

Individual Neurons in the Cingulate Cortex Encode Action Monitoring, Not Selection, during Adaptive Decision-Making

Yin S. Li,¹ Matthew R. Nassar,² Joseph W. Kable,³ and Joshua I. Gold⁴

¹Department of Psychiatry, University of Pennsylvania, Philadelphia, Pennsylvania 19104, ²Department of Neuroscience, Brown University, Providence, Rhode Island 02906, Departments of ³Psychology, and ⁴Neuroscience, University of Pennsylvania, Philadelphia, Pennsylvania 19104

The cingulate cortex contributes to complex, adaptive behaviors, but the exact nature of its contributions remains unresolved. Proposals from previous studies, including evaluating past actions or selecting future ones, have been difficult to distinguish in part because of an incomplete understanding of the task-relevant variables that are encoded by individual cingulate neurons. In this study, we recorded from individual neurons in parts of both the anterior cingulate cortex (ACC) and posterior cingulate cortex (PCC) in 2 male rhesus monkeys performing a saccadic reward task. The task required them to use adaptive, feedback-driven strategies to infer the spatial location of a rewarded saccade target in the presence of different forms of uncertainty. We found that task-relevant, spatially selective feedback signals were encoded by the activity of individual neurons in both brain regions, with stronger selectivity for spatial choice and reward-target signals in PCC and stronger selectivity for feedback in ACC. Moreover, neurons in both regions were sensitive to sequential effects of feedback that partly reflected sequential behavioral patterns. However, neither brain region exhibited systematic modulations by the blockwise conditions that governed the reliability of the trial-by-trial feedback and drove adaptive behavioral patterns. There was also little evidence that single-neuron responses in either brain region directly predicted the extent to which feedback and contextual information were used to inform choices on the subsequent trial. Thus, certain cingulate neurons encode diverse, evaluative signals needed for adaptive, feedback-driven decision-making, but those signals may be integrated elsewhere in the brain to guide actions.

Key words: cingulate cortex; decision-making; nonhuman primate; reward learning

Significance Statement

Effective decision-making in dynamic environments requires adapting to changes in feedback and context. The anterior and posterior cingulate cortex have been implicated in adaptive decision-making, but the exact nature of their respective roles remains unresolved. Here we compare patterns of task-driven activity of subsets of individual neurons from parts of the two brain regions in monkeys performing a saccadic task with dynamically changing reward locations. We find evidence for regional specializations in neural representations of choice and feedback, including task-relevant modulations of activity that could be used for performance monitoring. However, we find little evidence that these neural representations are used directly to adjust choice behavior, which thus likely requires integration of these signals elsewhere in the brain.

Introduction

Organisms tend to repeat actions that lead to positive outcomes and avoid actions that lead to negative outcomes (Thorndike,

1911). This kind of adaptive behavior is particularly sensitive to errors, which can drive adjustments that ultimately lead to appropriate behavioral policies in uncertain environments (Bertsekas and Tsitsiklis, 1996; Sutton and Barto, 1998). However, errors can arise from different sources that may have different, even opposite, implications for effective learning. For example, a baseball batter may realize after a missed swing that the new pitcher throws harder than the previous one and should adjust accordingly. However, if the same batter, facing the same pitcher, again misses later in the game, the best policy is probably not to over-

Received Jan. 18, 2019; revised May 14, 2019; accepted June 13, 2019.

Author contributions: Y.S.L., M.R.N., J.W.K., and J.I.G. designed research; Y.S.L. and M.R.N. performed research; Y.S.L. analyzed data; Y.S.L. wrote the first draft of the paper; Y.S.L., M.R.N., J.W.K., and J.I.G. edited the paper; Y.S.L., M.R.N., J.W.K., and J.I.G. wrote the paper.

This work was supported by National Science Foundation NCS 1533623, National Institutes of Health R01 MH115557, and a Soros Fellowship for New Americans to Y.S.L. We thank Takahiro Doi, Long Ding, and Kamesh Krishnamurthy for valuable comments; Fred Letterio and Mike Suplick of the MINS Machine Shop for technical support; and Jean Zweigle for expert animal care and training.

The authors declare no competing financial interests.

Correspondence should be addressed to Joshua I. Gold at jigold@penmedicine.upenn.edu.

<https://doi.org/10.1523/JNEUROSCI.0159-19.2019>
Copyright © 2019 the authors

compensate. Put another way, errors that arise from either changes in or uncertainty about an environment likely represent good opportunities for learning and thus should cause changes in behavior. In contrast, errors that are an inevitable consequence of noise or other inherent limitations are often best ignored (Yu and Dayan, 2005; Courville et al., 2006; Adams and MacKay, 2007; Behrens et al., 2007; Fearnhead and Liu, 2007; Nassar et al., 2010; Wilson et al., 2013).

Such flexibility implies that, in addition to processing errors, the brain should also keep track of contextual information needed to interpret the source of those errors to guide learning. Accordingly, of particular interest are the anterior cingulate cortex (ACC) and posterior cingulate cortex (PCC), which can encode both error feedback and relevant contextual information, including the likelihood of a recent environmental change-point and subjective uncertainty about the current state of the environment (Vogt and Pandya, 1987; Ito et al., 2003; Behrens et al., 2007; Matsumoto et al., 2007; Quilodran et al., 2008; Jocham et al., 2009; Pearson et al., 2011; Heilbronner and Platt, 2013; Payzan-LeNestour et al., 2013; Heilbronner and Haber, 2014; McGuire et al., 2014). This feedback encoding has been identified at the single-neuron level in the ACC and PCC of monkeys (McCoy and Platt, 2005; Matsumoto et al., 2007; Seo and Lee, 2007; Heilbronner and Platt, 2013). However, the relevant contextual signals have generally been studied only using much lower-resolution imaging techniques (Behrens et al., 2007; Kolling et al., 2014; McGuire et al., 2014). Therefore, it is currently unknown whether and how such signals are combined at the single-neuron level in the cingulate cortex, and whether there are regional specializations in the computation of these variables between the anterior and PCC. Moreover, it is unclear whether these signals represent monitoring of past actions, selection of future actions, or both during error-driven learning (Alexander and Brown, 2011; Pearson and Platt, 2013; Shenhav et al., 2013; Botvinick and Cohen, 2014; Heilbronner and Hayden, 2016).

We recorded from individual ACC and PCC neurons in monkeys performing an adaptive-inference task in which the task context affected the degree to which particular errors should lead to behavioral adjustments. We focused on the relationship between noise and change-points. In one condition, small spatial errors were most likely a result of noise in an otherwise stable reward-generating process. In another condition, the same errors instead typically resulted from change-points in the process itself. We tested whether and how individual ACC and PCC neurons encode choice, feedback, and contextual variables required to monitor performance and make adjustments to guide future task-relevant actions.

Materials and Methods

Subjects

Two adult male rhesus monkeys (*Macaca mulatta*) were used for this study. All training, surgery, and experimental procedures were performed in accordance with the National Institutes of Health's *Guide for the care and use of laboratory animals* and were approved by the University of Pennsylvania Institutional Animal Care and Use Committee.

Behavioral task

On each trial, the reward target was drawn probabilistically from a distribution centered on the best target. The best target was selected uniformly from the 10 possible targets based on a change-point process governed by a flat hazard rate. Here, the hazard rate is defined as the probability, on any given trial, that the best target would be randomly reselected. In the no-noise, high-hazard ("unstable") condition, the best target was rewarded 100% of the time, but its location was unstable, changing with a hazard rate of 0.4–0.45 per trial. In the high-noise, low-hazard ("noisy") condition, the best target was rewarded 60% of the time, with adjacent pairs of targets rewarded at 15% and 5%, but the best location was more stable, changing with a hazard

rate of 0.02 per trial. Noise and hazard rate were counterbalanced to achieve similar overall performance for the two task conditions. The two environments were presented in blocks of 100–300 trials and cued explicitly by the color and size of the fixation point (large red disk for "unstable" and small gray disk for "noisy" trials).

Each trial began with the presentation of a central fixation point. The monkey had 500 ms to acquire fixation, after which the circular, 10-target array (radius 10° visual angle) appeared. The monkey needed to maintain fixation for a variable duration (~500–1000 ms) until the fixation point disappeared, signaling the monkey to indicate its prediction by making a saccadic eye movement to a single reward target. After the monkey had indicated its choice, visual feedback was given indicating the chosen target (small green dot) and the reward target (large white dot). This visual feedback was present for a variable duration (randomly chosen from 1200, 1350, 1500, 1650, and 1800 ms), during which time the monkey could freely view the screen. After the feedback was extinguished, the screen went blank; and, if the reward target had been chosen, the juice was delivered. The variable delay between feedback onset and juice delivery allowed us to differentiate between the effects of receiving feedback information and receive the juice reward itself.

Eye position was monitored using a video-based system (Eye-Link 2000; SR Research) sampled at 1000 Hz. Visual stimuli were generated using the Psychtoolbox (Brainard, 1997) and custom MATLAB software and presented on an LCD monitor (BenQ) located 60 cm from the monkey's eyes.

Electrophysiology

Each monkey was implanted with a head holder and two recording cylinders that targeted the ACC and PCC on either the left side (Monkey SP) or the right side (Monkey AT). ACC cylinders were placed at Horsley-Clark coordinates 33 mm AP, 8 mm L for Monkey SP and 43 mm AP, 8 mm L for Monkey AT (the differences in AP coordinates reflected differences in the animals' head sizes). PCC cylinders for both monkeys were placed at 0 mm AP, 5 mm L, and tilted at an angle of 8.5° along the ML plane to point downward toward the midline. For ACC recordings, we targeted area 24c along the dorsal bank of the anterior cingulate sulcus (~4–6 mm below cortical surface) because of prior work demonstrating decision-related neural activity in this area, which is also known as dACC (e.g., Seo and Lee, 2007; Hayden et al., 2009; Kennerly et al., 2011). For PCC recordings, we targeted areas 31 and 23 in the posterior cingulate gyrus (~7–11 mm below cortical surface) because of spatial- and decision-related neural activity that have been identified in this area, also known as CGp (e.g., Olson et al., 1996; Dean and Platt, 2006; Hayden et al., 2008; Heilbronner and Platt, 2013). Both brain regions were targeted using MRI and custom software (Kalwani et al., 2009) as well as by listening for characteristic patterns of white and gray matter during recordings. Neural recordings were conducted using either single-contact glass-coated tungsten electrodes (Alpha-Omega) or multicontact linear electrode arrays (V-probe, Plexon) and a Multichannel Acquisition System (Plexon). Spike waveforms were sorted offline to identify putative single units (here referred to as "neurons").

Behavioral analyses

Logistic regression for effects of trial history. We used a logistic regression model to quantify the relationship between feedback on previous trials and probability of switching on the next trial (P_{switch}) while controlling for the effects of error magnitude and noise on the current trial as follows:

$$\log \frac{P_{\text{switch}}}{P_{\text{stay}}} = \sum_{i=0, j=0}^{1,3} \alpha_{i,j} \delta_{\text{noise},i} \delta_{\text{errMag},j} + \sum_{j=1, k=1}^{3,5} \beta_{j,k} \delta_{\text{errMag},-k, j} + \sum_{j=1, k=1}^{3,5} \gamma_{j,k} \delta_{\text{noise},1} \delta_{\text{errMag},-k, j} + \sum_{i=0, k=1}^{1,5} \epsilon_{i,k} \delta_{\text{noise},i} \delta_{\text{switch},-k,1} \quad (1)$$

where $\alpha_{i,j}$, $\beta_{j,k}$, $\gamma_{j,k}$, and $\epsilon_{i,k}$ are coefficients fitted by MATLAB's "glmfit" using the binomial distribution, $\delta_{i,j}$ is the Kronecker delta func-

tion (such that $\delta_{i,j} = 1$ if and only if $i = j$), “noise” denotes the noise condition ($0 = \text{unstable}, 1 = \text{noisy}$), “errMag_{*t*}” denotes the spatial error magnitude on trial *t* (for which errors of >3 are coded as 3), and “switch_{*t*}” denotes whether the subject switches its choice on the trial following *t*. In this formulation, the coefficients $\alpha_{i,j}$ measure the influence of feedback on the current trial (*t*), the coefficients $\beta_{j,k}$ measure the average influence of feedback of a given error magnitude from *k* trials back, whereas the coefficients $\gamma_{j,k}$ measure the differential influence of feedback in noisy versus unstable trials from *k* trials back. The coefficients $\varepsilon_{i,k}$ account for the sequential effects of choice history. Significance for the logistic regression coefficients was determined by the Wald test ($p < 0.05$).

