Surprise! Neural correlates of Pearce–Hall and Rescorla–Wagner coexist within the brain

Matthew R. Roesch,[1,2] Guillem R. Esber,[3] Jian Li,[4,5] Nathaniel D. Daw[4,5] and Geoffrey Schoenbaum[3,6,7]

[1]Department of Psychology, University of Maryland College Park, College Park, MD, USA
[2]Program in Neuroscience and Cognitive Science, University of Maryland College Park, College Park, MD, USA
[3]Department of Anatomy and Neurobiology, University of Maryland School of Medicine, Baltimore, MD, USA
[4]Psychology Department, New York University, New York, NY, USA
[5]Center for Neural Science, New York University, New York, NY, USA
[6]Department of Psychiatry, University of Maryland School of Medicine, Baltimore, MD, USA
[7]NIDA Intramural Research Program, 251 Bayview Drive, Baltimore, 21224 MD, USA

Abstract

Learning theory and computational accounts suggest that learning depends on errors in outcome prediction as well as changes in processing of or attention to events. These divergent ideas are captured by models, such as Rescorla–Wagner (RW) and temporal difference (TD) learning on the one hand, which emphasize errors as directly driving changes in associative strength, vs. models such as Pearce–Hall (PH) and more recent variants on the other hand, which propose that errors promote changes in associative strength by modulating attention and processing of events. Numerous studies have shown that phasic firing of midbrain dopamine (DA) neurons carries a signed error signal consistent with RW or TD learning theories, and recently we have shown that this signal can be dissociated from attentional correlates in the basolateral amygdala and anterior cingulate. Here we will review these data along with new evidence: (i) implicating habenula and striatal regions in supporting error signaling in midbrain DA neurons; and (ii) suggesting that the central nucleus of the amygdala and prefrontal regions process the amygdalar attentional signal. However, while the neural instantiations of the RW and PH signals are dissociable and complementary, they may be linked. Any linkage would have implications for understanding why one signal dominates learning in some situations and not others, and also for appreciating the potential impact on learning of neuropathological conditions involving altered DA or amygdalar function, such as schizophrenia, addiction or anxiety disorders.

Introduction

Over the past four decades, the development of animal learning models has been constrained by the need to account for certain cardinal traits of associative learning, such as cue selectivity (e.g. blocking) and cue interactions (e.g. overexpectation). One effective way of accounting for these traits has been to implement Kamin's dictum that, for learning to occur, the outcome of the trial must be surprising (Kamin, 1969). Formally, this dictum is captured in many influential models by the notion of 'prediction error', or discrepancy between the actual outcome and the outcome predicted on that trial (Rescorla & Wagner, 1972; Pearce & Hall, 1980; Pearce *et al.*, 1982; Sutton, 1988; Lepelley, 2004; Pearce & Mackintosh, 2010; Esber & Haselgrove, 2011).

Despite their widespread use, prediction errors do not always serve the same function across different models. For example, the Rescorla–Wagner (RW; Rescorla & Wagner, 1972) and temporal difference learning models (TD; Sutton, 1988) use errors to drive associative changes in a direct fashion. Large errors will result in correspondingly large changes in associative strength, but no change

will occur if the error is zero (i.e. the outcome is already predicted). Furthermore, the sign of the error determines the kind of learning that takes place. If the error is positive, as when the outcome is better than predicted, the association between the cues and the outcome will strengthen. If the error is negative, on the other hand, as when the outcome is worse than expected, this association will weaken or even become negative.

In contrast to these models others, such as the Pearce–Hall (PH; Pearce & Hall, 1980; Pearce *et al.*, 1982), utilize the absolute value of the prediction error to modulate the amount of attention devoted to the cues on subsequent trials, which in turn dictates how much will be learned about them. Large errors will result in a boost in the attention paid to the cues that accompanied the errors, thereby facilitating subsequent learning, whereas small errors will weaken that attention, hampering learning. (Because the cue-specific attention learned by PH modulates subsequent associative learning about the cue, it is also known as the cue's 'associability'. We use the terms interchangeably in this review.) On this view, therefore, by modulating associability, 'unsigned' prediction errors ultimately determine the amount of learning.

Although initially conceived as mutually exclusive, evidence that prediction errors may well promote learning directly and by altering the attention to cues has gradually rendered these views compatible, as

reflected in more recent learning models (Lepelley, 2004; Pearce & Mackintosh, 2010; Esber & Haselgrove, 2011). Indeed, the two mechanisms may serve complementary roles. For example, the degree to which associations in the RW and TD models are updated following a prediction error is scaled by free 'learning rate' parameters, but the theories contain no explicit account of what factors or mechanisms influence them. The attentional allocation learned by PH could address this point, as the associabilities of PH play the same role as learning rates, but are learned from experience rather than entirely free. Indeed, a separate family of Bayesian theories has studied how learning rules of this sort can be derived from principles of sound statistical reasoning; these considerations typically lead to models containing both a RW-like error-driven update, scaled by a PH-like associability mechanism (Sutton, 1992; Dayan & Long, 1998; Kakade & Dayan, 2002; Courville *et al.*, 2006; Behrens *et al.*, 2007; Bossaerts & Preuschoff, 2007). Yet even if RW and PH learning are viewed as competing rather than complementary, it might still be useful for the brain to employ multiple mechanisms to promote flexibility in different situations and robustness in the face of damage or other interference affecting either mechanism.

Consistent with this proposal, numerous studies have shown that phasic firing of midbrain dopamine (DA) neurons carries a RW- or TD-like error signal (see Bromberg-Martin *et al.*, 2010b for review), and recently we have shown that this signal can be dissociated from a PH attentional correlate in the basolateral amygdala (ABL; Roesch *et al.*, 2010). Below we will describe this evidence along with new data implicating habenula and striatal regions in supporting error signaling in midbrain DA neurons and central nucleus and prefrontal regions in processing this attentional signal. We will also suggest that while the RW and PH signals are dissociable, they are likely linked. Thus, while the brain utilizes both signals, they may not be fully independent. Interdependence has implications for understanding why one signal dominates learning in some situations vs. others, and also for appreciating the potential impact of pathological conditions such as schizophrenia, anxiety disorders or even addiction on learning mechanisms.

DA neurons signal bidirectional prediction errors

According to the influential RW (Rescorla & Wagner, 1972) model, prediction errors are calculated from the difference between the outcome predicted by all the cues available on that trial ($\sum V$) and the outcome that is actually received (λ). If the outcome is underpredicted, so that the value of λ is greater than that of $\sum V$, the error will be positive and learning will accrue to those stimuli that happened to be present. Conversely, if the outcome is overpredicted, the error will be negative and a reduction in learning will take place. Thus, the magnitude and sign of the resulting change in the cues' associations (ΔV) is directly determined by prediction error according to the following equation –

$$\Delta V = \alpha\beta(\lambda - \Sigma V) \qquad (1)$$

wherein α and β are constants referring to the salience of the cue and the reinforcer, respectively, included to control the learning rate. This basic idea is also captured in temporal difference learning models (TD), which extend the idea to apply recursively to learning about cues from other cues' previously learned associations (Sutton, 1988).

Dopamine neurons of the midbrain have been widely reported to signal signed errors in reward, and more recently punishment, predictions that are consistent with RW and TD models (Mirenowicz

& Schultz, 1994; Houk *et al.*, 1995; Montague *et al.*, 1996; Hollerman & Schultz, 1998; Waelti *et al.*, 2001; Fiorillo *et al.*, 2003; Tobler *et al.*, 2003; Ungless *et al.*, 2004; Bayer & Glimcher, 2005; Pan *et al.*, 2005; Morris *et al.*, 2006; Roesch *et al.*, 2007; D'Ardenne *et al.*, 2008; Matsumoto & Hikosaka, 2009). Although this idea is not without critics, particularly regarding the precise timing of the phasic response (Redgrave & Gurney, 2006) and the classification of these neurons as dopaminergic (Ungless *et al.*, 2004; Margolis *et al.*, 2006), the evidence supporting this proposal is strong, deriving from multiple labs, species and tasks. These studies demonstrate that a large proportion of putative DA neurons exhibit bidirectional changes in activity in response to rewards that are better or worse than expected. These neurons fire to an unpredicted reward, and firing declines when the reward becomes predicted and is suppressed when the predicted reward is not delivered. This is illustrated in the single unit example in Fig. 1. Moreover, the same neurons display activity in response to unpredicted cues, if those cues had been previously associated with reward. Although these cue-evoked responses might initially seem like they carry qualitatively different information from the error-related responses to primary rewards, TD models explain them both as
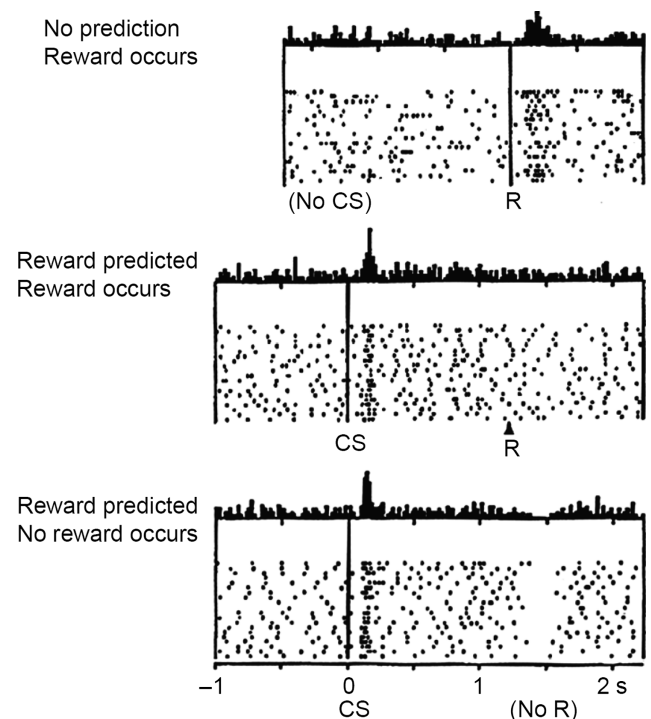

