

Surprise Signals in Anterior Cingulate Cortex: Neuronal Encoding of Unsigned Reward Prediction Errors Driving Adjustment in Behavior

Benjamin Y. Hayden,^{1,2} Sarah R. Heilbronner,^{1,2} John M. Pearson,^{1,2} and Michael L. Platt^{1,2,3}

¹Department of Neurobiology, Duke University School of Medicine, ²Center for Cognitive Neuroscience, and ³Department of Evolutionary Anthropology, Duke University, Durham, North Carolina 27710

In attentional models of learning, associations between actions and subsequent rewards are stronger when outcomes are surprising, regardless of their valence. Despite the behavioral evidence that surprising outcomes drive learning, neural correlates of unsigned reward prediction errors remain elusive. Here we show that in a probabilistic choice task, trial-to-trial variations in preference track outcome surprisingness. Concordant with this behavioral pattern, responses of neurons in macaque (*Macaca mulatta*) dorsal anterior cingulate cortex (dACC) to both large and small rewards were enhanced when the outcome was surprising. Moreover, when, on some trials, probabilities were hidden, neuronal responses to rewards were reduced, consistent with the idea that the absence of clear expectations diminishes surprise. These patterns are inconsistent with the idea that dACC neurons track signed errors in reward prediction, as dopamine neurons do. Our results also indicate that dACC neurons do not signal conflict. In the context of other studies of dACC function, these results suggest a link between reward-related modulations in dACC activity and attention and motor control processes involved in behavioral adjustment. More speculatively, these data point to a harmonious integration between reward and learning accounts of ACC function on one hand, and attention and cognitive control accounts on the other.

Introduction

Learning theory seeks to describe how we and other animals form associations between stimuli, actions, and their consequences for reward or punishment. Reinforcement learning (RL) holds that learning depends primarily on reward prediction errors (RPEs), the difference between the reward expected and the reward received (Rescorla and Wagner, 1972; Sutton and Barto, 1998). RL accounts for a wide array of phenomena, including Pavlovian and instrumental conditioning, and appears to be supported by subcortical and cortical signaling systems including striatum, midbrain, orbitofrontal cortex, and anterior cingulate cortex (Lauwereyns et al., 2002; Barraclough et al., 2004; Samejima et al., 2005; Schultz, 2006; Lee and Seo, 2007; Matsumoto and Hikosaka, 2007; Rushworth et al., 2007).

Aside from the reward prediction error, which is a signed quantity, behavior sometimes depends on the degree to which outcomes are surprising, independent of their sign (Mackintosh, 1975; Hall and Pearce, 1979; Pearce and Hall, 1980; Courville et al., 2006). Corresponding models include a term called surprisingness or associability, which is a function of the absolute value of the difference between observed and expected outcomes

(Mackintosh, 1975; Hall and Pearce, 1979; Pearce and Hall, 1980). These attentional models of learning posit that surprising events marshal neural resources to enhance their processing, thereby driving learning. Several experiments provide empirical support for this idea (Kaye and Pearce, 1984; Swan and Pearce, 1988; Courville et al., 2006).

Much evidence links the amygdala and its target, the basal forebrain, with surprise-based learning (Holland and Gallagher, 1999, 2006; Belova et al., 2007; Lin and Nicolelis, 2008). The connections between the ACC and parts of the amygdala (Morecraft and Van Hoesen, 1993) invite the hypothesis that the ACC plays a complementary role. The dorsal anterior cingulate cortex (dACC) is a primary target of dopamine neurons (Paus, 2001), tracks reward outcomes of choices (Procyk et al., 2000; Shidara and Richmond, 2002; Williams et al., 2004; Amiez et al., 2006; Quilodran et al., 2008), and is involved in action and outcome monitoring (Shima and Tanji, 1998; Gehring and Willoughby, 2002; Holroyd and Coles, 2002; Matsumoto et al., 2007). However, the form of the reward outcome signal carried by dACC neurons remains unknown.

Here we show that, in a probabilistic choice task (Hayden et al., 2010), dACC neurons signal the surprisingness of reward outcomes. On each trial of the task, monkeys chose between two targets offering stochastic (large or small) juice rewards with probabilities specified by symbolic cues. We thus obtained neural responses to large and small rewards cued by a large range of probabilities. On a subset of trials, we introduced an occluder that obscured information about reward probabilities. These ambiguous cues provide no cue to the likelihood of an outcome, so

Received Sept. 3, 2010; revised Jan. 18, 2011; accepted Jan. 24, 2011.

This work was supported by National Institutes of Health Grants R01 EY013496 (to M.L.P.) and K99 DA027718 (to B.Y.H.), and a fellowship from the Tourette Syndrome Association (to B.Y.H.). We thank Karli Watson for help in training the animals and Steve Chang for useful comments in the analysis.

Correspondence should be addressed to Benjamin Y. Hayden, Department of Neurobiology, Duke University School of Medicine, Durham, NC 27710. E-mail: hayden@neuro.duke.edu.

DOI:10.1523/JNEUROSCI.4652-10.2011

Copyright © 2011 the authors 0270-6474/11/314178-10\$15.00/0

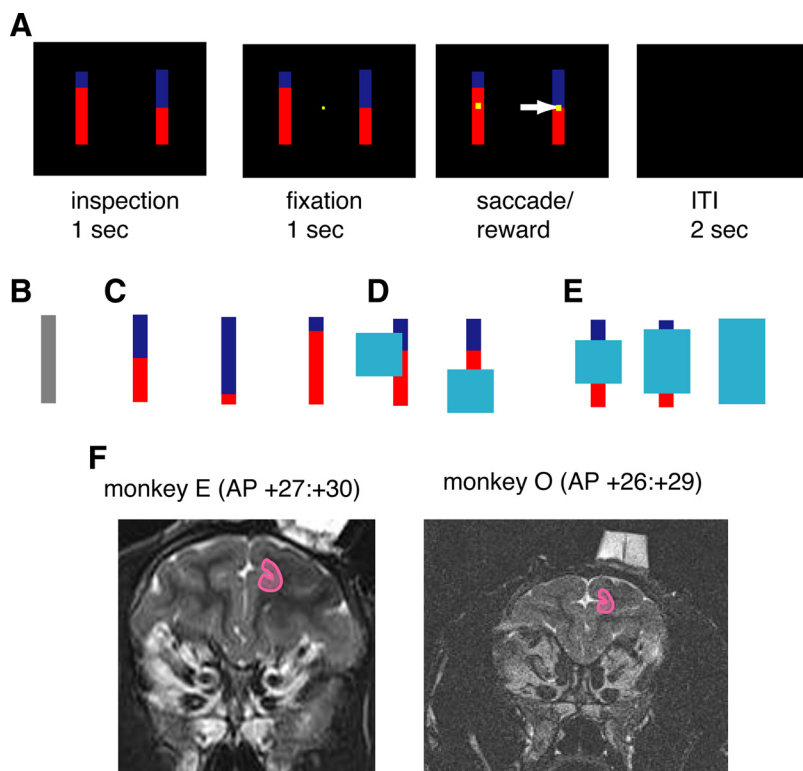


Figure 1. Stimuli, task, and recording location. **A**, Task design (details in Materials and Methods). **B**, Gray bar, safe stimulus, yields 200 μ l of juice. **C**, Examples of risky bars. Blue/red bars yield either 333 or 67 μ l of juice; $p(\text{Large})$ varied from 0 to 1. In this example, $p(\text{Large})$ are 0.5, 0.88, and 0.17. **D**, Example of risky bar with a partially concealing occluder that did not render probabilities uncertain. **E**, Ambiguous stimuli. The size of the occluder rendered the probability associated with the bar ambiguous (three sizes or occluders were used; data were averaged over all three). **F**, Approximate recording site locations within the anterior cingulate cortex shown on a coronal MRI section in the two subjects. Approximate rostrocaudal recording locations shown above; approximate mediolateral and dorsoventral recording positions shown in pink.

any particular reward can be considered unsurprising. We therefore predicted—and indeed observed—weaker modulation of reward-related neuronal responses on ambiguous trials. These observations endorse the idea that a population of neurons in the dACC signal an unsigned error in reward prediction and are thus consistent with the idea that dACC contributes to the attentional component of learning.

