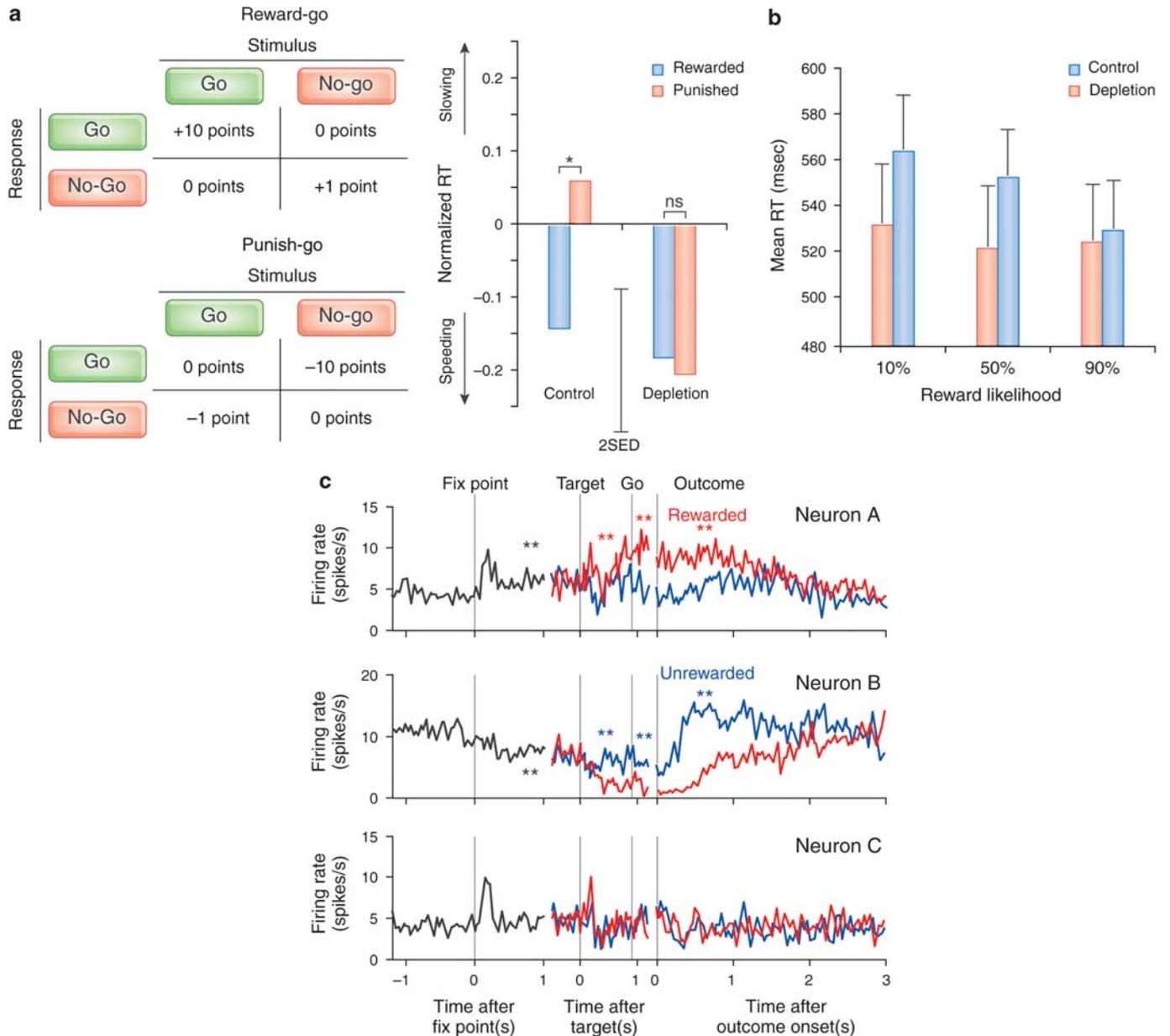


Figure 2. Preliminary empirical evidence for a role for serotonin in the interaction between vigor and negative reward signals. (a) *Left panel:* experimental paradigm employed by Crockett *et al* (2009). In the reward-go condition, subjects received large rewards for correct go responses and small rewards for correct nogo responses. In the punish-go condition, subjects received large punishments for incorrect nogo responses and small punishments for incorrect go responses. The complementary reward-nogo and punish-nogo conditions are not depicted here. *Right panel:* effect of tryptophan depletion on correct go reaction times in punished conditions relative to rewarded conditions. Tryptophan depletion abolished punishment-induced slowing. Reproduced with permission from Crockett *et al* (2009). (b) Effect of tryptophan depletion on response vigor as a function of reward likelihood (Cools *et al*, 2005). In this experiment, subjects responded as fast as possible to a target that was preceded by one of three reward cues, signaling 10, 50, or 90% reward likelihood. After placebo, subjects responded more slowly in response to low reward cues relative to high reward cues. Tryptophan depletion speeded reaction times in response to cues signaling low reward likelihood (depletion \times reward cue interaction; $P = 0.009$). Error bars represent standard errors of the mean. (c) Three types of modulation of activity of primate dorsal raphe neurons during a memory-guided saccade task. Histograms are aligned to fixation point onset (left), target onset (middle), and outcome onset (right). Lines indicate mean firing rate of all trials (black), large-reward trials (red), and small-reward trials (blue). Black asterisks indicate significant difference in activity during the 500–900 ms after fixation point onset compared with a 400 ms pre-fixation period ($P < 0.005$, rank-sum test). Red and blue asterisks indicate significant difference between two reward conditions during 150–450 ms after target onset, go onset, or outcome onset. *Top panel:* example of a neuron that increased its activity during the task and fired more for large- than small-reward trials after target onset and outcome onset. *Middle panel:* example of a neuron that decreased its activity during the task and fired more for small- than large-reward trials after target onset and outcome onset. *Bottom panel:* example of a neuron that did not change its activity during the task and did not show a significant reward effect after target onset and outcome onset. Panel c reproduced with permission from Bromberg-Martin *et al* (2010).

Performance- and reward-related activity has also been reported in behaving monkeys performing a rewarded saccade task (Nakamura *et al*, 2008). A significant proportion of recorded DRN neurons (20%) exhibited modulation of activity after the presentation of the target and/or after delivery of the reward, and this activity was proportional to the expected and/or received (large *vs* small) reward. Some neurons showed stronger activity during expectation and/or receipt of the large reward, whereas other neurons showed stronger activity during expectation and/or receipt of the small reward, the latter possibly reflecting a negative reward signal. Often, the activity pattern was characterized by long-lasting, tonic modulation. Furthermore, whereas putative DA neurons recorded on the same task followed the classic reward prediction error pattern, the DRN neurons faithfully followed expected or received reward value during the performance of the tasks (Nakamura *et al*, 2008).

This latter observation highlights one important distinction between the methods adopted to study recordings from dopaminergic and serotonergic nuclei. Both nuclei contain a number of different types of nonserotonergic and nondopaminergic units that are likely to also be recorded, and isolating the neuromodulatory units is presently at best imperfect in the awake, behaving preparation. In response to this problem, neurons in the dopaminergic midbrain nuclei are generally screened carefully for physiological and sometimes functional properties, with only those units carrying a quantitatively interpretable 'prediction error' signal being reported as putative DA neurons. Although it is quite doubtful that these screens are either necessary or sufficient to identify DA neurons (Ungless *et al*, 2004; Fields *et al*, 2007; Brischoux *et al*, 2009; Matsumoto and Hikosaka, 2009), they do isolate a highly homogenous and computationally important population. In contrast, recordings from serotonergic nuclei have not yet reached a similar degree of precise targeting—typically, a wide range of units is encountered and reported—hence, discovering any potential counterpart to the prediction error population may require further subselection of raphé neurons.