Optimal Bayesian inference. Optimal Bayesian inference for the change-point task involves computing the posterior distribution over the best target μ_t on trial *t* after observing the *t*th reward target from the history of all reward targets $x_1, \dots, x_p, P(\mu_t|x_{1:t})$. We developed a recursive, online algorithm to compute this discrete posterior distribution (Adams and MacKay, 2007; Behrens et al., 2007; Fearnhead and Liu, 2007; Nassar et al., 2010; Wilson et al., 2013). Let *H* be the flat hazard rate determining whether the best target is the same as from the previous trial or repicked randomly from 1 of the 10 possible targets. The prior probability over the best targets on trial *t* is then given by a weighted sum of the posterior on the best target from the previous trial and the flat uniform distribution, where *H* governs the weighting as follows:

$$P(\mu_t|x_{1:t-1}) = P(\mu_{t-1}|x_{1:t-1})(1 - H) + \frac{1}{10}H \quad (2)$$

The posterior probability $P(\mu_t|x_{1:t})$ after observing the *t*th outcome x_t is given by Bayes’ rule as follows:

$$P(\mu_t|x_{1:t}) = \frac{P(x_t|\mu_t)P(\mu_t|x_{1:t-1})}{P(x_t|x_{1:t-1})} \quad (3)$$

where $P(x_t|\mu_t)$ is the known likelihood, or noise, distribution (e.g., for the unstable condition, $P(x_t|\mu_t) = 1$ for $x_t = \mu_t$, and 0 otherwise), and $P(x_t|x_{1:t-1})$ acts as a normalization term. Combining Equations 2 and 3 provides a recursive algorithm for computing the posterior on the best target given knowledge of the hazard rate *H* and the noise distribution $P(x_t|\mu_t)$ as follows:

$$P(\mu_t|x_{1:t}) = \frac{P(x_t|\mu_t) \left(P(\mu_{t-1}|x_{1:t-1})(1 - H) + \frac{1}{10}H \right)}{P(x_t|x_{1:t-1})} \quad (4)$$

where $P(\mu_{t-1}|x_{1:t-1})$ equals 1/10 for $t = 1$.

Pure Bayesian behavioral model. For the pure Bayesian model, we generated a probability distribution of picking target *k* on trial $t + 1$, $P_{\text{Bayes}}(k, t + 1)$, by taking the softmax transformation of the optimal Bayesian inference algorithm’s posterior probability on trial *t* as computed in Equation 4 as follows:

$$P_{\text{Bayes}}(k, t + 1) = \frac{\exp(\beta \times P(\mu_t = k|x_{1:t}))}{\sum_k \exp(\beta \times P(\mu_t = k|x_{1:t}))} \quad (5)$$

where β is known as the inverse-temperature term, with $\beta = 0$ resulting in all targets being chosen with equal probability and higher β increasing the probability of picking the most probable target.

We modeled the (discrete) noise distribution after the von Mises distribution for circular variables (Fisher, 1995) as follows:

$$P(x_t|\mu_t) = \frac{\exp\left(\kappa \cos\left(\frac{\pi}{5}(x_t - \mu_t)\right)\right)}{\sum_{\mu_t} \exp\left(\kappa \cos\left(\frac{\pi}{5}(x_t - \mu_t)\right)\right)} \quad (6)$$

where the parameter κ is inversely related to the width of the noise distribution. The pure Bayes model thus had five free parameters: two noise-dependent subjective hazard rates *H*, two noise-dependent inverse-noise parameters κ , and the inverse temperature term β .

Hybrid Bayesian behavioral model. The hybrid Bayesian model used the same optimal Bayesian inference algorithm, but without the softmax transformation and instead with three additional terms that describe the excessive tendency of the monkeys to use stay-switch heuristics. These heuristics were encoded as the probability (bounded between 0 and 1) that they persisted in picking the same target (stay) after correct feedback (s_{pos}), error feedback in the unstable condition ($s_{\text{neg,unstable}}$), and error feedback in the noisy condition ($s_{\text{neg,noisy}}$). The probability of picking target *k* on the *t*th trial was then given by the following:

$$P(k, t) = P(\mu_{t-1} = k|x_{1:t-1}) \times (1 - s) + s \quad (7)$$

if *k* had been chosen on trial $t - 1$

$$P(k, t) = P(\mu_{t-1} = k|x_{1:t-1}) \times (1 - s) \quad (7)$$

if *k* had not been chosen on trial $t - 1$

where $P(\mu_{t-1} = k|x_{1:t-1})$ is the posterior on best targets as computed in Equation 4, and *s* is the noise- and feedback-dependent stay term that enhances the probability of the monkey to choose the same target as the previous trial. The hybrid Bayes model thus had seven free parameters: two noise-dependent subjective hazard rates *H*, two noise-dependent inverse-noise parameters κ , and three noise- and feedback-dependent stay terms *s*.

Reinforcement learning (RL) models. We applied a fixed learning rate RL model to the monkeys’ trial \times trial choice behavior (Sutton and Barto, 1998). The value $V_k(t)$ of a choice target *k* on trial *t* was updated according to the following delta rule:

$$V_k(t + 1) = V_k(t) + \alpha(R_k(t) - V_k(t)) \quad (8)$$

where R_k is 1 if target *k* was the reward target and 0 otherwise. The learning rate α differed depending on whether or not *k* was the chosen target: the chosen target was associated with an “experienced learning rate” α_{exp} , and all nonchosen targets were associated with a “fictive learning rate” α_{fic} (Hayden et al., 2009). In the “pure RL” model, the probability of choosing target *k*, $P(k)$, was given by the softmax transformation of V_k as in Equation 5, resulting in three free parameters. In the “hybrid RL” model, additional choice heuristics as in Equation 7 was applied, resulting in six free parameters: two choice-dependent learning rates α , an inverse-temperature term β , and three stay terms *s*.

Heuristic models. We tested two heuristic models (see Table 2). “Heuristic 1” consists of only the stay-switch heuristic used in the hybrid Bayes and hybrid RL models (Eq. 7). “Heuristic 2” captures the observation that monkeys almost always switch for large errors and tend to stay for small errors, unless the same target has been consistently rewarded. This heuristic model always switches to the rewarded target when the error magnitude is >2 . For error magnitude ≤ 2 , the model’s choice function is a probabilistic mixture of a model that stays with probability *s* and an RL model with a single fictive learning rate (α , Eq. 8) and an inverse-temperature parameter (β) for soft-max choice selection (Eq. 5).

Model fitting and comparison. Behavioral models were fit using MATLAB’s “fmincon” function to find values of free parameters that maximized the likelihood of the monkeys’ actual sequences of choices for each session. To compare goodness-of-fit for the models while accounting for differences in the number of free parameters, we used both Akaike information criterion (AIC) and Bayesian information criterion (BIC) (Burnham and Anderson, 2004) as follows:

$$\text{AIC} = -2\log L + 2m + 2m(m + 1)/(n - m - 1), \quad (9)$$

$$\text{BIC} = -2\log L + m \log n, \quad (10)$$

where *L* is the likelihood of the monkey’s choices given model parameters, *m* is the number of model parameters (see Table 2), and *n* is the number of trials. Lower AIC and BIC are preferred.

Neural analyses

Neurons were included in the database for further analyses only if we recorded ≥ 350 total trials, of which ≥ 125 were low-noise trials and ≥ 150 were high-noise trials. This selection criterion yielded a database

consisting of 102 neurons from ACC ($n = 67$ from Monkey SP, 35 from Monkey AT), of which 77 were considered high-isolation quality single neurons ($n = 49$ from Monkey SP, 28 from Monkey AT), and 235 neurons from PCC ($n = 75$ from Monkey SP, 160 from Monkey AT), of which 188 were considered high-isolation quality single neurons ($n = 58$ from Monkey SP, 130 from Monkey AT). On average, 710 trials (308 low noise, 402 high noise) were recorded for ACC neurons, and 832 trials (329 low noise, 503 high noise) were recorded for PCC neurons. Isolation quality did not affect our main conclusions (whether an effect was statistically significant or not), so here we present analyses that include both low- and high-isolation quality neurons.

GLM. We used a GLM to examine the relationship between trial-by-trial neural activity and task and behavioral variables of interest while accounting for important potential confounds. We fit firing rates, r , of individual neurons to the following equation:

$$r = \beta_0 + \beta_1 \text{choice}_x + \beta_2 \text{choice}_y + \beta_3 \text{rwdTarg}_x + \beta_4 \text{rwdTarg}_y + \beta_5 \text{fdbk} + \beta_6 \text{errMag} + \beta_7 \text{noise} + \beta_8 \text{fdbkPrev} + \beta_9 \text{fdbk*noise} + \beta_{10} \text{fdbk*fdbkPrev} + \beta_{11} \text{switch} + \beta_{12} \text{postDelta}_x + \beta_{13} \text{postDelta}_y + \beta_{14} \text{rewRate} + \beta_{15} \text{trialNum} + \beta_{16} r\text{Prev}, \quad (11)$$

where β^* are regression coefficients fit by MATLAB's "glmfit" function using "identity" link function. "choice_x" and "choice_y" denote the horizontal and vertical coordinates, respectively, of the chosen target (normalized to radius 1), "rwdTarg_x" and "rwdTarg_y" denote the horizontal and vertical coordinates, respectively, of the rewarded target (normalized to radius 1), "fdbk" denotes categorical feedback on the current trial (0 = error, 1 = correct), "errMag" is the spatial error magnitude (mean-centered for error trials), "noise" denotes the noise condition (0 = unstable, 1 = noisy), and "fdbkPrev" denotes categorical feedback on the previous trial (0 = error, 1 = correct). The interaction terms "fdbk*noise" and "fdbk*fdbkPrev" are the mean-subtracted products of the indicated variables. The term "switch" denotes whether the monkey stays (0) or switches (1) its choice on the subsequent trial. The position terms "postDelta_x" and "postDelta_y" reflect the difference in horizontal/vertical coordinates of the postfeedback saccadic target relative to the original chosen target and were included to account for the confounding role of postfeedback saccades on the apparent encoding of choice and feedback variables. The remaining terms were included as potential confounds for the apparent neural selectivity for noise: "rewRate" denotes the average reward rate in the previous 10 trials, "trialNum" is the trial number during a session (accounting for slow rise or fall in neural activity over time), and the autoregressive term "rPrev" denotes the firing rate from the previous trial. Significance for individual coefficients was determined by the t test ($p < 0.05$). A neuron was considered selective for choice if either β_1 or β_2 were significantly different from 0. Likewise, a neuron was considered selective for rewarded target location if either β_3 or β_4 were significantly different from 0.

For sliding-window analyses, we performed the above GLM for firing rates computed in 250 ms time bins advanced in 50 ms steps. For epoched analysis, we defined a prefixation epoch (500 ms before fixation point onset), a postsaccade epoch (500 ms after saccade onset), and a feedback epoch (250–750 ms after feedback onset).

Spatial tuning. To determine whether spatially tuned neurons were better characterized by unimodal or bimodal tuning, we fit spatial tuning curves to cosine functions with spatial frequency of 1 or 2 and assessed which was the better least-squares fit (Hayden and Platt, 2010). To determine tuning widths, we fit spatial tuning curves to von Mises distribution and report the width of the tuning curve at half-maximal height. Tuning strength was measured in two different ways: (1) an receiver operating characteristic (ROC)-based selectivity index (the three targets closest to the preferred direction are grouped as "preferred" and the three targets in the opposite direction are designated as "null"); and (2) the norm of the GLM coefficients (i.e., square root of $\beta_1^2 + \beta_2^2$ for choice and square root of $\beta_3^2 + \beta_4^2$ for reward target). The results are presented in Table 3.

Results

The monkeys learned adaptively from feedback

We trained 2 monkeys to perform a novel change-point task that required them to make a saccadic eye movement to predict, based on the outcomes of previous trials, the likely location of a single, stochastically determined "reward target" from among a circular array of 10 visually identical circles (Fig. 1*a*). The monkeys performed this task under two different, explicitly cued conditions presented in blocks of 100–300 trials. These blocks differed in terms of the reliability with which feedback signaled change-points in the most likely location of the reward target (Fig. 1*b,c*). In zero-noise, high-hazard ("unstable") blocks, a single target was always rewarded until a change-point occurred, governed by a relatively high hazard rate (0.4–0.45). The hazard rate was defined as the probability of a change-point occurring on any given trial. To maximize reward in these blocks, the monkeys should always choose the previously rewarded target and thus immediately shift choices after errors resulting from change-points. In contrast, in high-noise, low-hazard ("noisy") blocks, the location of the reward target fluctuated around a relatively stable best target that changed with a hazard rate of 0.02. To maximize reward in these blocks, the monkey should identify the best target by integrating noisy feedback information across trials and thus typically suppress behavioral adjustments to small errors.

We analyzed the performance of the 2 monkeys from all sessions from which we made neural recordings ($n = 93,278$ trials over 67 sessions for Monkey SP; 71,422 trials over 65 sessions for Monkey AT). Overall, the monkeys chose the reward target well above chance (=10%) in both the unstable (49% for Monkey SP, 50% for Monkey AT) and noisy (42% for Monkey SP, 43% for Monkey AT) conditions. However, they adjusted to true change-points of the best target differently in the two conditions (Fig. 2*a,b*). In the unstable condition, they tended to choose the best target >50% of the time after just one trial after a change-point and >95% of the time after ~10 trials. In the noisy condition, they tended to choose the best target >50% of the time within 5–6 trials after a change-point, then after ~15 trials reliably chose the best target more frequently than the reward rate for that target (60%). These results imply that they were not just choosing the most recent reward target or matching reward probabilities (Sugrue et al., 2004; Lau and Glimcher, 2005), but rather selecting the target inferred to be maximally rewarding.