Materials and Methods

Some of the behavioral data collected for this experiment have been published previously (Hayden et al., 2010). However, all of the figures and analyses presented here are new and the physiological data have not been previously published.

Surgical procedures. All animal procedures were approved by the Duke University Institutional Animal Care and Use Committee and were designed and conducted in compliance with the Public Health Service's Guide for the Care and Use of Animals. Two male rhesus monkeys (*Macaca mulatta*) served as subjects. A small prosthesis for holding the head was used. Animals were habituated to laboratory conditions and then trained to perform oculomotor tasks for liquid reward. A stainless steel recording chamber (Crist Instruments) was placed over anterior cingulate cortex and verified by MRI (Hayden et al., 2009). Animals received appropriate analgesics and antibiotics after all procedures. Throughout both behavioral and physiological recording sessions, the chamber was kept sterile with regular antibiotic washes and sealed with sterile caps.

Behavioral techniques. Monkeys were placed on controlled access to fluid outside of experimental sessions. Eye position was sampled at 1000 Hz by an infrared eye-monitoring camera system (SR Research). Stimuli were controlled by a computer running Matlab (Mathworks) with Psy-

chtoolbox (Brainard, 1997) and Eyelink Toolbox (Cornelissen et al., 2002). Visual stimuli were colored rectangles on a computer monitor placed directly in front of the animal and centered on his eyes (Fig. 1). A standard solenoid valve controlled the duration of juice delivery. Reward volume was 67, 200, or 333 μ l in all cases. The accuracy and linearity of volume as a function of time of the solenoid valve was checked immediately before and after the experiment.

Every trial began when two bars and one occluder appeared (Fig. 1A). The monkey had one second to inspect these stimuli. Casual observation showed that monkeys reliably looked at both bars during this period. Next, a small yellow fixation point appeared at the center of the monitor. Once fixation was acquired ($\pm 0.5^\circ$), the monkey had to maintain fixation for one second. Any failure led to an “incorrect” sign (a large green square) and a timeout period (3 s). The fixation point was then extinguished, and two eccentric small yellow squares appeared overlaid on the center of the probability bars. The monkey then had to select one of these bars by shifting gaze to the square superimposed on the bar ($\pm 3^\circ$). Following the saccade, the gamble was immediately resolved by the computer and the appropriate reward was provided. Then all stimuli disappeared. No visual cue indicated the reward outcome, nor were the reward probabilities for the ambiguous stimuli ever revealed.

On 10% of trials (chosen randomly), one of the targets was safe and provided a deterministic reward with 100% probability. On most trials, both targets were risky (risky trials), and both offered a gamble between a large (0.333 ml) and a small (0.067 ml) squirt of juice at a fixed, fully specified, probability. On one-third of trials (ambiguous trials), one of the two risky targets was occluded, rendering its reward probabilities uncertain, or, formally speaking, ambiguous. These two stochastic processes were independent, so that 1/30 trials pitted an ambiguous against a safe option. The number of safe trials was too low to fully analyze on a neuron by neuron basis, and neural activity on these trials was not studied.

On all trials, a cyan occluder appeared somewhere on the screen. On two-thirds of trials (risky trials), the occluder appeared at a random location on the screen, and sometimes covered part of the bar but did not obscure information about probabilities. On one-third of trials (ambiguous trials), the occluder overlapped the center of one of the bars. The probability that the ambiguous option would provide a large reward was drawn from a uniform distribution of the probabilities within the range of those obscured by the occluder, and an outcome was chosen accordingly. This is mathematically equivalent to a 50% probability of a large reward on all ambiguous trials. On a small minority of risky trials, the occluder covered only part of the bar; on all such trials, the border between the blue and red regions was visible, and these trials were considered risky (rather than ambiguous) in all analyses. Ambiguous stimuli were equally likely to have small, medium, or large size (i.e., one-third likelihood of each). There were too few trials of each size to detect firing rate differences, so these three trial types were combined in all analyses.

The safe bar was colored gray. The risky and ambiguous bars were divided into blue and red portions. The blue portion, always on top, indicated the probability that choosing the bar would yield the large reward. The red portion indicated the probability that choosing the bar would yield the small reward. All probability bars were 80 pixels wide and 300 pixels tall. The occluder was 150, 225, or 300 pixels tall and always 200 pixels wide. The horizontal position of the occluder was randomly

jittered so as to emphasize the idea that the bar just happened to be covering the stimulus, and that the ambiguous was not a single distinct stimulus; in all cases, information about probabilities was obscured. On ambiguous trials, the vertical position of the occluder was always centered on the center of the bar.

Microelectrode recording techniques. We recorded action potentials from single neurons in two monkeys during the performance of the task. Single electrodes (Frederick Haer) were lowered with a hydraulic micro-drive (Kopf) until the waveform of a single (1–3) neuron(s) was isolated. Individual action potentials were identified by standard criteria and isolated on a Plexon system. Neurons were selected for recording on the basis of the quality of isolation only, and not on task-related response properties.

We approached dACC through a standard recording grid. Dorsal ACC was identified by structural MRI images taken before the experiment. Neuroimaging was performed at the Center for Advanced Magnetic Development at Duke University Medical Center, on a 3T Siemens Medical Systems Trio MR Imaging Instrument using 1 mm slices. We confirmed that we were in dACC by using stereotactic measurements, as well as by listening for characteristic sounds of white and gray matter during recording. Our recordings likely came from area 24 and the dorsal and ventral banks of the anterior cingulate sulcus. Before recording, we performed several exploratory recording sessions to map out the physiological response properties of the tissue accessible through our recording chamber. We were able to distinguish white from gray matter by the presence of neural activity and by the distinct sounds associated with gray matter. During these mapping sessions, we were able to identify both the dorsoventral and mediolateral extent of the cingulate sulcus.

Behavioral and neuronal analyses. Peristimulus time histograms (PSTHs) were constructed by aligning spike rasters to trial events and averaging firing rates across multiple trials. Firing rates were calculated in 1 ms bins. For display, PSTHs were Gaussian-smoothed (SD, 100 ms). Data were aligned to the saccade that ended the trial (time 0 on plots). Statistical comparisons were performed on binned, unsmoothed, firing rates of single neurons in a 1 s postreward epoch beginning 0.5 s after the end of the choice saccade (Student's *t* test on individual trials counts). All statistical tests were confirmed with nonparametric statistical tests, and similar results were obtained in all cases.

Results

Monkeys are risk-seeking and avoid ambiguous options

We examined behavior of two monkeys in a probabilistic choice task (Fig. 1A–E). As reported previously (Hayden et al., 2010), when choosing between two risky options, monkeys reliably discriminated bars that differed by as little as 0.04 in the probability of a large reward (i.e., 12 pixels, $p = 0.026$ for 4%, $p < 0.001$ for larger differences, binomial test). Previously published work using a variant of this task that included a gray bar to indicate probability of medium sized rewards demonstrated that monkeys make their choices by considering the size of both the red and blue portions of the bar to compute the expected value of each option (Hayden et al., 2010). In the standard variant of the task considered here, both monkeys preferred risky options to safe options with the same expected value ($p < 0.001$ in both cases). Each risky option can be defined by the probability it would provide a large reward, and we will use the term $p(\text{Large})$ to denote this quantity. As the $p(\text{Large})$, and thus the expected value, of risky options fell, monkeys continued to prefer them over safe options until $p(\text{Large})$ fell to ~ 0.23 . In terms of expected value, monkeys sacrificed up to 71 μl of juice (± 7.4 SE) to choose the risky option instead of the safe option.