Indeed, further analyses of the Nakamura data, breaking the neurons down by functional properties, have begun to discriminate some regularities and clearer functional classes (Bromberg-Martin *et al*, 2010). In particular, some DRN neurons exhibited activity reflecting reward value in a consistent manner both after task initiation and after the trial's value was revealed. Neurons that were tonically excited during the task period before the receipt of rewards also predominantly carried positive reward signals, firing more following the receipt of a large than a small reward. Neurons that were tonically inhibited during the task period before the receipt of rewards predominantly carried inhibitory reward signals (Figure 2c). This work represents a first step in parsing the raphé population into more functionally discrete classes; indeed the sustained, tonic reward-inhibited responses exhibited there might provide a substrate for the average punishment signal envisioned in this article. Of course, the same figure also illustrates

a mirror-image class of reward-activated neurons, and there is at present no evidence to guide the identification of serotonergic status with either (or both) of these populations.

INTERTEMPORAL CHOICE

So far, we have discussed modeling showing how the concept of an opportunity cost (together with the effects of average reward and punishment rates on this cost) helps to unite the aversive and inhibitory associations of 5-HT, and, similarly the appetitive and activational functions of DA. In fact, this computational concept also captures several additional, potentially distinct, domains of function of these neurotransmitters: time discounting, perseveration versus switching, and risk (Figures 1c–e).

Time discounting is the subject of another prominent computational theory of serotonergic function (Doya, 2002), which posits that 5-HT controls (im)patience in intertemporal choice: the degree of preference for immediate rewards over delayed rewards. Specifically, Doya proposed that 5-HT controls a parameter common to many decision models known as the temporal discount factor according to which delayed rewards are viewed as less valuable than immediate ones, with higher 5-HT promoting greater patience.

Colloquially, impatience is another form of impulsivity—although in principle potentially different from the more motoric sorts of impulsivity discussed so far—and so this proposal seems at least broadly related to the behavioral withholding functions of 5-HT. This is formally the case under Niv *et al*'s model, in which the opportunity cost of time (the variable purported to be signaled by tonic 5-HT and DA) should control impatience in intertemporal choice in the same manner, and for the same reason, that it controls vigor of motor responding. Indeed, Niv *et al*'s original analysis of the activational problem of deciding how vigorously (ie, when) to press a lever actually treated this problem formally as an intertemporal choice problem: whether to push it faster (getting the outcome, eg, reward, sooner but incurring more energetic cost) or slower (getting the outcome later but at lower cost). A typical intertemporal choice problem also involves choosing between earlier and later rewards, although in this case, they differ in magnitudes rather than costs. Here, just as in the vigor case, the degree to which a subject might be willing to wait should, in the Niv *et al*'s model, be controlled by the opportunity cost of time, which has a role analogous to the temporal discount factor in the Doya model. This is because whether it is worth waiting for a larger reward depends essentially on trading off the value of that reward against the cost of the delay, which can be measured by the rewards (minus punishments) that would, on average, be foregone by waiting, that is, the opportunity cost or average reward (Figure 1c).

Thus, the theory sketched here resolves the seeming contradiction between the earlier 5-HT models of Daw *et al* (2002) and Doya (2002), as it proposes a common role in these functions and in particular contains the Doya model as, in effect, a special case. More empirically, if 5-HT participates in reporting the opportunity cost that controls this tradeoff, then it should have common effects both on behavioral vigor and on choice between immediate and delayed rewards. Indeed there is considerable evidence implicating 5-HT in intertemporal choice, which of course was what prompted the Doya proposal initially. Briefly, studies with experimental rodents have shown that depleting forebrain 5-HT leads to consistent choices of small, immediate rewards over large, delayed rewards, possibly reflecting hypersensitivity to the delay (Wogar *et al*, 1993; Mobini *et al*, 2000; Cardinal *et al*, 2004; Denk *et al*, 2005; Cardinal, 2006) (but see Winstanley *et al*, 2003). Conversely, increasing 5-HT function with the 5-HT indirect agonist fenfluramine decreases impulsive choice (Poulos *et al*, 1996; Bizot *et al*, 1999); and 5-HT efflux was found to be increased in the medial PFC (though not OFC) during delay discounting, as measured with microdialysis (Winstanley *et al*, 2006b). In line with this proposal and animal work, Schweighofer *et al* (2008) have recently shown that ATD also steepens delayed reward discounting in humans, resulting in increased choice of the more immediate small rewards (but see Crean *et al* (2002), who used hypothetical rather than experiential choices). These findings are reminiscent of other results obtained by the same group showing that ATD impaired learning when actions were followed by delayed punishment (Tanaka *et al*, 2009).

Thus, consistent with the proposal's predictions, manipulations of 5-HT have common effects both on the balance between behavioral withholding and vigor (as exemplified by premature responding on the 5CSRTT, see above, as well as passive avoidance) and on choice between immediate and delayed rewards.

Another implication of the theoretical view on discounting presented here is that, insofar as tonic DA is also thought to be involved in reporting appetitive components of the opportunity cost, it should also have effects on intertemporal choice that parallel its effects on vigor and oppose those of 5-HT. Time discounting has not had as prominent a role in computational models of dopaminergic function, and, empirically, the answer is not so straightforward. Similar to 5-HT depletion, amphetamine administration increases impulsive, premature responding on the 5CSRTT in a DA-dependent fashion (Cole and Robbins, 1987; Harrison *et al*, 1999; Van Gaalen *et al*, 2006)—this is another instance of the overall involvement of DA in behavioral activation with which this article began. However, effects of DA-enhancing psychostimulants on intertemporal choice have varied, with some studies reporting that they promote choice of delayed reinforcers (Wade *et al*, 2000; de Wit *et al*, 2002), consistent with its beneficial effect on clinical impulsivity in ADHD, whereas

others have found the opposite effect (Logue *et al*, 1992; Charrier and Thiebot, 1996; Evenden and Ryan, 1999). Only the latter set of findings is consistent with the model presented here.

An important issue to consider is the degree to which effects of psychostimulants are mediated by DA and/or 5-HT. For example Winstanley *et al* (2003) have found that effects of amphetamine, which also increases 5-HT transmission (Kuczenski *et al*, 1987), on intertemporal choice are attenuated by 5-HT depletion. One implication of this observation is that (some of) the calming, anti-impulsive effects of amphetamine administration in ADHD might be related to the drugs' enhancing effect on 5-HT transmission.

One other way to reconcile the contradictory data on amphetamine with the current model is by considering the possible role of intervening events during the delay (Lattal, 1987), which might acquire conditioned reinforcing properties of their own. For example, consistent with the current model, Cardinal *et al* (2000) have observed that amphetamine promoted choice of the small, immediate reinforcer if the large, delayed reinforcer was not signaled, whereas the same treatment promoted choice of the large, delayed reinforcer if it was signaled with a stimulus spanning the gap. It is possible that the impulsivity-reducing effects of amphetamine reflect effects on conditioned reinforcement (Hill, 1970; Robbins, 1976) rather than effects on the appetitive component of the opportunity cost or waiting *per se*. Conditioned reinforcement is closely linked to the learning functions of (presumably phasic) DA, as traditionally posited in RL models such as the actor/critic (Balleine *et al*, 2008; Balleine and O'Doherty, 2010; Maia, 2010), and effects of amphetamine on this function might have masked the additional, performance-related effects of the opportunity cost posited by Niv *et al*.

PERSEVERATION AND SWITCHING

This brings us back to the hypothesized role of DA and, potentially, 5-HT in reinforcement. RL models have traditionally envisioned that the prediction error carried by phasic DA (and, in the Daw *et al* (2002) model, a hypothesized aversive prediction error tentatively identified with phasic 5-HT), has a role in reinforcement, by which better-than-expected outcomes increase the propensity to take the actions that led to them, and worse-than-expected outcomes decrease it (Houk *et al*, 1995; Balleine *et al*, 2008; Maia, 2010).

What are the implications for reinforcement and choice of a model like Niv *et al*'s that incorporates opportunity costs? Might these changes introduced by Niv *et al* help us conceptualize further aspects of the neuromodulators' function? The same average reward (and average punishment) terms that furnish the opportunity cost and are supposed to control vigor and time discounting also appear in the prediction error learning rule associated with these models (Daw *et al*, 2002; Daw and Touretzky, 2002;

Niv *et al*, 2007). There, they have the role of a ‘comparison term’ or baseline against which obtained rewards and punishments are weighed, before their being used to drive learning (Figure 1d). In particular, in this class of models, the average reward is subtracted from the obtained one (and similarly for punishments). The intuition for this is that the average rewards represent a sort of ‘aspiration level’: a particular reward is only ‘good enough’ if it is better than the average reward that would have been expected anyway; otherwise it is, comparatively, a loss.