Fig. 1. Changes in firing in response to positive and negative reward prediction errors (PEs) in a primate DA neuron. The figure shows spiking activity in a putative DA neuron recorded in the midbrain of a monkey performing a simple task in which a conditioned stimulus (CS) is used to signal reward. Data are displayed in a raster format in which each tic mark represents an action potential and each row a trial. Average activity per trial is summarized in a peri-event time histogram at the top of each panel. The top panel shows activity to an unpredicted reward (+PE). The middle panel shows activity to the reward when it is fully predicted by the CS (no PE). The bottom panel shows activity on trials in which the CS is presented but the reward is omitted (−PE). As described in the text, the neuron exhibits a bidirectional correlate of the reward prediction error, firing to unexpected but not expected reward, and suppressing firing on omission of an expected reward. The neuron also fires to the CS; in theory such activity is thought to reflect the error induced by unexpected presentation of the valuable CS. This feature distinguishes a TDRL signal from the simple error signal postulated by RW. This figure is adapted from Schultz *et al.* (1997).

instances of a common prediction error signal. In TD, unexpected reward-predictive cues induce a prediction error because they reflect a change from the expected value, just as delivery of a primary reward does earlier in learning. Of course, signed prediction errors do not necessarily have to be represented in the form of bidirectional phasic activity within the same neuron; however, the occurrence of such a firing pattern in some DA neurons does serve to rule out other competing interpretations such as encoding of salience, attention or motivation, or even simple habituation.

Although much of this evidence has come from primates, similar results have been reported in other species. For example, functional imaging results from many groups indicate that the blood oxygen level-dependent (BOLD) response in the human ventral striatum also displays all the hallmarks of prediction errors seen in primate DA recordings, such as elevation for unexpected rewards and suppression when expected rewards are omitted (McClure *et al.*, 2003; O'Doherty *et al.*, 2003; Hare *et al.*, 2008). Although the BOLD response is a metabolic signal and not specific to any particular neurochemical, several results support the inference that BOLD correlates of prediction errors in the striatum may in part reflect the effect of its prominent dopaminergic input (Pessiglione *et al.*, 2006; Day *et al.*, 2007; Knutson & Gibbs, 2007). Also, although technically more challenging to image unambiguously, BOLD activity in the human midbrain dopaminergic nuclei also appears to reflect a prediction error (D'Ardenne *et al.*, 2008).

Similarly, Hyland and colleagues have reported signaling of reward prediction errors in a simple Pavlovian conditioning and extinction task (Pan *et al.*, 2005) in rodents. Putative DA neurons recorded in rat ventral tegmental area (VTA) initially fired to the reward. With learning, this reward-evoked response declined and the same neurons developed responses to the preceding cue. During extinction, activity was suppressed on reward omission. Parallel modeling revealed that the changes in activity closely paralleled theoretical error signals in their task.

Prediction error signaling has also been reported in VTA DA neurons in rats performing a more complicated choice task. In this task, reward was unexpectedly delivered or omitted by altering the timing or number of rewards delivered (Roesch *et al.*, 2007). This task is illustrated in Fig. 2, along with a heat plot showing population activity on the subset of trials in which reward was delivered unexpectedly. DA activity increased when a new reward was instituted and then declined as the rats learned to expect that reward. Activity in these same neurons was suppressed later, when the expected reward was omitted. In addition, activity was also high for delayed reward, consistent with recent reports in primates (Fiorillo *et al.*, 2008; Kobayashi & Schultz, 2008). Furthermore, the same DA neurons that fired to unpredictable reward also developed phasic responses to preceding cues with learning, and this activity was higher when the cues predicted the more valuable reward. These features of DA firing are entirely consistent with prediction error encoding (Mirenowicz & Schultz, 1994; Montague *et al.*, 1996; Hollerman & Schultz, 1998; Waelti *et al.*, 2001; Fiorillo *et al.*, 2003; Tobler *et al.*, 2003; Bayer & Glimcher, 2005; Pan *et al.*, 2005; Morris *et al.*, 2006; Roesch *et al.*, 2007; D'Ardenne *et al.*, 2008; Matsumoto & Hikosaka, 2009).

Where do DA neurons get information relevant to predicting reward and where do these signals go?

The latter question is easier to address, as the working hypothesis for all intents and purposes has been 'everywhere'. Error signals are found prominently in both the substantia nigra and VTA, and because DA neurons project widely and nearly every brain region is thought to be involved in some form of associative learning, a reasonable hypothesis might be that these signals could be important anywhere.

However, while this may be a reasonable starting point, recent and not-so-recent evidence suggests it is an (obvious) over-simplification. First, while DA neurons may project widely, their inputs are particularly concentrated on striatal regions, which are heavily implicated in associative learning. This combined with converging evidence from imaging and computational neuroscience has led to the suggestion that these regions are particularly receptive to the dopaminergic teaching signals. Second, there has been increasing emphasis on whether the dopaminergic error signal (and DA neuron population it is contained in) is truly homogenous. While the signal seems to be present in the majority of DA neurons identified by waveform, this still leaves a substantial number of these neurons that are not coding errors. Further traditional waveform criteria may fail to identify at least some TH+ neurons (Margolis *et al.*, 2006, 2008); indeed there may be a particular sampling bias against those that project to prefrontal regions (Lammel *et al.*, 2008). Given that prediction errors do not appear to be present in neurons that fail to show the classical waveform criteria (Roesch *et al.*, 2007), this suggests an interesting possibility that dopaminergic prediction error signals may have their primary impact on subcortical associative learning nodes, such as striatum and amygdala, and much less impact on prefrontal executive regions. Of course this is largely speculative; until it is possible to directly tag specific neurons during extracellular recording, in order to identify their projection targets, it will be difficult to conclusively identify which projections carry these signals and for what purposes.

More developed ideas exist regarding where information to support dopaminergic error signals originates. One set of ideas has focused on the information content of this signal. In order to generate a prediction error, one must have a prediction – the overall expected value based on the summed values of the cues present in the environment. This summed value is similar to the overall expected value in RW. Based on anatomical and imaging data, it has been suggested that the ventral striatum serves this role, providing information about the summed expected value given current circumstances to the midbrain (O'Doherty *et al.*, 2004). Other accounts have suggested that value predictions may also be derived from the amygdala (Belova *et al.*, 2008). Notably, these are both limbic areas that are clearly implicated in signaling cue values. An interesting outstanding question is whether information impacting the DA signal also comes from prefrontal regions (Balleine *et al.*, 2008), thought to represent task structure in so-called model-based reinforcement learning models (Daw *et al.*, 2005). Some evidence exists suggesting that DA signals (or error-related BOLD responses in striatal regions) may reflect this type of information (Bromberg-Martin *et al.*, 2010d; Daw *et al.*, 2011; Simon & Daw, in press). Consistent with this, we have recently shown that input from the orbitofrontal cortex, a key region in encoding such model-based associative structures (Pickens *et al.*, 2003; Izquierdo *et al.*, 2004; Ostlund & Balleine, 2007; Burke *et al.*, 2008), is necessary for expectancy-related changes in phasic firing in midbrain DA neurons (Takahashi *et al.*, 2011).