Both monkeys preferred risky options to ambiguous options. As the $p(\text{Large})$, and thus the expected value, of risky options fell, monkeys continued to prefer them over ambiguous ones until $p(\text{Large})$ fell to ~ 0.30 . In terms of expected value, monkeys sacrificed 53 μl of juice to choose the risky over the ambiguous

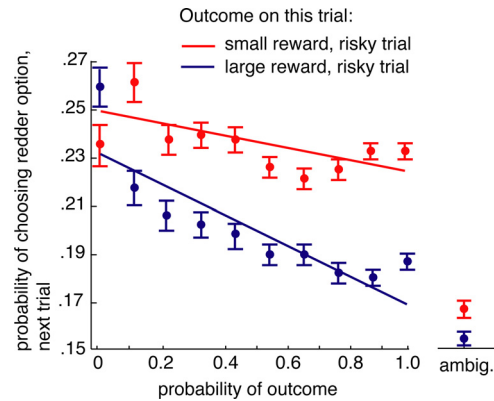


Figure 2. Likelihood of adjusting strategy depends on reward size and surprise. Plot of probability that monkey chose inferior (i.e., option with lower probability of large reward) option on next trial as a function of reward size (color) and its associated probability (abscissa) on current trial. Likelihood of switching to this strategy was greatest following unexpected small outcomes and smallest following expected large outcomes. Data for all sessions are combined into single plot. ambig., Ambiguous.

option. These data demonstrate that monkeys distinguish between risky and ambiguous forms of uncertainty, and, like humans, are reluctant to choose gambles with uncertain probabilities (Ellsberg, 1961; Fox and Tversky, 1995; Hsu et al., 2005; Huettel et al., 2006; Hayden et al., 2010). Moreover, we found that ambiguity aversion gradually disappeared during additional training periods that occurred after physiological recordings were terminated (Hayden et al., 2010). This slow learning effect demonstrates that monkeys gradually learn to associate accurate expected values with ambiguous options, and that ambiguity aversion reflects discomfort with unfamiliar stimuli.

Switching behavior reflects the surprisingness of rewards

When choosing between two risky options in this task, monkeys should not, in principle, change their strategy as a function of the outcome of the previous trial, since trials were independent. Typically, monkeys sensibly chose the option offering a higher probability of large reward (81.8% of trials). However, monkeys also exhibited a weak but reliable increase in willingness to choose the option with the lower $p(\text{Large})$ (the redder of the two targets) following small rewards (Fig. 2). They also exhibited a weak but reliable increase in willingness to choose the option with the lower $p(\text{Large})$ following surprising (that is, statistically unlikely) outcomes, regardless of the size of the outcome. This main effect of reward size is unlikely to reflect trial-to-trial effects based on the choices made, since reward was largely independent of choice made: small rewards were almost as likely to come from choices of the bluer option (79.0% of smaller rewards) as were larger rewards (83.9% of larger rewards).

Improbable large rewards biased monkeys toward choosing the less probable (i.e., redder) option on the next trial (regression of switch likelihood against win probability, coefficient = -0.0014 , $p < 0.001$). Nearly identical results were observed for both monkeys individually (coefficient = 0.0013 for monkey E and 0.0015 for monkey O). One intuitive explanation for this increased willingness to choose the suboptimal strategy is that a surprising large reward (i.e., one obtained from a redder bar) suggests, despite the monkey's extensive experience, that something about the meaning of the stimulus has changed, and that perhaps the redder bars are now more rewarding, so a change in strategy could be worth trying. In contrast, an unsurprising large reward indicates that the environment is predictable and that the

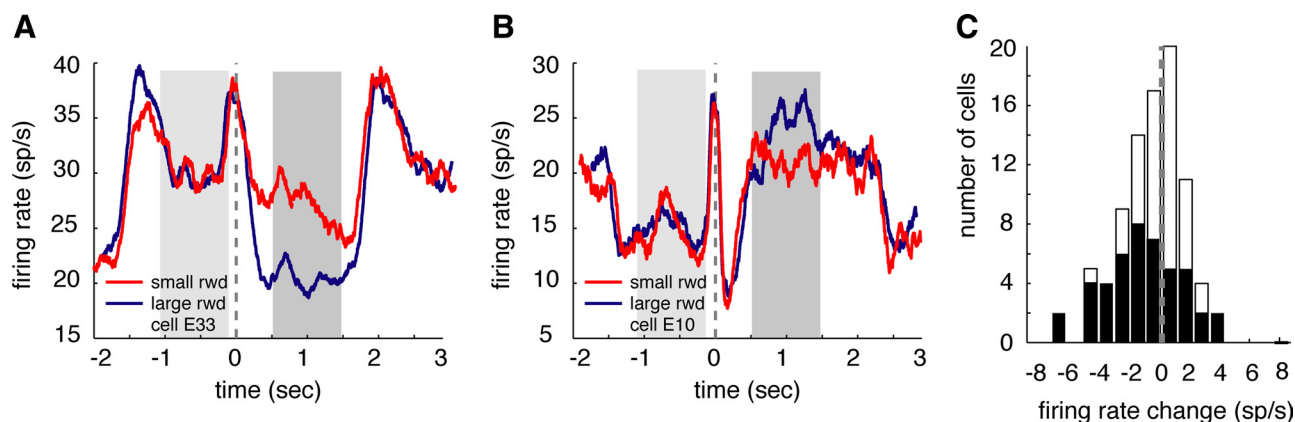


Figure 3. dACC neurons signal reward (rwd) outcomes in probabilistic choice task. **A**, PSTH for a single dACC neuron showing average responses to large and small outcomes, aligned to end of saccade and beginning of reward (time = 0). Responses were larger following small rewards than following large rewards. Epoch of analysis begins 0.5 s after reward and lasts 1 s (dark gray box). Epoch used in the fixation period analysis (Table 1) is shown as well (light gray box). **B**, PSTH for another dACC neuron showing the opposite response pattern. **C**, Average difference in response to large and small outcomes of all neurons during postreward epoch (dark gray regions). Responses were more often smaller following large rewards than following small rewards (left of zero).

current strategy is working. In the framework of Bayesian reinforcement learning, increases in decisional uncertainty should drive faster learning and more rapid changes in behavior and exploration (Courville et al., 2006). In this vein, we suggest that choosing the redder option may be a form of exploratory choice.

We found that unexpected small rewards promoted larger willingness to choose the redder option than expected small rewards and all large rewards (regression of switch likelihood against win probability, coefficient = -0.0019 , $p < 0.001$). Nearly identical results were observed for both monkeys individually (coefficient = 0.0021 for monkey E and 0.0018 for monkey O). One intuitive explanation for this effect is that unexpected small rewards imply that the meanings of the stimuli may have changed, just as in the case of the surprising large reward, but that the situation is more urgent because circumstances have become less favorable, more forcefully calling for an adjustment of choice strategy (cf. Courville et al., 2006; Yu and Dayan, 2005).

Monkeys were relatively unlikely to choose the less probable redder option after they had chosen the ambiguous option, regardless of the reward obtained ($p < 0.001$ for monkey E and $p < 0.01$ for monkey O, binomial test) (Fig. 2, right). These effects are consistent with the idea that, because ambiguous options do not strongly predict either a large or a small reward, neither outcome is particularly surprising. Since no prediction was made, no expectation was violated, and there is little reason to change strategy following either reward from an ambiguous choice. These ideas are distinct from the notion that ambiguous options are simply valued less than risky options, or that they are perceived as having a lower probability of reward—both of which would simply cause them to be treated as low probability rewards (our interpretation is also supported by the neural data; see below).