One consequence of this view is that, if we consider any experimental treatment that modulates these average comparison terms (putatively, tonic 5-HT or DA), while leaving more phasic prediction error signaling relatively intact, such a treatment should essentially function to modulate the overall baseline or aspiration level against which all other outcomes are measured. Making this baseline more appetitive (increasing tonic DA or decreasing tonic 5-HT) should render rewards, effectively, less good and punishments worse; the opposite manipulations should have the opposite effect. Through reinforcement, then, the effect of this should be to promote switching away from an action or option when the baseline is good (and outcomes look worse in comparison), as in the case of high tonic DA and low tonic 5-HT, and perseverating in it when the baseline is bad (and outcomes look better in comparison), as in the case of low tonic DA and high tonic 5-HT.

These predictions may relate to a number of results concerning how neuromodulatory manipulations encourage either perseveration or switching in various dynamic learning tasks such as reversals. For example, modest reduction of background 5-HT with ATD impairs choice during probabilistic reversal learning (Murphy *et al*, 2002), in which the correct choice is rewarded on 80% of trials, but punished on 20% of trials. The hypothesis that this effect of ATD, which might well have a selective effect on tonic 5-HT, reflects enhanced switching in response to poor outcomes concurs with the observation that a single dose of the selective 5-HT reuptake inhibitor (SSRI) citalopram increased the likelihood of inappropriate switching after probabilistic punishment (Chamberlain *et al*, 2006). Acute SSRI administration has been hypothesized to reduce 5-HT transmission through action at presynaptic receptors, leading to a net reduction in activity of the 5-HT system (Artigas, 1993; Blier and de Montigny, 1999), and the enhanced impact of poor outcomes on switching could reflect this net reduction in 5-HT activity. Indeed, enhanced impact of poor outcomes during probabilistic reversal learning was also found after ATD in terms of a potentiation of blood oxygenation level-dependent signals, measured with fMRI during the receipt of punishment in this task (Evers *et al*, 2005). Recent genetic data have confirmed that the tendency to switch after punishment during probabilistic reversal learning is sensitive to 5-HT transmission by showing that subjects homozygous for the long allele of the 5-HT transporter polymorphism, associated with increased expression of the 5-HT transporter, exhibit increased

similar tendency to switch after punishment relative to carriers of the short allele (Den Ouden *et al*, 2010). The hypothesis that decreasing tonic 5-HT with ATD renders punishments worse by making the baseline more appetitive also fits with other recent data showing that ATD enhances the ability to predict punishment in an observational outcome prediction task (Cools *et al*, 2008a).

However, again the results are not clean. A series of studies with nonhuman primates (marmosets) has shown that depletion of 5-HT by injection of 5,7-DHT actually increases perseveration on reversal learning (Clarke *et al*, 2004; Clarke *et al*, 2005; Clarke *et al*, 2007) and detour reaching tasks (Walker *et al*, 2006), while also inducing stimulus-bound responding in tests of conditioned reinforcement and extinction (Walker *et al*, 2009). Of course it remains to be determined how the relationship between putative tonic and phasic 5-HT might be affected by manipulation of 5,7-DHT, which has much more profound effects on 5-HT levels, thus also possibly affecting phasic transmission than the more modest manipulations of ATD (and possibly than acute administration of low doses of SSRIs). Resolution of similar uncertainty about mechanisms of action in terms of tonic versus phasic transmission will be necessary for interpreting effects on punishment-based switching of dopaminergic drugs (Frank *et al*, 2004; Cools *et al*, 2006; Clatworthy *et al*, 2009; Cools *et al*, 2009).

‘Switching’ as discussed above refers literally to changing from one option to another, as with a rat moving from one lever to another in a multiple operant task. The concept is that the organism learns to assign values to the choice of different options, and the effect of the comparison term on this learning promotes switching or perseveration in the action. Such an account could also be extended to more abstract sorts of switching associated with cognitive control—such as switching between task sets, or between rules in a task such as the Wisconsin Card Sorting test. In particular, the former type of switching between task sets, at least when they are well learnt, is highly sensitive to dopaminergic drugs in patients with PD as well as in healthy volunteers (Kehagia *et al*, 2010; Cools, 2006; Robbins, 2007). Recent genetic imaging studies have shown that task set switching also varies as a function of individual genetic differences in DA function, particularly when subjects are expecting to be rewarded (Aarts *et al*, 2010). The latter study revealed that this DA-dependent effect of reward on task set switching was accompanied by modulation of the dorsomedial part of the striatum (Aarts *et al*, 2010), further highlighting that effects of DA on task set switching might occur via modulation of different dopaminergic target regions in more dorsal parts of the striatum than those associated with reversal learning, which rather implicates the ventral striatum (Cools *et al*, 2001).

The potential computational bridge between physical and cognitive switching is recent modeling work (O’Reilly and Frank, 2006; Todd *et al*, 2008) that has conceptualized more abstract, regulatory decisions of this sort (specifically, what

task set to maintain) as RL problems about internal or cognitive ‘actions’ (such as gating contents in or out of working memory). This viewpoint places issues of regulation and action control on a common conceptual footing: regulatory decisions are conceived as being controlled by RL processes entirely analogous to those for decisions about physical actions, although operating over distinct networks such as prefrontal cognitive control systems. Thus, the operation of a comparison term on this hypothesized learning about which internal actions to favor—say, the choice of which task set to activate at a given trial—might produce perseverative or switch-promoting effects analogous to learning about different external options. Consonant with the genetic imaging data discussed above, learning about cognitive versus physical actions is envisioned to involve dopaminergic action at different target areas (O’Reilly and Frank, 2006).

RISK

A third domain of function captured by the computational concepts presented here is risk. Risk seeking is a tendency in decision making to favor options with more variable payoffs compared with more stable ones, even if this is disadvantageous on average. As with impatience, although this preference might broadly be considered a form of impulsivity, it has no obvious mechanistic link to motor impulsivity or behavioral vigor. However, here again, the concept that obtained rewards and punishments are weighed relative to the comparison term helps to bring this function under a common umbrella with the others discussed here.

To develop the relationship, standard models of risk sensitivity must be considered. In economics, the predominant account of risk sensitivity is nonlinearity in the subjective value of outcomes. For instance, if \$2 is not worth twice as much to you as \$1, then you’d be better off taking \$1 for sure than gambling for \$2, with 50% probability (and \$0 otherwise)—thus, you are risk averse for gains. Conversely, if the prospect of losing \$2 hurts you less than twice as much as losing \$1, you’re better off gambling on a 50/50 shot at losing nothing (*vs* \$2), than losing \$1 for sure—you are risk seeking for losses.

This basic pattern—of risk aversion for gains and risk seeking for losses—is typical in human economic decisions (Kahneman and Tversky, 1979). What connects this to comparison terms—and thus, potentially, to DA and 5-HT—is that what counts as a gain versus a loss is relative to some measure of the *status quo*. The idea that outcomes are weighed relative to some reference point, with risk aversion above it and risk seeking below it due to nonlinear valuation of gains and losses, is central to prospect theory, a predominant account of risk-sensitive choice in humans (Kahneman and Tversky, 1979). The proposed dopaminergic and serotonergic average reward and punishment signals discussed here are candidate neural substrates for

this baseline. Although there is relatively little work in behavioral economics on how the reference point might be determined from experience, there is a study of choices in the televised game show ‘Deal or No Deal’, investigating how contestants’ risk sensitivity fluctuates following events in the game (Post *et al*, 2008). The results suggest that the contestants’ reference points follow a weighted average of past (paper) wealth states, substantially similar to proposals from DA and 5-HT models for tracking the average reward by averaging previous rewards or prediction errors (Daw *et al*, 2002; Daw and Touretzky, 2002; Niv *et al*, 2007).