Both monkeys used adaptive, feedback-driven strategies to solve the task. The monkeys clearly attended to the visual feedback that they received at the end of each trial, indicating the location of the rewarded target, which on error trials tended to elicit eye movements to that location in the free-viewing period at the end of the trial (Fig. 3). Subsequently, they were more likely to switch their choice after error versus correct feedback (Fig. 2*c,d*). However, they did not treat all errors equally. They were more likely to switch after making larger spatial errors (e.g., when the rewarded target was far from the chosen target). Moreover, for a given spatial error magnitude, they were less likely to switch in the noisy versus unstable condition (Fig. 2*c,d*). This difference was largest for small spatial errors, which were more predictive of change-points in the unstable versus noisy condition. These adaptive features of behavior were present throughout the experimental sessions (switch rate depended on the interaction between spatial error magnitude and noise condition in each tertile of experimental sessions for each monkey; logistic regressions, t test, H_0 : regression coefficient = 0, tertile values of $t = -8.79$,

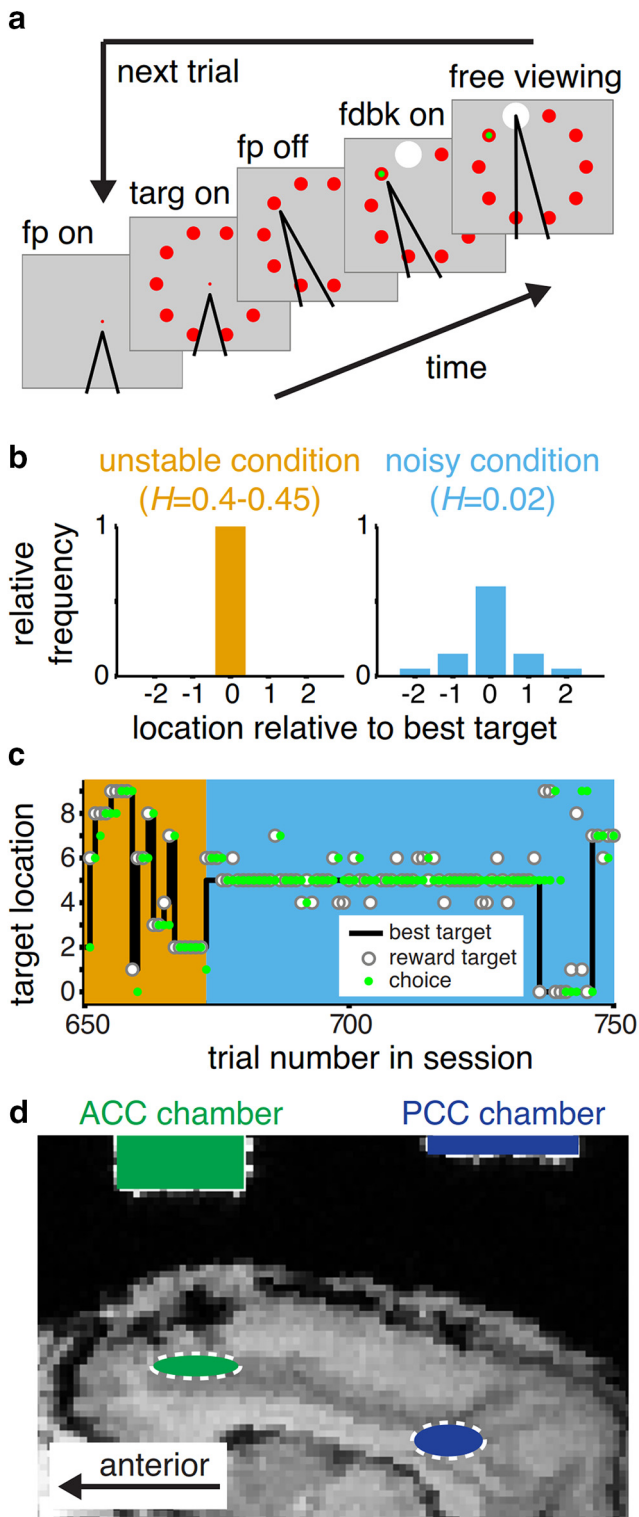


Figure 1. Task and recording sites. *a*, On each trial, the monkey chose 1 of 10 targets with a saccade. Visual feedback then indicated both the chosen target (small green dot) and the reward target (large white dot). Juice was then delivered if the monkey chose the reward target. *b*, The two task conditions were determined by the following: (1) the hazard rate (H) governing the rate at which the best target was repicked at random from among the 10 targets; and (2) the noise in the distribution of reward probability as a function of distance from the best target. *c*, An example sequence of trials for Monkey AT. Background color represents the task condition, as in *b*. *d*, Structural MRI for Monkey SP, showing recording chambers (shaded rectangles) and approximate recording locations (dashed), determined as in Kalwani et al. (2009).

–16.9, and –13.6 for Monkey SP, –4.23, –19.5, and –13.6 for Monkey AT, $p < 0.001$ in each case), even as both monkeys' overall performance tended to improve over time (Spearman's ρ of best-target choices vs experimental session; $\rho = 0.52$ and $\rho = 0.32$, for Monkey SP in unstable and noisy conditions, respectively, $p < 0.05$ in each case; and $\rho = -0.12$, $\rho = 0.32$, and $\rho = 0.44$, $p < 0.05$, for Monkey AT in unstable and noisy conditions, respectively). Subsequent analyses therefore combine data across recording sessions.

Their adaptive behavior reflected context-specific differences in the effects of recent trial history on the current choice (Fig. 4). For example, when encountering a spatial error of 1 (i.e., the rewarded target was immediately adjacent to the chosen target) in a noisy block, the monkeys were more likely to switch after several consecutive error trials than after several consecutive correct trials (Fig. 4*d,e*), but not when encountering the same spatial error of one in an unstable block (Fig. 4*a,b*). More generally, the relationship between feedback history and behavioral adjustment was stronger and extended further back in time for trials in which the feedback was more uncertain (e.g., noisy trials with small error magnitudes) than for trials in which the feedback was more informative (e.g., noisy trials with large error magnitudes or unstable trials). These adaptive dependencies on feedback history could not be explained simply by the history of choices: a logistic regression model that included both feedback history and coregulators representing choice history (the pattern of stay vs switch on previous trials) showed stronger feedback dependence in the noisy versus unstable conditions, particularly in terms of using feedback from at least one trial back (Fig. 4*g,h*).

Thus, the monkeys' behavioral adjustments to error feedback depended on both how predictive the error was of a change-point and the monkey's current uncertainty about the environment, integrated across multiple trials in a context-dependent manner. These behavioral characteristics are consistent with normative (Bayesian) theory and human behavior on comparable change-point tasks (Courville et al., 2006; Daw et al., 2006; Behrens et al., 2007; Krugel et al., 2009; Nassar et al., 2010; O'Reilly et al., 2013; Wilson et al., 2013; McGuire et al., 2014; Diederer and Schultz, 2015). The monkeys also exhibited non-normative tendencies to persist in choosing the same target, especially after positive feedback. Accordingly, a hybrid model with fully adaptive, Bayesian learning plus additional stay-switch sequential heuristics (Table 1) reproduced key features of the monkeys' behavior, including dependence on error magnitude and noise (Fig. 2*g,j*), as well as context-dependent influence of recent feedback (Fig. 4*c,f,i*). This model provided a better match to behavior than several other models, including: (1) a purely Bayesian model (Fig. 2*f,i*); (2) the heuristic model alone; (3) a standard, fixed-learning rate (i.e., not adaptive to task context) reinforcement-learning model that included the same heuristics; or (4) an alternative heuristic model that always switched for large errors and stayed for small errors unless a target has been consistently rewarded without adapting to task context (Table 2).

Spatial and feedback signals were represented differently in ACC and PCC

We recorded the activity of 102 ACC neurons ($n = 67$ from Monkey SP, 35 from Monkey AT) and 235 PCC neurons (75 from Monkey SP, 160 from Monkey AT) as the monkeys performed the change-point task.

Consistent with existing literature, individual neurons in both brain regions encoded spatial choice (the location of the chosen target on a given trial; Fig. 5*a,b*) and categorical feedback (correct

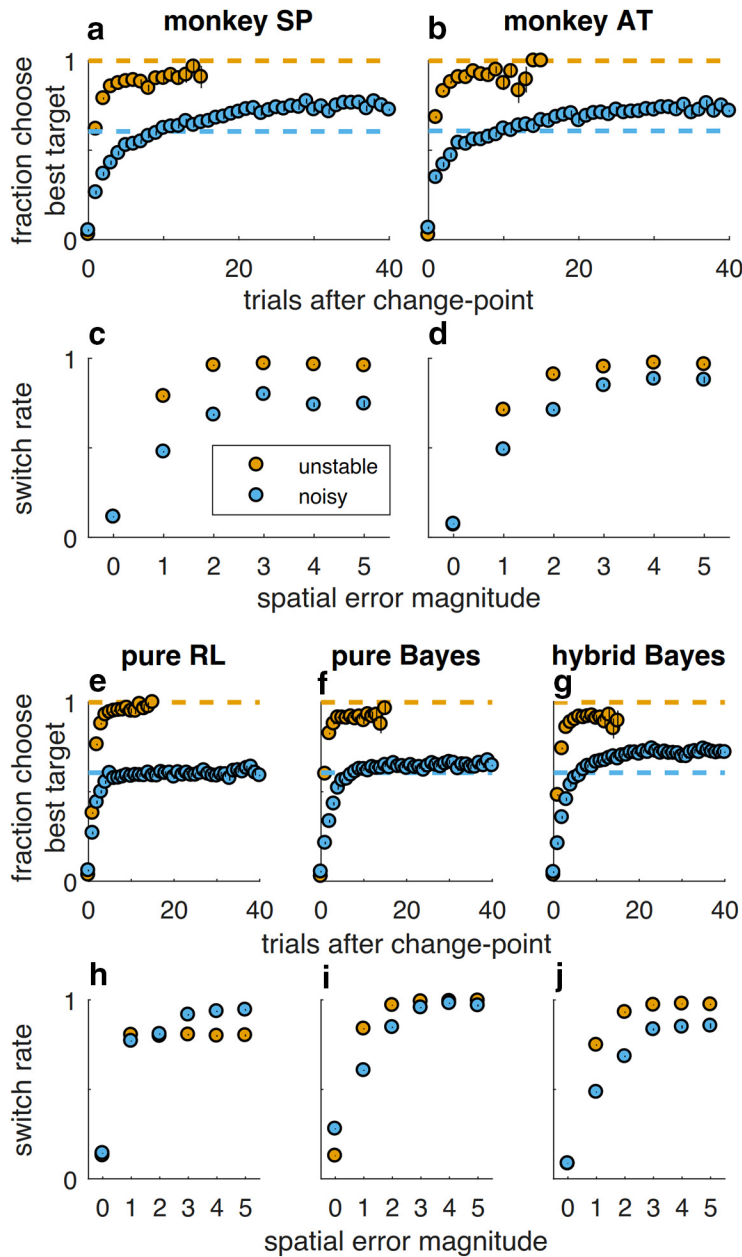


Figure 2. Adaptive, feedback-driven behavioral performance. Behavioral data from Monkeys SP and AT (top rows) and simulated behavioral data from median best-fit parameters of three behavioral models (bottom rows). **a, b, e–g**, Fraction of trials picking the best target as a function of trials after change-point, plotted separately for unstable (yellow) and noisy (blue) conditions. Dashed lines indicate the fraction of trials in which the best target was rewarded for each task conditions. **c, d, h–j**, Probability of switching to a different target on the following trial, per condition, as a function of the spatial error magnitude (the number of target locations between the chosen target and the reward target). Switch probability depended on spatial error magnitude (Spearman’s $\rho = 0.62$ for Monkey SP and 0.67 for Monkey AT, $p < 0.001$ for both) and task condition (partial ρ accounting for spatial error magnitude = -0.15 for Monkey SP, -0.10 for Monkey AT, $p < 0.001$ for both), particularly for small errors (median difference for errors of 1 and 2 = -0.31 vs $3+$ errors = -0.17 for Monkey SP, Wilcoxon signed-rank test, signed rank = 55.0 , $p < 0.001$; -0.24 vs -0.09 for Monkey AT, signed rank = 47.0 , $p < 0.001$). Colors as in **a**. In all panels, points/error bars indicate mean/SEM across all trials from recording sessions.

or error; Fig. 5*e,f*) (McCoy et al., 2003; Matsumoto et al., 2007; Hayden et al., 2008; Quilodran et al., 2008; Heilbronner and Platt, 2013). Individual neurons also encoded the location of the rewarded target (Fig. 5*c,d*) and graded signals reflecting spatial error magnitude (Fig. 5*e,f*), a type of prediction error that strongly influenced switching behavior (Fig. 2) and is similar to fictive learning signals identified previously in ACC (Hayden et al., 2009). None of these forms of selectivity could be explained solely

by postfeedback eye movements (Fig. 3), which were included as coregressors in the generalized linear model used to assess neuronal selectivity (Eq. 11). These spatial and feedback signals were represented in a broadly similar manner in ACC and PCC (Fig. 5*g,h*), but with important quantitative as well as qualitative differences (Figs. 6, 7; Table 3). ACC and PCC spatial selectivity tended to be unimodal for both choice- and reward-target tuning, with tuning widths that spanned ~ 4 – 5 visual targets (Table 3). The timing of the onset of selective responses for choice location, reward-target location, and correct versus error feedback was similar in the two brain regions and thus did not provide any evidence for sequential processing between the two. Moreover, selectivity for choice location tended to occur after the choice was made, implying that these spatial signals do not play a role in generating the choice but possibly helping to evaluate its outcome. However, there was substantially stronger selectivity for spatial choice and reward-target signals in PCC than in ACC (Fig. 6*a–d*; Table 3). In contrast, there was stronger selectivity for feedback in ACC than in PCC (Fig. 6*e,f*), but a preponderance of error-preferring neurons in PCC (mean β_5 for feedback = -0.666 , unpaired t test $p < 0.001$, $t = -6.18$; mean β_6 for error magnitude = 0.208 , $p < 0.001$, $t = 5.33$) that was not evident in ACC (mean $\beta_5 = -0.081$, $p = 0.8$, $t = -0.26$; mean $\beta_6 = -0.011$, $p = 0.8$, $t = -0.24$). These differences in selectivity imply possibly different roles for the two brain regions in postdecision processing.

