Neuron
Review

Twenty-Five Lessons from Computational Neuromodulation

Peter Dayan[1,*]
[1]Gatsby Computational Neuroscience Unit, 17 Queen Square, London WC1N 3AR, UK
*Correspondence: dayan@gatsby.ucl.ac.uk
http://dx.doi.org/10.1016/j.neuron.2012.09.027

Neural processing faces three rather different, and perniciously tied, communication problems. First, computation is radically distributed, yet point-to-point interconnections are limited. Second, the bulk of these connections are semantically uniform, lacking differentiation at their targets that could tag particular sorts of information. Third, the brain's structure is relatively fixed, and yet different sorts of input, forms of processing, and rules for determining the output are appropriate under different, and possibly rapidly changing, conditions. Neuromodulators address these problems by their multifarious and broad distribution, by enjoying specialized receptor types in partially specific anatomical arrangements, and by their ability to mold the activity and sensitivity of neurons and the strength and plasticity of their synapses. Here, I offer a computationally focused review of algorithmic and implementational motifs associated with neuromodulators, using decision making in the face of uncertainty as a running example.

Introduction

Who talks to whom, what they are allowed to say, and how the answers to these questions can change, are central to the design and operation of distributed computing systems. Brains adopt distributed computation to a prodigious degree and thus face critical issues with each of them. The problem with "who talks to whom" is that some sorts of information need to be broadcast rather widely, since they can affect many aspects of ongoing computations. However, the number of synapses made is severely limited compared with the number of possible targets. Unlike networks such as the internet, there is of course no opportunity for packets of information to be routed indirectly.

The problem with "what they are allowed to say" is that the preponderant forms of synaptic communication are severely restricted. For instance, short of architectural specializations or complex neural activity codes, postsynaptic cells cannot distinguish separate sorts of presynaptic activation or inhibition, even though different sorts of information need to have radically different effects. Equivalently, different inputs lack intrinsic tags to their sources. This is particularly important for signals that are broadcast in order to address the problems of distribution. Of course, there are many architectural specializations but this does not preclude other, more direct, solutions.

The issue raised by the question of "how the answers… can change," is that anatomy is relatively stable, and yet different conditions can require dynamics or information integration that may need to change in characteristic ways to short order. For instance, the collective behavior of neurons comprising central pattern generators involved in creating rhythmic motion needs to alter in the light of different environmental challenges; equally, the strengths of different sources of data bearing on a sensory processing problem should optimally adjust with the relative reliabilities of those data. How can structurally fixed networks be endowed with the substantial degree of context dependence that seems to be required?

The organization and effects of neuromodulators, at least under a suitably catholic construal (including monoamines, acetylcholine, peptides, steroids, hormones, gases such as nitric oxide, and even conventional neurotransmitters such as glutamate in some of their modes of operation), appear to offer solutions to all these concerns.

Neuromodulators can be broadly distributed via the bloodstream, via volume transmission and diffusion from widespread release sites such as synaptic varicosities (Agnati et al., 2006), and via massive axonal arborizations having huge numbers of release sites. There are also more selective indirect pathways. Furthermore, neuromodulators luxuriate in a lush variety of targets. For the issues here, key to their effects are membrane-bound receptors. Such receptors can be highly specific for different neuromodulators, providing the "tagging" discussed above. As we will see, architectural and neuromodulatory specializations are frequently integrated. These observations jointly address the questions of "who talks to whom" and "what they are allowed to say."

Second, in terms of their effects, neuromodulators can manipulate neural processing over short and long timescales in many ways. The medium of modulation includes directly hyperpolarizing or depolarizing neurons, changing their responsivity to input, altering the strengths of synapses, and shaping the plasticity of those synapses. When integrated across a network of neurons, this can lead to dramatically different dynamics and input-output behavior. The influences can also interact—for instance, in Hebbian forms of long term potentiation and depression, plasticity is partly determined by activity and can be affected by neuromodulators both directly and indirectly through their effects on that activity. Neuromodulatory effects are remarkably strong—as evidenced by the actions of drugs on the global dynamics and processing of the brain. These are all ways by which neuromodulators realize context dependence and so address the issues of "how answers... can change."

Neuromodulation is a vast field to which it is impossible to do full justice in a short paper, and there are many excellent reviews of numerous of its facets. In order to scrutinize how neuromodulators solve the communication problems posed at the outset, a single class of computations associated with decision making in the face of uncertainty will be the focus. Neuromodulators are deeply and revealingly involved in decision making, albeit with many contentious issues remaining. I use decision making as a backdrop to highlight twenty-five general lessons from computational neuromodulation, as promised in the title (see Table 1).

Two main conceptual components of decision making are utility and uncertainty. Subjects should make choices by engaging in a form of planning to assess the expected long-run utility of possible actions based on a characterization of the current circumstance and then choose accordingly (note the term ''circumstance'' is used to refer to the detailed aspects of the current and past sensory environment that suffice to determine as best as possible the future effects of the subject's current choice). However, uncertainty permeates both the determination of the current circumstance, for instance because of sensory noise, and the evaluation of the utility of actions, for instance because of ignorance stemming from incomplete learning. As we will see, multiple, partially independent, systems are involved in the overall processes of choice and are thus tied up with utility and uncertainty, and all the systems are influenced by neuromodulators.

Our restriction to decision making leads to a concentration on the four major ascending neuromodulators: acetylcholine (ACh), dopamine (DA), norepinephrine (NE), and serotonin (5-HT). Even just for these four, there is not the space to discuss many of their operations or to provide the mathematical details of the models that underlie the analysis (as described in detail in the cited papers). The focus will be on data from rodents and primates, although there is substantial commonality of neuromodulator effects (if not always their identities) in invertebrates (Katz, 2011). This analysis is influenced by Doya (2002) and the contributions in Doya et al. (2002). It is important to note that almost none of the computationally richer cases discussed is yet universally accepted.

Utility

Utility or affective value is a central piece of information that influences behavior. In terms of reinforcement learning (RL; Sutton and Barto, 1998), predictions about future values are made based on the current circumstance to determine choice and action; and, at least when disconfirmed, command learning. Utility should be influenced by aspects of a subject's motivational state—the prospect of food is more valuable to a hungry than a thirsty animal. When choices can (perhaps also) avoid punishments, it is net utility that counts—it may not be worth stopping to collect either outcome in the face of mortal threat. Utility also plays roles other than determining the suitability of discrete choices. For instance, one can argue (Niv et al., 2007) that the average rate of (positive) utility quantifies the effective cost of the passage of time, in that the larger the expected rate, the more costly it is to deny oneself that much utility through failing to act for a given length of time. This can energize behavior (Guitart-Masip et al., 2011).

Table 1. Twenty-Five Lessons from Computational Neuromodulation, in Two Broad Categories

Organization	
#	Content
A	Neuromodulatory systems can report selective information.
B	This report can be over a very quick timescale.
C	This report can also be over multiple timescales, particularly tonic and phasic.
D	Via different affinities and time courses, different receptor types can respond selectively to separate characteristics of the signal.
E	Different receptor types can be localized on anatomically different pathways.
G	Autoreceptors play an important regulatory role.
H	Interactions among different neuromodulators are very widespread.
I	Of these interactions, opponency is especially prevalent.
J	Opponency is rarely simple or symmetric; but rather competition and cooperation are mixed and effects are frequently asymmetric.
K	There is a complex tapestry of structural heterogeneity within each system.
P	There are structured loops involving sensory and frontal cortical areas and neuromodulatory nuclei.
Q	There are limits to the structural and functional specificity of the neuromodulators.
R	Partially independent of the spiking of neuromodulatory neurons, there is local, presumably glutamatergic, control over release.
X	Neuromodulatory neurons can corelease other neurotransmitters, notably glutamate.
Y	Neuromodulators are vasoactive, affecting the interpretation of results from fMRI.
Effects	
#	Content
F	Neuromodulatory signals can be multiplexed—the same information turned to different uses.
L	Neuromodulators play a key role in regulating internally-directed computations such as gated working memory.
M	Neuromodulators can influence the course of activity by regulating which of a number of gross pathways determine the activity of neurons.
N	Neuromodulators regulate the nature and structure of oscillations.
O	Neuromodulators affect plasticity over many time scales.
S	Neuromodulators are involved in the regulation of energy utilization in the brain and body.
T	Many neuromodulators exhibit an inverted U-shaped (or Yerkes-Dodson) curve relating concentration or release to effect.
U	It takes significant time for changes in neuromodulatory activity to be reported to target sites, potentially limiting their effects.
V	Failures of neuromodulatory systems are tied to debilitating neurological and psychiatric diseases; they are also major therapeutic targets.
W	Individual differences in neuromodulatory receptors or transporters have observable effects on decision-making behavior.

Utility is of fundamental importance to a wide swath of information processing and raises all three communication problems discussed above: very many areas of the brain need to know about the utility of outcomes and predicted outcomes; the utilities are affected by a variety of different factors for which segregated informational channels would be ideal; and utilities can have immediate effects on structurally fixed networks. We will see that the involvement of neuromodulation in computations to do with utility illuminates all these issues and also highlights a number of other general properties.

One important complexity about utility is the parallel involvement of two different instrumental systems and also Pavlovian influences. These systems are subject to neuromodulation in partially different ways, and so are discussed individually below. The goal-directed, or model-based, instrumental system (Dickinson and Balleine, 2002), which involves frontal regions and the dorsomedial striatum (Balleine, 2005; Valentin et al., 2007), is believed to construct a model of the task and to use that model prospectively to predict outcomes consequent on choices (Tolman, 1948). One central mark of goal-directed control is its sensitivity to motivational state—predicted outcomes are evaluated under current (or possibly predicted; Raby et al., 2007) motivational states. The second instrumental control system is habitual, or model free (Dickinson and Balleine, 2002), and is more closely associated with a different set of regions that includes the dorsolateral striatum (Balleine, 2005; Tricomi et al., 2009). This learns what to do from direct experience of past actions and reward and so plans retrospectively (Thorndike, 1911). That planning is retrospective implies that it is the motivational state that pertained during learning that is important, and so model-free actions may be inappropriate for the current motivational state.

Finally, for instrumental systems, choices are ultimately contingent on the delivery of suitable outcomes. Conversely, under Pavlovian control, elicitation of preparatory and consummatory actions associated with predictions of, or the actual presence of, biologically significant reinforcers, appears to be automatic. Evidence for this is that the actions are still elicited even if they have deleterious consequences in terms of actually getting or preventing good or bad outcomes (Williams and Williams, 1969; Hershberger, 1986; Dayan et al., 2006). One interpretation is that Pavlovian actions are the result of evolutionary preprogramming, providing heuristic choices that are typically, though not always, appropriate. The predictions underlying Pavlovian control may be made in model-based or model-free ways.

Appetitive and aversive utilities act in rather distinct ways, a fact that is better understood for model-free control. Thus, reward and punishment are considered separately in the latter.

Reward in Model-Free Instrumental and Pavlovian Control

Dopamine is a key ascending neuromodulator. There is ample evidence that the phasic activity of DA neurons and the phasic release of DA in macaques (Bayer and Glimcher, 2005; Schultz et al., 1997; Morris et al., 2006; Satoh et al., 2003; Nakahara et al., 2004), rodents (Hyland et al., 2002; Roesch et al., 2007; Garris et al., 1999; Gan et al., 2010), and even humans (Zaghloul et al., 2009; Kishida et al., 2011) report a particular form of so-called temporal difference prediction error (Sutton, 1988) for long run future reward (Montague et al., 1996; Schultz et al., 1997; Barto, 1995). Note that "reward" here is defined as the sort of appetitive reinforcement that is objectively realized in terms of causing actions leading to it to be repeated (Thorndike, 1911) (i.e., "wanting," as distinct from "liking" [Berridge, 2004], which is more opioid than dopaminergically sensitive [Peciña et al., 2006]). The prediction error arises whenever there is an unexpected change in future reward, both positively (when either a reward arrives that was not expected or a stimulus arrives that was itself not expected but that predicts a future reward) and negatively (e.g., when an expected reward is withheld). The predictions are based on all aspects of the circumstances of the subject at the time they are made, but pertain to sequences of future reward. Usually, distal rewards are discounted, or downweighted in importance, compared with proximal ones.