Finally, then, if we identify the average reward with the reference point—or import prospect theory’s reference-dependent nonlinear values into the RL account developed here—then this couples an effect on risk sensitivity to the other factors discussed thus far (Figure 1e). In particular, we predict that a more appetitive baseline (high DA or low 5-HT) should promote risk seeking by making more outcomes look, relatively, like losses, and, conversely, more aversive baselines (low DA or high 5-HT) should promote risk aversion. Accordingly, DA replacement therapy for PD is associated with impulse control disorders including pathological gambling (Dodd *et al*, 2005). Genetic polymorphisms related to DA and 5-HT function also interact with risk sensitivity; notably, subjects homozygous for the short allele of the 5-HT transporter gene (associated with reduced transporter function and possibly enhanced 5-HT levels) are more risk averse than others (Kuhnen and Chiao, 2009). Finally, Murphy *et al* (2009) studied risk preference under dietary tryptophan loading (expected to increase 5-HT). They found that the treatment attenuates both risk aversion for gains and risk seeking for losses, but more consonant with the view here, that it also selectively attenuates discrimination between small and large rewards, consistent with the nonlinear valuation supposed to underlie risk effects, that is, diminishing sensitivity for rewards relative to a more aversive reference point.

SUMMARY

To advance the study of 5-HT’s complex role in behavior, we have leveraged current understanding of the role of DA in behavior. According to current theorizing, two seemingly separate affective and activational consequences of DA are necessarily and not accidentally related through a more fundamental role in trading off the costs and benefits of taking action for reward. Here, we suggest to extend this reasoning to 5-HT and argue that, although DA serves to promote behavioral activation to seek rewards, conversely 5-HT serves to inhibit actions when punishment may occur. This is hypothesized to result from an analogous fundamental role of 5-HT in trading off the costs and benefits of waiting to avoid punishment.

These functions, in turn, are proposed to follow from a more fundamental involvement of tonic DA and 5-HT in

representing the opportunity cost of time—measured by the average rates of reward and punishment—a variable that is expected to control the balance between behavioral activation and withholding. We have further shown how these same core quantities should have a host of other functional effects, including on time discounting, switching, and risk sensitivity. On the basis of the above, our working hypothesis is that 5-HT and DA should control neither reward or punishment nor behavioral activation or inhibition *per se*, but instead their interaction, and should further implicate a number of other functions.

Most existing theoretical accounts of DA and 5-HT have focused on the function of phasic changes in neurotransmission, for example, RL. Extrapolating these insights to the role of tonic neurotransmission and response vigor is critical not only for reconciling paradoxical laboratory observations and for directing future fundamental research but also for progress in the understanding and treatment of neuropsychiatric disorders. Indeed, the therapeutic benefit offered by dopaminergic and serotonergic drugs for disorders characterized by motor and cognitive control most likely reflects changes in tonic neurotransmission in addition, or even as opposed, to changes in phasic neurotransmission. The observation that alterations in the putative tonic average outcome signal can have a wide variety of functional consequences ranging from response slowing to cognitive inflexibility, impatience for reward, and risk seeking might account for the fact that these drugs show apparent nonspecific efficacy in the treatment of a wide variety of abnormalities ranging from PD to pain, depression, and impulse control disorder. However, the framework also provides a theoretical basis for more broadly defined specificity of drug effects observed clinically, with dopaminergic and serotonergic drugs having opposite effects in the domains of motor and cognitive impulsivity and flexibility. According to this framework, these wide ranging effects might stem from the modulation of a common signal, but the precise direction of effects will depend critically on the degree to which treatments affect phasic and/or tonic neurotransmission.

FUTURE RESEARCH DIRECTIONS

Although our review of the extant literature from the perspective of the model outlined here has identified numerous anomalous or confusing findings, we do find, at minimum, a great deal of evidence that the numerous behavioral factors that we identify are all clearly sensitive to manipulations of both neuromodulators. Therefore, although we think it highly unlikely that our simple working model will survive future experiments unscathed, we advocate a systematic assessment of these key factors, and particularly their relationships and interactions, at a variety of levels to clarify in exactly what respects this account breaks down.

One ambiguity pervading the interpretation and comparability of the data is the actual effect of different

experimental treatments, including their differential effects on the two neuromodulators, on tonic versus phasic activity, and even in some cases the overall direction of their net effect. Thus, the finding of clear effects, but sometimes in unexpected directions, may suggest that our account captures essential functions of the neuromodulators but what is lacking is an understanding of the experimental treatments. In this respect, as the functional framework here predicts a clear clustering of effects due to their hypothesized common underlying cause, it may be useful to assess covariation across all these measures under a common neuromodulatory manipulation. For instance, an increased average reward signal should speed operant behavior, decrease patience in temporal discounting, decrease perseveration, and promote risk seeking, (Figure 1b–e).

At the same time, it should be possible to pursue both more precise methods and more understanding of the existing toolbox. For instance, in order to fully understand these neuromodulatory effects, it will be particularly important to consider their timescale (tonic or phasic). Specifically, it will be important to obtain better insights in the degree to which commonly used 5-HT manipulations affect phasic versus tonic transmission, thus highlighting the necessity of combining temporally precise methods in freely behaving animals, such as neurophysiological recording of single 5-HT and DA neurons, electrochemical voltammetric approaches (Hashemi *et al*, 2009), and/or optogenetics with procedures used to study the effects of 5-HT, for example, 5,7-DHT lesions, ATD, SSRI administration, and the 5-HT transporter gene-linked polymorphism (5HTTLPR).

In addition, in terms of neurophysiological recording from serotonergic nuclei, progress in discovering any potential counterpart to the DA neuron population will depend on the development of a similar degree of precise targeting by neurochemical means (Ungless *et al*, 2004; Fields *et al*, 2007) or functional procedures for subselection of 5-HT neurons. We also identify the average reward and punishment as functionally and computationally important signals, quantitatively defined and easily manipulable, for which neural correlates might usefully be directly tested in electrophysiology, voltammetry, or dialysis.

Finally, it will be important to take into account the regional specificity of neuromodulatory effects, not only given receptor specificity but also given that differential processing in distinct target regions will likely influence the behavioral expression of the common function proposed here. Thus, as is the case for DA, 5-HT might have distinct effects in the ventral striatum, the amygdala, and the OFC (Clarke *et al*, 2008; Boulougouris and Robbins, 2010), or on functions associated with ventral versus dorsal frontostriatal circuitry (Tanaka *et al*, 2007). Crucial insights will also derive from an understanding of the neural mechanisms that control the activity of 5-HT neurons, such as the medial prefrontal cortex (Amat *et al*, 2005) and/or lateral habenula (Hikosaka *et al*, 2008; Hikosaka, 2010).

ACKNOWLEDGEMENTS

This work was supported by a Research grant from the Human Frontiers Science Program to KN, RC, and NDD (RGP0036/2009-C). RC is also supported by a VIDI grant from the Innovational Research Incentives Scheme of the Netherlands Organisation for Scientific Research (NWO). NDD is also supported by a Scholar Award from the McKnight Foundation, a Young Investigator Award from NARSAD, and NIH grant R01MH087882-01, part of the CRCNS program. KN is supported by the Precursory for Embryonic Science and Technology, Takeda Foundation, the Nakayama Foundation, a Grant-in-Aid for Scientific Research B, and a Grant-in-Aid for Scientific Research on Priority Areas. We are grateful to our collaborators and colleagues Peter Dayan, Ben Seymour, Yael Niv, Y-Lan Boureau, Trevor Robbins, and Daniel Campbell-Meiklejohn for many useful discussions and ideas.

DISCLOSURE

The authors declare no conflict of interest.