Consistent with this idea, we found dynamic encoding of spatial signals by certain error-sensitive neurons in PCC but not ACC (Fig. 7). Specifically, because of the spatial nature of our task, effective, error-driven decision-making required knowing the locations of both the choice and the reward target, which occurred at two different times in a trial. The example ACC and PCC neurons in Figure 7*a, b* demonstrate different ways in which choice and reward target locations were encoded by individual neurons. The ACC neuron retained similar selectivity for choice location throughout the postsaccade and feedback epochs (note similar, vertically oriented firing patterns in Fig. 7*a*), with some additional modulation by reward-target location in the feedback epoch. By contrast, the PCC neuron was selective for a particular location that corresponded to the monkey’s choice in the postsaccade epoch (Fig. 7*b*, left, vertically oriented patterns) but that corresponded to the location of the reward target in the feedback epoch (more horizontally oriented firing patterns in Fig. 7*b*, right).

These different patterns of spatial and feedback selectivity were representative of the populations of neurons we recorded in ACC and PCC. Spatial selectivity was fairly common in the two brain regions. Of 102 recorded ACC neurons, 29 (28.4%) exhibited choice selectivity, 27 (26.5%) exhibited reward-target selectivity, and 13 (12.7%) exhibited selectivity for both during the feedback epoch, as computed by the GLM (Eq. 11). Of 235 recorded PCC neurons, 117 (49.8%) exhibited choice selectivity, 103 (43.8%) exhibited reward-target selectivity, and 76 (32.3%) exhibited selectivity for both during the feedback epoch. To illustrate the encoding schemes seen in single neurons (Fig. 7*a,b*) across the population of recorded neurons, we: (1) constructed 2D tuning curves for each neuron, using z-scored firing rates for each combination of choice and reward-target location; (2) aligned both dimensions relative to the preferred choice location during the postsaccadic epoch; and (3) averaged across populations of error- or correct-prefering neurons in each brain region. In the postsaccadic epoch (Fig. 7*c–f*, top row), as expected by design, the population tuning curves reflected primarily the choice selectivity of the individual neurons. In the feedback epoch (Fig. 7*c–f*, bottom row), the ACC and PCC populations showed strikingly different encoding schemes. For ACC neurons, spatial selectivity with respect to choice was somewhat maintained from the postsaccadic to the feedback epoch, albeit weaker in strength (Fig. 7*c,d*, compare top and bottom panels). For PCC neurons, particularly those that responded more strongly to error versus correct feedback (“PCC-”), choice selectivity in the postsaccadic epoch was transformed to reward-target selectivity at the same location in the feedback epoch (Fig. 7*e,f*).

To quantify these relationships for individual neurons, we computed correlation coefficients between the marginal tuning curves for choice and reward target for different time epochs. These marginal tuning curves were computed using only error trials to avoid inflating correlations (because choice and reward target locations were, by definition, the same on correct trials). For the population of ACC neurons, choice tuning in the postsaccadic epoch tended to be positively correlated with choice tuning in the feedback epoch (Fig. 7*g,h*, abscissa; median correlation coefficients = 0.394 and 0.406 for error- and correct-prefering neurons, respectively, H_0 : median = 0, Wilcoxon signed-rank test, signed rank = 1298.5 and 1039.5, respectively, $p < 0.001$ for both), but not reward-target tuning (Fig. 7*g,h*, ordinate; median correlation coefficients = -0.055 and 0.109, signed rank = 581.5 and 637.5, respectively, $p > 0.1$ for both). By contrast, for the population of PCC neurons, choice tuning in the postsaccadic epoch tended to be positively correlated with both choice tuning during the feedback epoch (Fig. 7*i,j*, abscissa; median correlation coefficients = 0.182 and 0.200 for error- and correct-prefering neurons, respectively, with signed rank = 9162 and 1933.5, $p < 0.001$ for both) and reward-target tuning during the feedback epoch (Fig. 7*i,j*, ordinate; median correlation coefficients = 0.248 and 0.261, signed rank = 10,932.5 and

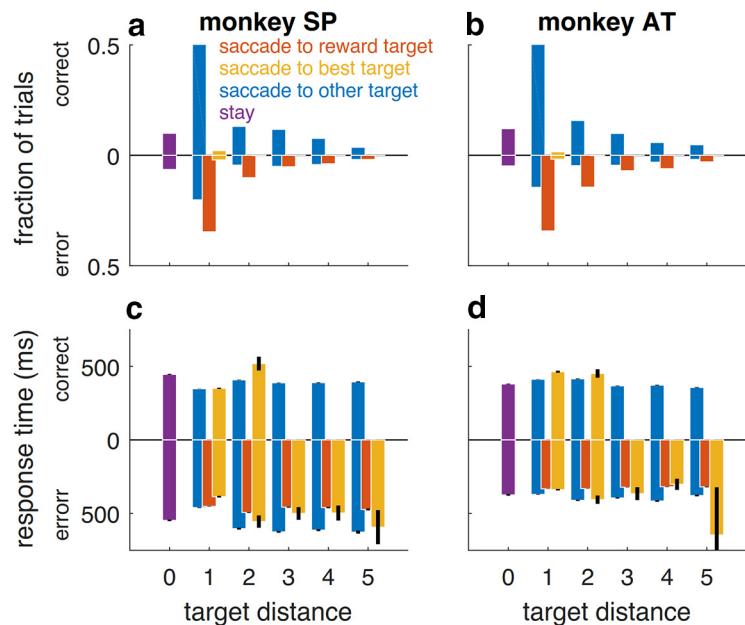


Figure 3. Postfeedback saccades. Summary of the location (top row) and timing (bottom row) of the first saccade made by Monkey SP (left column) and Monkey AT (right column) in the free-viewing period on each trial, after the saccadic choice was made and feedback was presented (Fig. 1*a*). In all panels, data from correct trials are shown as positive values on the ordinate, and data from error trials are shown as negative values on the ordinate. *a, b*, Histograms indicating the relative fractions of trials from all recording sessions in which the first free-viewing saccade was made to the location indicated. For error trials, both monkeys made the first saccade to the rewarded target (red) on most trials. *c, d*, Response times, which were approximately similar for all conditions. “Stay” response times were measured from trials in which the monkey made a small, postfeedback eye movement that landed near the original choice target.

1970.5, $p < 0.001$ for both). These relationships reflected systematically varying encoding of choice versus reward-target encoding in the feedback epoch across individual neurons, particularly in negative-prefering neurons: neurons that tended to lose their choice selectivity in the feedback epoch (smaller values along the abscissa in Fig. 7*g–j*) were also the ones that became selective for reward-target tuning at that time (larger values along the ordinate; Pearson’s correlation coefficient of the scatterplots shown in Fig. 7*g–j*; -0.310, $p = 0.024$; 0.051, $p = 0.73$; -0.408, $p < 0.001$; -0.159, $p = 0.185$, respectively).

Error encoding in cingulate was modulated by sequential, but not blockwise, task context

A primary finding from our behavioral analyses was that the monkeys used different strategies in the noisy versus stable block to integrate feedback information across trials to determine the location of the rewarded target (Figs. 2, 4). However, we found very little evidence for an associated blockwise modulation of feedback signals in either ACC or PCC. Specifically, the GLM (Eq. 11) identified selectivity for noise condition that was independent of recent reward rate and certain forms of slow drift and appeared to persist throughout each trial for ~20% of units in both brain regions (Fig. 8*a,b,e,f*). However, in most cases, this selectivity could not be distinguished from nonspecific, slow fluctuations in firing rate over the course of the recording session: only 3 of 29 neurons in ACC (10.3%, 95% CI = [2.2%, 27.4%]) and 11 of 84 neurons in PCC (13.1%, 95% CI = [6.7%, 22.2%]) that were classified as having selectivity for task condition had statistically reliable effects when using a permutation test that scrambled the identity of task conditions but maintained the overall temporal distribution of the task blocks ($p < 0.05$). There also were a small number of individual neurons in both brain

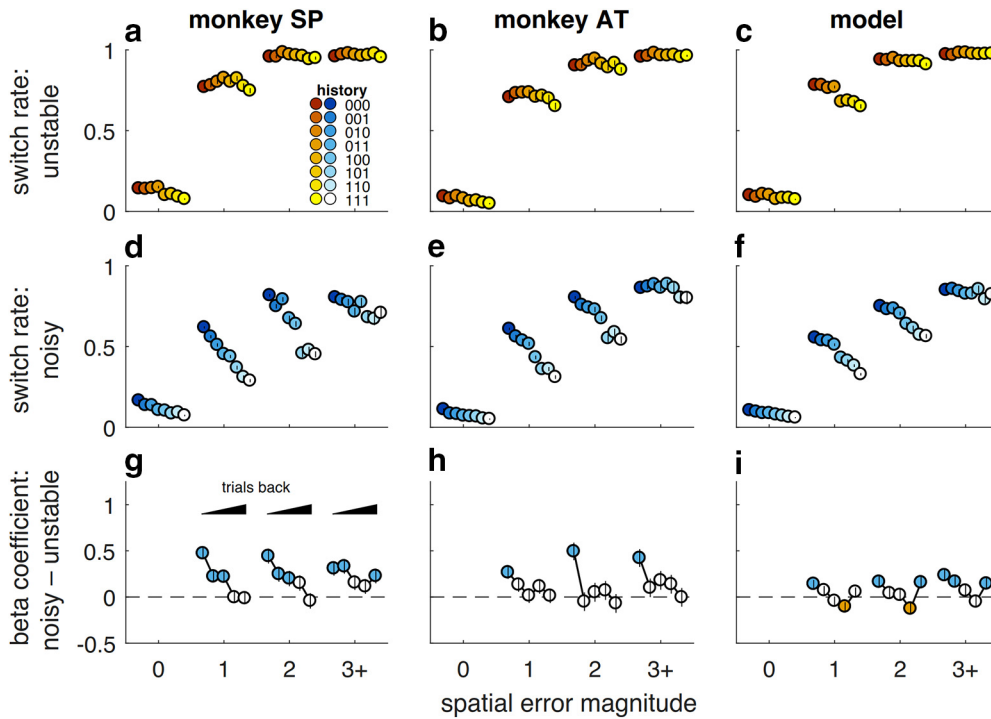


Figure 4. Adaptive integration of feedback history. Behavioral data from Monkeys SP (left column) and AT (middle column) and simulated data from the hybrid Bayesian model (right column). **a–c**, Probability of switching in the unstable condition. For each spatial error magnitude on the current trial (spatial errors ≥ 3 are grouped together as 3+), behavior is plotted separately for different combinations of feedback history for up to 3 trials before the current trial. In the legend, 0 = error and 1 = correct, such that, for example, 001 indicates a correct choice 3 trials back, followed by two consecutive error trials. Error bars indicate SEM across all trials from all recording sessions for the monkeys, or simulated sessions for the model. **d–f**, Same conventions as in **a–c**, but for the noisy condition. **g–i**, Difference in logistic regression coefficients for the influence of previous error trials on switching after the current trial in the noisy versus unstable condition. For each error magnitude, the influence of error feedback from up to 5 trials back is plotted (going from left to right) from more recent to more distant trials (the first point in each series represents one trial back, etc.). Positive coefficients imply that an error from that many trials back promoted more switching in the noisy versus unstable condition, above what was expected for that particular error magnitude on the current trial. Error bars indicate 95% CI. Filled symbols represent the Wald test for $H_0 =$ no difference in coefficients ($p < 0.05$).