Monkeys' increased willingness to choose the redder option following surprising outcomes might reflect an increased sampling of the alternative following an unexpected small reward. Another possibility is that monkeys followed a strategy of random guessing following an unexpectedly small reward and that this led them to choose the less probable option more often. Either way, the monkeys demonstrably responded to unexpected outcomes, good or bad, by adjusting some aspect of their strategy. This strategy-switching phenomenon is reminiscent of a win—stay/lose—shift heuristic (Barracough et al., 2004; Hayden et al., 2008), and also has intuitive connections with an explore (as opposed to exploit) strategy (Daw et al., 2006; Pearson et al., 2009).

We looked for other possible signatures of attentional effects on neural responses. Such possible signatures include changes in the likelihood of error commission, initiation time for the next trial, and duration of cue sampling. The change in the likelihood of loss of fixation was weak and not statistically significant (regression coefficient = 0.001 following large rewards and 0.003 following small rewards, $p > 0.05$ in both cases). There was no effect on trial initiation time (coefficient = 0.000 following large rewards, and -0.001 following small rewards, $p > 0.05$ in both cases). The duration of cue sampling also showed no effect (coefficient = 0.010 following large rewards, -0.002 following small rewards, $p > 0.05$ in both cases). One reason these effects may be elusive is that monkeys were so well trained in the task that performance was virtually asymptotic.

Collectively, these data indicate that monkeys, like humans, are sensitive not only to reward outcomes but how much these outcomes deviate from expectations. The observed adjustments follow two principles: first, that small rewards lead to greater adjustments than large rewards; and second, that adjustment is increased for surprising rewards, regardless of sign. This second effect indicates that behavioral changes are influenced by the unsigned reward prediction error of the outcome, similar to that predicted in attentional theories of learning (Courville et al., 2006; Lin and Nicolelis, 2008; Roesch et al., 2010). In such theories, the rate of learning is influenced by the amount of attention elicited by the outcomes, which in turn depends on the absolute value of the error of the prediction (Pearce and Hall, 1980). This learning-related form of attention is distinguished from other notions of attention that describe improvements in sensory detection or discrimination performance associated with prior knowledge or voluntary control (Desimone and Duncan, 1995; Egeth and Yantis, 1997).

dACC neurons signal unsigned reward prediction errors

We recorded the activity of 92 dACC neurons (61 in monkey E and 31 in monkey O) while monkeys performed the task (minimum 600 trials per neuron, mean 871 trials per neuron). Activity of an example neuron is shown in Figure 3A. This neuron had a baseline firing rate of ~ 20 spikes/s (sp/s) and showed clear enhancements around the time of the saccade that began the task (~ 1.25 s before time 0) and the choice saccade (time 0). In the epoch following the receipt of the reward, neuronal activity was 25.7% (0.9 sp/s) greater following small outcomes (red line) than

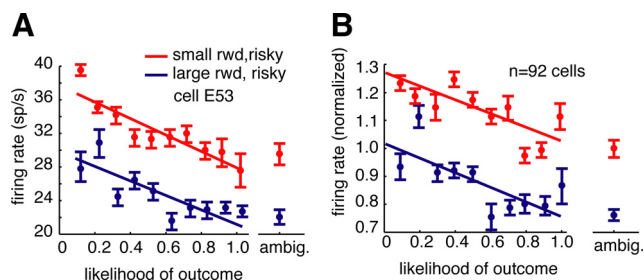


Figure 4. dACC neurons signal unsigned reward (rwd) prediction errors. **A**, Average firing rate of a single neuron during postreward epoch (Fig. 3A, dark gray box) separated by outcome (large or small reward, blue and red lines) and reward likelihood (divided into 10 equal bins, abscissa). Responses following ambiguous (ambig.) choices are shown to the right. Firing rates following large and small rewards are larger when they are unexpected. **B**, Average response of population of neurons, normalized to global average firing rate for the neuron during 0.5 s pretrial epoch.

following large ones (blue line) ($p < 0.001$, bootstrap t test on raw firing rates). Another neuron is shown in Figure 3B. This neuron showed a greater firing rate in response to large outcomes than to small ones. The activity of 55% of neurons ($n = 51/92$, 35 in monkey E, 16 in monkey O) signaled the size of the reward delivered following risky choices. The majority of significantly modulated neurons (71%, $n = 36/51$) showed greater responses to smaller rewards than to larger rewards (average modulation 2.2 sp/s) (Fig. 3C).

Neuronal activity on any given trial varied with reward size, but it also depended on the prior probability associated with that reward. For both the example neurons and the population, firing rates in the postreward epoch were greater following highly improbable large rewards than following probable large rewards (regression of firing rate against probability, coefficient 0.032 for the neuron and 0.018 for the population, $p < 0.02$ in both cases) (Fig. 4). Figure 4 shows the firing rate of a single neuron and the population following large and small rewards (blue and red lines, respectively), separated by the cued probability of those rewards (separated into deciles for ease of viewing). Similarly, firing rates were greater following improbable small rewards than following probable ones (regression of firing rate against probability, coefficient 0.022 for the neuron and 0.015 for the population, $p < 0.02$ in both cases). Thus, neuronal activity reflected the unsigned difference between predicted reward size and observed reward size. Interestingly, neuronal responses did not solely represent unsigned reward prediction error. Instead, they were greater following small outcomes than following large outcomes, suggesting that neural responses reflect the sum of reward size and surprise. This pattern closely mirrors the pattern observed for the likelihood of abandoning the optimal behavioral strategy and instead choosing the suboptimal strategy on the next trial as a function of the surprisingness of the reward (Fig. 2).

To estimate the prevalence of these effects across the population, we performed regressions of firing rate against reward probability separately for large and small rewards for each neuron in the population. We found a statistically significant regression coefficient ($p < 0.05$) following small rewards for 41% of neurons ($n = 38$; 27 from monkey E, 11 from monkey O) and a statistically significant coefficient following large rewards for 33% ($n = 30$; 21 from monkey E, 9 from monkey O) of neurons in the population. Of the neurons with a significant effect following large rewards, 83% (25/30) also showed a significant effect following small rewards. These findings demonstrate a systematic, largely monotonic relationship between violation of expect-

tations and firing rate for a given reward size. Regression coefficients were predominantly negative. For small rewards, regression coefficients were negative in 71% of significantly modulated neurons (27/38); for large rewards, they were negative in 67% (20/30). These frequencies are more than would be expected by chance (binomial test, $p < 0.01$) and are consistent with the idea that dACC neurons signal the surprisingness of reward outcomes.

One of the most intriguing and puzzling aspects of these data is that firing rates associated with surprising large rewards and fully expected small rewards were nearly identical. Indeed, a direct comparison of firing rates in the top quartile (i.e., most surprising) of large reward trials and bottom quartile (least surprising) of small responses revealed no statistical difference ($p > 0.5$ for the single neuron and for the population as a whole). Prima facie, this overlap appears to pose a problem for the downstream decoding system—how do readout neurons know whether a large or small reward was obtained? We conjecture that dACC responses do not need to be decoded, because it tracks outcome-related variables related to the likelihood of altering behavior, regardless of the cause (see Discussion, below) (Hayden et al., 2009). Consistent with this idea, we found no evidence that likelihood of strategy switching was different following the most surprising (top quartile) large rewards and least surprising (bottom quartile) small rewards ($p > 0.3$).

Neuronal responses to ambiguous options reflect reduced behavioral shifting

On one-third of trials, monkeys chose between a risky option and an ambiguous option. Following ambiguous trials, monkeys were less likely to follow the suboptimal strategy of choosing the redder option than following risky trials (Fig. 2). This behavioral effect suggests that learning is reduced on such trials, and this may be because the ambiguous stimuli do not make a strong prediction about a reward. We conjecture that because ambiguous options do not provide explicit information about reward likelihood, any outcome, large or small, may be treated as less surprising than the same outcome predicted by an equivalent risky option, in which reward probabilities are explicitly cued.