At least three roles have been postulated for this dopaminergically encoded prediction error. First, it should inspire learning to make accurate predictions based on the current circumstance and, depending on the precise interpretation, learning to choose actions in that circumstance that lead to greater reward (Sutton and Barto, 1998) or to avoid actions that lead to smaller reward. Many regions of the brain are involved in making predictions; and indeed DA can influence synaptic plasticity in various ways (see Tritsch and Sabatini, 2012, this issue of *Neuron*).

The striatum is a particularly important target for dopaminergic neuromodulation. One major anatomical feature of this structure is the existence of separated direct and indirect pathways, defined by their output targets. Neurons in the direct or "go" pathway are influenced largely by D1 dopamine receptors and are involved in the initiation and inspiration of action. D1 receptors have been suggested as being sensitive to phasic increases in the concentration of dopamine consequence on bursts and so boosting the future propensity to perform actions found to have surprisingly good outcomes (Frank, 2005; Frank et al., 2004; Frank and O'Reilly, 2006; Cohen and Frank, 2009; Kravitz et al., 2012).

Conversely, neurons in the indirect or "no-go" pathway are subject to D2 dopamine receptors and influence the inhibition of action (Gerfen et al., 1990; Smith et al., 1998). Dopamine normally suppresses the indirect pathway via D2 receptors; D2 receptors are more sensitive to dopamine than D1 receptors and so are more greatly affected by dips below baseline caused when reward are worse than expected. Activity-controlled plasticity would thus lead to a more intense or likely rejection of the disadvantageous action (Frank, 2005; Frank et al., 2004; Frank and O'Reilly, 2006; Cohen and Frank, 2009; Kravitz et al., 2012).

Temporal difference learning has the effect of transferring phasic activity from the time of occurrence of an unexpected reward to the time of occurrence of the earliest reliable predictor of that reward, without changing its magnitude. Thus, the long run average rate of the prediction error (which would be reflected in more tonic concentrations of dopamine) is just the long run average reward rate, which we argued above acts as an opportunity cost for the passage of time and determines measures of the vigor of responding (Niv et al., 2007). A role for dopamine in vigor is consistent with the effect of dopaminergic lesions on

effort costs (Salamone et al., 2009), the willingness of patients with Parkinson's disease (characterized by the loss of dopamine cells) to engage in effortful actions (Mazzoni et al., 2007), and even the way that dopamine levels in various parts of the striatum track changes in vigor induced by satiety (Ostlund et al., 2011). It is known, though, that the phasic and tonic activity of dopamine cells are at least partly separable (Grace, 1991; Goto and Grace, 2005), suggesting greater complexities in the relationship.

The third role for the phasic dopaminergic prediction error signal that arises when a predictor of future reward is presented is to liberate (or perhaps invigorate) Pavlovian responses associated with the prospect of reward (Panksepp, 1998; Ikemoto and Panksepp, 1999). Such predictors lead to Pavlovian boosting of instrumental responses (Satoh et al., 2003; Estes, 1943; Dickinson and Balleine, 2002; Nakamura and Hikosaka, 2006; Talmi et al., 2008), a process believed to involve the action of dopamine in the nucleus accumbens (Murschall and Hauber, 2006), potentially via D1 receptors (Frank, 2005; Surmeier et al., 2007, 2010). The phasic dopamine signal consequent on predictive cues provides a formal underpinning for the theory of incentive salience (McClure et al., 2003; Berridge and Robinson, 1998), which is concerned with motivational influences over the attention garnered by such stimuli.

A first group of the twenty-five general lessons about neuromodulation emerges from this focus on dopamine (Table 1, A–Y). Perhaps the most important are that (A) neuromodulatory neurons can report very selective information (i.e., reward prediction errors for dopamine) on a (B) very quick timescale. To put it another way, there is no reason why anatomical breadth should automatically be coupled with either semantic or temporal breadth. Nevertheless (C), neuromodulators can also signal over more than one timescale, with at least partially separable tonic and phasic activity, and different receptor types may be sensitive to the different timescales; additionally (D) by having different affinities (as do D1 and D2 receptors), different types can respond selectively to separate characteristics of the signal (Frank, 2005). Along with their different properties (E), different receptor types can be localized on different pathways, and these pathways are also potentially subject to modulation from a variety of other systems, such as the local, tonically active interneurons in the striatum that release ACh (Aosaki et al., 1994; Aosaki et al., 1995; Kaneko et al., 2000; Higley et al., 2009).

In addition (F), observe the multiplexing inherent in having a neuromodulator report a signal (here a reward prediction error) that has a variety of important, but distinct, functions, we will see some further putative functions of this phasic dopamine signal below. This can make interpretation very complicated—particularly for experiments which manipulate dopamine or its receptors systemically.

It is also known (G) that a key role is played by autoreceptors that are typically inhibitory to the release of the neuromodulator concerned, e.g., dopamine receptors on dopamine neurons and their terminals. An obvious role for these is feedback control. However, this can pose a problem for interpretation—the semantics of vigorous activity of dopamine in terms of a prediction error would become contingent on the nature of the current set point; it is also a confound for pharmacological investigation and treatment. Autoinhibition is a way for tonic signaling to set a baseline for phasic signaling, an issue whose computational implications have been explored a little both for reward (Daw et al., 2002; Boureau and Dayan, 2011) and, as we discuss later, uncertainty (Aston-Jones and Cohen, 2005). There are other forms of short term plasticity in the release of dopamine in response to bursts, including facilitation as well as inhibition (Montague et al., 2004).

The dopaminergic prediction error is generally considered to be part of habitual and model-free Pavlovian systems, involved in retrospective control. One might think that dopamine signaling would therefore be insensitive to motivational state. However, there are various ways in which sensitivity can be imported. First, if the information about state forms part of the representation of the stimulus, so state can be treated also as circumstance, then regular learning that maps circumstances to predictions will endow dopamine activity with state dependence. Second, dopamine neurons themselves have receptors for neuromodulators such as orexins (or hypocretins) (Siegel, 2004; Aston-Jones et al., 2010). This would allow their activity to be directly sensitive to motivational state. Indeed (H), interneuromodulatory interactions, such as the influence of one set of neuromodulators on others are very widespread (Briand et al., 2007). Third, structures that drive dopamine activity might themselves be directly sensitive to motivational state—for instance, it has been suggested that the amygdala's sensitivity to the neuromodulator oxytocin will change its responding in the face of social threats or opportunities (De Dreu, 2012), and this could affect dopamine responding.

Punishment in Model-free Instrumental and Pavlovian Control

Avoiding or minimizing punishment and threats is also of critical importance, and the same considerations as for appetitive outcomes might lead one to expect neuromodulators to play a central role in learning in aversive conditions. One important complexity is that animals have a very extensive repertoire of species-specific defensive consummatory behaviors appropriate to the nature and imminence of frank threats, at least partly realized in the sophisticated structure of areas such as the periacqueductal gray (Bolles, 1970; McNaughton and Corr, 2004; Keay and Bandler, 2001). This makes it hard to understand the interplay between such inbuilt responses, Pavlovian preparatory responses such as behavioral inhibition that are tied via prediction (whose neuromodulatory basis is debated; McNally et al., 2011) to initially neutral stimuli, and fully-fledged instrumental responses in the light of aversion.

One long-standing and critical division is between passive and active avoidance (Konorski, 1967). Although exact definitions differ, passive avoidance involves not doing actions that lead to punishment, whereas active avoidance requires emitting specific responses to avoid deleterious outcomes. The abstinence in passive avoidance can be specific to particular, problematical, choices, or it can be general, as in behavioral inhibition or certain forms of freezing. Conversely active avoidance involves the emission of specific responses to obviate potential punishment. A key idea here is so-called two-factor learning (Mowrer, 1947) and safety signaling. This involves learning that circumstances which could be associated with punishment

have low values, and that the change in circumstance associated with removing the threat is appetitive. It can therefore reinforce the action concerned, just as in the previous section.

To the extent that unexpected punishments are coded in the inhibition of phasic dopamine responses below baseline (Ungless et al., 2004; Cohen et al., 2012), just like non-delivery of expected reward (Schultz et al., 1997), the indirect pathway through the striatum which is tonically inhibited by dopamine via D2 receptors is well-placed to realize specific passive avoidance (Frank et al., 2004). Indeed, selectively activating neurons in just this pathway has recently been shown to lead to place and action avoidance in spatial and operant paradigms (Kravitz et al., 2012), exactly opposite to the effect of activating neurons in the direct pathway.

However, suppression of phasic dopamine activity is not the whole story for passive avoidance, since serotonergic neuromodulation has also been implicated in behavioral inhibition (Gray and McNaughton, 2003; Crockett et al., 2009, 2012), including in the face of punishment. Apparently more problematic is the fact that dopamine neurons have been reported to be phasically excited by punishments (Mirenowicz and Schultz, 1996; Bromberg-Martin et al., 2010), albeit with evidence of the particular involvement of those neurons whose cell bodies lie in just one part of the dopamine system, the so-called mesocortical region of the ventral tegmental area (VTA) which projects to (pre)frontal cortex (Brischoux et al., 2009; Lammel et al., 2008, 2011). Consideration of safety signaling in active avoidance suggests a rationale for this activation, as arising from the temporal difference prediction error signal that we discussed in the context of reward. This error signal is based on differences in the predicted values of successive circumstances (Johnson et al., 2001; Moutoussis et al., 2008; Maia, 2010). Thus, dopamine would be phasically activated by the incompletely expected attainment of safety, or the prediction of the prospects of this. This would enable it to control learning of the appropriate avoidance response, for instance in the direct pathway of the striatum.

One might expect such predictions about future safety to be sensitive to the controllability of the punishment. Unfortunately, dopamine release (as measured by microdialysis) provides a somewhat mixed picture. It is known that the release of dopamine to aversive outcomes does not always persist in the face of uncontrollable (and thus unavoidable) contingencies (Cabib and Puglisi-Allegra, 2012; Cabib and Puglisi-Allegra, 1994; Puglisi-Allegra et al., 1991); although this may differ in different target regions (Bland et al., 2003a, 2003b), in particular with reports of inescapable, but not escapable, shock increasing dopamine levels in mPFC, at least during the provision of the punishment (Bland et al., 2003a). Further, whereas rats from a strain favoring active coping strategies show an increase in dopamine in medial prefrontal cortex (mPFC) in the face of stress, rats from a different strain that engages in more reactive or passive strategies, do not (Giorgi et al., 2003). These finer grain details at least militate against the suggestion based on the activation of dopamine in both appetitive and aversive circumstances that it codes primarily for salience (Redgrave et al., 1999; Horvitz, 2000), although it has been suggested that this is true for some selected groups of dopamine neurons (Lammel et al., 2011).

The next question becomes the mechanism for learning the prediction about the possible future punishment that is ultimately responsible for the safety signal. There is evidence that this does not depend only on dopamine—for instance blocking D2 dopamine receptors leaves learning about aversive contingencies intact, only impairing the learning of the avoidance actions (Beninger, 1983). One possibility is that one part of the system of serotonergic neuromodulation plays the role of an opponent to dopamine, being associated with aversive rather than appetitive outcomes (Deakin, 1983). This claim is subject to a range of complexities and contention (discussed at length in Cools et al., 2011; Boureau and Dayan, 2011).

In the scheme of safety signaling, the idea would be that predictions of future aversion (“unwanting” rather than subjective “disliking”; Berridge, 2004) associated with environmental circumstances would arise through, and have temporal difference prediction errors represented by, the activity of selected 5-HT neurons (Schweimer and Ungless, 2010). Then, as circumstances change when actions stave off the prospect of punishment, this would lead to an appetitive temporal difference prediction error (reported by the phasic activity of dopamine neurons) that would reinforce the avoidance action (Johnson et al., 2001; Moutoussis et al., 2008; Maia, 2010). Similarly, the tonic activity of dopamine neurons would include the average achievement of safety along with the average delivery of reward, and thus be able to inspire suitably vigorous avoidance actions (Dayan, 2012b).