Table 1. Best-fit model parameters for the hybrid Bayesian model

Parameter	Description	Median (IQR)	
		Monkey SP ($n = 67$ sessions)	Monkey AT ($n = 65$ sessions)
$H(\text{unstable})$	Hazard rate for unstable trials	0.70 (0.63, 0.76)	0.71 (0.63, 0.79)
$H(\text{noisy})$	Hazard rate for noisy trials	0.30 (0.23, 0.36)	0.36 (0.29, 0.43)
$\kappa(\text{unstable})$	Inverse-noise parameter for unstable trials ^a	4.27 (2.78, 5.34)	4.87 (3.62, 5.76)
$\kappa(\text{noisy})$	Inverse-noise parameter for noisy trials ^a	2.33 (1.25, 3.02)	2.44 (1.78, 3.01)
$s(\text{pos})$	Likelihood of stay after correct feedback	0.66 (0.59, 0.72)	0.78 (0.74, 0.81)
$s(\text{neg, noisy})$	Likelihood of stay after error feedback in unstable trials	0.01 (0.00, 0.03)	0.03 (0.01, 0.06)
$s(\text{neg, unstable})$	Likelihood of stay after error in noisy trials	0.16 (0.09, 0.20)	0.12 (0.08, 0.19)

^aFor reference, a von Mises distribution with κ of 4 has a half-maximal width of 68.5 degrees, spanning ~ 2 visual targets, whereas a distribution with κ of 2 correspond to a half-maximal width of 97.0 degrees, spanning ~ 3 visual targets.

Table 2. Comparison of behavioral models^a

Model	Description	Model parameters	ΔAIC mean (t statistic)	ΔBIC mean (t statistic)
Pure RL	RL model with fixed learning rates	3	711.62 (26.68)	691.46 (26.03)
Hybrid RL	RL model (above) with additional win-stay/lose-stay heuristics	6	377.45 (23.83)	372.42 (23.5)
Pure Bayes	Bayesian model with adaptive, context-dependent learning	5	120.57 (12.68)	110.50 (11.65)
Hybrid Bayes	Bayesian model (above) with additional win-stay/lose-stay heuristics	7	NA	NA
Heuristic 1	Heuristic model using win-stay/lose-stay strategies	3	1318.90 (29.71)	1298.73 (29.33)
Heuristic 2	Heuristic model that always switches for large errors and switches for small errors if the same target is consistently rewarded	3	5026.12 (19.55)	5005.95 (19.48)

^a ΔAIC and ΔBIC are computed relative to the hybrid Bayesian model as the standard model; positive differences indicate that the respective model is a worse match to observed data than the hybrid Bayes model. $p < 0.001$ for each pairwise t test comparison (H_0 : mean $\Delta = 0$).

regions that exhibited feedback encoding (different firing patterns on correct vs error trials) that was different in the unstable versus noisy block (Fig. 8c–f). However, across the population, there was no systematic relationship between feedback encoding

and task condition, in contrast to what might be expected if feedback encoding were more sensitive to errors in unstable compared with noisy conditions: the GLM coefficients for feedback encoding were not consistently correlated with GLM coefficients

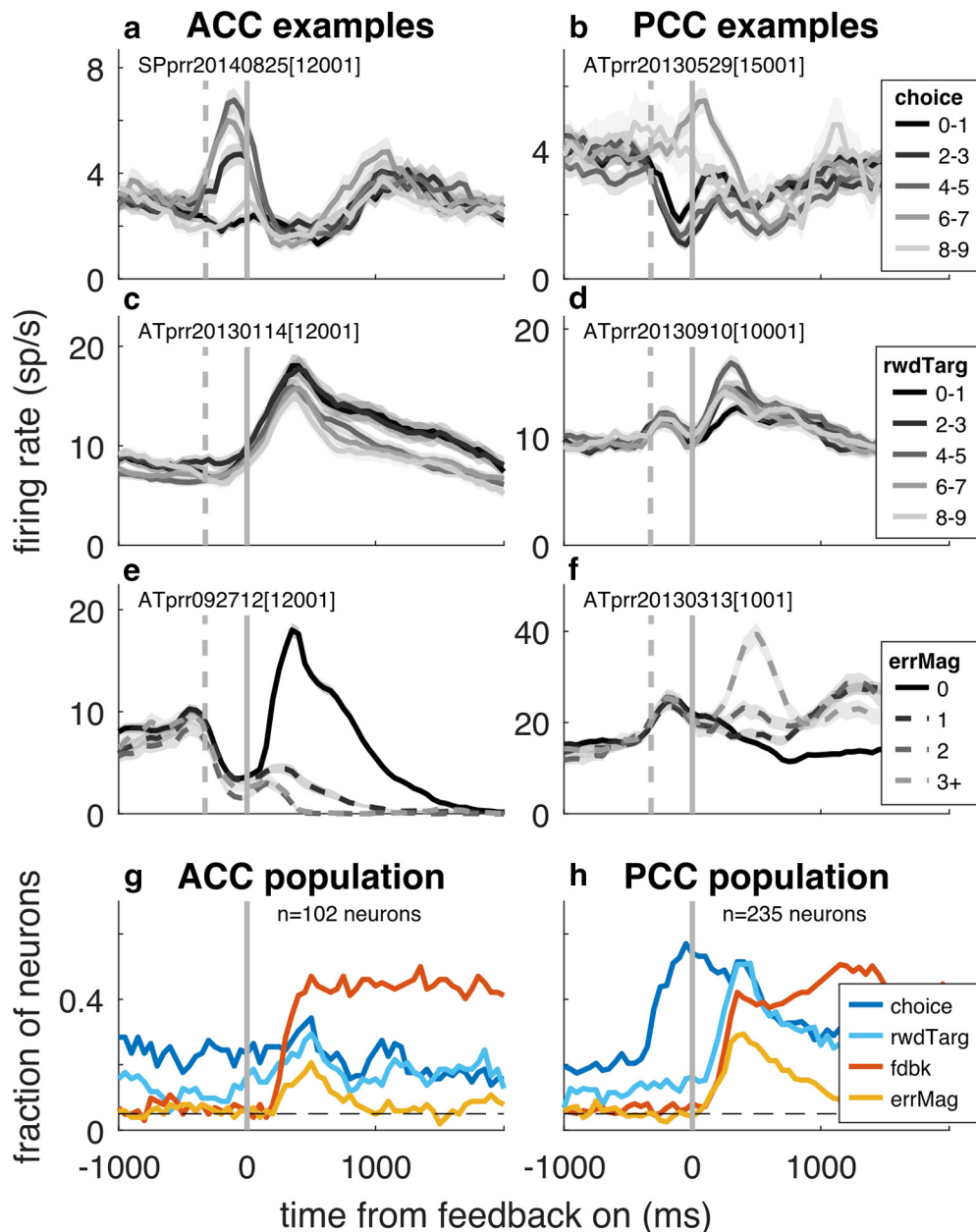


Figure 5. Neuronal selectivity for choice and feedback in ACC (left) and PCC (right). *a–f*, Example neurons illustrating neural selectivity for the following: *a, b*, the spatial location of the chosen target (numbered 0–9, indicating the location clockwise from directly right of fixation) determined from all correct and error trials; *c, d*, the spatial location of the rewarded target, numbered as in *a, b* and determined from all correct and error trials; and *e, f*, spatial error magnitude. Thick lines/ribbons indicate mean/SEM firing rates in 250 ms sliding windows. Dashed vertical gray lines indicate median times of saccade onset. Solid vertical lines indicate time of feedback onset. *g, h*, Population selectivity for ACC and PCC for the indicated task variables, based on GLM (Eq. 11) applied to firing rates in 250 ms bins stepped in 50 ms increments ($p < 0.05$). All panels represent data aligned to feedback onset.

for either noise (Fig. 8*g,h*) or noise \times feedback interaction term (Fig. 8*i,j*) for both monkeys. An alternative analysis comparing ROC-based selectivity index (ROCI) for each neuron assessed separately for the two block conditions also failed to find consistent blockwise differences: the slopes of linear regressions comparing an ROC-based selectivity index for each neuron assessed separately for the two conditions were not consistently different from unity for ACC (slope = 1.07 [95% CI = 0.89, 1.25] for Monkey SP, slope = 1.31 [1.13, 1.81] for Monkey AT, permutation test) or for PCC (slope = 1.04 [0.75, 1.30] for Monkey SP, slope = 0.64 [0.51, 0.82] for Monkey AT).

Similarly, we failed to find a systematic difference in spatial selectivity, for either choice or reward target, as a function of

noisy versus unstable conditions. For ACC, the median difference in ROCI for choice during the postsaccadic epoch in the noisy versus unstable conditions was 0.003 (H_0 : median = 0, signed-rank test, $p = 0.8$, signed rank = 2633) and reward target during the feedback epoch was -0.007 ($p = 0.8$, signed rank = 2557). For PCC, the median difference for ROCI for choice was -0.001 ($p = 0.6$, signed rank = 13,276) and for reward-target was 0.010 ($p = 0.056$, signed rank = 15,861).

In contrast, the monkeys' behavioral sensitivity to sequential feedback patterns was reflected in both ACC and PCC firing patterns. Neurons in both brain regions exhibited robust selectivity for feedback on the previous trial, which has been reported for both brain regions (Seo and Lee, 2007; Hayden et al., 2008; Ken-

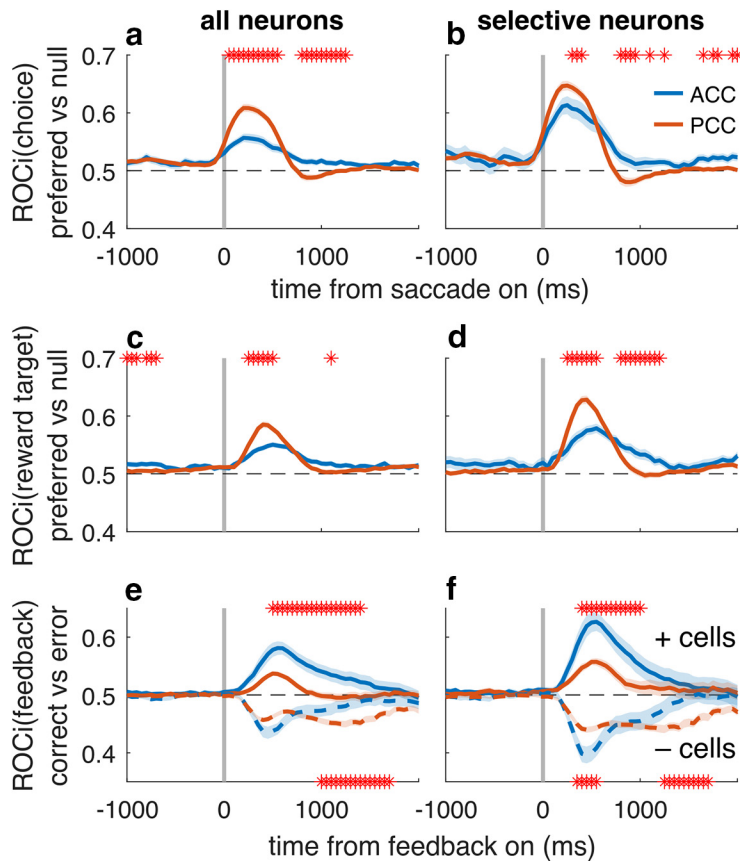


Figure 6. Differential encoding of spatial and feedback signals in ACC (blue) and PCC (red). Left column represents ROC-based analyses (values $\neq 0.5$ imply selectivity computed in 250 ms sliding windows) applied to all recorded neurons. Right column represents the same analyses applied to just the subset of neurons that exhibited selectivity for the given conditions (during the saccade epoch for choice selectivity and during the feedback epoch for reward target and feedback selectivity). Lines/ribbons are mean/SEM across neurons. Asterisks indicate time bins in which ACC and PCC selectivity differed from each other (Wilcoxon rank-sum test for H_0 : equal median ROCi values, $p < 0.05$). **a, b**, Selectivity for the spatial location of the chosen target, computed with respect to each neuron's preferred direction (i.e., the saccade direction eliciting the highest average response in the feedback epoch) versus the opposite (null) direction. **c, d**, Selectivity for the spatial location of the rewarded target, computed with respect to each neuron's preferred direction (i.e., the saccade direction eliciting the highest average response in the feedback epoch) versus the opposite (null) direction. **e, f**, Selectivity for correct versus error feedback, computed separately for neurons that tended to respond more strongly to correct feedback (solid lines) or to error feedback (dashed lines).

nerley et al., 2011). In our recordings, this selectivity tended to be slightly more prevalent at the start of the trial than at the end of the current trial (Fig. 9*a,b,e,f*). Across the population, there was no overall preference for correct or error feedback on the previous trial (mean GLM coefficient for previous feedback = 0.142, H_0 : mean = 0, unpaired t test, $t = 1.26$, $p > 0.1$ for ACC and mean GLM coefficient = 0.039, $t = 0.74$, $p > 0.1$ for PCC). However, there was a systematic, negative relationship between selectivity for feedback on the current trial versus selectivity for feedback on the previous trials for both ACC and PCC in both monkeys, such that error-preferring neurons responded more when the previous trial was correct, and correct-preferring neurons responded more when the previous trial was an error (Fig. 9*g,h*).