A second set of ideas has focused on locating the error signal itself in upstream structures. In particular, considerable focus has been on the lateral habenula (LHb; Matsumoto & Hikosaka, 2007; Bromberg-Martin *et al.*, 2010c; Hikosaka, 2010), which is thought to receive error information from the globus pallidus (GP; Hong & Hikosaka, 2008). Activity in the LHb reflects signed prediction errors the same as activity of DA neurons but, remarkably, in the opposite direction. That is, activity is inhibited and excited by
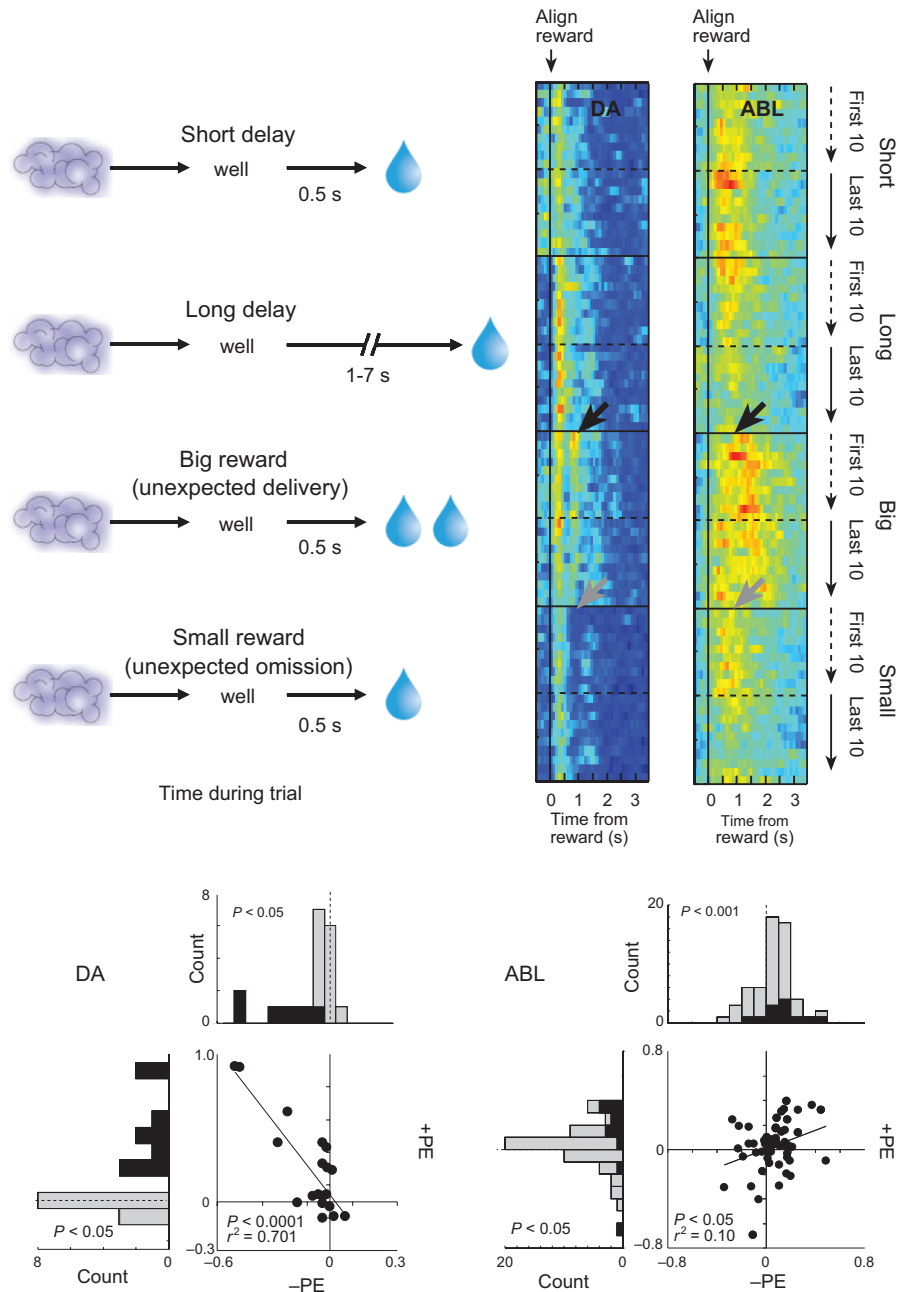
FIG. 2. Changes in firing in response to positive and negative reward prediction errors (PEs) in rat dopamine (DA) and basolateral amygdala (ABL) neurons. Rats were trained on a simple choice task in which odors predicted different rewards. During recording, the rats learned to adjust their behavior to reflect changes in the timing or size of reward. As illustrated in the left panel, this resulted in delivery of unexpected rewards (+PE) and omission of expected rewards (−PE). Activity in reward-responsive DA and ABL neurons is illustrated in the heat plots to the right, which show average firing synchronized to reward in the first and last 10 trials of each block, and in the scatter/histograms below, which plot changes in firing for each neuron in response to +PEs and −PEs. As described in the text, neurons in both regions fired more to an unexpected reward (+PE, black arrow). However, only the DA neurons also suppressed firing on reward omission; ABL neurons instead increased firing (−PE, gray arrows). This is inconsistent with the bidirectional error signal postulated by RW or TDRL, and instead is more like the unsigned error signal utilized in attentional theories, such as PH. This figure is adapted from Roesch *et al.* (2007, 2010).

positive and negative prediction errors, respectively. Consistent with the idea that the LHb is feeding DA neurons this information, activity of DA neurons is inhibited by LHb stimulation and prediction error signaling in LHb occurs earlier than those in DA neurons (Matsumoto & Hikosaka, 2007; Bromberg-Martin *et al.*, 2010a). DA neurons are likely to receive this information via indirect connections through midbrain γ-aminobutyric acid (GABA) neurons in the VTA and the adjacent rostromedial tegmental nucleus, which has similar response properties as LHb neurons and heavy inhibitory

projections to midbrain DA neurons (Ji & Shepard, 2007; Jhou *et al.*, 2009; Kaufling *et al.*, 2009; Omelchenko *et al.*, 2009; Brinschwitz *et al.*, 2010; Hong *et al.*, 2011).

How to integrate these two stories is not clear. One possibility is that the GP–LHb–midbrain circuit forms an 'error-signaling axis' that information relevant to determining the value expected in a particular state, provided by limbic and perhaps prefrontal regions, feeds into at multiple levels. Whatever the case, the emerging picture suggests a widespread error-signaling circuit, in position to receive a torrent of

information regarding value, arising from nearly all the brain systems we would currently think of as important for valuation and decision-making.

Amygdala neurons signal shifts in attention

While the amygdala has often been viewed as critical for learning to predict aversive outcomes (Davis, 2000; LeDoux, 2000), the last two decades have revealed a more general involvement in associative learning (Gallagher, 2000; Murray, 2007; Tye *et al.*, 2008). Although the mainstream view holds that the amygdala is important for acquiring and storing associative information (LeDoux, 2000; Murray, 2007), there have been hints in the literature that the amygdala may also support other functions related to associative learning and error signaling. For example, damage to the central nucleus disrupts orienting and increments in attentional processing after changes in expected outcomes (Gallagher *et al.*, 1990; Holland & Gallagher, 1993b, 1999), and other studies have proposed a critical role for amygdala – particularly central nucleus output to subcortical areas – in mediating vigilance or surprise (Davis & Whalen, 2001).

Consistent with this idea, Salzman and colleagues (Belova *et al.*, 2007) have reported that amygdala neurons in monkeys are responsive to unexpected outcomes. However, this study showed minimal evidence of negative prediction error encoding or transfer of positive prediction errors to conditioned stimuli (CS) in amygdala neurons, and many neurons fired similarly to unexpected rewards and punishments. This pattern of firing does not meet the criteria for a RW or temporal difference reinforcement learning (TDRL) prediction error signal.

Similar firing patterns have also been reported in rat ABL. For example, Janak and colleagues examined the activity of neurons in ABL during unexpected omission of reward during the extinction of reward-seeking behavior (Tye *et al.*, 2010). Rats were first trained to respond at a nosepoke operandum for partial reinforcement (50%). After several training sessions, the sucrose reward was withheld during extinction. Rats quickly learned to stop nosepoking during extinction, and many neurons in ABL started to fire to the empty port. Importantly, these changes were correlated with the response intensity and the extinction resistance of the rat, and fit well with several lesion studies demonstrating the importance of the ABL in altering behavior when expected values change (Corbit & Balleine, 2005; McLaughlin & Floresco, 2007; Ehrlich *et al.*, 2009).

Modulation of neural activity in lateral and basal nucleus of the amygdala by expectation has also been described during fear conditioning in rats (Johansen *et al.*, 2010). In this procedure shock to the eye lid was induced after presentation of an auditory CS. Recordings were conducted during initial CS–unconditioned stimulus pairings, and showed that shocks elicited stronger neural responses early during learning than later after learning. For these cells, responses evoked by the shock were inversely correlated with freezing to the CS, suggesting that diminution of shock-evoked activity was related to the changes in expectation. Overall, activity was stronger for unsignaled vs. signaled shock, consistent with the amygdala encoding the surprise induced by unexpected shock.