REFERENCES

- Aarts E, Roelofs A, Franke B, Rijkema M, Fernandez G, Helmich RC *et al* (2010). Striatal dopamine mediates the interface between motivational and cognitive control in humans: evidence from genetic imaging. *Neuropsychopharmacology* **35**: 1943–1951.
- Alex KD, Pehek EA (2007). Pharmacological mechanisms of serotonergic regulation of dopamine neurotransmission. *Pharmacol Ther* **113**: 296–320.
- 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.
- Arbuthnott G, Wickens J (2007). Space, time and dopamine. *Trends Neurosci* **30**: 62–69.
- Artigas F (1993). 5-HT and antidepressants: new views from microdialysis studies. *Trends Pharmacol Sci* **14**: 262.
- Balleine B, Daw N, O'Doherty J (2008). Multiple forms of value learning and the function of dopamine. In: Glimcher P, Fehr E, Camerer C, Poldrack R (eds) *Neuroeconomics: Decision-Making and the Brain*. pp 367–385.
- 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.
- Bari A, Eagle DM, Mar AC, Robinson ES, Robbins TW (2009). Dissociable effects of noradrenaline, dopamine, and serotonin uptake blockade on stop task performance in rats. *Psychopharmacology (Berl)* **205**: 273–283.
- Barton DA, Esler MD, Dawood T, Lambert EA, Haikerwal D, Brenchley C *et al* (2008). Elevated brain serotonin turnover in patients with depression: effect of genotype and therapy. *Arch Gen Psychiatry* **65**: 38–46.
- Bayer HM, Glimcher PW (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47**: 129–141.
- Berridge KC (2007). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology (Berl)* **191**: 391–431. **Review of longstanding controversies about what psychological aspects of reward dopamine might subserv, including hedonics, learning and motivation.**
- Berridge KC, Robinson TE (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res Brain Res Rev* **28**: 309–369.
- Bizot J, Le Bihan C, Puech AJ, Hamon M, Thiebot M (1999). Serotonin and tolerance to delay of reward in rats. *Psychopharmacology (Berl)* **146**: 400–412.
- Blier P, de Montigny C (1999). Serotonin and drug-induced therapeutic responses in major depression, obsessive-compulsive and panic disorders. *Neuropsychopharmacology* **21**: 91S–98S.
- Booij L, Van der Does AJ, Riedel WJ (2003). Monoamine depletion in psychiatric and healthy populations: review. *Mol Psychiatry* **8**: 951–973.
- Boulougouris V, Robbins T (2010). Enhancement of spatial reversal learning by 5-HT_{2c} receptor antagonism is neuroanatomically specific. *J Neurosci* **30**: 930–938.
- Boureau Y-L, Dayan P (2010). Opponency revisited: competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology*. **Review in the current issue taking a complementary approach, offering in particular a more detailed discussion of the nature of interactions between DA and 5-HT and between reward and punishment in the context of conflicts that arise between Pavlovian and instrumental responses.**
- Brischoux F, Chakraborty S, Brierley DI, Ungless MA (2009). Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. *Proc Natl Acad Sci USA* **106**: 4894–4899.
- Bromberg-Martin E, Hikosaka O, Nakamura K (2010). Coding of task reward value in the dorsal raphe nucleus. *J Neurosci* **30**: 6262–6272. **Empirical single unit recording work showing that dorsal raphe neurons encode task performance in terms of its future motivational consequences. One class of neurons exhibited tonic reward-inhibited responses, which could correspond to the average punishment signal conceptualized in this article.**
- Cardinal R, Winstanley C, Robbins T, Everitt B (2004). Limbic corticostriatal systems and delayed reinforcement. *Ann NY Acad Sci* **1021**: 33–50.
- Cardinal RN (2006). Neural systems implicated in delayed and probabilistic reinforcement. *Neural Netw* **19**: 1277–1301. **Comprehensive review of the contributions to delay and uncertainty discounting of neuromodulators including serotonin, dopamine, and noradrenaline, and of specific neural structures.**
- Cardinal RN, Robbins TW, Everitt BJ (2000). The effects of d-amphetamine, chlordiazepoxide, alpha-flupenthixol and behavioural manipulations on choice of signalled and unsignalled delayed reinforcement in rats. *Psychopharmacology (Berl)* **152**: 362–375.
- Chamberlain SR, Muller U, Blackwell AD, Clark L, Robbins TW, Sahakian BJ (2006). Neurochemical modulation of response inhibition and probabilistic learning in humans. *Science* **311**: 861–863.
- Charrier D, Thiebot MH (1996). Effects of psychotropic drugs on rat responding in an operant paradigm involving choice between delayed reinforcers. *Pharmacol Biochem Behav* **54**: 149–157.
- Clark L, Chamberlain SR, Sahakian BJ (2009). Neurocognitive mechanisms in depression: implications for treatment. *Annu Rev Neurosci* **32**: 57–74.
- Clark L, Roiser JP, Cools R, Rubinsztein DC, Sahakian BJ, Robbins TW (2005). Stop signal response inhibition is not modulated by tryptophan depletion or the serotonin transporter polymorphism in healthy volunteers: implications for the 5-HT theory of impulsivity. *Psychopharmacology (Berl)* **182**: 570–578.
- Clarke H, Robbins T, Roberts AC (2008). Lesions of the medial striatum in monkeys produce perseverative impairments during reversal learning similar to those produced by lesions of the orbitofrontal cortex. *J Neurosci* **28**: 10972–10982.
- Clarke H, Dalley J, Crofts H, Robbins T, Roberts A (2004). Cognitive inflexibility after prefrontal serotonin depletion. *Science* **304**: 878–880.
- 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.
- Clarke HF, Walker SC, Crofts HS, Dalley JW, Robbins TW, Roberts AC (2005). Prefrontal serotonin depletion affects reversal learning but not attentional set shifting. *J Neurosci* **25**: 532–538.
- Clatworthy PL, Lewis SJ, Richard L, Hong YT, Izquierdo D, Clark L *et al* (2009). Dopamine release in dissociable striatal subregions predicts the different effects of oral methylphenidate on reversal learning and spatial working memory. *J Neurosci* **29**: 4690–4696.
- Cole BJ, Robbins TW (1987). Amphetamine impairs the discriminative performance of rats with dorsal noradrenergic bundle lesions on a 5-choice serial reaction time task: new evidence for central dopaminergic-noradrenergic interactions. *Psychopharmacology (Berl)* **91**: 458–466.
- Cools R (2006). Dopaminergic modulation of cognitive function-implications for L-DOPA treatment in Parkinson's disease. *Neurosci Biobehav Rev* **30**: 1–23.
- Cools R, Altamirano L, D'Esposito M (2006). Reversal learning in Parkinson's disease depends on medication status and outcome valence. *Neuropsychologia* **44**: 1663–1673.
- Cools R, Robinson O, Sahakian BJ (2008a). Acute tryptophan depletion in healthy volunteers enhances punishment prediction but does not affect reward prediction. *Neuropsychopharmacology* **33**: 2291–2299. **A study in healthy volunteers showing effects of tryptophan depletion on punishment, but not reward prediction.**
- Cools R, Roberts AC, Robbins TW (2008b). Serotonergic regulation of emotional and behavioural control processes. *Trends Cogn Sci* **12**: 31–40. **Review highlighting the apparently paradoxical role of 5-HT in both aversion and response inhibition.**