Some neurons in each brain region ($\sim 20\%$) also exhibited more complex forms of selectivity for interactions between feedback on the previous versus current trial (Fig. 9*c-f*). In some cases, this interaction reflected behavioral sensitivity to consecutive correct or error trials (e.g., the correct-preferring ACC neuron in Fig. 9*c* responded most strongly to two consecutive correct trials and least to two consecutive error trials), whereas in other neurons, the interaction effect led to an opposite effect (e.g., the

error-preferring PCC neuron in Fig. 9*d* responded more to feedback on the current trial when it differed from that in the prior trial). Across the population, there was no consistent relationship between selectivity for feedback on the current trial and the previous \times current feedback interaction term (Fig. 9*i,j*).

Consistent with the paucity of adaptive, integrative signals in the neurons that we recorded, there was no strong link between the activity of individual neurons and behavioral choices. In particular, in the feedback epoch, activity that predicted the monkey's subsequent choice (stay vs switch) was present in about as many neurons as would be expected by chance: 6 ACC neurons (5.9% [95% CI = 2.2%, 12.4%] of neurons; H_0 : β_{11} in Eq. 11 = 0, t test, $p < 0.05$) and 17 PCC neurons (7.2% [4.3%, 11.3%] of neurons). Similar results were obtained when using other metrics of behavioral adjustments, including switch to reward target versus no, switch to a nearby versus distant target, or switch in unstable versus noisy conditions.

Discussion

Effective decision-making in dynamic and unpredictable environments requires adaptive forms of inference that use contextual information to interpret new data (Yu and Dayan, 2005; Courville et al., 2006; Adams and MacKay, 2007; Behrens et al., 2007; Fearnhead and Liu, 2007; Nassar et al., 2010; Wilson et al., 2013). Here we showed that rhesus monkeys can use this kind of contextual adaptation to solve a reward-seeking task in which the reliability of feedback differed in different task conditions. We further showed that individual neurons in certain parts of the ACC and PCC encoded various spatial choice and feedback signals that could be used to monitor actions. However, these neurons did not exhibit strong modulation by contextually relevant task conditions and had little direct relationship to subsequent behavioral adjustments. Together, these results imply that at least some parts of ACC and PCC are actively engaged in monitoring adaptive, feedback-driven decisions without contributing directly to the processes that uses that information to drive behavior.

Our novel task included two blockwise task conditions that differed in terms of how error feedback should be interpreted to reduce uncertainty about the spatial location of the rewarded target. The monkeys used adaptive strategies that were based on both trialwise and blockwise information to perform the task. Their tendency to adjust their behavior (switch choices) depended on both the spatial error magnitude and recent patterns of trial-by-trial feedback. These dependencies differed substantially for the noisy versus unstable blocks, indicating that the monkeys were not just performing a simple instructed-choice task based on the explicit feedback about the rewarded location on the previous trial. We accounted for the behavioral sensitivity

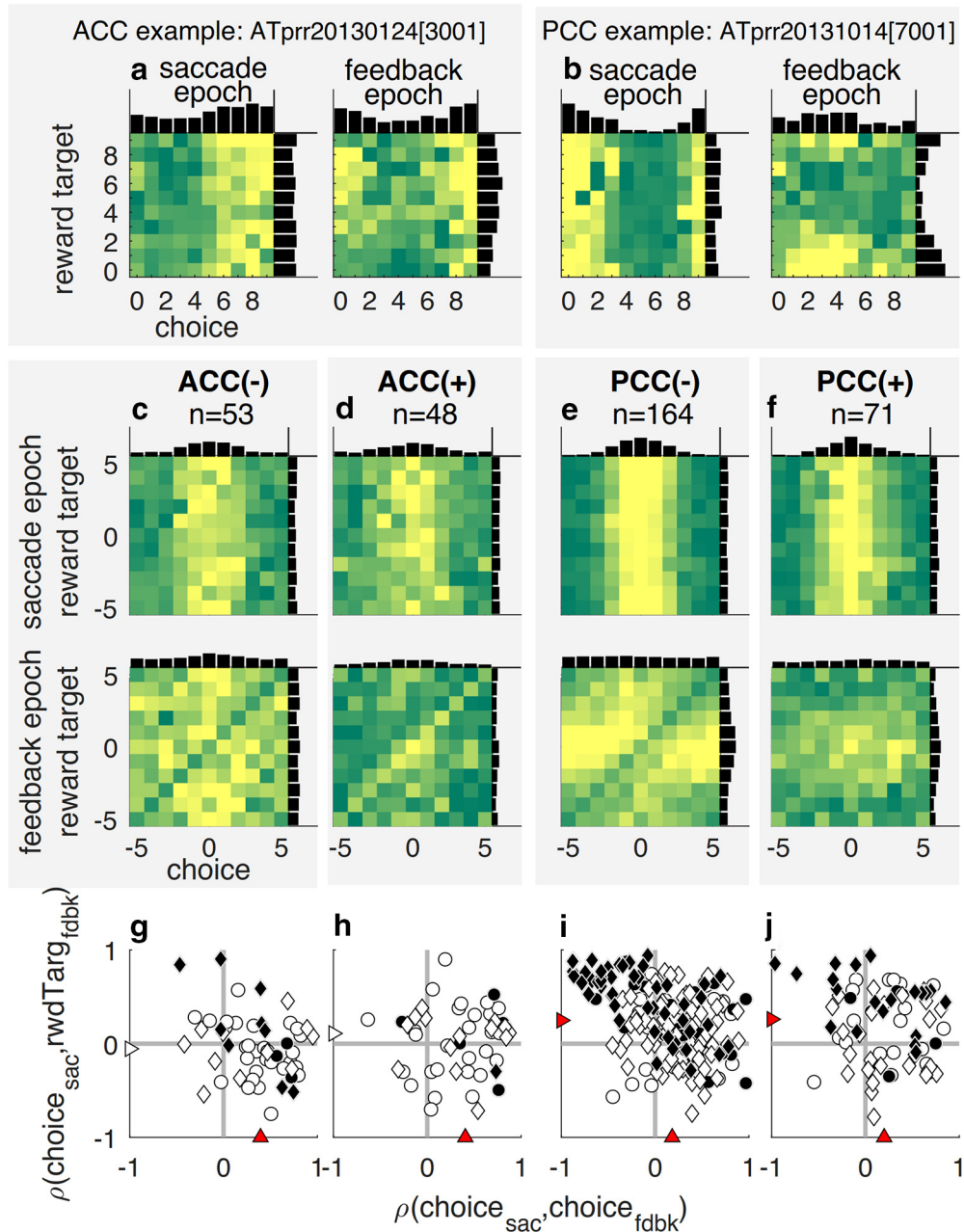


Figure 7. Neuronal encoding of spatially selective choice and reward-target signals. *a, b*, Example neurons from ACC (*a*) and PCC (*b*) illustrating spatial neural selectivity, with heatmaps showing joint histograms of firing rate as a function of both choice (horizontal axis) and reward target (vertical axis) during the postsaccade and feedback epochs. *c–f*, Population joint histograms for ACC (left two columns) and PCC (right two columns) neurons, averaged over z-scored firing rates of error-preferring (–, *c, e*) and correct-preferring (+, *d, f*) neurons. Abscissa indicates the location of the chosen target on the given trial, aligned with respect to each neuron’s preferred choice–target location. Ordinate indicates the location of the reward (feedback) target, aligned with respect to each neuron’s preferred choice–target location. Negative/positive values are counterclockwise/clockwise with respect to the preferred target. The color of each pixel is proportional to the mean firing rate across neurons (yellow represents higher firing rates) for the given choice (abscissa)/reward–target (ordinate) combination, measured in the postsaccadic (top row) or feedback (bottom row) epoch. Thus, a vertically aligned yellow streak indicates consistent choice tuning, whereas a horizontally aligned yellow streak indicates consistent reward–target tuning. Neurons were identified as (–) or (+) based on the sign of the coefficient for feedback (β_2) in GLM (Eq. 11). Marginal distributions are plotted along the edges in black in *a–f*. *g–j*, Relationship between choice tuning during the postsaccadic epoch and reward–target tuning during the feedback epoch (ordinate; positive/negative values of Spearman’s ρ imply that the given neuron tended to have matching/opposite selectivity for choice and reward feedback) is plotted against the relationship between choice tuning during the postsaccadic epoch and choice tuning during the feedback epoch (abscissa; positive/negative values imply that the given neuron tended to have matching/opposite choice selectivity at the two time points). Triangles along the axes represent the median correlation coefficients. Red triangles represent H_0 : median = 0 (Wilcoxon signed–rank test, $p < 0.05$). Circle represents Monkey SP; diamond represents Monkey AT. Filled symbols represent neurons selective for both choice and reward–target locations during the feedback epoch, according to GLM (Eq. 11).

of both monkeys to task condition, error magnitude, and feedback history using a mixture model that included both normative principles that operated over short (trial-by-trial) and long (block-by-block) timescales and non-normative heuristics that operated over short timescales (e.g., win-stay).

In contrast to the robust behavioral sensitivity to task conditions, individual neurons in ACC and PCC were relatively insensitive to those same blockwise task conditions. Consider, for example, the feedback-selective neurons that have been reported previously and in our study constituted approximately half of the

Table 3. Spatial tuning properties

	Selective <i>n</i> (%)	Unimodal <i>n</i> (%)	Tuning width median (IQR)	Tuning strength	
				ROCI mean (SEM)	Norm GLM coefficients mean (SEM)
Choice-target tuning^a					
ACC	28/102 (27.5%)	22 (78.6%)	167° (154°,171°)	0.645 (0.020)	1.745 (0.290)
PCC	131/235 (55.7%)	98 (74.8%)	161° (150°,170°)	0.689 (0.010)	2.311 (0.156)
ACC versus PCC (all neurons)			$p < 0.001$ (rank-sum = 19,533)	$p < 0.001$ (rank-sum = 13,465)	$p < 0.001$ (rank-sum = 13,343)
ACC versus PCC (selective neurons)			$p = 0.167$ (rank-sum = 2546)	$p = 0.055$ (rank-sum = 1815)	$p = 0.042$ (rank-sum = 1789)
Reward-target tuning^b					
ACC	27/102 (26.5%)	18 (66.7%)	167° (163°,172°)	0.578 (0.012)	1.188 (0.145)
PCC	103/235 (43.8%)	74 (71.8%)	163° (150°,170°)	0.630 (0.009)	1.584 (0.126)
ACC versus PCC (all neurons)			$p = 0.067$ (rank-sum = 18,743)	$p = 0.003$ (rank-sum = 14,771)	$p = 0.020$ (rank-sum = 15,114)
ACC versus PCC (selective neurons)			$p = 0.143$ (rank-sum = 2024)	$p = 0.002$ (rank-sum = 1235)	$p = 0.324$ (rank-sum = 1596)

^aChoice-target tuning computed for postsaccade epoch.

^bReward-target tuning computed for feedback epoch.

sampled populations in both brain regions (McCoy et al., 2003; Matsumoto et al., 2007; Hayden et al., 2008; Quilodran et al., 2008; Heilbronner and Platt, 2013). Based on both monkeys' behavioral patterns, which included normative tendencies to switch more often following small spatial errors in the unstable condition (in which any error indicated a change in the location of the best target) than in the noisy condition (in which small errors were expected even when the location of the best target did not change), we expected errors to be processed differently under those conditions. We found some neurons with different response patterns for correct versus error feedback that were modulated by the task block. However, these neurons were rare, and in general there was no systematic difference in response to errors as a function of task condition. More generally, overall fluctuations in spiking activity that appeared to covary with task block in most cases could not be distinguished from slow fluctuations that naturally occur over similar timescales.

These results contrast with previous reports of adaptive processing by individual cingulate neurons to guide action selection (Hayden et al., 2008; Quilodran et al., 2008; Sheth et al., 2012; Heilbronner and Platt, 2013) as well as contextual signals in more aggregate measures of ACC activity, such as fMRI (Behrens et al., 2007; Krugel et al., 2009; Jessup et al., 2010; O'Reilly et al., 2013; Payzan-LeNestour et al., 2013; McGuire et al., 2014). It is possible that the monkeys' use of stay/switch heuristics obscured our ability to detect adaptive integrative signals in our neural population. However, the discrepancy between the strong behavioral signatures of adaptive decision-making, and the paucity of these signatures in neural activity is striking. Overall, these results support the notion that at least some subsets of feedback-encoding neurons in both ACC and PCC are insensitive to certain relatively long-term contextual factors that are needed to interpret feedback appropriately and guide adaptive behaviors.