The patterns observed in these studies do not appear to be fully consistent with what is predicted by RW or TD, and instead seem more like a correlate of surprise or attention. However, it is unclear if they fully comply with predictions of the PH theory for an attentional correlate, as in most cases they were recorded in settings that make it difficult to identify whether they are unsigned, either because recordings were done across days or without both types of errors or due to potential confounds between aversive positive errors with

appetitive negative errors (or vice versa). In order to fully address this question, it would be useful to record these neurons in a task already applied to isolate RW- or TD-like signed errors in midbrain DA neurons.

Data from such a study are presented in Fig. 2. In this experiment, neural activity was recorded from ABL neurons in rats during performance of the same choice task used to isolate prediction error signaling in rat DA neurons (Roesch *et al.*, 2010). As in monkeys, many ABL neurons increased firing when reward was delivered unexpectedly. However, such activity differed markedly in its temporal specificity from what was observed in VTA DA neurons. This is evident in Fig. 2, where the increased firing in ABL occurs somewhat later and is much broader than that in DA neurons.

Moreover, activity in these ABL neurons was not inhibited by omission of an expected reward. Instead, activity was actually stronger during reward omission, and those neurons that fired most strongly for unexpected reward delivery also fired most strongly after reward omission.

Activity in ABL also differed from that in VTA in its onset dynamics. While firing in VTA DA neurons was strongest on the first encounter with an unexpected reward and then declined, activity in the ABL neurons continued to increase over several trials after a shift in reward. These differences and the overall pattern of firing in the ABL neurons are inconsistent with signaling of a signed prediction error as envisioned by RW and TD, at least at the level of individual single-units. Instead, such activity in ABL appears to relate to an unsigned error signal or, more particularly, to an associability or attention variable derived from unsigned errors.

That is, theories of associative learning have traditionally employed unsigned errors to drive changes in stimulus processing or attention, operationalized as an associability parameter controlling learning rate (Mackintosh, 1975; Pearce & Hall, 1980). According to this idea, the attention that a cue receives is equal to the weighted average of the unsigned error generated across the past few trials, reflecting the idea that cues that have been accompanied by either positive or negative errors in the past are more likely to be responsible for errors in the future, and thus should be updated preferentially. Mathematically, according to one variant of PH (Pearce *et al.*, 1982), the associability of a cue on trial n (α_n) is updated relative to its previous level (α_{n-1}) using the absolute value of the prediction error on that trial. Thus –

$$\alpha_n = \gamma|\lambda_{n-1} - \sum V_{n-1}| + (1-\gamma)\alpha_{n-1} \qquad (2)$$

where ($\lambda_{n-1} - \sum V_{n-1}$) is defined as the difference between the value of the reward predicted by all cues in the environment ($\sum V_{n-1}$) and the value of the reward that was actually received (λ_{n-1}), and γ is a weighting factor ranging between 0 and 1, which serves to account for the observed gradual change in attention or associability. The resultant quantity – termed attention (α) – can then be used to scale the update of a cue's value. PH's version of this rule is –

$$\Delta V = \alpha S \lambda \qquad (3)$$

where S and λ represent the intrinsic salience (e.g. intensity) of the cue (S) and the reward, respectively.

Of course, the same associability term can be used, in principle, with RW (Eqn 1) to modulate learning rate (Lepelley, 2004). Interestingly, hybrid RW/PH mechanisms of this sort also emerge in Bayesian models that attempt to explain learning in conditioning as arising from normative principles of statistical reasoning. In these models, learning about a cue is often driven by prediction error (as in

RW), but the learning rate is not a free parameter. Instead, it should be determined by uncertainty about the cue's associative strength – all else equal, if you are more uncertain about a cue's value, you should be more willing to update it. In these models, then, uncertainty plays a role like associability in controlling learning rates. In these models, uncertainty is determined by a rule strikingly similar to Eqn 2. This is because an important determinant of uncertainty is the variance of the prediction error, of which the unsigned (in this case, squared rather than absolute) error is a sample (Dayan & Long, 1998; Courville et al., 2006; Bossaerts & Preuschoff, 2007). Thus, Bayesian models offer a normative interpretation for Eqn 2 and the attentional factor it describes, and motivate its use in hybrid models together with RW-like mechanisms.

The activity of the ABL neurons in Fig. 2 appears to provide an associability signal similar to that proposed by PH and these other models to underlie surprise and attentional modulation in the service of associative learning. As predicted by these theoretical accounts, this pattern is clearly unlike that of a signed error contained in the firing of the VTA DA neurons. A similar dissociation between the BOLD signal in ventral striatum and amygdala has also been reported in humans during aversive learning (Li et al., 2011). In this study, subjects performed a reversal task in which two visual cues were initially partially paired with an electrical shock or nothing, respectively, and then the associations were reversed. Changes in skin conductance response during presentation of the two cues indicated that subjects learned not only the value of the cues, but also represented the cues associability or salience as predicted by the PH model. Subsequent analyses of BOLD response at the time of cue termination (when the outcome was delivered or not) showed that activity in the ventral striatum was positively correlated with the aversive prediction error (Fig. 3), whereas activity in the amygdala was positively correlated with associability (Fig. 3). Given results implicating the DA system in signaling aversive errors (Ungless et al., 2004; Matsumoto & Hikosaka, 2009; Mileykovskiy & Morales, 2011), these results are at least consistent with a neural dissociation between dopaminergic error signaling and amygdalar signaling of attention.

Where do amygdala neurons get information relevant to modulating attention and where do these signals go?

As with the DA neurons, it is simpler to address where the signal might go and what its impact might be than to identify where the information to support the signal originates. Two major circuits appear to be the most likely recipients of the attentional signal from the ABL. The first is a circuit running through the central nucleus of the amygdala (CeA). The central nucleus receives information from the basolateral areas, and it has been strongly implicated in surprise and vigilance (Davis & Whalen, 2001) as well as in modulating attention for learning during unblocking procedures (Holland & Gallagher, 1993a,b). These and other related findings (Holland & Gallagher, 1993a,b; Holland & Kenmuir, 2005) demonstrate that the CeA is essential for the enhancement of attention that results from the unexpected omission of an outcome (negative error). The specificity of this effect suggests a special role of the CeA in processing the absence of an expected event; consistent with this we have found that neurons in the central nucleus provide a signal that specifically increases when an expected reward is omitted (Calu et al., 2010). Whether the partial representation of a PH signal in the central nucleus reflects a filtering of the fuller signal from ABL will require a causal disconnection to demonstrate, although recent data from Holland and colleagues implicating the ABL in attentional function during unblocking procedures would favor this hypothesis (Chang et al., 2010).

The second circuits that may receive information from the ABL regarding attentional modulation are prefrontal areas. As outlined above, these regions may not receive a dopaminergic error signal, as it has been reported that prefrontal-projecting DA neurons lack the characteristic waveform features that characterize error-signaling dopaminergic neurons in our work (Roesch et al., 2007; Lammel et al., 2008). Yet, prefrontal regions – particularly the orbital and cingulate regions that receive input from the amygdala – are generally implicated in learning and processes that would benefit from input highlighting particularly cues for attention.

The anterior cingulate (ACC) is a particularly likely candidate to receive error information from the ABL; it has strong reciprocal connections with the ABL (Sripanidkulchai et al., 1984; Cassell & Wright, 1986; Dziewiatkowski et al., 1998) and has already been shown to be involved in a number of functions related to error processing and attention (Carter et al., 1998; Scheffers & Coles, 2000; Paus, 2001; Holroyd & Coles, 2002; Ito et al., 2003; Rushworth et al., 2004, 2007; Walton et al., 2004; Amiez et al., 2005, 2006; Kennerley et al., 2006, 2009; Magno et al., 2006; Matsumoto et al., 2007; Oliveira et al., 2007; Sallet et al., 2007; Quilodran et al., 2008; Rudebeck et al., 2008; Rushworth & Behrens, 2008; Kennerley & Wallis, 2009; Totah et al., 2009; Hillman & Bilkey, 2010; Wallis & Kennerley, 2010; Hayden et al., 2011; Rothe et al., 2011).
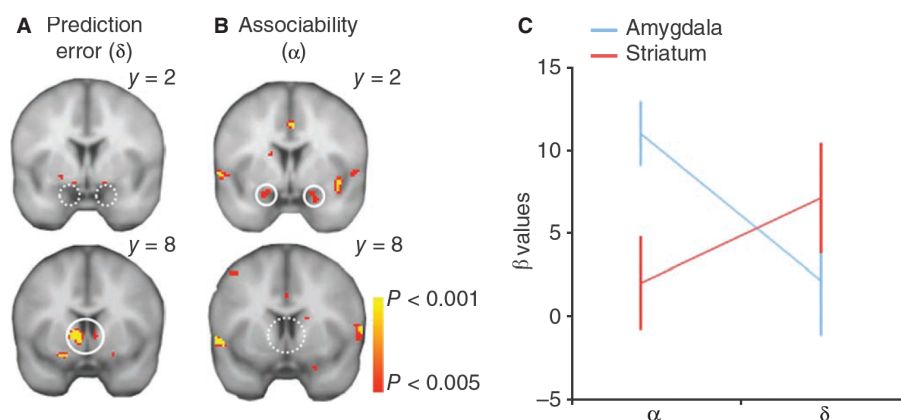

FIG. 3. Neural correlates of associability and prediction error term. (A) BOLD in the ventral striatum, but not amygdala, correlated with prediction error. (B) BOLD in the bilateral amygdala, but not ventral striatum, correlated with associability regressor (P < 0.05, SVC). (C) Differential representations of associability (α) and prediction error (δ) in striatum and amygdala BOLD (± SEM). This figure is adapted from Li et al. (2011).