We thus predicted that neural responses to both large and small rewards following ambiguous choices would be lower than those following risky choices. This is indeed what we observed (Fig. 4). Firing rates following small rewards on ambiguous trials were significantly lower than those observed when $p(\text{Large})$ was 0.4 or higher, even though, in practice, all ambiguous options had $p(\text{Large})$ of 0.5 ($p < 0.005$ in each case, bootstrap t test). By the same token, firing rates following large rewards when monkeys chose the ambiguous option were significantly lower than those observed when $p(\text{Large})$ was 0.5 or lower ($p < 0.01$ in each case, bootstrap t test). Although preference for the ambiguous option gradually rose over the course of training (Hayden et al., 2010), we were not able to detect a corresponding change in neural modulation. It is difficult to draw any conclusions from this lack of an observed effect, however, as the change in preference was noisy, the neural effects were subtle, and the corresponding analyses must be conducted on separate samples of neurons collected on different days over the course of the study.

We considered the possibility that monkeys were simply pessimistic about ambiguous outcomes, and treated them as if they had a low probability of providing a large reward. However, our data provide two pieces of evidence against this interpretation. First, at a behavioral level, switching likelihood following ambiguous outcomes is significantly lower than following even the most likely small outcome trials (Student's t test, $p < 0.001$ in for

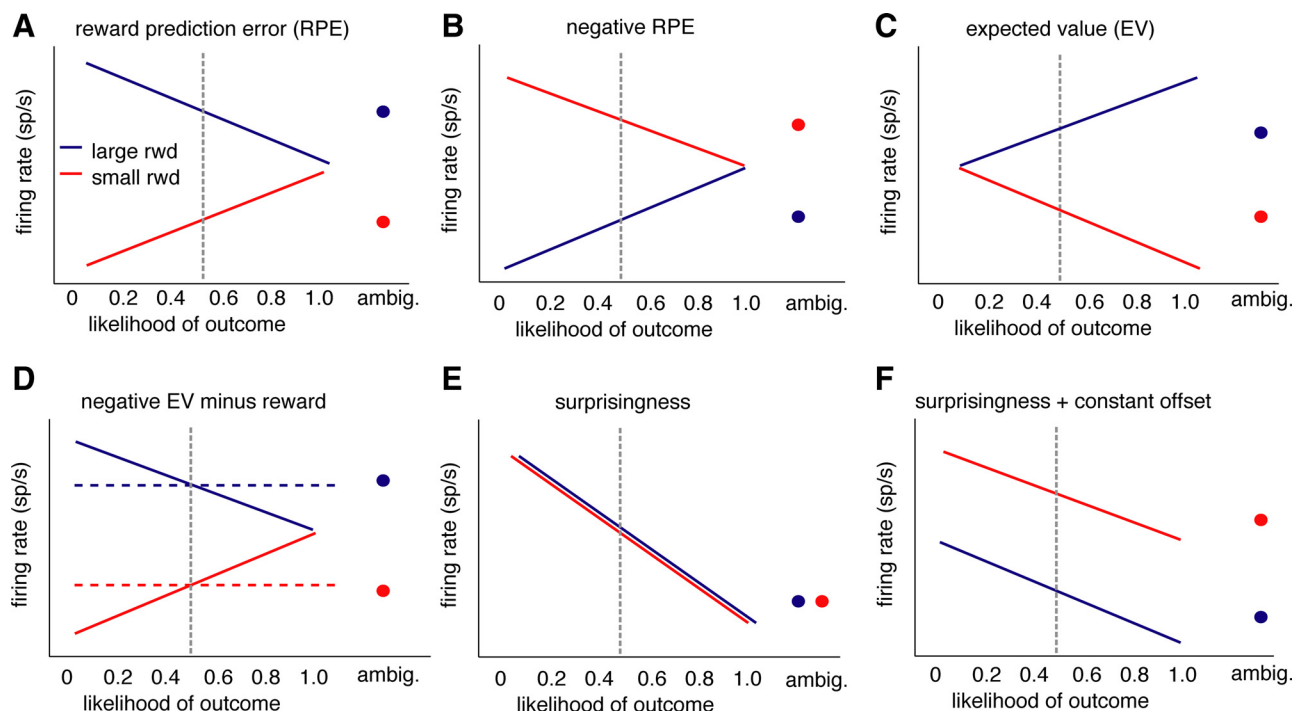


Figure 5. Schematic plots of behavior and neuronal responses expected from various learning models. **A**, A reward prediction error. **B–D**, The negative reward prediction error (**B**), a measure of expected value (EV) alone (**C**), and a negative expected value minus the reward (**D**) are other plausible models that fail to predict behavior in this task. **E, F**, Instead, behavior appears to match the surprisingness of the outcome (**E**) plus a constant term that is larger for negative than for positive outcomes (**F**).

both large and small ambiguous outcomes) (Fig. 2). Second, if monkeys treated ambiguous outcomes as having a low probability, then large outcomes would be surprising and would evoke a correspondingly high firing rate response. However, neural responses to large rewards from ambiguous trials were among the lowest observed, and significantly lower than those obtained from the five least likely deciles of large rewards for risky rewards (Fig. 4). Therefore, these data are inconsistent with the idea that monkeys were simply pessimistic about ambiguous options, and assigned them a lower expected value than they deserved. Our results are more consistent with the idea that monkeys treat ambiguous options as providing less information about outcomes than with the idea that they mistakenly assign them a lower probability.

Surprisingness uniquely accounts for these data

Collectively, these results demonstrate that the firing rates of neurons in dACC are influenced by both the size of the most recent reward and the cued probability that that reward would be delivered. This finding is inconsistent with the idea that dACC neurons encode either a pure value signal or a signed reward prediction error (i.e., the signed difference between expected and received rewards, RPE) (Fig. 5A). These results are also inconsistent with a simple transform of RPE, such as a negative RPE or a rectified RPE (Fig. 5B,C). They are also inconsistent with the idea that dACC signals the difference between the expected value of the outcome and the obtained outcome (Fig. 5D). In contrast, firing rates in dACC closely mirror outcome surprisingness (i.e., an unsigned prediction error) (Fig. 5E), combined with a constant offset for losses relative to gains (Fig. 5F). This pattern of firing rates is closely correlated with the observed effect of rewards on changes in behavior in our task, consistent with the idea that firing rates track the likelihood of changing behavior (Fig. 2).

If neurons reflect probability during the fixation period, then

Table 1. dACC neurons do not encode expected value or informational uncertainty (H) during the fixation period

	% Correlated	% Anticorrelated	Total	<i>p</i> Value, uncorrected
Probability left	3.26	3.26	6.52	0.799
Probability right	4.35	0	4.35	0.169
Probability of chosen	4.34	1.09	5.44	0.369
H left	2.17	6.52	8.69	0.047
H right	2.17	1.09	3.26	0.669
H chosen	3.26	2.17	5.44	0.881

We correlated firing rate with six parameters related to the bars on the screen. Firing rates of ACC neurons did not track five of these six values, and just barely tracked the sixth (H left). None of these values is significant when corrected for multiple comparisons. Significance tests were performed using a χ^2 .

these effects may influence outcome signals in the reward epoch. Another possibility is that neurons encode informational entropy, a measure which is maximal when probability is 0.5 and minimal when probabilities are 0 and 1. To test for these confounds, we examined, for each neuron in our sample, the correlation between firing rate in the fixation period (the 1 s period beginning with the acquisition of fixation and the extinction of the fixation spot) and these variables to determine how many neurons exhibited significant correlations. (Because firing rate may be anticorrelated with any of these variables, the test was two-tailed.) We tested the probability and entropy on the left, the right, and on the chosen target. The table below shows the proportion of neurons with significant modulation during the fixation period ($p < 0.05$) (Table 1). Given this significance cutoff, the proportion of neurons expected to exhibit each correlation by chance is 0.05 (4.6/92). These data indicate that dACC neurons do not encode reward probability during the fixation period, and only weakly encode entropy, if at all.