In contrast to their insensitivity to blockwise context, both brain regions contained feedback-sensitive neurons that were modulated by a more short-term contextual factor: the feedback received on the previous trial, a form of sequential dependency that has correlates in components of event-related potentials that have been linked to the ACC (Holroyd and Coles, 2002; Walsh and Anderson, 2012) and that has been identified in single-neuron activity in both ACC and PCC (Seo and Lee, 2007; Hayden et al., 2008; Kennerley et al., 2011). In our study, feedback selectivity in both brain regions tended to reverse on sequential trials: neurons that preferred errors on the current trial tended to prefer correct feedback on the previous trial, and vice versa. Such an encoding scheme is sensitive to differences in sequential outcomes, akin to an "edge-detector," which can be eff-

ective for detecting changes in the unstable condition but is less adaptive in the noisy condition of our task. This result was consistent with our other analyses that identified only minimal relationships between neural activity and subsequent behavioral choices, suggesting that the parts of cingulate cortex that we studied are more involved in monitoring previous choices than generating future ones.

These findings build on a host of prior studies that identified similar evaluative processes in cingulate cortex that support learning but, unlike our study, in some (but not all) cases also reported evidence for more direct roles in action selection (Amiez et al., 2006; Hayden et al., 2008; Quilodran et al., 2008; Donahue et al., 2013). The ACC has long been implicated in high-order cognitive functions, including cognitive control, conflict monitoring, pain and emotional processing, and reward processing (Behrens et al., 2007; Walton et al., 2007; Cole et al., 2009; Alexander and Brown, 2011; Rushworth et al., 2011; Shenhav et al., 2013). Lesion studies further suggest causal contributions to both the evaluation of recent feedback and use of that evaluation to guide subsequent decisions (Kennerley et al., 2006; Rudebeck et al., 2008; Buckley et al., 2009). Our findings emphasize that the varied and integrative characteristics ascribed to the ACC as a whole are not necessarily reflected in the activity patterns of individual neurons located throughout its rather extensive anatomical borders, which likely includes subregions that may process feedback- and decision-related information in different ways (Vogt and Pandya, 1987; Vogt et al., 2006; Heilbronner and Haber, 2014). In our case, area 24 along the dorsal bank of the anterior cingulate sulcus, ~4–6 mm below the cortical surface, exhibited strong feedback and spatial selectivity, but little modulation by task conditions.

The PCC, which has been less well studied, has traditionally been designated as a node of the default-mode network, based largely on its functional correlations with other regions when subjects are not engaged in an experimental task (Raichle, 2015). More recently, it has been implicated in evaluating decision options and outcomes in the context of reward-driven behaviors (McCoy et al., 2003; Kable and Glimcher, 2007; Hayden et al., 2008; Bartra et al., 2013; Heilbronner and Platt, 2013; Pearson and Platt, 2013). Our results add to these findings by showing that many PCC neurons have a particular form of spatial encoding not as apparent in ACC. In particular, both regions included neurons that encoded the spatial location of the chosen target. ACC neurons often maintained this selectivity throughout the feedback period (e.g., a neuron that responded most strongly on a trial with an upward choice tended to maintain that strong response until feedback was provided). In contrast, many PCC

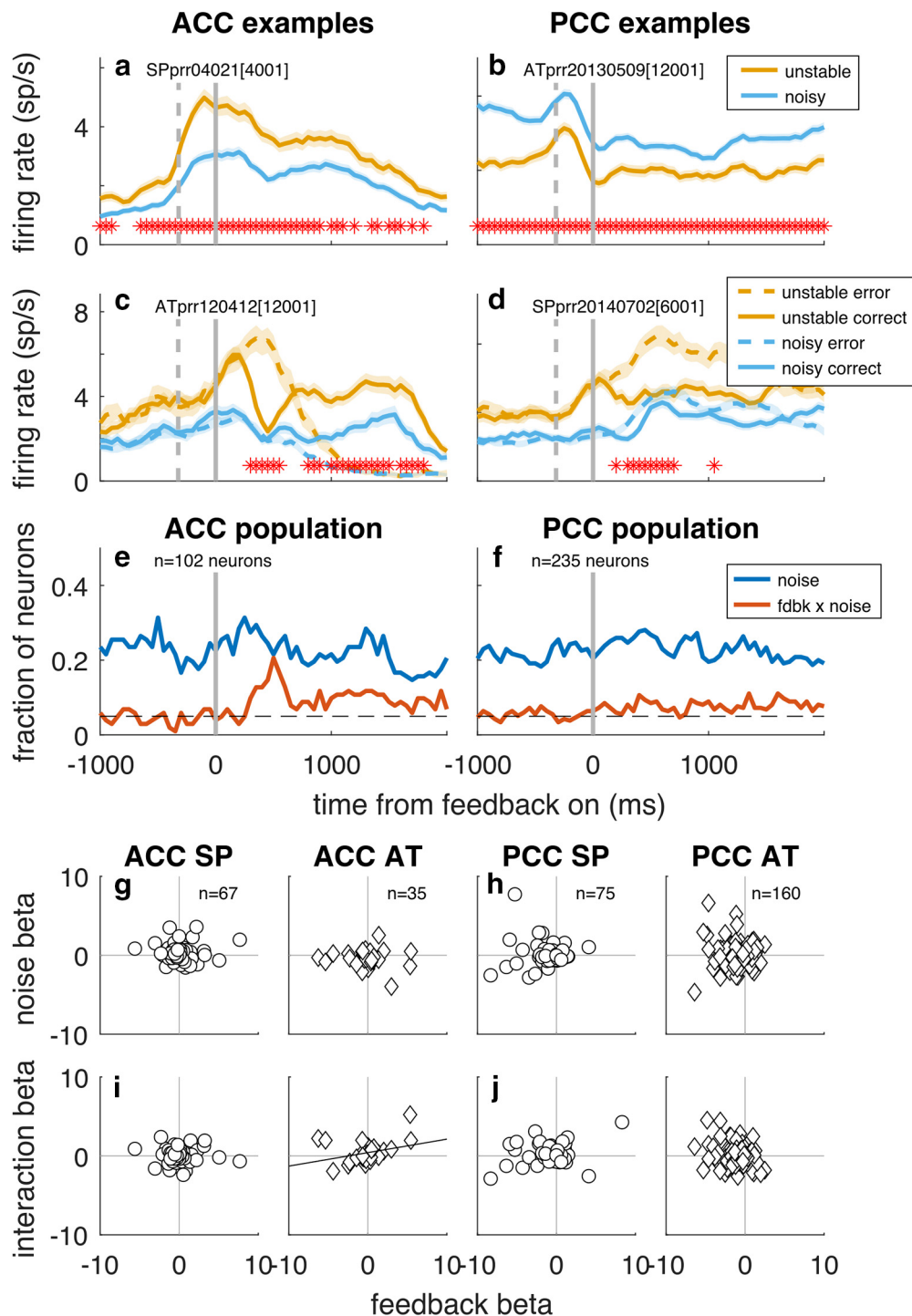


Figure 8. Little systematic encoding of blockwise noise condition in ACC (left) and PCC (right). **a–d**, Example neurons that responded differently in unstable versus noisy blocks either overall (**a, b**) or interacting with feedback responses (**c, d**). Thick lines/ribbons indicate mean/SEM firing rates in 250 ms sliding windows. Dashed vertical gray lines indicate median times of saccade onset. Solid vertical lines indicate time of feedback onset. Asterisks indicate time bins with different responses in the two blockwise conditions. **e, f**, Population selectivity for ACC and PCC for the indicated task variables as computed by GLM (Eq. 11), using firing rates in 250 ms bins stepped in 50 ms increments ($p < 0.05$). **a–f**, Data aligned to feedback onset. **g–j**, Population scatter plots showing the absence of a consistent relationship between GLM coefficients for feedback and either noise (**g, h**) or noise \times feedback interaction (**i, j**), shown separately for each monkey (circle represents Monkey SP; diamond represents Monkey AT). For each scatter plot, the least-squares line is shown if there was a significant correlation coefficient for the associated scatter ($p < 0.05$, H_0 : Spearman's $\rho = 0$).

neurons did not maintain those choice-selective responses, which were replaced by selectivity for error feedback presented at that same location. Thus, whereas many cingulate neurons tended to encode feedback with respect to the choice that was made, certain PCC neurons tended instead to encode feedback with respect to the choice that should have been made.

Overall, our results suggest that these two strongly reciprocally connected brain regions may play somewhat distinct, although overlapping, roles in tracking variables needed for effective, adaptive behavior. However, the observed regional differences should be interpreted with caution. ACC and PCC are large anatomical regions and thus are likely to have within-region specializations, including

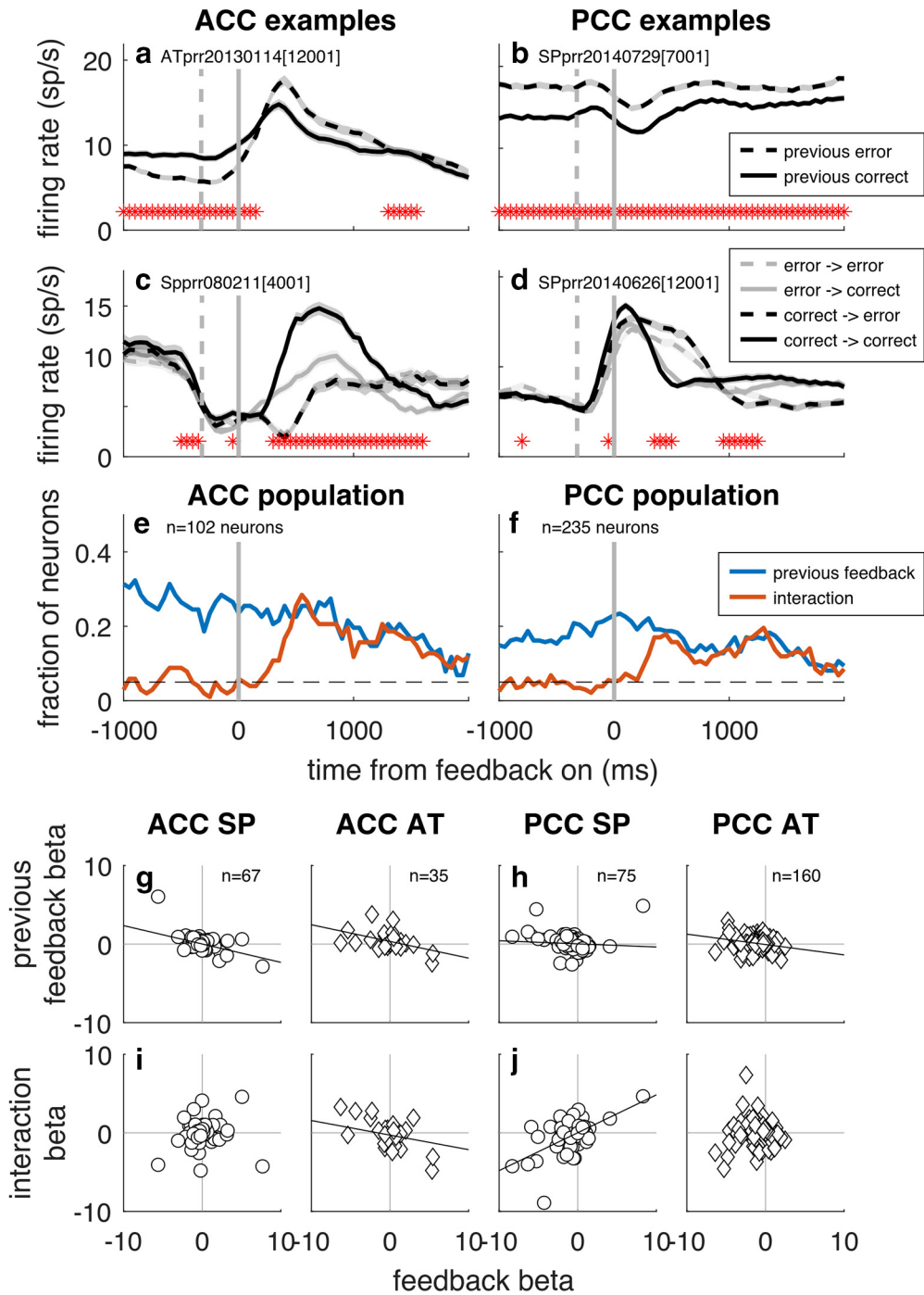


Figure 9. Sequential feedback encoding in ACC (left) and PCC (right). **a–d**, Example neurons that responded differently when the previous trial was correct (solid lines) or an error (dashed lines) either overall (**a, b**) or interacting with feedback responses on the current trial (**c, d**). Thick lines/ribbons indicate mean/SEM firing rates in 250 ms sliding windows. Dashed vertical gray lines indicate median times of saccade onset. Solid vertical lines indicate time of feedback onset. Asterisks indicate time bins with different responses in the two blockwise conditions. **e, f**, Population selectivity for ACC and PCC for the indicated task variables, using spike counts in 250 ms bins stepped in 50 ms increments ($p < 0.05$). **a–f**, Data aligned to feedback onset. **g–j**, Population scatter plots showing a consistently negative relationship between GLM coefficients for feedback on the current trial and feedback on the previous trial (**g, h**) but not current \times previous feedback interaction (**i, j**), shown separately for each monkey (circle represents Monkey SP; diamond represents Monkey AT). For each scatter plot, the least-squares line is shown if there was a significant correlation coefficient for the associated scatter ($p < 0.05$, H_0 : Spearman’s $\rho = 0$).

three distinct motor areas in the ACC (e.g., Heilbronner and Hayden, 2016). It is possible that our choice of where to record (intended to target subregions previously shown to have decision-making signals) resulted in sampling from a more “cognitive” portion of the ACC and more “motoric” portion of the PCC. Furthermore, cell-type differences in the two brain areas (predominantly agranular for ACC and granular for PCC) may have biased

the subsets of neurons sampled by our microelectrodes (Vogt et al., 2005). The fact that regional differences persist, even when we restrict our analyses to functionally selective subsets of the neurons, somewhat mitigates against these concerns. Future studies involving simultaneous recordings from ACC and PCC could shed further light on the respective roles of the two regions and the direction of information flow between the two.