Although most of this research has focused on the role of the ACC in the detection of errors of commission (Ito *et al.*, 2003; Amiez *et al.*, 2005; Quilodran *et al.*, 2008; Rushworth & Behrens, 2008; Totah *et al.*, 2009), more recent work has suggested that the ACC does not simply detect errors, but is important for signaling other aspects of behavioral feedback (Kennerley *et al.*, 2006; Oliveira *et al.*, 2007; Rothe *et al.*, 2011), including prediction error encoding similar to what we find in the ABL. For example, Hayden and colleagues showed that activity in monkey ACC was high when rewards were delivered, and omitted unexpectedly in a task in which rewards were delivered at predetermined probabilities. These changes in ACC firing occurred regardless of valence at the single-cell level, consistent with encoding of surprise or attentional variables (Hayden *et al.*, 2011).

To more directly address how activity related to changes in reward in the ACC compares to that in other areas, we have recently recorded in the ACC in rats in the same task in which we characterized firing in ABL and VTA. Importantly this task allows us to examine neural activity in the ACC during trials after reward prediction error signals occur in a setting in which animals learn from detection of such errors. Consistent with previous findings, we found that the ACC detects errors of commission and reward prediction at the time of their occurrence (Bryden *et al.*, 2011). However, we also found that activity in the ACC was elevated in anticipation of and during cue presentation on trials after reward contingencies changed. Such an elevation, which was not observed in the ABL, could reflect increased processing of cues on subsequent trials, as predicted by the PH theory. Consistent with this proposal, changes in firing were correlated with behavioral measures of attention to the cues evident on these trials. These effects are illustrated in a single-cell example and across the population in Fig. 4. Thus, the ACC is not only involved in detecting errors, but in utilizing that information to drive attention and learning on subsequent trials. Detection of prediction errors and the subsequent changes in attention critical for learning might be dependent on connections between the ABL and ACC.

Where information supporting these signals originates is again a more complex question, particularly because these signals have only been recently uncovered. Like RW or TD error signals, the attentional signals carried by neurons in basolateral and central amygdala also depend on information about expected value, in order to calculate the prediction error that forms the basis of the PH effect. The amygdala is a major associative learning node, thus it is well positioned to compute an error signal internally, utilizing associative information coded locally and sent to it by other regions, such as prefrontal regions or the hippocampus.

However, an alternative, intriguing possibility is that the PH signal in the amygdala is dependent on an external prediction error signal. An obvious candidate to provide this signal would be the midbrain DA neurons. These neurons already have access to the information necessary to compute this quantity and project strongly into the ABL. Indeed, although learning as in Eqn 2 could in principle be implemented using a signed error signal of the sort associated with DA (e.g. if the rules governing plasticity at target neurons effect the absolute value), recent work suggests that some DA neurons may provide an unsigned error signal. Specifically, in monkeys, some putative DA neurons increase firing when either an appetitive (juice) or aversive (air-puff) outcome is delivered unexpectedly (Matsumoto & Hikosaka, 2009). These results have yet to be shown clearly in rats, but would be consistent with a linkage between dopaminergic and amygdalar teaching signals.

Although speculative, we recently repeated our earlier experiment, recording in both controls and rats with ipsilateral 6-OHDA lesions (personal communication). These rats performed normally on the task,

presumably utilizing the intact hemisphere, and amygdala neurons recorded in these rats showed robust reward-evoked activity. However, unlike neurons in controls and in our prior experiment, neurons deprived of DA input failed to modulate firing in response to surprising upshifts or downshifts in reward. Thus, removal of DA input disrupted error-related modulations in the activity of these neurons. This result provides evidence linking the RW signal computed by the DA neurons with the PH attentional signal in ABL neurons.

The significance for normal and pathological learning

If single-unit activity in the amygdala contributes to attentional changes, then the role of this region in a variety of learning processes may need to be reconceptualized or at least modified to include this function. This is particularly true for ABL. ABL appears to be critical for encoding associative information properly; associative encoding in downstream areas requires input from the ABL (Schoenbaum *et al.*, 2003; Ambroggi *et al.*, 2008). This has been interpreted as reflecting an initial role in acquiring the simple associative representations (Pickens *et al.*, 2003). However, an alternative account – not mutually exclusive – is that ABL may also augment the allocation of attentional resources to directly drive acquisition of the same associative information in other areas. As noted above, the signal in ABL, identified here, may serve to augment or amplify the associability of cue-representations in downstream regions, so that these cues are more associable or salient on subsequent trials. Such an effect may be evident in neural activity prior to and during cue sampling reported here in the ACC.

A role in attentional function for the ABL would also affect our understanding of how this region is involved in neuropsychiatric disorders. For example, the amygdala has long been implicated in anxiety disorders such as post-traumatic stress disorder (Davis, 2000); while this involvement may reflect abnormal storage of information in the ABL, it might also reflect altered attentional signaling, affecting storage of information not in the ABL but in other brain regions. This would be consistent with accounts of amygdala function in fear that have emphasized vigilance (Davis & Whalen, 2001).

Understanding how RW and PH signaling are linked in the brain will also be important. One possibility, which we have suggested several times here and elsewhere (Lepelley, 2004; Courville *et al.*, 2006; Li *et al.*, 2011), is that the RW/TD and PH mechanisms effectively comprise the associative learning and learning rate control components of a single hybrid learner. Thus, the associabilities learned by PH may serve to scale the associative TD updates driven by dopaminergic prediction errors. The demonstration that signaling related to associability in ABL depends on a dopaminergic input would support such an integrated view, as it is consistent with the possibility that the error terms in Eqns 1 and 2 derive from a common source.