Equally, the behavioral inhibition mentioned above as the Pavlovian response to predictions of punishment would be mediated by serotonin, which has indeed been implicated in this function (Gray and McNaughton, 2003; Crockett et al., 2009, 2012). This would complement the role of dips below baseline in the activity of dopamine neurons that we also described previously. Serotonin plays a rich role in various forms of inhibition, not only for punishments as mentioned above, but also being involved when animals have to wait for a period before being allowed to act to get a reward (Fletcher, 1995; Miyazaki et al., 2011, 2012). This suggests that the interactions among multiple timescales that we noted above for the dopamine system will be even richer for serotonin; but there is unfortunately as yet rather little evidence. The serotonin system is notably more diverse than the dopamine system, with a particularly large set of receptors with different properties, and only one part may be involved in aversion.

According to this opponency view, low levels of 5-HT are associated with impulsivity because of serotonin’s association with inhibiting behavior. We should note an alternative idea about serotonin’s role that starts from impulsivity, suggesting that this comes from a decrease in the importance of distant affective outcomes compared with proximal ones, i.e., a change in a discount rate (Doya, 2000). If 5-HT is responsible for setting this rate, then impulsivity would indeed arise from low levels of this neuromodulator, with subjects being tempted by small immediate reward, ignoring large punishments (or delays) that might subsequently ensue (Cardinal, 2006; Schweighofer et al., 2008; Mobini et al., 2000). Although it is not a ubiquitous behavioral finding, neural signals associated with discounted values are indeed affected by 5-HT levels (Tanaka et al., 2007).

These accounts remain rather speculative; however, they again teach some general lessons about neuromodulation. First (I), forms of opponency between different neuromodulators are a common motif, both in the central nervous system and indeed in the periphery. However (J), this opponency is rarely simple or symmetric: for instance, although it appears as if the dominant influence of 5-HT on behaviors associated with dopamine in practice is inhibitory (Alex and Pehek, 2007; Hervé et al., 1979; Fletcher et al., 1999; Grottick et al., 2000), there are many types of serotonin receptor that have an excitatory net effect on dopamine (Alex and Pehek, 2007; Boureau and Dayan, 2011). In fact, an excitatory effect would actually be appropriate in some circumstances if the account about safety signaling is correct, as dopamine should respond to the prospect of future safety engendered by the serotonergic report of possible aversion. Distinctions such as this may provide a route for helping understand part of the multiplicity of serotonin receptors (Cooper et al., 2002; Hoyer et al., 2002). As mentioned, whether the safety is achievable depends on the degree of controllability of the environment (Maier and Watkins, 2005; Huys and Dayan, 2009); how controllability is represented is not clear. In terms of the asymmetry, dopamine appears not to exert nearly such strong effects on 5-HT as vice-versa. Finally (K), a complex tapestry of heterogeneity is revealed, particularly within the serotonin system. We have also noted substructure in the dopamine system such as the mesocortical dopamine neurons that are excited rather than inhibited by punishment (Brischoux et al., 2009; Lammel et al., 2011).

Goal-Directed Control, Working Memory, and Prefrontal Cortex

Neuromodulatory representations of utility appear to play a central role in habitual control, not the least by controlling learning directly. Since goal-directed control is based more on predictions of specific outcomes, one might expect different neuromodulatory issues to arise. Indeed, there is direct evidence that dopamine plays little role in evaluation in the goal-directed system (Dickinson et al., 2000). Nevertheless, it can still influence the vigor of the execution of the responses which it mandates (Palmiter, 2008).

We noted that goal-directed (Dickinson and Balleine, 2002; Balleine, 2005) or model-based (Daw et al., 2005; Doya, 2002) control exhibits fuller flexibility in the face of factors such as changes in motivational state. This requires that the utility of predicted outcomes can be assessed under the current motivational state. In turn, this suggests a role for direct and/or indirect neuromodulatory influences over neural structures such as gustatory insular cortex or possibly the basolateral nucleus of the amygdala involved in such evaluation (Balleine, 2005, 2011) as providing information about that state. However, although we may be able to predict the values of some outcomes under expected future motivational states, there appear to be definite limits to such predictions (Loewenstein and O'Donoghue, 2004), perhaps because of constraints on the subjunctive determination of neuromodulatory state. This would limit any such prospective somatic marker (Damasio, 1994).

For goal-directed control to make predictions about the current (or future) values of future outcomes, it would seem that these predictions must be made in the moment, or "online." One obvious way to do this is to enumerate possible future outcomes explicitly, and sum or average their motivational state-sensitive utilities. There is some more or less direct evidence for this (Fermin et al., 2010; Daw et al., 2011; Wunderlich et al., 2012a; Huys et al., 2012). However, if one views enumeration as depending on a set of internal actions that control mechanisms such as working memory (Hazy et al., 2006), one might expect them to be learned using, and influenced by, the same neuromodulatory machinery as externally directed actions (Dayan, 2012a). It has been suggested, for instance, that the Pavlovian mechanisms that lead to approach or withdrawal to external appetitive and aversive outcomes and predictors might influence the way that enumeration works. States associated with reward could be more likely to be enumerated than those with punishments, under the influence of dopamine (Smith et al., 2006) and serotonin (Dayan and Huys, 2008; Huys et al., 2012). If the process of enumeration is influenced by value, then its predictions will be biased, typically in an optimistic direction if possible aversive outcomes are suppressed but appetitive ones boosted.

Much of the mechanics of enumeration is wrapped up with the adaptive use of working memory. In fact, working memory is a much more general concern, even for habitual control. This is because the habit system takes a representation of the current circumstance and either predicts its value or that of actions that can be performed, or reports which action is preferred. In many cases, there is insufficient information in the current sensory input to determine these quantities, but if selected aspects of past input can be stored, then it will collectively suffice (Peshkin et al., 2001; Todd et al., 2009; Kaelbling et al., 1998; Nakahara et al., 2004). Control over working memory can have both instrumental and Pavlovian components. From an instrumental perspective, the basal ganglia could acquire policies that control the gating of information into working memory using reinforcement learning (O'Reilly and Frank, 2006). From a Pavlovian perspective, rather as we argued for enumeration, the phasic release of dopamine associated with a stimulus that predicts future reward or future safety, could directly influence the storage of this stimulus in working memory (Cohen and Servan-Schreiber, 1993; Durstewitz et al., 2000; O'Reilly et al., 2002), via dopamine's known effects in prefrontal cortex (Williams and Goldman-Rakic, 1995).

In total, there is an intricate set of dopaminergically influenced interactions between prefrontal regions and the striatum (Cools, 2011). It turns out that both phasic and tonic dopamine are important. For example of the latter, there is a battle for supremacy of control between goal-directed and habitual systems, and perhaps contrary to naive expectation, suppressing dopamine increases the influence of habits (de Wit et al., 2012), and boosting dopamine decreases their influence (Hitchcott et al., 2007; Wunderlich et al., 2012b); there are also powerful Pavlovian effects (Guitart-Masip et al., 2012). These might arise via dopamine's hegemony over prefrontal-striatal interactions, possibly through the medium of parts of the dopamine system that are separable from those involved in functions such as signaling reward prediction errors. It is certainly a general notion that (L) neuromodulators can play an important role in

regulating internally directed computations (Robbins and Arnsten, 2009; Cools et al., 2011), and working memory has been a particular focus for this.

Serotonin also influences the activity of prefrontal neurons in rather complicated ways (Puig and Gulledge, 2011), potentially enabling it to influence executive operations such as working memory. The relationship between this and other possible functions of 5-HT such as predictions about punishment, is not yet clear. It is known that serotonin in the orbitofrontal cortex is important for rapid adaption of behavior in paradigms in which inhibition of (possibly learned) prepotent responses is required (Roberts, 2011); and this can also be considered to be part of the regulation of internally directed computation. We discuss further aspects of this below.

Synthesis

Utility is a poster child for the way that neuromodulators solve the communication problems raised in the introduction. It also shows well the scope and force of neuromodulation, which is very deeply embedded in the very structure of decision making. It is perhaps the intricacy of the interacting systems of modulation that is most conspicuous, with many of the general lessons reflecting combinations of architectural and receptor specificity, and also the substantial interdependence among the various parts.

Uncertainty

The representation, updating and use of uncertainty, have become major foci of computational treatments of neural information processing (Dayan et al., 2000; Doya et al., 2007; Ma and Pouget, 2008; Deneve, 2008, Körding, 2007; Fiser et al., 2010), with Bayesian analyses dominating. At a coarse time scale, organisms suffer from ignorance about their environments, both because of limited opportunities to observe it, and because it changes in partly unpredictable ways. At a finer timescale, organisms have to take noisy and partial observations from multiple sensory systems to estimate their circumstance in the world. This in turn influences the evaluation (and thus the execution) of actions, as we have just discussed. All of these facets lead to uncertainty, which in turn places severe constraints on what computations are normatively appropriate.

Strict Bayesians admit no qualitative distinction between different sorts of uncertainty. However, strict Bayesian computations are usually radically intractable, and heuristics and approximations are necessary. Based on the evidence described below, it turns out to be appropriate to make the approximation of separating issues of uncertainty into learning, which takes place over a coarse timescale, and inference, which takes place over two successively finer timescales. We also distinguish between expected and unexpected uncertainty (Yu and Dayan, 2005b), with the former, often called risk in economics and neuroeconomics (Glimcher, 2010), quantifying what is known not to be known within the current conception of the organism's circumstance, and the latter capturing what lies outside these bounds—crudely, radical, unpredicted, changes indicating substantial failings in this current conception, and sharing some features with economics' notion of ambiguity.

The original communication issues that neuromodulators address also apply to uncertainty. For instance, it is clear that if unexpected uncertainty leads to the need for a dramatic revision of current computations, then many neural systems will need to know this fact. Equally, as we will see, expected uncertainty should control plasticity, and there are reasons to seek a tag which might label the sort of uncertainty involved. Finally, uncertainty regulates the way that different sources of information should be combined; this is a form of systemic adaptation of structurally fixed connections. There is evidence that the neuromodulators acetylcholine and norepinephrine play confined, but critical roles in both forms of uncertainty; with phasic and tonic delivery potentially distinguishing between inference and learning (Bouret and Sara, 2005; Dayan and Yu, 2006).

Uncertainty will first be considered in the context of learning, and then of inference. Most of the computational models are Bayesian, or at least approximately Bayesian, in character.

Learning

The only reason to learn is because of ignorance. In (Bayesian) statistical terms, ignorance is quantified by uncertainty, which is why uncertainty should control aspects of the nature and course of learning. Autoassociative memory provides a first example; then conditioning, which involves richer forms of expected uncertainty; and finally issues of unexpected uncertainty induced by change are discussed.

One case of the link between ignorance and learning arises in the context of auto-associative memory models of the hippocampus (Hasselmo, 2006; Hasselmo and Bower, 1993). Here, the idea is that an input should be assessed to see how familiar it is. If it is deemed novel, (i.e., the subject is suitably ignorant of it), it should be stored; if the input is familiar, then recall processes should remove noise from it and/or recall relevant context or associated information. Thus, on top of the assessment of familiarity, there are two implementational requirements for an input deemed to be novel: preventing attempts at recall from corrupting it and plasticizing appropriate synapses to store it. Within the particular connection patterns of the hippocampus, with anatomically and functionally segregated pathways from the main input structure, the entorhinal cortex, there is evidence that cholinergic neuromodulation can exert both these effects (Hasselmo, 2006). The notion is that in area CA3, synapses forming the recurrent connection from other area CA3 pyramidal cells, and the perforant path input from the entorhinal cortex have their effective strengths reduced, but are rendered more labile. The ability (M) of neuromodulators to control the course of activity by regulating which of a number of gross pathways determines the activity of neurons is a common scheme. There are also other potential neuromodulatory routes for this influence: for instance, ACh helps regulate oscillations ([N], a critical dynamical effect of neuromodulators in many circumstances) that simultaneously affect multiple sub-regions of the hippocampal formation (Buzsáki, 2002). It has been suggested that different pathways between these regions are dominant at different phases of theta (Hasselmo et al., 2002), providing a route for neuromodulatory effects. ACh is also capable of influencing shorter-term storage in working memory (Klink and Alonso, 1997; Hasselmo, 2006). The (O) effects of neuromodulators on various timescales of plasticity are among their most influential.