More studies are also needed to identify where in the brain the kinds of feedback and spatial information we identified, along with other contextual factors, are integrated to drive behavior. One possibility is that these computations occur within the cingulate but involve different subsets of neurons than the ones we targeted or are based on emergent, population codes that are not directly evident in the response properties of single neurons (Karlsson et al., 2012). Another possibility is that aspects of these computations have already been performed upstream of the cingulate. Context-dependent prediction errors have been recorded in the midbrain dopaminergic system, a prominent input to ACC, and may contribute to the activity of the subset of cingulate neurons exhibiting context-dependent feedback signals (Nakahara et al., 2004; Nieuwenhuis et al., 2005; Tobler et al., 2005; Diederer and Schultz, 2015). However, the fact that feedback and contextual signals are separable in our cingulate recordings argues somewhat against this possibility and suggests that there are multiple parallel pathways for the computation and implementation of adaptive signals in the brain. These other pathways may involve the insula, orbitofrontal cortex, posterior parietal cortex, and even further downstream oculomotor planning areas, as have been suggested by imaging studies (O'Reilly et al., 2013; McGuire et al., 2014; Chau et al., 2015). An additional, not mutually exclusive, possibility is that at least some adaptive variables, such as those related to error magnitude and uncertainty, might be computed in cingulate but then passed to neuromodulatory systems, such as the locus ceruleus-norepinephrine system, which can, in turn, affect many other brain systems (Aston-Jones and Cohen, 2005; Nassar et al., 2012). Exactly how these systems interact to produce adaptive behavior remains to be tested.

References

- Adams RP, MacKay DJ (2007) Bayesian online changepoint detection. Cambridge, UK: University of Cambridge Technical Report.
- Alexander WH, Brown JW (2011) Medial prefrontal cortex as an action-outcome predictor. *Nat Neurosci* 14:1338–1344.
- 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.
- Bartra O, McGuire JT, Kable JW (2013) The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage* 76:412–427.
- Behrens TE, Woolrich MW, Walton ME, Rushworth MF (2007) Learning the value of information in an uncertain world. *Nat Neurosci* 10:1214–1221.
- Bertsekas DP, Tsitsiklis JN (1996) *Neuro-dynamic programming*. Belmont, MA: Athena Scientific.
- Botvinick MM, Cohen JD (2014) The computational and neural basis of cognitive control: charted territory and new frontiers. *Cogn Sci* 38:1249–1285.
- Brainard DH (1997) The psychophysics toolbox. *Spat Vis* 10:433–436.
- Buckley MJ, Mansouri FA, Hoda H, Mahboubi M, Browning PG, Kwok SC, Phillips A, Tanaka K (2009) Dissociable components of rule-guided behavior depend on distinct medial and prefrontal regions. *Science* 325:52–58.
- Burnham KP, Anderson DR (2004) Multimodel inference understanding AIC and BIC in model selection. *Soc Methods Res* 33:261–304.
- Chau B, Sallet J, Papageorgiou GK, Noonan MP, Bell AH, Walton ME, Rushworth MF (2015) Contrasting roles for orbitofrontal cortex and amygdala in credit assignment and learning in macaques. *Neuron* 87:1106–1118.
- Cole MW, Yeung N, Freiwald WA, Botvinick M (2009) Cingulate cortex: diverging data from humans and monkeys. *Trends Neurosci* 32:566–574.
- Courville AC, Daw ND, Touretzky DS (2006) Bayesian theories of conditioning in a changing world. *Trends Cogn Sci* 10:294–300.
- Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ (2006) Cortical substrates for exploratory decisions in humans. *Nature* 441:876–879.
- Dean HL, Platt ML (2006) Allocentric spatial referencing of neuronal activity in macaque posterior cingulate cortex. *J Neurosci* 26:1117–1127.
- Diederer KM, Schultz W (2015) Scaling prediction errors to reward variability benefits error-driven learning in humans. *J Neurophysiol* 114:1628–1640.
- Donahue CH, Seo H, Lee D (2013) Cortical signals for rewarded actions and strategic exploration. *Neuron* 80:223–234.
- Fearnhead P, Liu Z (2007) On-line inference for multiple changepoint problems. *J R Stat Soc B* 69:589–605.
- Fisher NI (1995) *Statistical analysis of circular data*. Cambridge: Cambridge UP.
- Hayden BY, Platt ML (2010) Neurons in the anterior cingulate cortex multiplex information about reward and action. *J Neurosci* 30:3339–3346.
- 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.
- Heilbronner SR, Haber SN (2014) Frontal cortical and subcortical projections provide a basis for segmenting the cingulum bundle: implications for neuroimaging and psychiatric disorders. *J Neurosci* 34:10041–10054.
- Heilbronner SR, Platt ML (2013) Causal evidence of performance monitoring by neurons in posterior cingulate cortex during learning. *Neuron* 80:1384–1391.
- Heilbronner SR, Hayden BY (2016) Dorsal anterior cingulate cortex: a bottom-up view. *Annu Rev Neurosci* 39:149–170.
- 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.
- Ito S, Stuphorn V, Brown JW, Schall JD (2003) Performance monitoring by the anterior cingulate cortex during saccade countermanding. *Science* 302:120–122.
- Jessup RK, Bussemeyer JR, Brown JW (2010) Error effects in anterior cingulate cortex reverse when error likelihood is high. *J Neurosci* 30:3467–3472.
- Jocham G, Neumann J, Klein TA, Danielmeier C, Ullsperger M (2009) Adaptive coding of action values in the human rostral cingulate zone. *J Neurosci* 29:7489–7496.
- Kable JW, Glimcher PW (2007) The neural correlates of subjective value during intertemporal choice. *Nat Neurosci* 10:1625–1633.
- Kalwani RM, Bloy L, Elliott MA, Gold JJ (2009) A method for localizing microelectrode trajectories in the macaque brain using MRI. *J Neurosci Methods* 176:104–111.
- Karlsson MP, Tervo DG, Karpova AY (2012) Network resets in medial prefrontal cortex mark the onset of behavioral uncertainty. *Science* 338:135–139.
- Kennerly SW, Walton ME, Behrens TE, Buckley MJ, Rushworth MF (2006) Optimal decision making and the anterior cingulate cortex. *Nat Neurosci* 9:940–947.
- Kennerly SW, Behrens TE, Wallis JD (2011) Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nat Neurosci* 14:1581–1589.
- Kolling N, Wittmann M, Rushworth MF (2014) Multiple neural mechanisms of decision making and their competition under changing risk pressure. *Neuron* 81:1190–1202.
- Krugel LK, Biele G, Mohr PN, Li SC, Heekeren HR (2009) Genetic variation in dopaminergic neuromodulation influences the ability to rapidly and flexibly adapt decisions. *Proc Natl Acad Sci U S A* 106:17951–17956.
- Lau B, Glimcher P (2005) Dynamic response-by-response models of matching behavior in rhesus monkeys. *J Exp Anal Behav* 84:555–579.
- Matsumoto M, Matsumoto K, Abe H, Tanaka K (2007) Medial prefrontal cell activity signaling prediction errors of action values. *Nat Neurosci* 10:647–656.
- McCoy AN, Platt ML (2005) Risk-sensitive neurons in macaque posterior cingulate cortex. *Nat Neurosci* 8:1220–1227.
- McCoy AN, Crowley JC, Haghhighian G, Dean HL, Platt ML (2003) Saccade reward signals in posterior cingulate cortex. *Neuron* 40:1031–1040.
- McGuire JT, Nassar MR, Gold JJ, Kable JW (2014) Functionally dissociable influences on learning rate in a dynamic environment. *Neuron* 84:870–881.
- Nakahara H, Itoh H, Kawagoe R, Takikawa Y, Hikosaka O (2004) Dopa-

- mine neurons can represent context-dependent prediction error. *Neuron* 41:269–280.
- Nassar MR, Wilson RC, Heasly B, Gold JI (2010) An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *J Neurosci* 30:12366–12378.
- Nassar MR, Rumsey KM, Wilson RC, Parikh K, Heasly B, Gold JI (2012) Rational regulation of learning dynamics by pupil-linked arousal systems. *Nat Neurosci* 15:1040–1046.
- Nieuwenhuis S, Heslenfeld DJ, von Geusau NJ, Mars RB, Holroyd CB, Yeung N (2005) Activity in human reward-sensitive brain areas is strongly context dependent. *Neuroimage* 25:1302–1309.
- Olson CR, Musil SY, Goldberg ME (1996) Single neurons in posterior cingulate cortex of behaving macaque: eye movement signals. *J Neurophysiol* 76:3285–3300.
- O'Reilly JX, Jbabdi S, Rushworth MF, Behrens TE (2013) Brain systems for probabilistic and dynamic prediction: computational specificity and integration. *PLoS Biol* 11:e1001662.
- Payzan-LeNestour E, Dunne S, Bossaerts P, O'Doherty JP (2013) The neural representation of unexpected uncertainty during value-based decision making. *Neuron* 79:191–201.
- Pearson JM, Platt ML (2013) Change detection, multiple controllers, and dynamic environments: insights from the brain. *J Exp Anal Behav* 99:74–84.
- Pearson JM, Heilbronner SR, Barack DL, Hayden BY, Platt ML (2011) Posterior cingulate cortex: adapting behavior to a changing world. *Trends Cogn Sci* 15:143–151.
- Quilodran R, Rothé M, Procyk E (2008) Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron* 57:314–325.
- Raichle ME (2015) The brain's default mode network. *Annu Rev Neurosci* 38:433–447.
- Rudebeck PH, Behrens TE, Kennerley SW, Baxter MG, Buckley MJ, Walton ME, Rushworth MF (2008) Frontal cortex subregions play distinct roles in choices between actions and stimuli. *J Neurosci* 28:13775–13785.
- Rushworth MF, Noonan MP, Boorman ED, Walton ME, Behrens TE (2011) Frontal cortex and reward-guided learning and decision-making. *Neuron* 70:1054–1069.
- 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.
- Shenhav A, Botvinick MM, Cohen JD (2013) The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron* 79:217–240.
- Sheth SA, Mian MK, Patel SR, Asaad WF, Ziv M, Williams ZM, Dougherty DD, Bush G, Eskandar EN (2012) Human dorsal anterior cingulate cortex neurons mediate ongoing behavioural adaptation. *Nature* 488:218–221.
- Sugrue LP, Corrado GS, Newsome WT (2004) Matching behavior and the representation of value in the parietal cortex. *Science* 304:1782–1787.
- Sutton RS, Barto AG (1998) Reinforcement learning: an introduction. Cambridge, MA: Massachusetts Institute of Technology.
- Thorndike EL (1911) *Animal intelligence: experimental studies*. New York: MacMillan.
- Tobler PN, Fiorillo CD, Schultz W (2005) Adaptive coding of reward value by dopamine neurons. *Science* 307:1642–1645.
- Vogt BA, Pandya DN (1987) Cingulate cortex of the rhesus monkey: II. Cortical afferents. *J Comp Neurol* 262:271–289.
- Vogt BA, Vogt L, Farber NB, Bush G (2005) Architecture and neurocytology of monkey cingulate gyrus. *J Comp Neurol* 485:218–239.
- Vogt BA, Vogt L, Laureys S (2006) Cytology and functionally correlated circuits of human posterior cingulate areas. *Neuroimage* 29:452–466.
- Walsh MM, Anderson JR (2012) Learning from experience: event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neurosci Biobehav Rev* 36:1870–1884.
- Walton M, Crosson P, Behrens TE, Kennerley S, Rushworth MF (2007) Adaptive decision making and value in the anterior cingulate cortex. *Neuroimage* 36:142–154.
- Wilson RC, Nassar MR, Gold JI (2013) A mixture of delta-rules approximation to Bayesian inference in change-point problems. *PLoS Comput Biol* 9:e1003150.
- Yu AJ, Dayan P (2005) Uncertainty, neuromodulation, and attention. *Neuron* 46:681–692.