- Cools R, Barker RA, Sahakian BJ, Robbins TW (2001). Enhanced or impaired cognitive function in Parkinson's disease as a function of dopaminergic medication and task demands. *Cereb Cortex* **11**: 1136–1143.
- Cools R, Blackwell A, Clark L, Menzies L, Cox S, Robbins TW (2005). Tryptophan depletion disrupts the motivational guidance of goal-directed behavior as a function of trait impulsivity. *Neuropsychopharmacology* **30**: 1362–1373.
- Cools R, Frank M, Gibbs S, Miyakawa A, Jagust W, D'Esposito M (2009). Striatal dopamine synthesis capacity predicts dopaminergic drug effects on flexible outcome learning. *J Neurosci* **29**: 1538–1543.
- Cowen PJ, Parry-Billings M, Newsholme EA (1989). Decreased plasma tryptophan levels in major depression. *J Affect Disord* **16**: 27–31.
- Cragg S, Rice M (2004). DANCing past the DAT at a DA synapse. *Trends Neurosci* **27**: 270–277.
- Crean J, Richards J, 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.
- 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. **Empirical study in healthy volunteers, manipulating affective and activational factors independently in a task that did not require learning. Tryptophan depletion abolished punishment-induced inhibition without affecting overall motor response inhibition or the ability to adjust response bias in line with punishment contingencies.**
- D'Ardenne K, McClure SM, Nystrom LE, Cohen JD (2008). BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science* **319**: 1264–1267.
- 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.
- Daw N, Kakade S, Dayan P (2002). Opponent interactions between serotonin and dopamine. *Neural Netw* **15**: 603–616. **Early computational model positing that 5-HT might serve as simply a mirror image to the dopaminergic reward prediction error signal, an idea roughly consonant with the aversive processing aspects of 5-HT function.**
- Daw ND, Touretzky DS (2002). Long-term reward prediction in TD models of the dopamine system. *Neural Comput* **14**: 2567–2583.
- Day JJ, Roitman MF, Wightman RM, Carelli RM (2007). Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nat Neurosci* **10**: 1020–1028.
- Dayan P, Huys QJ (2008). Serotonin, inhibition, and negative mood. *PLoS Comput Biol* **4**: e4.
- de Wit H, Enggasser JL, Richards JB (2002). Acute administration of d-amphetamine decreases impulsivity in healthy volunteers. *Neuropsychopharmacology* **27**: 813–825.
- Deakin J (1983). Roles of serotonergic systems in escape, avoidance and other behaviours. In: Cooper S (ed). *Theory in Psychopharmacology*. Academic Press: London and New York.
- Deakin J (1998). The role of serotonin in panic, anxiety and depression. *Int Clin Psychopharmacol* **13**: S1–S5. **Review on the role of 5-HT in anxiety, panic and depression, hypothesizing that these distinct disorders arise from serotonergic modulation of distinct neural systems (eg the dorsal raphe projection to the amygdala, the brainstem and the median raphe projection to hippocampus, respectively), implicating different receptors.**
- Deakin J, Graeff F (1991). 5-HT and mechanisms of defence. *J Psychopharmacol* **5**: 305–315. **Review presenting the idea that brain 5-HT is concerned with adaptive responses to aversive events.**
- Delgado PL, Charney DS, Price LH, Aghajanian GK, Landis H, Heninger GR (1990). Serotonin function and the mechanism of antidepressant action. Reversal of antidepressant-induced remission by rapid depletion of plasma tryptophan. *Arch Gen Psychiatry* **47**: 411–418.
- Den Ouden H, Elshout J, Rijpkema M, Franke B, nandéz G, Cools R (2010). Dissociable effects of serotonin and dopamine transporter polymorphisms on probabilistic reversal learning. 7th Forum of European Neuroscience, 3–7 July 2010, Amsterdam.
- Denk F, Walton ME, Jennings KA, Sharp T, Rushworth MF, Bannerman DM (2005). Differential involvement of serotonin and dopamine systems in cost-benefit decisions about delay or effort. *Psychopharmacology (Berl)* **179**: 587–596.
- Dodd ML, Klos KJ, Bower JH, Geda YE, Josephs KA, Ahlskog JE (2005). Pathological gambling caused by drugs used to treat Parkinson disease. *Arch Neurol* **62**: 1377–1381.
- Doya K (2002). Metalearning and neuromodulation. *Neural Netw* **15**: 495–506.
- Dray A, Gonye TJ, Oakley NR, Tanner TI (1976). Evidence for the existence of a raphe projection to the substantia nigra in rat. *Brain Res* **113**: 45–57.
- Drevets W, Frank E, Price J, Kupfer D, Holt D, Greer P *et al* (1999). PET imaging of serotonin 1A receptor binding in depression. *Biol Psychiatry* **46**: 1375–1387.
- Eagle DM, Lehmann O, Theobald DE, Pena Y, Zakaria R, Ghosh R *et al* (2009). Serotonin depletion impairs waiting but not stop-signal reaction time in rats: implications for theories of the role of 5-HT in behavioral inhibition. *Neuropsychopharmacology* **34**: 1311–1321.
- Esher N, Roiser J (2010). Reward and punishment processing in depression. *Biol Psychiatry* **68**: 118–124.
- Evenden J (1999). Varieties of impulsivity. *Psychopharmacology* **146**: 348–361.
- Evenden JL, Robbins TW (1984). Effects of unilateral 6-hydroxydopamine lesions of the caudate-putamen on skilled forepaw use in the rat. *Behav Brain Res* **14**: 61–68.
- Evenden JL, Ryan CN (1999). The pharmacology of impulsive behaviour in rats VI: the effects of ethanol and selective serotonergic drugs on response choice with varying delays of reinforcement. *Psychopharmacology (Berl)* **146**: 413–421.
- Evers EA, van der Veen FM, van Deursen JA, Schmitt JA, Deutz NE, Jolles J (2006). The effect of acute tryptophan depletion on the BOLD response during performance monitoring and response inhibition in healthy male volunteers. *Psychopharmacology (Berl)* **187**: 200–208.
- Evers EA, Cools R, Clark L, van der Veen FM, Jolles J, Sahakian BJ *et al* (2005). Serotonergic modulation of prefrontal cortex during negative feedback in probabilistic reversal learning. *Neuropsychopharmacology* **30**: 1138–1147.
- Fields HL, Hjelmstad GO, Margolis EB, Nicola SM (2007). Ventral tegmental area neurons in learned appetitive behavior and positive reinforcement. *Annu Rev Neurosci* **30**: 289–316. **Review highlighting that neurons in the ventral tegmental area can be divided into distinct subpopulations that participate in different circuits mediating different behaviors, and the importance of determining their neurotransmitter content, eg by making use of cytochemical markers such as tyrosine hydroxylase, and their projection targets for interpreting in vivo single unit recording studies.**
- 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.
- 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, Seeberger LC, O'Reilly RC (2004). By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* **306**: 1940–1943.
- Gallistel CR, Stellar JR, Bubis E (1974). Parametric analysis of brain stimulation reward in the rat: I. The transient process and the memory-containing process. *J Comp Physiol Psychol* **87**: 848–859. **Early report of the dual reinforcing and activational aspects of (putatively dopaminergic) brain stimulation reward.**
- Garris P, Christensen J, Rebec G, Wightman R (1997). Real-time measurement of electrically evoked extracellular dopamine in the striatum of freely moving rats. *J Neurochem* **68**: 152–161.
- Geisler S, Wise RA (2008). Functional implications of glutamatergic projections to the ventral tegmental area. *Rev Neurosci* **19**: 227–244.
- Geisler S, Derst C, Veh RW, Zahm DS (2007). Glutamatergic afferents of the ventral tegmental area in the rat. *J Neurosci* **27**: 5730–5743.
- Gervais J, Rouillard C (2000). Dorsal raphe stimulation differentially modulates dopaminergic neurons in the ventral tegmental area and substantia nigra. *Synapse* **35**: 281–291.
- 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.
- 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.
- Harrison A, Everitt BL, Robbins TW (1997a). Central 5-HT depletion enhances impulsive responding without affecting the accuracy of attentional performance: interactions with dopaminergic mechanisms. *Psychopharmacology* **133**: 329–342.
- Harrison A, Everitt BL, Robbins TW (1997b). Double dissociable effects of median- and dorsal-raphe lesions on the performance of the five-choice serial reaction time test of attention in rats. *Behav Brain Res* **89**: 135–149.
- 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 E, Petrovic J, Keithley R, Wightmann R (2009). Voltammetric detection of 5-hydroxytryptamine release in the rat brain. *Anal Chem* **81**: 9462–9471.
- Hikosaka O (2010). The habenula: from stress evasion to value-based decision-making. *Nat Rev Neurosci* **11**: 503–513.