Another obvious issue for memory is whether or not an input actually merits long term storage. One way to assess this is to consider its affective consequences, bearing in mind that they may only be evident after some time has passed. Given the evidence adduced above, it should come as no surprise to find that dopamine is implicated in the later phases of hippocampal storage (Lisman et al., 2011), although this is a rather different function from the plasticity engendered by dopaminergically coded prediction errors that we discussed above as underpinning the learning of appetitive predictions. The extended timescale over which such assessments might be relevant could result in findings such as that patterns that are only incidentally correlated with the delivery of unexpected reward are also preferentially stored (Wittmann et al., 2005). Boosted storage can perhaps be seen as an instance of internal, cognitive, ''approach'' to a stimulus based on the reward it predicts (Adcock et al., 2006), matching the internal action of storage in working memory to the externally directed engagement actions that we mentioned above.

An informationally more complex case for neuromodulatory influences on plasticity comes in the context of animal conditioning experiments (Gallistel and Gibbon, 2000; Pearce and Hall, 1980), which have particularly centered on the model-free Pavlovian case. Psychological notions, such as that the associability of a stimulus varies with the degree of surprise with which it is endowed (Pearce and Hall, 1980), can be translated into computational terms as the relative learning rate of a stimulus being determined by its predictive uncertainty (Dayan et al., 2000), and then into neural constructs such as the operation of cholinergic neuromodulation influenced by regions in the amygdala, hippocampus, anterior cingulate cortex and beyond (Nassar et al., 2010; Behrens et al., 2007; Yu and Dayan, 2005b; Holland and Gallagher, 1999). Critical for the computational treatments is that learning depends on the product of the prediction error (putatively mediated by a dopaminergic signal, as discussed in the previous section on habitual control) and the learning rate (mediated by ACh)—so it is again an example of interneuromodulatory interactions. How this works biophysically is not completely clear. Similarly, model-based predictions and plans are dependent on learning about the structure of the environment in terms of transitions between circumstances and outcome contingencies. These should also be regulated by predictive uncertainty.

Unlike the unfamiliarity of a whole input, uncertainties about the relationship between conditioned and unconditioned stimuli or indeed between circumstances and outcomes, are not simple scalar quantities. They are computationally complex constructs that depend on rich aspects of present and past circumstances and the way that these are expected to change over time (Dayan et al., 2000; Behrens et al., 2007; Nassar et al., 2010). Learning can be characterized in Bayesian terms using exact or approximate forms of a Kalman filter. In particular, subjects can be differentially uncertain about different parts of the relationship, and this poses a key algorithmic problem for the representation and manipulation of uncertainty.

Although (P) there is structure in the loops connecting cholinergic nuclei to sensory processing and prefrontal cortices (Zaborszky, 2002), as indeed with other loops between prefrontal regions and neuromodulatory nuclei (Aston-Jones and Cohen, 2005; Robbins and Arnsten, 2009), there is only rather little work (Yu and Dayan, 2005a) as to how the relatively general forms of uncertainty that could be represented even by a wired neuromodulatory system might interact with the much more specific uncertainty that could be captured in, say, a cortical population code (Zemel et al., 1998; Ma et al., 2006). Certainly (Q), limits to the structural and functional specificity of neuromodulators must be acknowledged, given the relative paucity of neurons concerned, although it is worth noting that ACh and 5-HT appear to be rather more heterogeneous than DA and NE. There may be many distinct cholinergic systems, including the one mentioned above involving tonically active neurons in the striatum, which might set the stage for plasticity (Aosaki et al., 1994, 1995). There is (R) evidence for local, presumably glutamatergic, control of the release of neuromodulators in the cortex, independent of the spiking activity of the neuromodulatory neurons themselves (Marrocco et al., 1987), which could allow for more specificity in their local effects, but the computational implications of this in practice are not clear.

These cases of learning fit comfortably into a scheme of expected uncertainty (Yu and Dayan, 2005b), in that the unfamiliarity and associability are assessed within a given framework or, to adopt a term from the cognitive control literature, task set (Koechlin and Summerfield, 2007). As we mentioned, in some cases, the whole framework itself may be found to be inadequate, implying that a new one needs to be inferred (Collins and Koechlin, 2012). Such dramatic changes to the environment are considered to be forms of unexpected uncertainty, measured for instance by forms of model mismatch. They pose a critical requirement (and opportunity) for acquiring new information (Yu and Dayan, 2005b), and thus for exploration (Aston-Jones and Cohen, 2005). They may also be times of significant threat.

When a whole framework proves inadequate, a very wide set of neural systems might need to be adjusted, and so a neuromodulatory report of the inadequacy seems ideal. Indeed, there is evidence that tonic activity or levels of norepinephrine do indeed increase with unpredictable reversals in a simple reaction time task (Aston-Jones et al., 1991), and that boosting NE can speed the course of reversal learning (Devauges and Sara, 1990). Reversals, which are a popular way of inducing change, are normally signaled when actions or choices that used to be rewarded become unproductive or less productive; and actions that were formerly punished or nugatory become worthwhile. Thus, given their putative roles in providing information about, and inspiring actions associated with, reward and punishment, one might expect dopamine and serotonin to be involved directly in the assessment and realization of reversals. Rapid change is normally a feature of a model-based or goal-directed system, however, complexities associated with the competition between Pavlovian and instrumental control could ensue—the tendency of the original affective values of the stimuli to cause the cognitive equivalents of approach and withdrawal, would make it hard for these stimuli to be rejected and embraced as appropriate to their new values. Indeed, along with norepinephrine, the projections of serotonin and dopamine to the striatum and prefrontal regions have been implicated in forms of behavioral flexibility such as reversal learning and set shifting (Homberg,

2012; Robbins and Arnsten, 2009; Kehagia et al., 2010; Clark et al., 2004; Cools, 2011), with depletion or destruction leading to detriments in performance. However, there are interesting subtleties in this involvement—for instance reversal learning for reward in marmosets is impaired by either dopamine depletion in the caudate region of the striatum, or serotonin depletion in the orbitofrontal cortex, but not vice-versa (Clarke et al., 2011).

Ignorance about the framework provides an opportunity if there are rewards that could be exploited given suitable learning. However, it may also pose an escapable threat, if dangers that can be avoided could lurk. In both cases, ignorance is associated with expected uncertainty and hints that one might expect profound links between NE and ACh, as another example of interneuromodulatory interaction and also partial opponency (Yu and Dayan, 2005b). For the case of opportunities, exploration is mandated by the (initially unexpected) potential gain, and this may be treated as a form of appetitive prediction error known as an exploration bonus. One, presumably model-free, realization of such a bonus is phasic dopaminergic activity (Kakade and Dayan, 2002). Strictly speaking, the potential gain arises as a result of the expected uncertainty that follows from the unexpected change; how dopamine is coupled to ACh and/or NE in expressing this is not yet clear. The mechanism by which exploration bonuses arise in model-based calculations is also unknown.

In terms of potential threats, norepinephrine has long been linked with anxiety (Bremner et al., 1996a, 1996b). Environments associated with excessive unexpected uncertainty are highly stressful, since they lack stable relationships and pose substantial potential problems for safe exploitation. NE helps organize a massive response to stress, notably in conjunction with cortisol, a steroid hormone that acts as another neuromodulator (Wolf, 2008). This involves everything from changing energy storage and usage, via glucocorticoids (Nieuwenhuizen and Rutters, 2008) (involvement [S] with energy regulation is itself a more general principle of neuromodulation; Ellison, 1979; Tops et al., 2009; Montague, 2006), to changing the actual style of information processing. For instance, goal-directed or model-based calculations, which are typically slow, could be suppressed in favor of habitual or model-free ones, which are typically faster, though possibly less accurate, especially in the face of the quick changes associated with unexpected uncertainty. It has been suggested that suppression arises via functional inhibition wrought by two particular classes of NE receptor in the prefrontal cortex (α_1 and $\widehat{\beta}$) whose affinities make them sensitive to high levels of NE; Arnsten, 2011). This combines two previous general principles—selective affinities of different receptors, and neuromodulatory manipulation of gross pathways.

Inference

Information about the circumstance an agent occupies in its environment has to be combined from multiple sources of noisy and partial information and integrated over time as it progressively arises. The same turns out to be true for information stored in working memory, since neural activity has to be communicated to relevant targets progressively, through activity. It also arises for reading information out of synapses, for which presynaptic activity is necessary to extract their values, for instance using generic, background, activity (Mongillo et al., 2008). These processes can all fruitfully be seen as involving statistical inference, based on partial and noisy information, and so are all controlled or influenced by uncertainty (Fiser et al., 2010; Ma and Pouget, 2008; Gold and Shadlen, 2002; Ratcliff and Smith, 2004).

In particular, again from a Bayesian viewpoint, uncertainty determines just how modalities with low signal to noise ratios should be downweighted against those that are more useful. Uncertainty also determines how new pieces of information should be combined with data from the recent past, depending on factors such as the rate of change in the environment. This amounts to a form of selective attention. As for the case of exploration bonuses in learning, the impact of uncertainty should be governed by the utility associated with what can be discovered; and indeed important links have been found between reward and at least some forms of sensory attention (Gottlieb and Balan, 2010). We will consider two different timescales of the inferential effects of uncertainty, one acting across the length of the many trials that define a single task set; the other acting within the typically second or subsecond duration of each single trial as circumstances change.

Just as for conditioning, one might expect that much of the inferential uncertainty should be highly specific to the circumstances of the task, and so outside the realm of relatively coarse neuromodulatory systems. However, as also for conditioning, there is evidence for the involvement of both ACh and NE in controlling critical aspects of inference, at both the timescales mentioned above. Rather as we saw for the case of learning, a key phenomenon at the coarser time-scale appears to be controlling the strength of stimulus-bound information (relayed in this case by thalamocortical pathways), relative to that of what one might think of as prior- or model-bound information associated with the current task set (Hasselmo, 2006; Yu and Dayan, 2005b; Hasselmo and Sarter, 2011).

Take the paradigm known as the endogenous cue version of Posner's attentional task (Posner et al., 1978). In this, subjects have to respond according to a visual stimulus presented on one side of a display. Prior to the stimulus, a cue is presented at the center of the display indicating on which side the stimulus might appear. The cue can be valid (i.e., pointing to the correct side) or invalid. The percentage of trials on which the cue is valid is called its validity. Subjects pay attention to the cue in a manner that appears to be graded by its validity—the amount by which they are faster and more accurate on validly than invalidly cued trials scales with the cue's validity. In our terms, the validity of the cue determines its statistical quality. Subjects correctly set their inferential strategy to reflect this quality, and this underpins the effect of validity on behavior. There is evidence in rodents (Phillips et al., 2000) and humans (Bentley et al., 2004; Thiel and Fink, 2008) that ACh mediates this effect; a potential substrate is the combined nicotinic and muscarinic mechanism mentioned above by which thalamocortical pathways are boosted and intracortical pathways suppressed by ACh (Gil et al., 1997; Kimura et al., 1999), although the muscarinic effect may be dominant in humans (Thiel and Fink, 2008). Cholinergic influence over the interactions between bottom-up and top-down processing are also evident from the effects of

iontophoresing ACh or the muscarinic antagonist scolopamine on boosting or suppressing attentional effects on firing rates of neurons in area V1 of macaques while they perform a visually demanding task (Herrero et al., 2008). Also, stimulating the basal forebrain (where one population of ACh neurons lives) reduces the correlation between visual neurons reporting on natural scenes via a muscarinic mechanism (Goard and Dan, 2009).

Looking over a range of shorter timescales, cholinergic neuromodulation has also been implicated in aiding signal detection in rodents in tasks soliciting forms of sustained attention (McGaughy and Sarter, 1995; Parikh et al., 2007). For instance, Parikh et al. (2007) used amperometry to measure changes in the concentration of ACh in medial prefrontal (mPFC) cortex over various timescales in a Pavlovian task. Here, a cue was provided on each trial, predicting a reward after a delay of around 2 s or 6 s; the mark of attentional engagement was a cue-evoked shift in behavior, which then led to hastened reward acquisition. Cue detection in the task was impaired by removing cholinergic inputs from the mPFC, suggesting that performance was sensitive to ACh. For normal animals, ACh was substantially released over a short timescale on trials on which animals successfully detected the cue (but not when they failed); successful detection was associated with a decreasing rather than an increasing trend in ACh over the 20 s preceding the cue; and higher tonic levels of ACh concentration (measured over minutes) were tied to larger phasic ACh signals associated with the cue, and faster (Pavlovian) actions.