Conflict does not explain dACC neuron activity in this task

Several fMRI studies have implicated dACC in signaling response conflict (Carter et al., 2000; Botvinick et al., 2001; van Veen et al.,

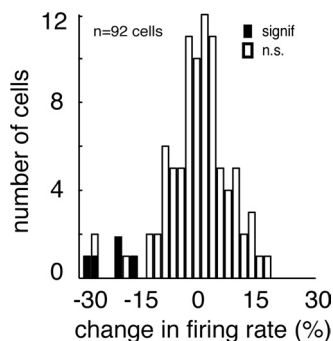


Figure 6. Conflict does not explain firing rate modulations of dACC neurons in this task. Bar histogram indicating average size of modulation of firing rates associated with conflict in dACC neurons. Horizontal axis indicates difference in firing rate during fixation epoch of task on high-conflict trials (in which targets were nearly indistinguishable and preferences were close to neutral) and low-conflict trials (in which targets were readily distinguishable). Very few neurons exhibited any significant modulation, and those that did (black bars) had greater activation on low-conflict trials (left side of graph). signif, Significant; n.s., not significant.

2001; Weissman et al., 2003; Kerns et al., 2004). Thus, the task-dependent changes in firing rate we observed may be a consequence of the different levels of response conflict associated with various options. We performed a control analysis to examine this question, reasoning that conflict would covary with the similarity of the two risky options. We compared firing rates on trials when two risky options had similar $p(\text{Large})$ (≥ 0.05 difference) to those on trials when the $p(\text{Large})$ of the two risky options differed substantially (≤ 0.10 difference). Supporting this categorization, reaction time (rt) on high conflict trials (rt = 236.3 ms) was significantly slower than on low conflict trials (rt = 211.7 ms, $p < 0.001$, Student's t test). Of the 92 neurons in our sample, reward-related responses of only five (5.4%) were modulated by conflict (Fig. 6)—no different from what would be expected by chance ($p = 0.848$, binomial test). We repeated these analyses with several other probability cutoffs instead of 0.05 and 0.1, and obtained virtually identical results.

Discussion

The relationship between stimuli, actions, and reward outcomes drives associative learning. Although many previous studies have emphasized the importance of reward prediction errors in associative learning (Montague et al., 1995, 1996; Schultz et al., 1997; Fiorillo et al., 2003; Bayer and Glimcher, 2005; Schultz, 2006; Matsumoto and Hikosaka, 2007, 2009), the absolute value of the reward prediction error, known as either surprisingness or associability, drives learning in many situations (Mackintosh, 1975; Hall and Pearce, 1979; Pearce and Hall, 1980; Courville et al., 2006). Here we show that this variable is explicitly calculated by the brain and represented in modulations of the firing rates of single neurons in the dACC in a probabilistically rewarded choice task. Although the magnitudes of these neural effects was small, they correspond to the similarly small effect of rewards on behavior in our task. These effects are inconsistent with the idea that dACC neurons solely carry a signed reward prediction error signal, as might be inferred based on the strong dopamine projections to ACC. Given the existence of a robust network specialized for computing and broadcasting RPE signals (i.e., the dopamine system), our data endorse the notion that the brain uses multiple complementary learning systems to guide behavior (White and McDonald, 2002; Poldrack and Packard, 2003; Courville et al., 2006; Yin and Knowlton, 2006).

More generally, these results strongly suggest that reward coding in dACC is not of a labeled line type, but is highly context-dependent (Adrian and Matthews, 1927; Barlow, 1972). This idea is consistent with the notion that dACC is positioned late in the sequence of processes that serves to convert sensory information to motor plans. If the goal of the brain is to use information arriving from the environment to make adaptive decisions, then one would expect neurons positioned late in the processing stream to represent decision variables, like whether a switch is needed, but not to represent constituent information relating to that decision—such as probability and reward size. Thus, we conjecture that the overlap in neural responses for unexpected large rewards and expected small rewards (Fig. 4) does not pose a decoding problem, because the only information that needs to be decoded is the need to alter behavior—which is matched in the two conditions.

Although we find here that firing rates are, on average, negatively correlated with reward value, we previously reported that dACC responses were positively correlated with reward value—both experienced and observed—in a different task (Hayden et al., 2009). Because this study and the previous one were performed on the same two animals using the same grid positions within a 6-month period, we believe this difference reflects task demands rather than nonoverlapping sets of neurons. Although there are many differences between the risky choice task used in this study and the fictive learning task used in the previous study, one difference looms largest. In the present study, strategic adjustments in behavior were promoted by both small rewards and by surprising ones (Fig. 2); in the prior study, strategic adjustments were promoted by large rewards, both fictive and experienced. Because the two tasks are quite different, we are necessarily defining strategic adjustments somewhat differently in the two tasks, so the comparison is imperfect. If the comparison can still be made, however, then it appears that firing rates in dACC vary more closely with likelihood of behavioral adjustment than with any specific aspect of reward *per se*. Instead, these two datasets, together, are consistent with our hypothesis that dACC neurons do not universally transmit a labeled line representation of reward size, or RPE, or even unsigned RPE. Instead, dACC neurons seem to carry a signal specifying the need to adjust behavior adaptively. Consequently, we conjecture that the role of dACC in learning is distinct from that of more central reward areas, such as the dopamine system, and that it occupies a later, more output-centric stage in the network that transforms values into actions.

The results we present here are consistent with a growing body of literature linking ACC in general and dACC in particular with associating actions and their reward outcomes (Procyk et al., 2000; Shidara and Richmond, 2002; Williams et al., 2004; Amiez et al., 2005; Sallet et al., 2007; Seo and Lee, 2007; Kennerley et al., 2009; Quilodran et al., 2008; Kennerley and Wallis, 2009). Neuronal activity in dACC varies with expected value (Amiez et al., 2005; Seo and Lee, 2007; Kennerley and Wallis, 2009), is normalized to global reward context (Sallet et al., 2007), and provides signals useful for learning (Procyk et al., 2000; Matsumoto et al., 2007) and behavioral adjustments (Shima and Tanji, 1998; Kennerley et al., 2006; Quilodran et al., 2008). Our results go beyond these earlier ones by parametrically manipulating expectation so as to fully characterize its effects on neuronal activity.

The dACC is unlikely to be the only brain area that signals outcome surprisingness. Our findings are consistent with a broader picture of the amygdala—which is directly connected to the ACC (Morecraft and Van Hoesen, 1993; Paus, 2001)—playing a fundamental role in the attentional component of learning

(Holland and Gallagher, 1999, 2006; Holland et al., 2000; McGaugh, 2004; Belova et al., 2007). Similar motivational salience signals have been observed in the basal forebrain, a major target of the amygdala (Lin and Nicolelis, 2008). Norepinephrine (NE) neurons also signal unexpected and attentionally salient events (Aston-Jones et al., 1986, 1994, 1997; Aston-Jones and Cohen, 2005). NE neurons project strongly to the ACC and to the posterior cingulate cortex, a major input to ACC. Thus, ACC may convolve reward signals derived from dopamine neurons with attentional signals emanating from the locus ceruleus and/or the amygdala. Notably, the area of our recordings, area 24, does not, to our knowledge, connect directly to the central amygdala. This suggests that the linkage between our recordings and results found in central amygdala do not reflect direct, monosynaptic connections.

The present results complement those from a previous study from our laboratory investigating firing rates of neurons in the posterior cingulate cortex (CGp) (Hayden et al., 2008). In that study, we found that neuronal activity was larger for small rewards than for large rewards obtained from simple 50/50 gambles, and that these responses were influenced by rewards obtained from the most recent few trials. Whereas the effects we observed in dACC lasted only 500 ms after the reward was given, the effects we observed in CGp persisted for several seconds and extended across trials. These results suggest that dACC and CGp may play complementary roles in learning, with dACC generating an immediate surprise signal, and CGp maintaining that information in a working memory buffer.