- Hikosaka O, Sesack SR, Lecourtier L, PD S (2008). Habenula: crossroad between the basal ganglia and the limbic system. *J Neurosci* **28**: 11825–11829.
- Hill R (1970). Facilitation of conditioned reinforcement as a mechanism of psychomotor stimulation. In: Costa E, Garratini S (eds). *Amphetamines and Related Compounds*. Raven Press: New York. pp 781–795.
- Hollander E, Rosen J (2000). Impulsivity. *J Psychopharmacol* **14**: S39–S44.
- Hollerman JR, Schultz W (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat Neurosci* **1**: 304–309.
- Homberg JR, Pattij T, Janssen MC, Ronken E, De Boer SF, Schoffemeer AN et al (2007). Serotonin transporter deficiency in rats improves inhibitory control but not behavioural flexibility. *Eur J Neurosci* **26**: 2066–2073.
- Houk J, Adams J, Barto A (1995). A model of how the basal ganglia generate and predict reinforcement. In: Houk J, Davis J, Beiser D (eds). *Models Processing in the Basal Ganglia*. MIT Press: Cambridge, MA. pp 249–270.
- 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.
- Jacobs BL, Fornal CA (1993). 5-HT and motor control: a hypothesis. *Trends Neurosci* **16**: 346–352.
- Kahneman D, Tversky A (1979). Prospect theory: an analysis of decision under risk. *Econometrica* **47**: 263–292.
- Kehagia AA, Murray GK, Robbins TW (2010). Learning and cognitive flexibility: frontostriatal function and monoaminergic modulation. *Curr Opin Neurobiol* **20**: 199–204.
- Kranz GS, Kasper S, Lanzenberger R (2010). Reward and the serotonergic system. *Neuroscience* **166**: 1023–1035.
- Kuczenski R, Segal DS, Leith NJ, Applegate CD (1987). Effects of amphetamine, methylphenidate, and apomorphine on regional brain serotonin and 5-hydroxyindole acetic acid. *Psychopharmacology (Berl)* **93**: 329–335.
- Kuhnen C, Chiao J (2009). Genetic determinants of financial risk taking. *PLoS One* **4**: e4362.
- Lattal K (1987). Considerations in the experimental analysis of reinforcement delay. In: Commons M (ed). *Quantitative Analyses of Behavior. V. The Effect of Delay and of Intervening Events on Reinforcement Value*. Lawrence Erlbaum: Hillsdale, NJ. pp 107–123.
- LeMarquand D, Benkelfat C, Pihl R, Palmour R, Young S (1999). Behavioral disinhibition induced by tryptophan depletion in nonalcoholic young men with multigenerational family histories of paternal alcoholism. *Am J Psychiatr* **156**: 1771–1779.
- Leyton M, Okazawa H, Diksic M, Paris J, Rosa P, Mzengeza S et al (2001). Brain Regional alpha-[11C]methyl-L-tryptophan trapping in impulsive subjects with borderline personality disorder. *Am J Psychiatry* **158**: 775–782.
- Linnoila M, Virkkunen M, Scheinin M, Nuutila A, Rimon R, Goodwin F (1983). Low cerebrospinal fluid 5-hydroxyindoleacetic acid concentration differentiates impulsive from nonimpulsive violent behavior. *Life Sci* **33**: 2609–2614.
- Linnoila V, Virkkunen M (1992). Aggression, suicidality, and serotonin. *J Clin Psychiatry* **53**: 46–51.
- Ljungberg T, Apicella P, Schultz W (1991). Responses of monkey midbrain dopamine neurons during delayed alternation performance. *Brain Res* **567**: 337–341.
- Logue AW, Tobin H, Chelonis JJ, Wang RY, Geary N, Schachter S (1992). Cocaine decreases self-control in rats: a preliminary report. *Psychopharmacology (Berl)* **109**: 245–247.
- Long AB, Kuhn CM, Platt ML (2009). Serotonin shapes risky decision making in monkeys. *Soc Cogn Affect Neurosci* **4**: 346–356.
- Lorens SA (1978). Some behavioral effects of serotonin depletion depend on method: a comparison of 5,7-dihydroxytryptamine, p-chlorophenylalanine, p-chloroamphetamine, and electrolytic raphe lesions. *Ann NY Acad Sci* **305**: 532–555.
- Lyon M, Robbins TW (1975). The action of central nervous system stimulant drugs: a general theory concerning amphetamine effects. In: Essman W (ed). *Current Developments in Psychopharmacology*. Spectrum: New York. pp 79–163.
- Maia TV (2010). Two-factor theory, the actor-critic model, and conditioned avoidance. *Learn Behav* **38**: 50–67.
- Matsumoto M, Hikosaka O (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* **459**: 837–841.
- McClure S, Berns G, Montague P (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron* **38**: 339–346.
- Milner P (1977). Theories of drive, reinforcement and motivation. In: Iversen L, Iversen S, Snyder S (eds). *Handbook of Psychopharmacology*. Plenum Press: New York, pp 181–200.
- Mink JW (1996). The basal ganglia: focused selection and inhibition of competing motor programs. *Prog Neurobiol* **50**: 381–425.
- Mobini S, Chiang T-J, Ho M-Y, Bradshaw CM, Szabadi E (2000). Effects of central 5-hydroxytryptamine depletion on sensitivity to delayed and probabilistic reinforcement. *Psychopharmacology* **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 P, Dayan P, Sejnowski T (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. **Paper describing the insight that the phasic firing of dopamine neurons quantitatively resembles a 'reward prediction error' signal used in computational algorithms for reinforcement learning.**
- Morris G, Nevet A, Arkadir D, Vaadia E, Bergman H (2006). Midbrain dopamine neurons encode decisions for future action. *Nat Neurosci* **9**: 1057–1063.
- Murphy FC, Smith K, Cowen P, Robbins TW, Sahakian BJ (2002). The effects of tryptophan depletion on cognitive and affective processing in healthy volunteers. *Psychopharmacology* **163**: 42–53.
- Murphy SE, Longhitano C, Ayres RE, Cowen PJ, Harmer CJ, Rogers RD (2009). The role of serotonin in nonnormative risky choice: the effects of tryptophan supplements on the 'reflection effect' in healthy adult volunteers. *J Cogn Neurosci* **21**: 1709–1719.
- Nakamura K, Matsumoto M, Hikosaka O (2008). Reward-dependent modulation of neuronal activity in the primate dorsal raphe nucleus. *J Neurosci* **28**: 5331–5343.
- Nauta HJ (1979). A proposed conceptual reorganization of the basal ganglia and telencephalon. *Neuroscience* **4**: 1875–1881.
- Nauta WJ (1982). Limbic innervation of the striatum. *Adv Neurol* **35**: 41–47.
- Niv Y, Daw ND, Joel D, Dayan P (2007). Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl)* **191**: 507–520. **Computational theory of dopamine, describing the concept of an opportunity cost of time, offering a common explanation for two seemingly distinct (affective and activational) properties of dopamine.**
- O'Doherty J, Dayan P, Friston K, Critchley H, Dolan R (2003). Temporal difference models and reward-related learning in the human brain. *Neuron* **38**: 329–337.
- O'Reilly RC, Frank MJ (2006). Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. *Neural Comput* **18**: 283–328.
- Pattij T, Vanderschuren LJ (2008). The neuropharmacology of impulsive behaviour. *Trends Pharmacol Sci* **29**: 192–199.
- 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.
- Phillips P, Stuber G, Heien M, Wightman R, Carelli R (2003). Subsecond dopamine release promotes cocaine seeking. *Nature* **422**: 614–618.
- Post T, Van den Assem M, Baltussen G, Thaler R (2008). Deal or no deal? Decision making under risk in a large-payoff game show. *Am Econ Rev* **98**: 38–71.
- Poulos CX, Parker JL, Le AD (1996). Dextfenfluramine and 8-OH-DPAT modulate impulsivity in a delay-of-reward paradigm: implications for a correspondence with alcohol consumption. *Behav Pharmacol* **7**: 395–399.
- Puumala T, Sirvio J (1998). Changes in activities of dopamine and serotonin systems in the frontal cortex underlie poor choice accuracy and impulsivity of rats in an attention task. *Neuroscience* **83**: 489–499.
- 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.
- Robbins T (1976). Relationship between reward-enhancing and stereotypical effects of psychomotor stimulant drugs. *Nature* **264**: 57–59.
- Robbins TW (2007). Shifting and stopping: fronto-striatal substrates, neurochemical modulation and clinical implications. *Philos Trans R Soc Lond B Biol Sci* **362**: 917–932.
- Robbins TW, Sahakian BJ (1979). Paradoxical effects of psychomotor stimulant drugs in hyperactive children from the standpoint of behavioural pharmacology. *Neuropharmacology* **18**: 931–950.
- Robbins TW, Everitt BJ (1982). Functional studies of the central catecholamines. *Int Rev Neurobiol* **23**: 303–365.
- Robbins TW, Everitt BJ (1992). Functions of dopamine in the dorsal and ventral striatum. *Semin Neurosci* **4**: 119–127.
- Robbins TW, Everitt BJ (2007). A role for mesencephalic dopamine in activation: commentary on Berridge (2006). *Psychopharmacology (Berl)* **191**: 433–437. **Comment on Berridge (2007) highlighting a more general activational account stressing both a performance-based energetic component to dopamine as well as reinforcement-related functions more akin to those posited in the computational reinforcement learning models.**
- Robinson OJ, Sahakian BJ (2008). Recurrence in major depressive disorder: a neurocognitive perspective. *Psychol Med* **38**: 315–318.
- Robinson OJ, Sahakian BJ (2009). A double dissociation in the roles of serotonin and mood in healthy subjects. *Biol Psychiatry* **65**: 89–92.