The various interactions with the medium term (20 s) and longer term (minutes) averages of the ACh concentration remind us of complexities surrounding a commonly reported finding for neuromodulators, (T) namely a inverted U-shaped curve of efficacy (Yerkes and Dodson, 1908). An example finding is that drugs that boost a neuromodulator such as dopamine have a beneficial effect for subjects whose baseline levels are low, but a harmful effect for subjects for whom these levels are high (Kimberg et al., 1997; Cools et al., 2011; Floresco and Magyar, 2006). Alternatively, increasing the tonic activity of a neuromodulator might have the same dual effects, as suggested for norepinephrine (Aston-Jones and Cohen, 2005; Berridge, 2008; Arnsten, 2011). It is certainly the case that if a neuromodulator represents a quantity such as expected uncertainty, then, for any environment, there will be an optimal level (together with optimal temporal modulation around this level) that leads to the most effective inference and learning: levels that are larger or smaller than this would produce computational inefficiencies. Subjects whose baseline levels are closer to this optimum value, perhaps because of their genetic endowments, past experience, or the interaction between the two, would perform best; subjects with too little or too much will be affected in obvious ways by boosting or suppressing the signal.

At the fine, subsecond, timescale of presentation of the cues, the phasic release of ACh is related to the expectation of a change in circumstance associated with the upcoming reward (which is when the phasic signal peaks; Parikh et al., 2007). This would arise as the subject's expectation about the possible change in circumstance rises following detection of the cue. Along these lines, (Sarter et al., 2009) measured ACh transients in a more complex task in which subjects had to detect and report a short signal whose delivery was designed to be highly unpredictable, or else report that the signal was not present. Given a cholinergic lesion, subjects were again more likely to miss the signal. In this task, significant ACh release in the mPFC on a trial only occurred if the subjects had both detected a signal on that trial and reported a non-signal on the previous trial. If one thinks of the signal circumstance in one trial as establishing a task set that lasts across subsequent trials (with a much shorter inter-trial interval than Parikh et al., 2007), there would therefore be little expected uncertainty when (detected) signal follows (detected) signal, and so ACh release would not be expected (Sarter et al., 2009).

This ACh transient can be seen as a phasic version of the expected uncertainty tonic signal suggested for ACh in the context of learning. Conversely, the phasic version of the NE signal would be to mark an unexpected defeat of the current circumstance. The inferential implication of such unexpected uncertainty or model failure is that existing inferences are made unsound, so, for instance, any ongoing integration of sensory information over time should be cancelled and reset, and that the subject should enjoy new, expected, uncertainty about its circumstance. The phasic activity of norepinephrine neurons in the locus coeruleus during signal processing tasks (Aston-Jones et al., 1994, 1997; Clayton et al., 2004; Rajkowski et al., 2004; Bouret and Sara, 2004), primarily in monkeys, has been interpreted as being consistent with this notion (Yu and Dayan, 2005b; Bouret and Sara, 2005). Along the same lines, NE plays a role in temporal alerting, for instance in the Posner paradigm when information is provided about when the target arrives rather than which side it arrives upon (Witte and Marrocco, 1997), and in another task called the stop-signal reaction time task (Bari et al., 2011) that is a popular way of assessing temporal aspects of the defeat of ongoing expectations.

The tasks used to examine phasic ACh (Parikh et al., 2007) and NE (Aston-Jones et al., 1994; Clayton et al., 2004) have some key points of similarity—notably relatively long and unpredictable delays before important events occur. One difference is that the tasks involving NE typically have rare targets (perhaps boosting unexpectedness), whereas those involving ACh have common targets. It would be interesting to record phasic NE and ACh signals simultaneously (perhaps indirectly in human subjects via pupil dilation; Gilzenrat et al., 2010)—one might expect that NE would be released to the cue, as a temporal alert, but that it is the phasic rise in ACh that prepares the ground for the (now expected) reward to be delivered.

Particularly for the case of DA (Servan-Schreiber et al., 1990) and NE (Brown et al., 2005), there has been work on how an effect of these neuromodulators on the input-output gain of neurons might influence overall network dynamics that implement inferences such as decision making. One of the simplest decision making networks involves effective mutual inhibition between two competing groups of neurons (Usher and McClelland, 2001), with action initiation occurring when the activity of one group reaches a threshold (Bogacz et al., 2006; Gold and Shadlen, 2002; Lo and Wang, 2006). Boosting the gain of the neurons in such a network can make it unstable and therefore allow whichever of the two groups currently has the greater activity to reach the threshold promptly, with barely any further

integration. This therefore controls a speed-accuracy tradeoff. Brown et al. (2005) considered the problem of decision making architectures in which one network determines the release of NE, which then modulates another network that is more directly responsible for initiating the decision. They pointed out what is a general issue for phasic activity (U), namely that the time it takes for the neuromodulator to be delivered to its site of action (norepinephrine fibers are not myelinated) appears to be at the margins of the period in which there is a chance to have a suitable effect on the on-going computation.

Synthesis

Unlike utility, which seems a natural candidate for neuromodulatory realizations, uncertainty does not, because of the exquisite selectivity that subjects should exhibit in their sensitivity to uncertainty. Nevertheless, substantial evidence suggests the involvement particularly of acetylcholine and norepinephrine in representing and acting on uncertainty, and we have also seen that there are rich links between these neuromodulators and also with dopamine. Many of the general lessons that we learnt for utility have been reiterated, and some new ones learned, particularly concerning the overall architecture of influences.

Discussion

This review has considered general properties of neuromodulators through the lens of effects on decision making. The latter is a critical competence, and we have seen the rich involvement of very many aspects of neuromodulation. Concomitantly, (V) problems or manipulations of neuromodulatory systems are tied to debilitating neurological and psychiatric diseases, such as addiction and Parkinson's disease, and they are also major therapeutic targets, as in schizophrenia, depression, Alzheimer's disease, and beyond (computational issues are discussed in Maia and Frank, 2011; Huys et al., 2011; Montague et al., 2012). Further (W), individual (e.g., genetic) differences in factors such as the properties of particular receptor types, or the efficacy of transporters controlling the longevity of neuromodulators following release, have been associated with differences in decision making behavior, such as the propensity to explore or to learn from positive or negative feedback (Frank et al., 2007, 2009).

We have seen many instances of the three communications problems reviewed in the introduction. However, these problems are rather generic, whereas the twenty five lessons discussed throughout the review have shown some of the peculiarities of the ways that neuromodulators help solve them. In Table 1, they are grouped into two broad categories, addressing issues of how neuromodulatory systems are organized and the consequences they have for information processing. For the first, we have seen common motifs such as heterogeneity in space (i.e., different receptor types with different affinities, some localized on different systems) and heterogeneity in time (with phasic and various scales of tonic release). There is a number of forms of control, including self-regulation by autoreceptors, complex forms of interneuromodulator interaction, and even the possibility of local glutamatergic control over release. Other, systemic, control mechanisms also exist, such as loops between prefrontal areas and neuromodulatory nuclei which exert mutual influence upon each other. These, and indeed other functions of the neuromodulators, may be complicated (X) by corelease of other neurotransmitters and other neuromodulators through the same axons (Stuber et al., 2010; Lavin et al., 2005).

Given the focus on decision making, the key neuromodulators were dopamine, serotonin, acetylcholine, and norepinephrine, which represent information about reward, punishments, and expected and unexpected uncertainty. However, these categories are, of course, crude, contentious, and incomplete. Even in the context of our discussion, issues such as the propensity of phasic dopamine activity to report a temporally sophisticated prediction error associated with the delivery of future reward (Sutton, 1988; Barto, 1995; Montague et al., 1996) illustrates some of the complexities: this signal resembles a prediction under certain circumstances rather than just a simple error; further, given a safety signaling interpretation of avoidance learning, it also represents predictions of the attainment of safety, which is not a conventional reward; further, the cumulative prediction error signal can report a measure of the long-run rate of reward, which is a signal with its own computational significance. Indeed, we have in general suggested ways of linking phasic and tonic interpretations of neuromodulatory activity, although they may act as partly independent information channels.

The second category of lessons in Table 1 concerns the effects of neuromodulators on neural processing. The two most important systemic effects are controlling plasticity (perhaps via controlling activity, under a Hebbian view) and controlling whole pathways, such as dopamine's influence over direct and indirect pathways through the striatum or over gated working memory, and acetylcholine's influence on thalamocortical versus intracortical interactions. In conjunction with suitable heterogeneity, manipulating pathways as a whole is perhaps of particular importance as a mechanism, influencing both external actions such as Pavlovian behaviors and instrumental vigor, but also internal actions, controlling the deployment of working memory or the expansion of a tree of possible future circumstances and actions that are being evaluated. There are also dynamical effects, such as changing the gain of competitive, decision making circuits, along with a substantial impact on central pattern generators that is best understood in invertebrate preparations (Harris-Warrick, 2011; Marder and Thirumalai, 2002).

For the future, one of the most immediately pressing issues concerns resolving the historical problems in recording from neuromodulatory neurons, measuring their local concentrations at target zones, and selectively manipulating their activity or that of particular receptor types. For instance, nuclei such as the ventral tegmental area or the dorsal raphe, which contain dopamine and serotonin neurons, also contain other neuron classes, and extracellular measures of facets such as spike shape are imperfect discriminators (Ungless et al., 2004). Many of these issues are on the cusp of being comprehensively addressed in animal studies through the use of new tools, including new and improved recording methods, molecular biology, and optogenetics. For instance, genetically encoded channelrhodopsin can be used to provide a functional tag for extracellular recordings (Cohen et al., 2012). Unfortunately, these advances have yet to provide help for work on humans. Although the new vogue for

psychosurgery is providing opportunities for recording (Zaghloul et al., 2009) and cyclic voltammetry (Kishida et al., 2011), the most important workhorse is functional magnetic resonance imaging (fMRI), perhaps combined with pharmacology (Honey and Bullmore, 2004). However, not only do we know very little about the coupling between activity and the blood oxygenation level-dependent (BOLD) signal that is measured in fMRI in areas such as the striatum that are the main targets of key neuromodulators, but also (Y) these neuromodulators might be able to affect local blood flow directly themselves (Peppiatt et al., 2006), further muddying the interpretation.

From a computational perspective, there is much work to do to understand the overall network and systems effects of the changes that we know different neuromodulators lead to in individual elements in those circuits. This may also help us understand aspects of various sorts of heterogeneity—e.g., what is achieved by the subtle differences within families of receptors, and also the rich intertwining of the neuromodulators. It may even help us unravel issues to do with pharmacological manipulation of the neuromodulators—for instance, helping explain the well-known fact that selective serotonin reuptake inhibitors have a rapid effect on serotonin transport but take weeks to have a stable effect on mood (Blier, 2003), perhaps partly because of effects on autoreceptors and negative feedback control mechanisms, and partly because any quick effect on (aversive) emotional processing has to be embedded through learning to affect dispositions (Harmer et al., 2009).

However, the most compelling computational issue is the one that has appeared in various places in this review, namely the relationship between specificity and generality and cortical versus neuromodulatory contributions to representation and processing. For utility, this issue centers on the interactions between model-free and model-based systems, with the former being substantially based on neuromodulators such as dopamine and serotonin, whereas the latter depends on cortical processing (albeit itself subject to modulation associated with specific stimulus values). For uncertainty, the question is how representations of uncertainty associated with cortical population codes, with their exquisite stimulus discrimination, interact with those associated with neuromodulators, with their apparent coarseness.

In sum, I have discussed how neuromodulators solve key problems associated with having a structurally languorous but massively distributed information processing system such as a brain. Neuromodulators both broadcast and narrowcast key information about the current character of the organism and its environment, and exert dramatic effects on processing by changing the dynamical properties of neurons, and the strengths and adaptability of selected of their synapses in both selected and dissipated targets.