One limitation of the present study is that the task we use here was not designed to test learning, and, indeed, punishes learning on risky trials. Any learning therefore occurs despite its associated costliness. Consequently, it is unsurprising that the learning effects—both behavioral and neural—we observed were weak. In any case, it will be important to confirm these ideas in the context of a task that more strongly drives learning. Indeed, a critical question is whether surprise signals will be observed in tasks where learning itself is rewarded or whether such unsigned reward prediction error signals in ACC are specific to probabilistic choice tasks. Another limitation is that we cannot determine whether the surprisingness of an outcome immediately influenced the neural response to the cue that had predicted that outcome because, in our task, cues disappeared before the outcome was given and were not immediately repeated.

Notably, our study failed to find any evidence that response conflict drives modulation of neuronal activity in dACC. Although conflict signals are robustly observed in neuroimaging studies, three previous electrophysiological studies of single neuron activity in dACC showed no effect of conflict (Ito et al., 2003; Nakamura et al., 2005; Amiez et al., 2006; for a broader discussion of some of the discrepancies in the ACC conflict literature, see Rushworth et al., 2005). We note that the form of conflict analyzed in our study is closely related to selection difficulty, and may be different from response conflict associated with multiple possible actions assessed in previous studies (Ito et al., 2003; Nakamura et al., 2005; Amiez et al., 2006). If so, our data contribute to and extend previous results showing that dACC neurons do not signal conflict or task difficulty; at least, not in the monkey. These results suggest that the conflict signals observed in human ACC may come from different anatomical regions (Stuphorn et al., 2000; Ito et al., 2003; Rushworth et al., 2005), may be a consequence of differences between BOLD signal and single unit activity (Naka-

mura et al., 2005), or may reflect species differences in information processing in ACC between monkeys and humans.

The surprise signal we observed is consistent with one postulated in so-called attentional theories of learning (Mackintosh, 1975; Hall and Pearce, 1979; Pearce and Hall, 1980; Courville et al., 2006). The term “attentional” is apt—these theories posit that surprising rewards, whatever their valence, recruit neural resources, and that attention itself promotes learning. The idea that dACC generates an attention-related signal is consistent with a large literature showing a role for ACC in attention and cognitive control (Mesulam, 1981, 1999; Posner and Petersen, 1990; Badgaiyan and Posner, 1998; Davis et al., 2000; Mesulam et al., 2001; Weissman et al., 2003; Kondo et al., 2004). Indeed, it often seems as if the attentional and reward-centric accounts of ACC function are working in parallel, with little or no overlap. The present results therefore point to a possible linkage between these otherwise disjoint literatures.

References

- Adrian ED, Matthews R (1927) The action of light on the eye. I. The discharge of impulses in the optic nerve and its relation to the electric changes in the retina. *J Physiol* 63:378–414.
- Amiez C, Joseph JP, Procyk E (2005) Anterior cingulate error-related activity is modulated by predicted reward. *Eur J Neurosci* 21:3447–3452.
- Amiez C, Joseph JP, Procyk E (2006) Reward encoding in the monkey anterior cingulate cortex. *Cereb Cortex* 16:1040–1055.
- Aston-Jones G, Cohen JD (2005) An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annu Rev Neurosci* 28:403–450.
- Aston-Jones G, Ennis M, Pieribone VA, Nickell WT, Shipley MT (1986) The brain nucleus locus coeruleus: restricted afferent control of a broad efferent network. *Science* 234:734–737.
- Aston-Jones G, Rajkowski J, Kubiak P, 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, Kubiak P (1997) Conditioned responses of monkey locus coeruleus neurons anticipate acquisition of discriminative behavior in a vigilance task. *Neuroscience* 80:697–715.
- Badgaiyan RD, Posner MI (1998) Mapping the cingulate cortex in response selection and monitoring. *Neuroimage* 7:255–260.
- Barlow HB (1972) Single units and sensation: a neuron doctrine for perceptual psychology? *Perception* 1:371–394.
- Barracough DJ, Conroy ML, Lee D (2004) Prefrontal cortex and decision making in a mixed-strategy game. *Nat Neurosci* 7:404–410.
- Bayer HM, Glimcher PW (2005) Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47:129–141.
- Belova MA, Paton JJ, Morrison SE, Salzman CD (2007) Expectation modulates responses to pleasant and aversive stimuli in primate amygdala. *Neuron* 55:970–984.
- Botvinick MM, Braver TS, Barch DM, Carter CS, Cohen JD (2001) Conflict monitoring and cognitive control. *Psychol Rev* 108:624–652.
- Brainard DH (1997) The Psychophysics Toolbox. *Spat Vis* 10:433–436.
- Carter CS, Macdonald AM, Botvinick M, Ross LL, Stenger VA, Noll D, Cohen JD (2000) Parsing executive processes: strategic vs. evaluative functions of the anterior cingulate cortex. *Proc Natl Acad Sci U S A* 97:1944–1948.
- Cornelissen FW, Peters EM, Palmer J (2002) The Eyelink Toolbox: eye tracking with MATLAB and the Psychophysics Toolbox. *Behav Res Methods Instrum Comput* 34:613–617.
- Courville AC, Daw ND, Touretzky DS (2006) Bayesian theories of conditioning in a changing world. *Trends Cogn Sci* 10:294–300.
- Davis KD, Hutchison WD, Lozano AM, Tasker RR, Dostrovsky JO (2000) Human anterior cingulate cortex neurons modulated by attention-demanding tasks. *J Neurophysiol* 83:3575–3577.
- Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ (2006) Cortical substrates for exploratory decisions in humans. *Nature* 441:876–879.
- Desimone R, Duncan J (1995) Neural mechanisms of selective visual attention. *Annu Rev Neurosci* 18:193–222.