Even if the two mechanisms are tightly interacting in many circumstances, they may also contribute separately or independently to other behaviors. Indeed, there do exist some behavioral phenomena that are difficult to explain by the simple composition of RW and PH mechanisms into a single hybrid learner, at least in the form described here, but instead may be consistent in various circumstances with the mechanism described by one or the other model can operate more or less independently. For example, when rewards are omitted unexpectedly, rats typically show increased responding to added cues (Holland & Gallagher, 1993a), which is consistent with the original PH models but not RW or TD learning (even when augmented by PH attentional updates; see Dayan & Long, 1998). However, if switches
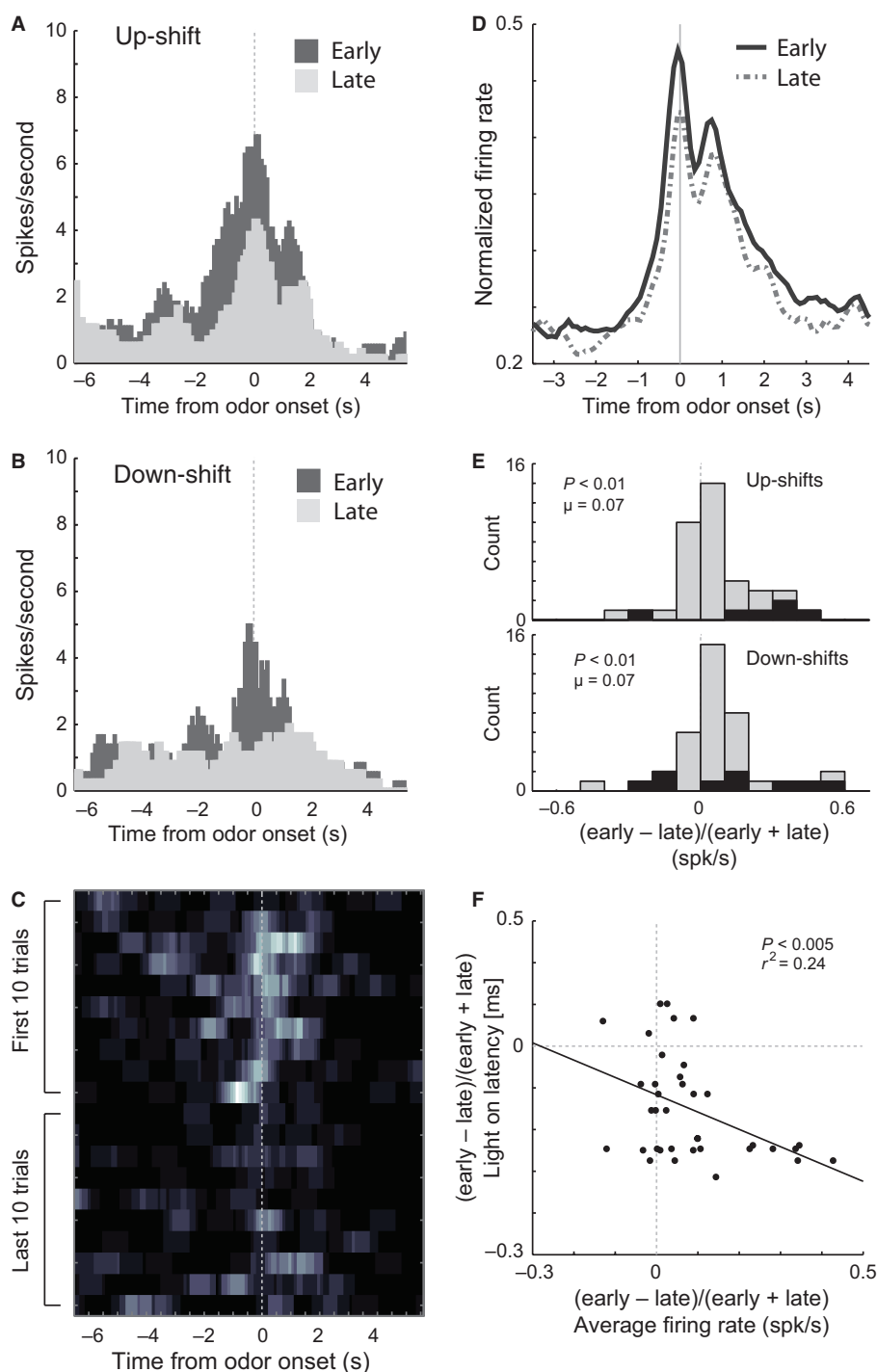
FIG. 4. Activity in ACC is stronger on trials after errors in reward prediction. (A, B) Histogram represents firing of one neuron during the first 10 (dark gray) and last 10 (light gray) trials after up- and down-shifts in value aligned on odor onset. (C) Heat plot shows the average firing, of the same neuron, across shifts during the first and last 10 trials after reward contingencies change. (D) Average normalized neural activity for all task-related neurons (odor onset to fluid well entry) comparing the first 10 trials in all blocks (early; solid black) to the last 10 trials in all blocks (late; dashed gray). (E) Distribution reflecting the difference in task-related firing rate between early and late in a trial block (early−late)/(early + late) following either up-shifts (top) or down-shifts (bottom). (F) Correlation between light-on latency (house light on until nosepoke; *y*-axis) and firing rate (*x*-axis) either early or late within a block [(early−late)/(early + late)]. *N* = 4 rats. This figure is adapted from Bryden *et al.* (2011).

involving omission or reward downshift occur repeatedly, rats show less or slower responding (Roesch *et al.*, 2007, 2010), consistent with RW but not with naked PH. These considerations suggest that the two classical conditioning mechanisms may contribute to different degrees in different circumstances, much as has been suggested for different

instrumental learning mechanisms (Dickinson & Balleine, 2002; Daw *et al.*, 2005). In the case of PH and RW, the factors governing the relative contribution of the mechanisms are not yet as well understood.

Finally, it is worth noting that a linkage (or lack thereof) between dopaminergic RW signals and PH signals in the ABL and elsewhere has

significance for our understanding of altered learning in neuropathological conditions. For example, some symptoms of schizophrenia have been proposed to reflect the spurious attribution of salience to cues (Kapur, 2003). These symptoms might be secondary to altered signaling in the ABL (Taylor *et al.*, 2005) under the influence of aberrant dopaminergic input. Such alterations could drive frank attentional problems as well as positive symptoms such as hallucinations and delusions. The amygdala has also been central to ideas about addiction, where it is proposed to mediate craving and the abnormal attribution of motivational significance to drug-associated cues and contexts. Again such aberrant learning may reflect, in part, disrupted attentional processing, potentially under the influence of altered DA function.

Acknowledgements

Abbreviations

ABL, basolateral amygdala; ACC, anterior cingulate; BOLD, blood oxygen level-dependent; CeA, central nucleus of the amygdala; CS, conditioned stimulus; DA, dopamine; GP, globus pallidus; LHb, lateral habenula; PH, Pearce–Hall; RW, Rescorla–Wagner; TD, temporal difference; TDRL, temporal difference reinforcement learning; VTA, ventral tegmental area.

References

Ambroggi, F., Ishikawa, A., Fields, H.L. & Nicola, S.M. (2008) Basolateral amygdala neurons facilitate reward-seeking behavior by exciting nucleus accumbens neurons. *Neuron*, **59**, 648–661.

Amiez, C., Joseph, J.P. & Procyk, E. (2005) Anterior cingulate error-related activity is modulated by predicted reward. *Eur. J. Neurosci.*, **21**, 3447–3452.

Amiez, C., Joseph, J.P. & Procyk, E. (2006) Reward encoding in the monkey anterior cingulate cortex. *Cereb. Cortex*, **16**, 1040–1055.

Balleine, B.W., Daw, N.D. & O'Doherty, J.P. (2008). Multiple forms of value learning and the function of dopamine. In Glimcher, P.W., Camerer, C.F., Fehr, E. & Poldrack, R.A. (Eds), *Neuroeconomics: Decision Making and the Brain*. Elsevier, Amsterdam, pp. 367–388.

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

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

Belova, M.A., Paton, J.J., Morrison, S.E. & Salzman, C.D. (2007) Expectation modulates neural responses to pleasant and aversive stimuli in primate amygdala. *Neuron*, **55**, 970–984.

Belova, M.A., Paton, J.J. & Salzman, C.D. (2008) Moment-to-moment tracking of state value in the amygdala. *J. Neurosci.*, **28**, 10023–10030.

Bossaerts, P. & Preuschoff, K. (2007) Adding prediction risk to the theory of reward learning in the dopaminergic system. *Ann. N.Y. Acad. Sci.*, **1104**, 135–146.

Brinschwitz, K., Dittgen, A., Madai, V.I., Lommel, R., Geisler, S. & Veh, R.W. (2010) Glutamatergic axons from the lateral habenula mainly terminate on GABAergic neurons of the ventral midbrain. *Neuroscience*, **168**, 463–476.

Bromberg-Martin, E.S., Matsumoto, M. & Hikosaka, O. (2010a) Distinct tonic and phasic anticipatory activity in lateral habenula and dopamine neurons. *Neuron*, **67**, 144–155.

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

Bromberg-Martin, E.S., Matsumoto, M., Hong, S. & Hikosaka, O. (2010d) A pallidus-habenula-dopamine pathway signals inferred stimulus values. *J. Neurophysiol.*, **104**, 1068–1076.

Bryden, D.W., Johnson, E.E., Tobia, S.C., Kashtelyan, V. & Roesch, M.R. (2011) Attention for learning signals in anterior cingulate cortex. *J. Neurosci.*, **31**, 18266–18274.

Burke, K.A., Franz, T.M., Miller, D.N. & Schoenbaum, G. (2008) The role of the orbitofrontal cortex in the pursuit of happiness and more specific rewards. *Nature*, **454**, 340–344.

Calu, D.J., Roesch, M.R., Haney, R.Z., Holland, P.C. & Schoenbaum, G. (2010) Neural correlates of variations in event processing during learning in central nucleus of amygdala. *Neuron*, **68**, 991–1001.

Carter, C.S., Braver, T.S., Barch, D.M., Botvinick, M.M., Noll, D. & Cohen, J.D. (1998) Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science*, **280**, 747–749.

Cassell, M.D. & Wright, D.J. (1986) Topography of projections from the medial prefrontal cortex to the amygdala in the rat. *Brain Res. Bull.*, **17**, 321–333.

Chang, S.E., Wheeler, D.S., McDannald, M. & Holland, P.C. (2010) The effects of basolateral amygdala lesions on unblocking. *Soc. Neurosci. Abs.*

Corbit, L.H. & Balleine, B.W. (2005) Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrumental transfer. *J. Neurosci.*, **25**, 962–970.

Courville, A.C., Daw, N.D. & Touretzky, D.S. (2006) Bayesian theories of conditioning in a changing world. *Trends Cogn. Sci.*, **10**, 294–300.