ACKNOWLEDGMENTS

I am very grateful to my many current and former collaborators in computational neuromodulation, notably Read Montague, Terry Sejnowski, Wolfram Schultz, Nathaniel Daw, Sham Kakade, Angela Yu, Yael Niv, Quentin Huys, Y-Lan Boureau, Ray Dolan, John O'Doherty, Ben Seymour, Debbie Talmi, Marc Guitart-Masip, Andrea Chiba, Chris Córdova, Alex Thiele, Jon Roiser, Diego Pizzgalli, Peter Shizgal, Daniel Salzman, Thomas Akam, and Mark Walton. I also thank Kenji Doya, Martin Sarter, Cindy Lustig, and William Howe for sharing unpublished data and thoughts. I particularly thank Y-Lan Boureau, Nathaniel Daw, Agnieszka Grabska-Barwinska, Marc Guitart-Masip, Quentin Huys, Zeb Kurth-Nelson, Yael Niv, Martin Sarter, Angela Yu, and two anonymous reviewers for very helpful comments on an earlier draft, and Meredith LeMasurier for philanthropic editing. Funding was from the Gatsby Charitable Foundation.

REFERENCES

Adcock, R.A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B., and Gabrieli, J.D.E. (2006). Reward-motivated learning: mesolimbic activation precedes memory formation. Neuron *50*, 507–517.

Agnati, L.F., Leo, G., Zanardi, A., Genedani, S., Rivera, A., Fuxe, K., and Guidolin, D. (2006). Volume transmission and wiring transmission from cellular to molecular networks: history and perspectives. Acta Physiol. (Oxf.) *187*, 329–344.

Alex, K.D., and Pehek, E.A. (2007). Pharmacologic mechanisms of serotonergic regulation of dopamine neurotransmission. Pharmacol. Ther. *113*, 296–320.

Aosaki, T., Tsubokawa, H., Ishida, A., Watanabe, K., Graybiel, A.M., and Kimura, M. (1994). Responses of tonically active neurons in the primate's striatum undergo systematic changes during behavioral sensorimotor conditioning. J. Neurosci. *14*, 3969–3984.

Aosaki, T., Kimura, M., and Graybiel, A.M. (1995). Temporal and spatial characteristics of tonically active neurons of the primate's striatum. J. Neurophysiol. *73*, 1234–1252.

Arnsten, A.F.T. (2011). Catecholamine influences on dorsolateral prefrontal cortical networks. Biol. Psychiatry *69*, e89–e99.

Aston-Jones, G., and Cohen, J.D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. Annu. Rev. Neurosci. *28*, 403–450.

Aston-Jones, G., Chiang, C., and Alexinsky, T. (1991). Discharge of noradrenergic locus coeruleus neurons in behaving rats and monkeys suggests a role in vigilance. Prog. Brain Res. *88*, 501–520.

Aston-Jones, G., Rajkowski, J., Kubiak, P., and Alexinsky, T. (1994). Locus coeruleus neurons in monkey are selectively activated by attended cues in a vigilance task. J. Neurosci. *14*, 4467–4480.

Aston-Jones, G., Rajkowski, J., and Kubiak, P. (1997). Conditioned responses of monkey locus coeruleus neurons anticipate acquisition of discriminative behavior in a vigilance task. Neuroscience *80*, 697–715.

Aston-Jones, G., Smith, R.J., Sartor, G.C., Moorman, D.E., Massi, L., Tahsili-Fahadan, P., and Richardson, K.A. (2010). Lateral hypothalamic orexin/hypocretin neurons: A role in reward-seeking and addiction. Brain Res. *1314*, 74–90.

Balleine, B.W. (2005). Neural bases of food-seeking: affect, arousal and reward in corticostriatolimbic circuits. Physiol. Behav. *86*, 717–730.

Balleine, B. (2011). Sensation, incentive learning, and the motivational control of goal-directed action. In Neurobiology of Sensation and Reward, Chapter 13, J. Gottfried, ed. (Boca Raton, FL: CRC Press).

Bari, A., Mar, A.C., Theobald, D.E., Elands, S.A., Oganya, K.C.N.A., Eagle, D.M., and Robbins, T.W. (2011). Prefrontal and monoaminergic contributions to stop-signal task performance in rats. J. Neurosci. *31*, 9254–9263.

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

Bayer, H.M., and Glimcher, P.W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron *47*, 129–141.

Behrens, T.E.J., Woolrich, M.W., Walton, M.E., and Rushworth, M.F.S. (2007). Learning the value of information in an uncertain world. Nat. Neurosci. *10*, 1214–1221.

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

Bentley, P., Husain, M., and Dolan, R.J. (2004). Effects of cholinergic enhancement on visual stimulation, spatial attention, and spatial working memory. Neuron *41*, 969–982.

Berridge, K.C. (2004). Motivation concepts in behavioral neuroscience. Physiol. Behav. *81*, 179–209.

Berridge, C.W. (2008). Noradrenergic modulation of arousal. Brain Res. Brain Res. Rev. *58*, 1–17.

Berridge, K.C., and Robinson, T.E. (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? Brain Res. Brain Res. Rev. *28*, 309–369.

Bland, S.T., Hargrave, D., Pepin, J.L., Amat, J., Watkins, L.R., and Maier, S.F. (2003a). Stressor controllability modulates stress-induced dopamine and serotonin efflux and morphine-induced serotonin efflux in the medial prefrontal cortex. Neuropsychopharmacology *28*, 1589–1596.

Bland, S.T., Twining, C., Watkins, L.R., and Maier, S.F. (2003b). Stressor controllability modulates stress-induced serotonin but not dopamine efflux in the nucleus accumbens shell. Synapse *49*, 206–208.

Blier, P. (2003). The pharmacology of putative early-onset antidepressant strategies. Eur. Neuropsychopharmacol. *13*, 57–66.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., and Cohen, J.D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. Psychol. Rev. *113*, 700–765.

Bolles, R.C. (1970). Species-specific defense reactions and avoidance learning. Psychol. Rev. *77*, 32–48.

Boureau, Y.-L., and Dayan, P. (2011). Opponency revisited: competition and cooperation between dopamine and serotonin. Neuropsychopharmacology *36*, 74–97.

Bouret, S., and Sara, S.J. (2004). Reward expectation, orientation of attention and locus coeruleus-medial frontal cortex interplay during learning. Eur. J. Neurosci. *20*, 791–802.

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

Bremner, J.D., Krystal, J.H., Southwick, S.M., and Charney, D.S. (1996a). Noradrenergic mechanisms in stress and anxiety: I. Preclinical studies. Synapse *23*, 28–38.

Bremner, J.D., Krystal, J.H., Southwick, S.M., and Charney, D.S. (1996b). Noradrenergic mechanisms in stress and anxiety: II. Clinical studies. Synapse *23*, 39–51.

Briand, L.A., Gritton, H., Howe, W.M., Young, D.A., and Sarter, M. (2007). Modulators in concert for cognition: modulator interactions in the prefrontal cortex. Prog. Neurobiol. *83*, 69–91.

Brischoux, F., Chakraborty, S., Brierley, D.I., and Ungless, M.A. (2009). Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. Proc. Natl. Acad. Sci. USA *106*, 4894–4899.

Bromberg-Martin, E.S., Matsumoto, M., and Hikosaka, O. (2010). Dopamine in motivational control: rewarding, aversive, and alerting. Neuron *68*, 815–834.

Brown, E., Gao, J., Holmes, P., Bogacz, R., Gilzenrat, M., and Cohen, J. (2005). Simple neural networks that optimize decisions. Int. J. Bifurcat. Chaos *15*, 803–826.

Buzsáki, G. (2002). Theta oscillations in the hippocampus. Neuron *33*, 325–340.

Cabib, S., and Puglisi-Allegra, S. (1994). Opposite responses of mesolimbic dopamine system to controllable and uncontrollable aversive experiences. J. Neurosci. *14*, 3333–3340.

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

Cardinal, R.N. (2006). Neural systems implicated in delayed and probabilistic reinforcement. Neural Netw. *19*, 1277–1301.

Clark, L., Cools, R., and Robbins, T.W. (2004). The neuropsychology of ventral prefrontal cortex: decision making and reversal learning. Brain Cogn. *55*, 41–53.

Clarke, H.F., Hill, G.J., Robbins, T.W., and Roberts, A.C. (2011). Dopamine, but not serotonin, regulates reversal learning in the marmoset caudate nucleus. J. Neurosci. *31*, 4290–4297.

Clayton, E.C., Rajkowski, J., Cohen, J.D., and Aston-Jones, G. (2004). Phasic activation of monkey locus ceruleus neurons by simple decisions in a forced-choice task. J. Neurosci. *24*, 9914–9920.

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

Cohen, J.D., and Servan-Schreiber, D. (1993). A theory of dopamine function and its role in cognitive deficits in schizophrenia. Schizophr. Bull. *19*, 85–104.

Cohen, J.Y., Haesler, S., Vong, L., Lowell, B.B., and Uchida, N. (2012). Neuron-type-specific signals for reward and punishment in the ventral tegmental area. Nature *482*, 85–88.

Collins, A., and Koechlin, E. (2012). Reasoning, learning, and creativity: frontal lobe function and human decision making. PLoS Biol. *10*, e1001293.

Cools, R. (2011). Dopaminergic control of the striatum for high-level cognition. Curr. Opin. Neurobiol. *21*, 402–407.

Cools, R., Nakamura, K., and Daw, N.D. (2011). Serotonin and dopamine: unifying affective, activational, and decision functions. Neuropsychopharmacology *36*, 98–113.

Cooper, J., Bloom, F., and Roth, R. (2002). The Biochemical Basis of Neuropharmacology, 8th Edition (New York: OUP).

Crockett, M.J., Clark, L., and Robbins, T.W. (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.

Crockett, M.J., Clark, L., Apergis-Schoute, A.M., Morein-Zamir, S., and Robbins, T.W. (2012). Serotonin modulates the effects of pavlovian aversive predictions on response vigor. Neuropsychopharmacology *37*, 2244–2252.

Damasio, A. (1994). Descartes' Error: Emotion, Reason, and the Human Brain (New York, NY: GP Putnam).

Daw, N.D., Kakade, S., and Dayan, P. (2002). Opponent interactions between serotonin and dopamine. Neural Netw. *15*, 603–616.

Daw, N.D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat. Neurosci. *8*, 1704–1711.

Daw, N.D., Gershman, S.J., Seymour, B., Dayan, P., and Dolan, R.J. (2011). Model-based influences on humans' choices and striatal prediction errors. Neuron *69*, 1204–1215.

Dayan, P. (2012a). How to set the switches on this thing. Curr. Opin. Neurobiol. Published online June 15, 2012. http://dx.doi.org/10.1016/j.conb.2012.05.011.

Dayan, P. (2012b). Instrumental vigour in punishment and reward. Eur. J. Neurosci. *35*, 1152–1168.

Dayan, P., and Huys, Q.J.M. (2008). Serotonin, inhibition, and negative mood. PLoS Comput. Biol. *4*, e4.

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

Dayan, P., Kakade, S., and Montague, P.R. (2000). Learning and selective attention. Nat. Neurosci. Suppl. *3*, 1218–1223.

Dayan, P., Niv, Y., Seymour, B., and Daw, N.D. (2006). The misbehavior of value and the discipline of the will. Neural Netw. *19*, 1153–1160.

De Dreu, C.K.W. (2012). Oxytocin modulates cooperation within and competition between groups: an integrative review and research agenda. Horm. Behav. *61*, 419–428.

de Wit, S., Standing, H.R., Devito, E.E., Robinson, O.J., Ridderinkhof, K.R., Robbins, T.W., and Sahakian, B.J. (2012). Reliance on habits at the expense of goal-directed control following dopamine precursor depletion. Psychopharmacology (Berl.) *219*, 621–631.

Deakin, J.F.W. (1983). Roles of brain serotonergic neurons in escape, avoidance and other behaviors. J. Psychopharmacol. *43*, 563–577.

Deneve, S. (2008). Bayesian spiking neurons II: learning. Neural Comput. *20*, 118–145.

Devauges, V., and Sara, S.J. (1990). Activation of the noradrenergic system facilitates an attentional shift in the rat. Behav. Brain Res. *39*, 19–28.

Dickinson, A., and Balleine, B. (2002). The role of learning in the operation of motivational systems. In Stevens' Handbook of Experimental Psychology, *Volume 3*, R. Gallistel, ed. (New York: Wiley), pp. 497–534.

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

Doya, K. (2000). Metalearning, neuromodulation, and emotion. Affective Minds *1*, 101–104.