- Egeth HE, Yantis S (1997) Visual attention: control, representation, and time course. *Annu Rev Psychol* 48:269–297.
- Ellsberg D (1961) Risk, ambiguity, and the savage axioms. *Q J Econ* 75:643–669.
- Fiorillo CD, Tobler PN, Schultz W (2003) Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299:1898–1902.
- Fox CR, Tversky A (1995) Ambiguity aversion and comparative ignorance. *Q J Econ* 110:585–603.
- Gehring WJ, Willoughby AR (2002) The medial frontal cortex and the rapid processing of monetary gains and losses. *Science* 295:2279–2282.
- Hall G, Pearce JM (1979) Latent inhibition of a CS during CS-US pairings. *J Exp Psychol Anim Behav Process* 5:31–42.
- Hayden BY, Nair AC, McCoy AN, Platt ML (2008) Posterior cingulate cortex mediates outcome-contingent allocation of behavior. *Neuron* 60:19–25.
- Hayden BY, Pearson JM, Platt ML (2009) Fictive reward signals in the anterior cingulate cortex. *Science* 324:948–950.
- Hayden BY, Heilbronner SR, Platt ML (2010) Ambiguity aversion in rhesus macaques. *Front Neurosci* 4:pii:166.
- Holland PC, Gallagher M (1999) Amygdala circuitry in attentional and representational processes. *Trends Cogn Sci* 3:65–73.
- Holland PC, Gallagher M (2006) Different roles for amygdala central nucleus and substantia innominata in the surprise-induced enhancement of learning. *J Neurosci* 26:3791–3797.
- Holland PC, Han JS, Gallagher M (2000) Lesions of the amygdala central nucleus impair performance on a selective attention task. *J Neurosci* 20:6701–6706.
- Holroyd CB, Coles MG (2002) The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol Rev* 109:679–709.
- Hsu M, Bhatt M, Adolphs R, Tranel D, Camerer CF (2005) Neural systems responding to degrees of uncertainty in human decision-making. *Science* 310:1680–1683.
- Huettel SA, Stowe CJ, Gordon EM, Warner BT, Platt ML (2006) Neural signatures of economic preferences for risk and ambiguity. *Neuron* 49:765–775.
- Ito S, Stuphorn V, Brown JW, Schall JD (2003) Performance monitoring by the anterior cingulate cortex during saccade countermanding. *Science* 302:120–122.
- Kaye H, Pearce JM (1984) The strength of the orienting response during blocking. *Q J Exp Psychol B* 36:131–144.
- Kennerley SW, Wallis JD (2009) Evaluating choices by single neurons in the frontal lobe: outcome value encoded across multiple decision variables. *Eur J Neurosci* 29:2061–2073.
- Kennerley SW, Walton ME, Behrens TE, Buckley MJ, Rushworth MF (2006) Optimal decision making and the anterior cingulate cortex. *Nat Neurosci* 9:940–947.
- Kennerley SW, Dahmubed AF, Lara AH, Wallis JD (2009) Neurons in the frontal lobe encode the value of multiple decision variables. *J Cogn Neurosci* 21:1162–1178.
- Kerns JG, Cohen JD, MacDonald AW 3rd, Cho RY, Stenger VA, Carter CS (2004) Anterior cingulate conflict monitoring and adjustments in control. *Science* 303:1023–1026.
- Kondo H, Osaka N, Osaka M (2004) Cooperation of the anterior cingulate cortex and dorsolateral prefrontal cortex for attention shifting. *Neuroimage* 23:670–679.
- Lauwereyns J, Takikawa Y, Kawagoe R, Kobayashi S, Koizumi M, Coe B, Sakagami M, Hikosaka O (2002) Feature-based anticipation of cues that predict reward in monkey caudate nucleus. *Neuron* 33:463–473.
- Lee D, Seo H (2007) Mechanisms of reinforcement learning and decision making in the primate dorsolateral prefrontal cortex. *Ann N Y Acad Sci* 1104:108–122.
- Lin SC, Nicolelis MA (2008) Neuronal ensemble bursting in the basal forebrain encodes salience irrespective of valence. *Neuron* 59:138–149.
- Mackintosh NJ (1975) A theory of attention: variations in the associability of stimuli with reinforcement. *Psychol Rev* 82:276–298.
- 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) Representation of negative motivational value in the primate lateral habenula. *Nat Neurosci* 12:77–84.
- Matsumoto M, Matsumoto K, Abe H, Tanaka K (2007) Medial prefrontal cell activity signaling prediction errors of action values. *Nat Neurosci* 10:647–656.
- McGaugh JL (2004) The amygdala modulates the consolidation of memories of emotionally arousing experiences. *Annu Rev Neurosci* 27:1–28.
- Mesulam MM (1981) A cortical network for directed attention and unilateral neglect. *Ann Neurol* 10:309–325.
- Mesulam MM (1999) Spatial attention and neglect: parietal, frontal and cingulate contributions to the mental representation and attentional targeting of salient extrapersonal events. *Philos Trans R Soc Lond B Biol Sci* 354:1325–1346.
- Mesulam MM, Nobre AC, Kim YH, Parrish TB, Gitelman DR (2001) Heterogeneity of cingulate contributions to spatial attention. *Neuroimage* 13:1065–1072.
- Montague PR, Dayan P, Person C, Sejnowski TJ (1995) Bee foraging in uncertain environments using predictive Hebbian learning. *Nature* 377:725–728.
- Montague PR, Dayan P, Sejnowski TJ (1996) A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci* 16:1936–1947.
- Morecraft RJ, Van Hoesen GW (1993) Frontal granular cortex input to the cingulate (M3), supplementary (M2) and primary (M1) motor cortices in the rhesus monkey. *J Comp Neurol* 337:669–689.
- Nakamura K, Roesch MR, Olson CR (2005) Neuronal activity in macaque SEF and ACC during performance of tasks involving conflict. *J Neurophysiol* 93:884–908.
- Paus T (2001) Primate anterior cingulate cortex: where motor control, drive and cognition interface. *Nat Rev Neurosci* 2:417–424.
- Pearce JM, Hall G (1980) A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol Rev* 87:532–552.
- Pearson JM, Hayden BY, Raghavachari S, Platt ML (2009) Neurons in posterior cingulate cortex signal exploratory decisions in a dynamic multioption choice task. *Current Biol* 19:1532–1537.
- Poldrack RA, Packard MG (2003) Competition among multiple memory systems: converging evidence from animal and human brain studies. *Neuropsychologia* 41:245–251.
- Posner MI, Petersen SE (1990) The attention system of the human brain. *Annu Rev Neurosci* 13:25–42.
- Procyk E, Tanaka YL, Joseph JP (2000) Anterior cingulate activity during routine and non-routine sequential behaviors in macaques. *Nat Neurosci* 3:502–508.
- Quilodran R, Rothé M, Procyk E (2008) Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron* 57:314–325.
- Rescorla RA, and Wagner AR (1972) A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: *Classical conditioning II: current research and theory* (Black AH, Prokasy WF, eds.). New York: Appleton-Century-Crofts.
- Roesch MR, Calu DJ, Esber GR, Schoenbaum G (2010) All that glitters: dissociating attention and outcome expectancy from prediction error signals. *J Neurophysiol* 104:587–595.
- Rushworth MF, Kennerley SW, Walton ME (2005) Cognitive neuroscience: resolving conflict in and over the medial frontal cortex. *Curr Biol* 15:R54–R56.
- Rushworth MF, Buckley MJ, Behrens TE, Walton ME, Bannerman DM (2007) Functional organization of the medial frontal cortex. *Curr Opin Neurobiol* 17:220–227.
- Sallet J, Quilodran R, Rothé M, Vezoli J, Joseph JP, Procyk E (2007) Expectations, gains, and losses in the anterior cingulate cortex. *Cogn Affect Behav Neurosci* 7:327–336.
- Samejima K, Ueda Y, Doya K, Kimura M (2005) Representation of action-specific reward values in the striatum. *Science* 310:1337–1340.
- Schultz W (2006) Behavioral theories and the neurophysiology of reward. *Annu Rev Psychol* 57:87–115.
- Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* 275:1593–1599.
- Seo H, Lee D (2007) Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *J Neurosci* 27:8366–8377.
- Shidara M, Richmond BJ (2002) Anterior cingulate: single neuronal signals related to degree of reward expectancy. *Science* 296:1709–1711.

- Shima K, Tanji J (1998) Role for cingulate motor area cells in voluntary movement selection based on reward. *Science* 282:1335–1338.
- Stuphorn V, Taylor TL, Schall JD (2000) Performance monitoring by the supplementary eye field. *Nature* 408:857–860.
- Sutton RS, Barto AG (1998) Reinforcement learning: an introduction. Cambridge, MA: MIT.
- Swan JA, Pearce JM (1988) The orienting response as an index of stimulus associability in rats. *J Exp Psychol Anim Behav Process* 14:292–301.
- van Veen V, Cohen JD, Botvinick MM, Stenger VA, Carter CS (2001) Anterior cingulate cortex, conflict monitoring, and levels of processing. *Neuroimage* 14:1302–1308.
- Weissman DH, Giesbrecht B, Song AW, Mangun GR, Woldorff MG (2003) Conflict monitoring in the human anterior cingulate cortex during selective attention to global and local object features. *Neuroimage* 19:1361–1368.
- White NM, McDonald RJ (2002) Multiple parallel memory systems in the brain of the rat. *Neurobiol Learn Mem* 77:125–184.
- Williams ZM, Bush G, Rauch SL, Cosgrove GR, Eskandar EN (2004) Human anterior cingulate neurons and the integration of monetary reward with motor responses. *Nat Neurosci* 7:1370–1375.
- Yin HH, Knowlton BJ (2006) The role of the basal ganglia in habit formation. *Nat Rev Neurosci* 7:464–476.
- Yu AJ, Dayan P (2005) Uncertainty, neuromodulation, and attention. *Neuron* 46:681–692.