D'Ardenne, K., McClure, S.M., Nystrom, L.E. & Cohen, J.D. (2008) BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science*, **319**, 1264–1267.

Davis, M. (2000). The role of the amygdala in conditioned and unconditioned fear and anxiety. In Aggleton, J.P. (Ed.), *The Amygdala: A Functional Analysis*. Oxford University Press, Oxford, pp. 213–287.

Davis, M. & Whalen, P.J. (2001) The amygdala: vigilance and emotion. *Mol. Psychiatry*, **6**, 13–34.

Daw, N.D., Niv, Y. & 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. & Dolan, R.J. (2011) Model-based influences on humans' choices and striatal prediction errors. *Neuron*, **69**, 1204–1215.

Day, J.J., Roitman, M.F., Wightman, R.M. & Carelli, R.M. (2007) Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nat. Neurosci.*, **10**, 1020–1028.

Dayan, P. & Long, T. (1998) Statistical models of conditioning. *Neural Information Processing Systems*, **10**, 117–123.

Dickinson, A. & Balleine, B.W. (2002). The role of learning in the operation of motivational systems. In Pashler, H. & Gallistel, R. (Eds), *Stevens' Handbook of Experimental Psychology (3rd Edition, Vol 3: Learning, motivation, and emotion)*. John Wiley & Sons, New York, NY, pp. 497–533.

Dziewiatkowski, J., Spodnik, J.H., Biranowska, J., Kowianski, P., Majak, K. & Morys, J. (1998) The projection of the amygdaloid nuclei to various areas of the limbic cortex in the rat. *Folia Morphol. (Warsz)*, **57**, 301–308.

Ehrlich, I., Humeau, Y., Grenier, F., Ciocchi, S., Herry, C. & Luthi, A. (2009) Amygdala inhibitory circuits and the control of fear memory. *Neuron*, **62**, 757–771.

Esber, G.R. & Haselgrove, M. (2011) Reconciling the influence of predictiveness and uncertainty on stimulus salience: a model of attention in associative learning. *Proc. Biol. Sci.*, **278**, 2553–2561.

Fiorillo, C.D., Tobler, P.N. & Schultz, W. (2003) Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, **299**, 1856–1902.

Fiorillo, C.D., Newsome, W.T. & Schultz, W. (2008) The temporal precision of reward prediction in dopamine neurons. *Nat. Neurosci.*, **11**, 966–973.

Gallagher, M. (2000). The amygdala and associative learning. In Aggleton, J.P. (Ed.), *The Amygdala: A Functional Analysis*. Oxford University Press, Oxford, pp. 311–330.

Gallagher, M., Graham, P.W. & Holland, P.C. (1990) The amygdala central nucleus and appetitive Pavlovian conditioning: lesions impair one class of conditioned behavior. *J. Neurosci.*, **10**, 1906–1911.

Hare, T.A., O'Doherty, J., Camerer, C.F., Schultz, W. & Rangel, A. (2008) Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *J. Neurosci.*, **28**, 5623–5630.

Hayden, B.Y., Heilbronner, S.R., Pearson, J.M. & Platt, M.L. (2011) Surprise signals in anterior cingulate cortex: neuronal encoding of unsigned reward prediction errors driving adjustment in behavior. *J. Neurosci.*, **31**, 4178–4187.

Hikosaka, O. (2010) The habenula: from stress evasion to value-based decision-making. *Nat. Rev. Neurosci.*, **11**, 503–513.

Hillman, K.L. & Bilkey, D.K. (2010) Neurons in the rat anterior cingulate cortex dynamically encode cost-benefit in a spatial decision-making task. *J. Neurosci.*, **30**, 7705–7713.

Holland, P.C. & Gallagher, M. (1993a) Amygdala central nucleus lesions disrupt increments, but not decrements, in conditioned stimulus processing. *Behav. Neurosci.*, **107**, 246–253.

Holland, P.C. & Gallagher, M. (1993b) Effects of amygdala central nucleus lesions on blocking an d unblocking. *Behav. Neurosci.*, **107**, 235–245.

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

Holland, P.C. & Kenmuir, C. (2005) Variations in unconditioned stimulus processing in unblocking. *J. Exp. Psychol. Anim. Behav. Process.*, **31**, 155–171.

Hollerman, J.R. & Schultz, W. (1998) Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat. Neurosci.*, **1**, 304–309.

Holroyd, C.B. & Coles, M.G. (2002) The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol. Rev.*, **109**, 679–709.

Hong, S. & Hikosaka, O. (2008) The globus pallidus sends reward-related signals to the lateral habenula. *Neuron*, **60**, 720–729.

Hong, S., Jhou, T.C., Smith, M., Saleem, K.S. & Hikosaka, O. (2011) Negative reward signals from the lateral habenula to dopamine neurons are mediated by rostromedial tegmental nucleus in primates. *J. Neurosci.*, **31**, 11457–11471.

Houk, J.C., Adams, J.L. & Barto, A.G. (1995). A model of how the basal ganglia generate and use neural signals that predict reinforcement. In Houk, J.C., Davis, J.L. & Beiser, D.G. (Eds), *Models of Information Processing in the Basal Ganglia*. MIT Press, Cambridge, pp. 249–270.

Ito, S., Stuphorn, V., Brown, J.W. & Schall, J.D. (2003) Performance monitoring by the anterior cingulate cortex during saccade countermanding. *Science*, **302**, 120–122.

Izquierdo, A.D., Suda, R.K. & Murray, E.A. (2004) Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. *J. Neurosci.*, **24**, 7540–7548.

Jhou, T.C., Fields, H.L., Baxter, M.G., Saper, C.B. & Holland, P.C. (2009) The rostromedial tegmental nucleus (RMTg), a GABAergic afferent to midbrain dopamine neurons, encodes aversive stimuli and inhibits motor responses. *Neuron*, **61**, 786–800.

Ji, H. & Shepard, P.D. (2007) Lateral habenula stimulation inhibits rat midbrain dopamine neurons through a GABA(A) receptor-mediated mechanism. *J. Neurosci.*, **27**, 6923–6930.

Johansen, J.P., Tarpley, J.W., LeDoux, J.E. & Blair, H.T. (2010) Neural substrates for expectation-modulated fear learning in the amygdala and periaqueductal gray. *Nat. Neurosci.*, **13**, 979–986.

Kakade, S. & Dayan, P. (2002) Acquisition and extinction in autoshaping. *Psychol. Rev.*, **109**, 533–544.

Kamin, L.J. (1969). Predictability, suprise, attention, and conditioning. In Campbell, B.A. & Church, R.M. (Eds), *Punishment and Aversive Behavior*. Appleton-Century-Crofts, New York, NY, pp. 242–259.

Kapur, S. (2003) Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *Am. J. Psychiatry*, **160**, 13–23.

Kaufling, J., Veinante, P., Pawlowski, S.A., Freund-Mercier, M.J. & Barrot, M. (2009) Afferents to the GABAergic tail of the ventral tegmental area in the rat. *J. Comp. Neurol.*, **513**, 597–621.

Kennerley, S.W. & Wallis, J.D. (2009) Evaluating choices by single neurons in the frontal lobe: outcome value encoded across multiple decision variables. *Eur. J. Neurosci.*, **29**, 2061–2073.

Kennerley, S.W., Walton, M.E., Behrens, T.E., Buckley, M.J. & Rushworth, M.F. (2006) Optimal decision making and the anterior cingulate cortex. *Nat. Neurosci.*, **9**, 940–947.

Kennerley, S.W., Dahmubed, A.F., Lara, A.H. & Wallis, J.D. (2009) Neurons in the frontal lobe encode the value of multiple decision variables. *J. Cogn. Neurosci.*, **21**, 1162–1178.

Knutson, B. & Gibbs, S.E.B. (2007) Linking nucleus accumbens dopamine and blood oxygenation. *Psychopharmacology*, **191**, 813–822.

Kobayashi, S. & Schultz, W. (2008) Influence of reward delays on responses of dopamine neurons. *J. Neurosci.*, **28**, 7837–7846.

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

LeDoux, J.E. (2000). The amygdala and emotion: a view through fear. In Aggleton, J.P. (Ed.), *The Amygdala: A Functional Analysis*. Oxford University Press, New York, NY, pp. 289–310.

Lepelley, M.E. (2004) The role of associative history in models of associative learning: a selective review and a hybrid model. *Q. J. Exp. Psychol.*, **57**, 193–243.

Li, J., Schiller, D., Schoenbaum, G., Phelps, E.A. & Daw, N.D. (2011) Differential roles of human striatum and amygdala in associative learning. *Nat. Neurosci.*, **14**, 1250–1252.

Mackintosh, N.J. (1975) A theory of attention: variations in the associability of stimuli with reinforcement. *Psychol. Rev.*, **82**, 276–298.