Doya, K. (2002). Metalearning and neuromodulation. Neural Netw. *15*, 495–506.

Doya, K., Dayan, P., and Hasselmo, M. (2002). Special issue: computational models of neuromodulation. Neural Netw. *16*, 5–6.

Doya K., Ishii S., Pouget A., and Rao R., eds. (2007). Bayesian brain: Probabilistic approaches to neural coding (Cambridge, MA: The MIT Press).

Durstewitz, D., Seamans, J.K., and Sejnowski, T.J. (2000). Dopamine-mediated stabilization of delay-period activity in a network model of prefrontal cortex. J. Neurophysiol. *83*, 1733–1750.

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

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

Fermin, A., Yoshida, T., Ito, M., Yoshimoto, J., and Doya, K. (2010). Evidence for model-based action planning in a sequential finger movement task. J. Mot. Behav. *42*, 371–379.

Fiser, J., Berkes, P., Orbán, G., and Lengyel, M. (2010). Statistically optimal perception and learning: from behavior to neural representations. Trends Cogn. Sci. *14*, 119–130.

Fletcher, P.J. (1995). Effects of combined or separate 5,7-dihydroxytryptamine lesions of the dorsal and median raphe nuclei on responding maintained by a DRL 20s schedule of food reinforcement. Brain Res. *675*, 45–54.

Fletcher, P.J., Korth, K.M., and Chambers, J.W. (1999). Selective destruction of brain serotonin neurons by 5,7-dihydroxytryptamine increases responding for a conditioned reward. Psychopharmacology (Berl.) *147*, 291–299.

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

Frank, M.J. (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, M.J., and O'Reilly, R.C. (2006). A mechanistic account of striatal dopamine function in human cognition: psychopharmacological studies with cabergoline and haloperidol. Behav. Neurosci. *120*, 497–517.

Frank, M.J., Seeberger, L.C., and O'reilly, R.C. (2004). By carrot or by stick: cognitive reinforcement learning in parkinsonism. Science *306*, 1940–1943.

Frank, M.J., Moustafa, A.A., Haughey, H.M., Curran, T., and Hutchison, K.E. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. Proc. Natl. Acad. Sci. USA *104*, 16311–16316.

Frank, M.J., Doll, B.B., Oas-Terpstra, J., and Moreno, F. (2009). Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. Nat. Neurosci. *12*, 1062–1068.

Gallistel, C.R., and Gibbon, J. (2000). Time, rate, and conditioning. Psychol. Rev. *107*, 289–344.

Gan, J.O., Walton, M.E., and Phillips, P.E.M. (2010). Dissociable cost and benefit encoding of future rewards by mesolimbic dopamine. Nat. Neurosci. *13*, 25–27.

Garris, P.A., Kilpatrick, M., Bunin, M.A., Michael, D., Walker, Q.D., and Wightman, R.M. (1999). Dissociation of dopamine release in the nucleus accumbens from intracranial self-stimulation. Nature *398*, 67–69.

Gerfen, C.R., Engber, T.M., Mahan, L.C., Susel, Z., Chase, T.N., Monsma, F.J., Jr., and Sibley, D.R. (1990). D1 and D2 dopamine receptor-regulated gene expression of striatonigral and striatopallidal neurons. Science *250*, 1429–1432.

Gil, Z., Connors, B.W., and Amitai, Y. (1997). Differential regulation of neocortical synapses by neuromodulators and activity. Neuron *19*, 679–686.

Gilzenrat, M.S., Nieuwenhuis, S., Jepma, M., and Cohen, J.D. (2010). Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. Cogn. Affect. Behav. Neurosci. *10*, 252–269.

Giorgi, O., Lecca, D., Piras, G., Driscoll, P., and Corda, M.G. (2003). Dissociation between mesocortical dopamine release and fear-related behaviours in two psychogenetically selected lines of rats that differ in coping strategies to aversive conditions. Eur. J. Neurosci. *17*, 2716–2726.

Glimcher, P. (2010). Foundations of Neuroeconomic Analysis (New York, NY: Oxford Univ Press).

Goard, M., and Dan, Y. (2009). Basal forebrain activation enhances cortical coding of natural scenes. Nat. Neurosci. *12*, 1444–1449.

Gold, J.I., and Shadlen, M.N. (2002). Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. Neuron *36*, 299–308.

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

Gottlieb, J., and Balan, P. (2010). Attention as a decision in information space. Trends Cogn. Sci. *14*, 240–248.

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

Gray, J.A., and McNaughton, N. (2003). The Neuropsychology of Anxiety, Second Edition (New York: OUP).

Grottick, A.J., Fletcher, P.J., and Higgins, G.A. (2000). Studies to investigate the role of 5-HT(2C) receptors on cocaine- and food-maintained behavior. J. Pharmacol. Exp. Ther. *295*, 1183–1191.

Guitart-Masip, M., Beierholm, U.R., Dolan, R., Duzel, E., and Dayan, P. (2011). Vigor in the face of fluctuating rates of reward: an experimental examination. J. Cogn. Neurosci. *23*, 3933–3938.

Guitart-Masip, M., Chowdhury, R., Sharot, T., Dayan, P., Duzel, E., and Dolan, R.J. (2012). Action controls dopaminergic enhancement of reward representations. Proc. Natl. Acad. Sci. USA *109*, 7511–7516.

Harmer, C.J., Goodwin, G.M., and Cowen, P.J. (2009). Why do antidepressants take so long to work? A cognitive neuropsychological model of antidepressant drug action. Br. J. Psychiatry *195*, 102–108.

Harris-Warrick, R.M. (2011). Neuromodulation and flexibility in central pattern generator networks. Curr. Opin. Neurobiol. *21*, 685–692.

Hasselmo, M.E. (2006). The role of acetylcholine in learning and memory. Curr. Opin. Neurobiol. *16*, 710–715.

Hasselmo, M.E., and Bower, J.M. (1993). Acetylcholine and memory. Trends Neurosci. *16*, 218–222.

Hasselmo, M.E., and Sarter, M. (2011). Modes and models of forebrain cholinergic neuromodulation of cognition. Neuropsychopharmacology *36*, 52–73.

Hasselmo, M.E., Bodelón, C., and Wyble, B.P. (2002). A proposed function for hippocampal theta rhythm: separate phases of encoding and retrieval enhance reversal of prior learning. Neural Comput. *14*, 793–817.

Hazy, T.E., Frank, M.J., and O'Reilly, R.C. (2006). Banishing the homunculus: making working memory work. Neuroscience *139*, 105–118.

Herrero, J.L., Roberts, M.J., Delicato, L.S., Gieselmann, M.A., Dayan, P., and Thiele, A. (2008). Acetylcholine contributes through muscarinic receptors to attentional modulation in V1. Nature *454*, 1110–1114.

Hershberger, W.A. (1986). An approach through the looking-glass. Anim. Learn. Behav. *14*, 443–451.

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

Higley, M.J., Soler-Llavina, G.J., and Sabatini, B.L. (2009). Cholinergic modulation of multivesicular release regulates striatal synaptic potency and integration. Nat. Neurosci. *12*, 1121–1128.

Hitchcott, P.K., Quinn, J.J., and Taylor, J.R. (2007). Bidirectional modulation of goal-directed actions by prefrontal cortical dopamine. Cereb. Cortex *17*, 2820–2827.

Holland, P.C., and Gallagher, M. (1999). Amygdala circuitry in attentional and representational processes. Trends Cogn. Sci. *3*, 65–73.

Homberg, J.R. (2012). Serotonin and decision making processes. Neurosci. Biobehav. Rev. *36*, 218–236.

Honey, G., and Bullmore, E. (2004). Human pharmacological MRI. Trends Pharmacol. Sci. *25*, 366–374.

Horvitz, J.C. (2000). Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. Neuroscience *96*, 651–656.

Hoyer, D., Hannon, J.P., and Martin, G.R. (2002). Molecular, pharmacological and functional diversity of 5-HT receptors. Pharmacol. Biochem. Behav. *71*, 533–554.

Huys, Q.J.M., and Dayan, P. (2009). A Bayesian formulation of behavioral control. Cognition *113*, 314–328.

Huys, Q.J.M., Moutoussis, M., and Williams, J. (2011). Are computational models of any use to psychiatry? Neural Netw. *24*, 544–551.

Huys, Q.J.M., Eshel, N., O'Nions, E., Sheridan, L., Dayan, P., and Roiser, J.P. (2012). Bonsai trees in your head: how the pavlovian system sculpts goal-directed choices by pruning decision trees. PLoS Comput. Biol. *8*, e1002410.

Hyland, B.I., Reynolds, J.N.J., Hay, J., Perk, C.G., and Miller, R. (2002). Firing modes of midbrain dopamine cells in the freely moving rat. Neuroscience *114*, 475–492.

Ikemoto, S., and 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.

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

Kaelbling, L., Littman, M., and Cassandra, A. (1998). Planning and acting in partially observable stochastic domains. Artif. Intell. *101*, 99–134.

Kakade, S., and Dayan, P. (2002). Dopamine: generalization and bonuses. Neural Netw. *15*, 549–559.

Kaneko, S., Hikida, T., Watanabe, D., Ichinose, H., Nagatsu, T., Kreitman, R.J., Pastan, I., and Nakanishi, S. (2000). Synaptic integration mediated by striatal cholinergic interneurons in basal ganglia function. Science *289*, 633–637.

Katz, P.S. (2011). Neural mechanisms underlying the evolvability of behaviour. Philos. Trans. R. Soc. Lond. B Biol. Sci. *366*, 2086–2099.

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

Kehagia, A.A., Murray, G.K., and Robbins, T.W. (2010). Learning and cognitive flexibility: frontostriatal function and monoaminergic modulation. Curr. Opin. Neurobiol. *20*, 199–204.

Kimberg, D.Y., D'Esposito, M., and Farah, M.J. (1997). Effects of bromocriptine on human subjects depend on working memory capacity. Neuroreport *8*, 3581–3585.

Kimura, F., Fukuda, M., and Tsumoto, T. (1999). Acetylcholine suppresses the spread of excitation in the visual cortex revealed by optical recording: possible differential effect depending on the source of input. Eur. J. Neurosci. *11*, 3597–3609.

Kishida, K.T., Sandberg, S.G., Lohrenz, T., Comair, Y.G., Sáez, I., Phillips, P.E.M., and Montague, P.R. (2011). Sub-second dopamine detection in human striatum. PLoS ONE *6*, e23291.

Klink, R., and Alonso, A. (1997). Muscarinic modulation of the oscillatory and repetitive firing properties of entorhinal cortex layer II neurons. J. Neurophysiol. *77*, 1813–1828.

Koechlin, E., and Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. Trends Cogn. Sci. *11*, 229–235.

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

Körding, K. (2007). Decision theory: what "should" the nervous system do? Science *318*, 606–610.

Kravitz, A.V., Tye, L.D., and Kreitzer, A.C. (2012). Distinct roles for direct and indirect pathway striatal neurons in reinforcement. Nat. Neurosci. *15*, 816–818.

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

Lammel, S., Ion, D.I., Roeper, J., and Malenka, R.C. (2011). Projection-specific modulation of dopamine neuron synapses by aversive and rewarding stimuli. Neuron *70*, 855–862.

Lavin, A., Nogueira, L., Lapish, C.C., Wightman, R.M., Phillips, P.E.M., and Seamans, J.K. (2005). Mesocortical dopamine neurons operate in distinct temporal domains using multimodal signaling. J. Neurosci. *25*, 5013–5023.

Lisman, J., Grace, A.A., and Duzel, E. (2011). A neoHebbian framework for episodic memory; role of dopamine-dependent late LTP. Trends Neurosci. *34*, 536–547.

Lo, C.-C., and Wang, X.-J. (2006). Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. Nat. Neurosci. *9*, 956–963.

Loewenstein, G.F., and O'Donoghue, T. (2004). Animal spirits: affective and deliberative processes in economic behavior. Working Papers 04-14, Cornell University, Center for Analytic Economics.

Ma, W.J., and Pouget, A. (2008). Linking neurons to behavior in multisensory perception: a computational review. Brain Res. *1242*, 4–12.