- Rogers RD, Blackshaw AJ, Middleton HC, Matthews K, Hawtin K, Crowley C *et al* (1999). Tryptophan depletion impairs stimulus-reward learning while methylphenidate disrupts attentional control in healthy young adults: implications for the monoaminergic basis of impulsive behaviour. *Psychopharmacology* **146**: 482–491.
- Roitman MF, Wheeler RA, Wightman RM, Carelli RM (2008). Real-time chemical responses in the nucleus accumbens differentiate rewarding and aversive stimuli. *Nat Neurosci* **11**: 1376–1377.
- Rubia K, Lee F, Cleare AJ, Tunstall N, Fu CH, Brammer M *et al* (2005). Tryptophan depletion reduces right inferior prefrontal activation during response inhibition in fast, event-related fMRI. *Psychopharmacology (Berl)* **179**: 791–803.
- Ruhe HG, Mason NS, Schene AH (2007). Mood is indirectly related to serotonin, norepinephrine and dopamine levels in humans: a meta-analysis of monoamine depletion studies. *Mol Psychiatry* **12**: 331–359.
- Salamone JD, Correa M, Farrar A, Mingote SM (2007). Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology (Berl)* **191**: 461–482.
- Samejima K, Ueda Y, Doya K, Kimura M (2005). Representation of action-specific reward values in the striatum. *Science* **310**: 1337–1340.
- Schultz W (2007). Multiple dopamine functions at different time courses. *Annu Rev Neurosci* **30**: 259–288.
- Schultz W, Dayan P, Montague PR (1997). A neural substrate of prediction and reward. *Science* **275**: 1593–1599.
- 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.
- Smith KA, Fairburn CG, Cowen PJ (1997). Relapse of depression after rapid depletion of tryptophan. *Lancet* **349**: 915–919.
- Somers LA, Beyene M, Carelli RM, Wightman RM (2009). Synaptic overflow of dopamine in the nucleus accumbens arises from neuronal activity in the ventral tegmental area. *J Neurosci* **29**: 1735–1742.
- Soubrié P (1986). Reconciling the role of central serotonin neurons in human and animal behavior. *Behav Brain Sci* **9**: 319–362. **Extensive review summarizing animal research linking serotonin to behavioral inhibition rather than anxiety per sé.**
- Staley JK, Malison RT, Innis RB (1998). Imaging of the serotonergic system: interactions of neuroanatomical and functional abnormalities of depression. *Biol Psychiatry* **44**: 534–549.
- Sutton R, Barto A (1990). Time-derivative models of Pavlovian reinforcement. In: Gabriel M, More J (eds). *Learning and Computational Neuroscience: Foundations of Adaptive Networks*. MIT Press: Cambridge, MA, pp 497–537.
- Sutton R, Barto A (1998). *Reinforcement learning*. MIT Press: Cambridge, MA.
- 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.
- Tanaka S, 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.
- Tanaka SC, Shishida K, Schweighofer N, Okamoto Y, Yamawaki S, Doya K (2009). Serotonin affects association of aversive outcomes to past actions. *J Neurosci* **29**: 15669–15674.
- Todd M, Niv Y, Cohen J (2008). Learning to use working memory in partially observable environments through dopaminergic reinforcement. In: Daphne K, Yoshua B, Dale S, Leon Bottou, Aron C (eds) *Twenty-First Annual Conference on Neural Information Processing Systems (NIPS) 2008*. Vancouver: Canada.
- Tops M, Russo S, Boksem MA, Tucker DM (2009). Serotonin: modulator of a drive to withdraw. *Brain Cogn* **71**: 427–436.
- Trent F, Tepper JM (1991). Dorsal raphe stimulation modifies striatal-evoked antidromic invasion of nigral dopaminergic neurons *in vivo*. *Exp Brain Res* **84**: 620–630.
- Tsai CT (1989). Involvement of serotonin in mediation of inhibition of substantia nigra neurons by noxious stimuli. *Brain Res Bull* **23**: 121–127.
- Tsai HC, 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.
- Ungerstedt U (1971). Stereotaxic mapping of the monoamine pathways in the rat brain. *Acta Physiol Scand (Suppl)* **367**: 1–48.
- Ungless M, Magill P, Bolam J (2004). Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. *Science* **303**: 2040–2042.
- Van Gaalen M, van Koten R, Schoffeleer A, Vanderschuren L (2006). Critical involvement of dopaminergic neurotransmission in impulsive decision making. *Biol Psychiatr* **60**: 66–73.
- Wade TR, de Wit H, Richards JB (2000). Effects of dopaminergic drugs on delayed reward as a measure of impulsive behavior in rats. *Psychopharmacology (Berl)* **150**: 90–101.
- Waelti P, Dickinson A, Schultz W (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature* **412**: 43–48.
- Walker SC, Robbins TW, Roberts AC (2009). Response disengagement on a spatial self-ordered sequencing task: effects of regionally selective excitotoxic lesions and serotonin depletion within the prefrontal cortex. *J Neurosci* **29**: 6033–6041.
- Walker SC, Mikheenko Y, Argyle L, Robbins T, Roberts A (2006). Selective prefrontal serotonin depletion impairs acquisition of a detour-reaching task. *Eur J Neurosci* **23**: 3119–3123.
- Ward BO, Wilkinson LS, Robbins TW, Everitt BJ (1999). Forebrain serotonin depletion facilitates the acquisition and performance of a conditional visual discrimination task in rats. *Behav Brain Res* **100**: 51–65.
- Winstanley C, Dalley J, Theobald D, Robbins T (2003). Global 5-HT depletion attenuates the ability of amphetamine to decrease impulsive choice on a delay-discounting task in rats. *Psychopharmacology* **170**: 320–331.
- Winstanley CA, Eagle DM, Robbins TW (2006a). Behavioral models of impulsivity in relation to ADHD: translation between clinical and preclinical studies. *Clin Psychol Rev* **26**: 379–395.
- Winstanley CA, Theobald DE, Dalley JW, Cardinal RN, Robbins TW (2006b). Double dissociation between serotonergic and dopaminergic modulation of medial prefrontal and orbitofrontal cortex during a test of impulsive choice. *Cereb Cortex* **16**: 106–114.
- Wise RA (2004). Dopamine, learning and motivation. *Nat Rev Neurosci* **5**: 483–494.
- Wogar M, Bradshaw C, Szabadi E (1993). Does the effect of central 5-hydroxytryptamine depletion on timing depend on motivational change? *Psychopharmacology* **112**: 86–92.
- Zaghloul KA, Blanco JA, Weidemann CT, McGill K, Jaggi JL, Baltuch GH *et al* (2009). Human substantia nigra neurons encode unexpected financial rewards. *Science* **323**: 1496–1499.