Magno, E., Foxe, J.J., Molholm, S., Robertson, I.H. & Garavan, H. (2006) The anterior cingulate and error avoidance. *J. Neurosci.*, **26**, 4769–4773.

Margolis, E.B., Lock, H., Hjelmstad, G.O. & Fields, H.L. (2006) The ventral tegmental area revisited: is there an electrophysiological marker for dopaminergic neurons? *J. Physiol.*, **577**, 907–924.

Margolis, E.B., Mitchell, J.M., Ishikawa, Y., Hjelmstad, G.O. & Fields, H.L. (2008) Midbrain dopamine neurons: projection target determines action potential duration and dopamine D(2) receptor inhibition. *J. Neurosci.*, **28**, 8908–8913.

Matsumoto, M. & Hikosaka, O. (2007) Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature*, **447**, 1111–1115.

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

Matsumoto, M., Matsumoto, K., Abe, H. & Tanaka, K. (2007) Medial prefrontal cell activity signaling prediction errors of action values. *Nat. Neurosci.*, **10**, 647–656.

McClure, S.M., Berns, G.S. & Montague, P.R. (2003) Temporal prediction errors in a passive learning task activate human striatum. *Neuron*, **38**, 339–346.

McLaughlin, R.J. & Floresco, S.B. (2007) The role of different subregions of the basolateral amygdala in cue-induced reinstatement and extinction of food-seeking behavior. *Neuroscience*, **146**, 1484–1494.

Mileykovskiy, B. & Morales, M. (2011) Duration of inhibition of ventral tegmental area dopamine neurons encodes a level of conditioned fear. *J. Neurosci.*, **31**, 7471–7476.

Mirenowicz, J. & Schultz, W. (1994) Importance of unpredictability for reward responses in primate dopamine neurons. *J. Neurophysiol.*, **72**, 1024–1027.

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

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

Murray, E.A. (2007) The amygdala, reward and emotion. *Trends Cogn. Sci.*, **11**, 489–497.

O'Doherty, J., Dayan, P., Friston, K.J., Critchley, H. & Dolan, R.J. (2003) Temporal difference learning model accounts for responses in human ventral striatum and orbitofrontal cortex during Pavlovian appetitive learning. *Neuron*, **38**, 329–337.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K.J. & Dolan, R.J. (2004) Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, **304**, 452–454.

Oliveira, F.T., McDonald, J.J. & Goodman, D. (2007) Performance monitoring in the anterior cingulate is not all error related: expectancy deviation and the representation of action-outcome associations. *J. Cogn. Neurosci.*, **19**, 1994–2004.

Omelchenko, N., Bell, R. & Sesack, S.R. (2009) Lateral habenula projections to dopamine and GABA neurons in the rat ventral tegmental area. *Eur. J. Neurosci.*, **30**, 1239–1250.

Ostlund, S.B. & Balleine, B.W. (2007) Orbitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental learning. *J. Neurosci.*, **27**, 4819–4825.

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

Paus, T. (2001) Primate anterior cingulate cortex: where motor control, drive and cognition interface. *Nat. Rev. Neurosci.*, **2**, 417–424.

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

Pearce, J.M. & Mackintosh, N.J. (2010). Two theories of attention: a review and a possible integration. In Mitchell, C.J. & LePelley, M.E. (Eds), *Attention and Associative Learning: From Brain to Behaviour*. Oxford University Press, Oxford, UK, pp. 11–39.

Pearce, J.M., Kaye, H. & Hall, G. (1982). Predictive accuracy and stimulus associability: development of a model for Pavlovian learning. In Commons, M.L., Herrnstein, R.J. & Wagner, A.R. (Eds), *Quantitative Analyses of Behavior*. Ballinger, Cambridge, MA, pp. 241–255.

Pessiglione, M., Seymour, P., Flandin, G., Dolan, R.J. & Frith, C.D. (2006) Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, **442**, 1042–1045.

Pickens, C.L., Setlow, B., Saddoris, M.P., Gallagher, M., Holland, P.C. & Schoenbaum, G. (2003) Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. *J. Neurosci.*, **23**, 11078–11084.

Quilodran, R., Rothe, M. & Procyk, E. (2008) Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron*, **57**, 314–325.

Redgrave, P. & Gurney, K. (2006) The short-latency dopamine signal: a role in discovering novel actions? *Nat. Rev. Neurosci.*, **7**, 967–975.

Rescorla, R.A. & Wagner, A.R. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In Black, A.H. & Prokasy, W.F. (Eds), *Classical Conditioning II: Current Research and Theory*. Appleton-Century-Crofts, New York, NY, pp. 64–99.

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

Roesch, M.R., Calu, D.J., Esber, G.R. & Schoenbaum, G. (2010) Neural correlates of variations in event processing during learning in basolateral amygdala. *J. Neurosci.*, **30**, 2464–2471.

Rothe, M., Quilodran, R., Sallet, J. & Procyk, E. (2011) Coordination of high gamma activity in anterior cingulate and lateral prefrontal cortical areas during adaptation. *J. Neurosci.*, **31**, 11110–11117.

Rudebeck, P.H., Bannerman, D.M. & Rushworth, M.F. (2008) The contribution of distinct subregions of the ventromedial frontal cortex to emotion, social behavior, and decision making. *Cogn. Affect. Behav. Neurosci.*, **8**, 485–497.

Rushworth, M.F. & Behrens, T.E. (2008) Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat. Neurosci.*, **11**, 389–397.

Rushworth, M.F., Walton, M.E., Kennerley, S.W. & Bannerman, D.M. (2004) Action sets and decisions in the medial frontal cortex. *Trends Cogn. Sci.*, **8**, 410–417.

Rushworth, M.F., Behrens, T.E., Rudebeck, P.H. & Walton, M.E. (2007) Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour. *Trends Cogn. Sci.*, **11**, 168–176.

Sallet, J., Quilodran, R., Rothe, M., Vezoli, J., Joseph, J.P. & Procyk, E. (2007) Expectations, gains, and losses in the anterior cingulate cortex. *Cogn. Affect. Behav. Neurosci.*, **7**, 327–336.

Scheffers, M.K. & Coles, M.G. (2000) Performance monitoring in a confusing world: error-related brain activity, judgments of response accuracy, and types of errors. *J. Exp. Psychol. Hum. Percept. Perform.*, **26**, 141–151.

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

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

Simon, D.A. & Daw, N.D. (in press) Neural correlates of forward planning in a spatial decision task in humans. *J. Neurosci.*, **31**, 5526–5539.

Sripanidkulchai, K., Sripanidkulchai, B. & Wyss, J.M. (1984) The cortical projection of the basolateral amygdaloid nucleus in the rat: a retrograde fluorescent dye study. *J. Comp. Neurol.*, **229**, 419–431.

Sutton, R.S. (1988) Learning to predict by the method of temporal difference. *Machine Learning*, **3**, 9–44.

Sutton, R.S. (1992) Adapting bias by gradient descent: an incremental version of delta-bar-delta. *Proceedings of the Tenth National Conference on Artificial Intelligence*, pp. 171–176.

Takahashi, Y.K., Roesch, M.R., Wilson, R.C., Toreson, K., O'Donnell, P., Niv, Y. & Schoenbaum, G. (2011) Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. *Nat. Neurosci.*, **14**, 1590–1597.

Taylor, S.F., Phan, K.L., Britton, J.C. & Liberzon, I. (2005) Neural responses to emotional salience in schizophrenia. *Neuropsychopharmacology*, **30**, 984–995.

Tobler, P.N., Dickinson, A. & Schultz, W. (2003) Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. *J. Neurosci.*, **23**, 10402–10410.

Totah, N.K., Kim, Y.B., Homayoun, H. & Moghaddam, B. (2009) Anterior cingulate neurons represent errors and preparatory attention within the same behavioral sequence. *J. Neurosci.*, **29**, 6418–6426.

Tye, K.M., Stuber, G.D., De Ridder, B., Bonci, A. & Janak, P.H. (2008) Rapid strengthening of thalamo-amygdala synapses mediates cue-reward learning. *Nature*, **453**, 1253–1257.

Tye, K.M., Cone, J.J., Schairer, W.W. & Janak, P.H. (2010) Amygdala neural encoding of the absence of reward during extinction. *J. Neurosci.*, **30**, 116–125.

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

Waelti, P., Dickinson, A. & Schultz, W. (2001) Dopamine responses comply with basic assumptions of formal learning theory. *Nature*, **412**, 43–48.

Wallis, J.D. & Kennerley, S.W. (2010) Heterogeneous reward signals in prefrontal cortex. *Curr. Opin. Neurobiol.*, **20**, 191–198.

Walton, M.E., Devlin, J.T. & Rushworth, M.F. (2004) Interactions between decision making and performance monitoring within prefrontal cortex. *Nat. Neurosci.*, **7**, 1259–1265.