Ma, W.J., Beck, J.M., Latham, P.E., and Pouget, A. (2006). Bayesian inference with probabilistic population codes. Nat. Neurosci. *9*, 1432–1438.

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

Maia, T.V., and Frank, M.J. (2011). From reinforcement learning models to psychiatric and neurological disorders. Nat. Neurosci. *14*, 154–162.

Maier, S.F., and Watkins, L.R. (2005). Stressor controllability and learned helplessness: the roles of the dorsal raphe nucleus, serotonin, and corticotropin-releasing factor. Neurosci. Biobehav. Rev. *29*, 829–841.

Marder, E., and Thirumalai, V. (2002). Cellular, synaptic and network effects of neuromodulation. Neural Netw. *15*, 479–493.

Marrocco, R.T., Lane, R.F., McClurkin, J.W., Blaha, C.D., and Alkire, M.F. (1987). Release of cortical catecholamines by visual stimulation requires

activity in thalamocortical afferents of monkey and cat. J. Neurosci. *7*, 2756–2767.

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

McClure, S.M., Daw, N.D., and Montague, P.R. (2003). A computational substrate for incentive salience. Trends Neurosci. *26*, 423–428.

McGaughy, J., and Sarter, M. (1995). Behavioral vigilance in rats: task validation and effects of age, amphetamine, and benzodiazepine receptor ligands. Psychopharmacology (Berl.) *117*, 340–357.

McNally, G.P., Johansen, J.P., and Blair, H.T. (2011). Placing prediction into the fear circuit. Trends Neurosci. *34*, 283–292.

McNaughton, N., and Corr, P.J. (2004). A two-dimensional neuropsychology of defense: fear/anxiety and defensive distance. Neurosci. Biobehav. Rev. *28*, 285–305.

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

Miyazaki, K., Miyazaki, K.W., and Doya, K. (2011). Activation of dorsal raphe serotonin neurons underlies waiting for delayed rewards. J. Neurosci. *31*, 469–479.

Miyazaki, K., Miyazaki, K.W., and Doya, K. (2012). The role of serotonin in the regulation of patience and impulsivity. Mol. Neurobiol. *45*, 213–224.

Mobini, S., Chiang, T.J., Al-Ruwaitea, A.S., Ho, M.Y., Bradshaw, C.M., and Szabadi, E. (2000). Effect of central 5-hydroxytryptamine depletion on intertemporal choice: a quantitative analysis. Psychopharmacology (Berl.) *149*, 313–318.

Mongillo, G., Barak, O., and Tsodyks, M. (2008). Synaptic theory of working memory. Science *319*, 1543–1546.

Montague, R. (2006). Why Choose This Book?: How We Make Decisions (New York: EP Dutton).

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

Montague, P.R., Hyman, S.E., and Cohen, J.D. (2004). Computational roles for dopamine in behavioural control. Nature *431*, 760–767.

Montague, P.R., Dolan, R.J., Friston, K.J., and Dayan, P. (2012). Computational psychiatry. Trends Cogn. Sci. *16*, 72–80.

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

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

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

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

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

Nakamura, K., and Hikosaka, O. (2006). Role of dopamine in the primate caudate nucleus in reward modulation of saccades. J. Neurosci. *26*, 5360–5369.

Nassar, M.R., Wilson, R.C., Heasly, B., and Gold, J.I. (2010). An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. J. Neurosci. *30*, 12366–12378.

Nieuwenhuizen, A.G., and Rutters, F. (2008). The hypothalamic-pituitary-adrenal-axis in the regulation of energy balance. Physiol. Behav. *94*, 169–177.

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

O'Reilly, R.C., and Frank, M.J. (2006). Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. Neural Comput. *18*, 283–328.

O'Reilly, R.C., Noelle, D.C., Braver, T.S., and Cohen, J.D. (2002). Prefrontal cortex and dynamic categorization tasks: representational organization and neuromodulatory control. Cereb. Cortex *12*, 246–257.

Ostlund, S.B., Wassum, K.M., Murphy, N.P., Balleine, B.W., and Maidment, N.T. (2011). Extracellular dopamine levels in striatal subregions track shifts in motivation and response cost during instrumental conditioning. J. Neurosci. *31*, 200–207.

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

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

Parikh, V., Kozak, R., Martinez, V., and Sarter, M. (2007). Prefrontal acetylcholine release controls cue detection on multiple timescales. Neuron *56*, 141–154.

Pearce, J.M., and Hall, G. (1980). A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. Psychol. Rev. *87*, 532–552.

Peciña, S., Smith, K.S., and Berridge, K.C. (2006). Hedonic hot spots in the brain. Neuroscientist *12*, 500–511.

Peppiatt, C.M., Howarth, C., Mobbs, P., and Attwell, D. (2006). Bidirectional control of CNS capillary diameter by pericytes. Nature *443*, 700–704.

Peshkin, L., Meuleau, N., and Kaelbling, L. (2001). Learning policies with external memory. arXiv, arXiv:cs/0103003. http://arxiv.org/abs/cs/0103003.

Phillips, J.M., McAlonan, K., Robb, W.G., and Brown, V.J. (2000). Cholinergic neurotransmission influences covert orientation of visuospatial attention in the rat. Psychopharmacology (Berl.) *150*, 112–116.

Posner, M., Nissen, M., and Ogden, W. (1978). Attended and unattended processing modes: the role of set for spatial location. In Modes of Perceiving and Processing Information, H. Pick and I. Saltzman, eds. (Hillsdale, NJ: Lawrence Erlbaum Associates), p. 137.

Puglisi-Allegra, S., Imperato, A., Angelucci, L., and Cabib, S. (1991). Acute stress induces time-dependent responses in dopamine mesolimbic system. Brain Res. *554*, 217–222.

Puig, M.V., and Gulledge, A.T. (2011). Serotonin and prefrontal cortex function: neurons, networks, and circuits. Mol. Neurobiol. *44*, 449–464.

Raby, C.R., Alexis, D.M., Dickinson, A., and Clayton, N.S. (2007). Planning for the future by western scrub-jays. Nature *445*, 919–921.

Rajkowski, J., Majczynski, H., Clayton, E., and Aston-Jones, G. (2004). Activation of monkey locus coeruleus neurons varies with difficulty and performance in a target detection task. J. Neurophysiol. *92*, 361–371.

Ratcliff, R., and Smith, P.L. (2004). A comparison of sequential sampling models for two-choice reaction time. Psychol. Rev. *111*, 333–367.

Redgrave, P., Prescott, T.J., and Gurney, K. (1999). Is the short-latency dopamine response too short to signal reward error? Trends Neurosci. *22*, 146–151.

Robbins, T.W., and Arnsten, A.F.T. (2009). The neuropsychopharmacology of fronto-executive function: monoaminergic modulation. Annu. Rev. Neurosci. *32*, 267–287.

Roberts, A.C. (2011). The importance of serotonin for orbitofrontal function. Biol. Psychiatry *69*, 1185–1191.

Roesch, M.R., Calu, D.J., and Schoenbaum, G. (2007). Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. Nat. Neurosci. *10*, 1615–1624.

Salamone, J.D., Correa, M., Farrar, A.M., Nunes, E.J., and Pardo, M. (2009). Dopamine, behavioral economics, and effort. Front. Behav. Neurosci. *3*, 13.

Sarter, M., Parikh, V., and Howe, W.M. (2009). nAChR agonist-induced cognition enhancement: integration of cognitive and neuronal mechanisms. Biochem. Pharmacol. *78*, 658–667.

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

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

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

Schweimer, J.V., and Ungless, M.A. (2010). Phasic responses in dorsal raphe serotonin neurons to noxious stimuli. Neuroscience *171*, 1209–1215.

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

Siegel, J.M. (2004). Hypocretin (orexin): role in normal behavior and neuropathology. Annu. Rev. Psychol. *55*, 125–148.

Smith, Y., Bevan, M.D., Shink, E., and Bolam, J.P. (1998). Microcircuitry of the direct and indirect pathways of the basal ganglia. Neuroscience *86*, 353–387.

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

Stuber, G.D., Hnasko, T.S., Britt, J.P., Edwards, R.H., and Bonci, A. (2010). Dopaminergic terminals in the nucleus accumbens but not the dorsal striatum corelease glutamate. J. Neurosci. *30*, 8229–8233.

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

Surmeier, D.J., Shen, W., Day, M., Gertler, T., Chan, S., Tian, X., and Plotkin, J.L. (2010). The role of dopamine in modulating the structure and function of striatal circuits. Prog. Brain Res. *183*, 149–167.

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

Sutton, R.S., and Barto, A.G. (1998). Reinforcement Learning: An Introduction (Cambridge, MA: MIT Press).

Talmi, D., Seymour, B., Dayan, P., and Dolan, R.J. (2008). Human pavlovian-instrumental transfer. J. Neurosci. *28*, 360–368.

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

Thiel, C.M., and Fink, G.R. (2008). Effects of the cholinergic agonist nicotine on reorienting of visual spatial attention and top-down attentional control. Neuroscience *152*, 381–390.

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

Todd, M.T., Niv, Y., and Cohen, J.D. (2009). Learning to use working memory in partially observable environments through dopaminergic reinforcement. In Advances in Neural Information Processing Systems 21, D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, eds. (Cambridge, MA: MIT Press), pp. 1689–1696.

Tolman, E.C. (1948). Cognitive maps in rats and men. Psychol. Rev. *55*, 189–208.

Tops, M., Russo, S., Boksem, M.A.S., and Tucker, D.M. (2009). Serotonin: modulator of a drive to withdraw. Brain Cogn. *71*, 427–436.

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

Tritsch, N.X., and Sabatini, B.L. (2012). Dopaminergic Modulation of Synaptic Transmission in Cortex and Striatum. Neuron *76*, this issue, 33–50.

Ungless, M.A., Magill, P.J., and Bolam, J.P. (2004). Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. Science *303*, 2040–2042.

Usher, M., and McClelland, J.L. (2001). The time course of perceptual choice: the leaky, competing accumulator model. Psychol. Rev. *108*, 550–592.

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

Williams, G.V., and Goldman-Rakic, P.S. (1995). Modulation of memory fields by dopamine D1 receptors in prefrontal cortex. Nature *376*, 572–575.

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

Witte, E.A., and Marrocco, R.T. (1997). Alteration of brain noradrenergic activity in rhesus monkeys affects the alerting component of covert orienting. Psychopharmacology (Berl.) *132*, 315–323.

Wittmann, B.C., Schott, B.H., Guderian, S., Frey, J.U., Heinze, H.J., and Düzel, E. (2005). Reward-related FMRI activation of dopaminergic midbrain is associated with enhanced hippocampus-dependent long-term memory formation. Neuron *45*, 459–467.

Wolf, O.T. (2008). The influence of stress hormones on emotional memory: relevance for psychopathology. Acta Psychol. (Amst.) *127*, 513–531.

Wunderlich, K., Dayan, P., and Dolan, R.J. (2012a). Mapping value based planning and extensively trained choice in the human brain. Nat. Neurosci. *15*, 786–791.

Wunderlich, K., Smittenaar, P., and Dolan, R.J. (2012b). Dopamine enhances model-based over model-free choice behavior. Neuron *75*, 418–424.

Yerkes, R., and Dodson, J. (1908). The relation of strength of stimulus to rapidity of habit-formation. J. Comp. Neurol. Psychol. *18*, 459–482.

Yu, A.J., and Dayan, P. (2005a). Inference, attention, and decision in a bayesian neural architecture. In Advances in Neural Information Processing Systems 17, L.K. Saul, Y. Weiss, and L. Bottou, eds. (Cambridge, MA: MIT Press), pp. 1577–1584.

Yu, A.J., and Dayan, P. (2005b). Uncertainty, neuromodulation, and attention. Neuron *46*, 681–692.

Zaborszky, L. (2002). The modular organization of brain systems. Basal forebrain: the last frontier. Prog. Brain Res. *136*, 359–372.

Zaghloul, K.A., Blanco, J.A., Weidemann, C.T., McGill, K., Jaggi, J.L., Baltuch, G.H., and Kahana, M.J. (2009). Human substantia nigra neurons encode unexpected financial rewards. Science *323*, 1496–1499.

Zemel, R.S., Dayan, P., and Pouget, A. (1998). Probabilistic interpretation of population codes. Neural Comput. *10*, 403–430.