Behavioral/Cognitive

# Neuronal Activity in the Posterior Cingulate Cortex Signals Environmental Information and Predicts Behavioral Variability during Trapline Foraging

David L. Barack[1] and Michael L. Platt[2]

[1]Departments of Philosophy, Neuroscience, and Economics, Columbia University, New York, New York 10027, and [2]Departments of Neuroscience, Psychology, and Marketing, University of Pennsylvania, Philadelphia, Pennsylvania 19104

Animals engage in routine behavior to efficiently navigate their environments. This routine behavior may be influenced by the state of the environment, such as the location and size of rewards. The neural circuits tracking environmental information and how that information impacts decisions to deviate from routines remain unexplored. To investigate the representation of environmental information during routine foraging, we recorded the activity of single neurons in posterior cingulate cortex (PCC) in 2 male monkeys searching through an array of targets in which the location of rewards was unknown. Outside the laboratory, people and animals solve such traveling salesman problems by following routine traplines that connect nearest-neighbor locations. In our task, monkeys also deployed traplining routines; but as the environment became better known, they deviate from them despite the reduction in foraging efficiency. While foraging, PCC neurons tracked environmental information but not reward and predicted variability in the pattern of choices. Together, these findings suggest that PCC may mediate the influence of information on variability in choice behavior.

*Key words:* cognition; exploration; foraging; information; posterior cingulate

### Significance Statement

Many animals seek information to better guide their decisions and update behavioral routines. In our study, subjects visually searched through a set of targets on every trial to gather two rewards. Greater amounts of information about the distribution of rewards predicted less variability in choice patterns, whereas smaller amounts predicted greater variability. We recorded from the posterior cingulate cortex, an area implicated in the coding of reward and uncertainty, and discovered that these neurons signaled the expected information about the distribution of rewards instead of signaling expected rewards. The activity in these cells also predicted the amount of variability in choice behavior. These findings suggest that the posterior cingulate helps direct the search for information to augment routines.

## Introduction

Imagine you are at a horse race, and there are 6 horses, with Local Field Potential (LFP) the underdog, facing 100:1 odds against. When LFP wins, a $1 bet will pay out $100. But in addition to the reward received from this bet, learning that, of the 6 horses, LFP is the winner reduces your uncertainty about the

outcome. Hence, LFP crossing the finish line first yields both reward and information.

Similar problems are often faced by organisms in their environment. Animals are adept at learning not only the sizes of rewards but also their locations, timing, or other properties. For example, hummingbirds will adapt their nectar foraging in response to unexpected changes in reward timing (Garrison and Gass, 1999). Similarly, monkeys will adapt their foraging routines on receiving information that a highly valued resource has become available (C. R. Menzel, 1991). In general, animals can make better decisions by tracking such reward information. Perhaps once a reward has been received, it no longer pays to wait for more because the resource is exhausted or the time between rewards is too great (McNamara, 1982), as occurs for some foraging animals. Or, perhaps receiving a reward also resolves any remaining uncertainty about an environment (Stephens and Krebs, 1986). Keeping track of reward information independent of reward size thus serves as an important input into animals' decision processes.

We designed an experiment to probe this oft-neglected informational aspect of reward-based decision-making. Our experiment is based on the behavior of animals that exploit renewable resources by following an efficient foraging path, a strategy known as traplining (Freeman, 1968; Berger-Tal and Bar-David, 2015). Trapline foraging has a number of benefits, including reducing the variance of a harvest and thereby attenuating risk (Possingham, 1989), efficiently capitalizing on periodically renewing resources (Possingham, 1989; Bell, 1990; Ohashi et al., 2008), and helping adapt to changes in competition (Ohashi et al., 2013). Many animals trapline, including bats (Racey and Swift, 1985), bees (Manning, 1956; Janzen, 1971), butterflies (Boggs et al., 1981), hummingbirds (Gill, 1988), and an array of primates, including rhesus macaques (E. W. Menzel, 1973), baboons (Noser and Byrne, 2010), vervet monkeys (Cramer and Gallistel, 1997), and humans (Hui et al., 2009). Wild primates foraging for fruit (E. W. Menzel, 1973; Noser and Byrne, 2010), captive primates searching for hidden foods (Gallistel and Cramer, 1996; Desrochers et al., 2010), and humans moving through simulated (MacGregor and Chu, 2011) and real (Hui et al., 2009) environments all use traplining to minimize total distance traveled and thereby maximize resource intake rates.

Although many primates trapline, information about the state of the environment, such as weather (Janmaat et al., 2006), the availability of new foods (C. R. Menzel, 1991), or possible feeding locations (Hemmi and Menzel, 1995; C. R. Menzel, 1996), can influence choices made while foraging. Such detours result in longer search distances and more variable choices (Hui et al., 2009; Noser and Byrne, 2010) but allow animals to identify new resources (C. R. Menzel, 1991) and engage in novel behaviors (Noser and Byrne, 2010). These benefits are consistent with computer simulations that show traplining with variation in routes yields better long-term returns than traplining without variation by uncovering new resources or more efficient routes (Ohashi and Thomson, 2005). In this way, environmental information may improve foraging efficiency during routine foraging over the longer term.

The neural mechanisms that track, update, and regulate the impact of environmental information on decision-making remain unknown. Neuroimaging studies have revealed that the posterior cingulate cortex (PCC) is activated by a wide range of cognitive phenomena that involve rewards, including prospection (Benoit et al., 2011), value representation (Kable and Glimcher, 2007; Clithero and Rangel, 2014), strategy setting (Wan et al., 2015), and cognitive control (Leech et al., 2011). Intracranial recordings in monkeys have found that PCC neurons signal reinforcement learning strategies (Pearson et al., 2009), respond to novel stimuli during conditional visuomotor learning (Heilbronner and Platt, 2013), represent value (McCoy et al., 2003), risk (McCoy and Platt, 2005), and task switches (Hayden and Platt, 2010), and stimulation there can induce shifts away from a default option (Hayden et al., 2008). Together, these observations suggest that the PCC mediates the effect of environmental information on variability in routine behavior. However, no studies to date have attempted to disentangle hedonic value from the informational value of rewards in PCC.

Previously, we reported that, in our traplining task, neurons in PCC increased their firing rates during choices before decisions to diverge from the typical trapline, the most common circular pattern of choices (Barack et al., 2017). We reported decisions to diverge from typical traplines were driven by the salience of the pattern of total rewards during foraging. PCC neuron firing rates predicted decisions to diverge from typical traplines and signaled the interaction between foraging decision salience, reward, and time. Finally, these cells displayed a large transient increase in activity before decisions to diverge that was especially marked in low reward rate environments.

Here, we explore how information influenced decisions to deviate from traplines (circular patterns of choices) and test the hypothesis that PCC tracks reward information. We recorded the activity of PCC neurons in monkeys foraging through an array of targets in which environmental information, operationalized as the pattern of rewards, was partially decorrelated from reward size. Monkeys developed traplines in which they moved directly between nearest neighbor targets in a circle. When they expected more information about the state of the environment, their trapline foraging behavior was less variable. While foraging, PCC neurons tracked environmental information but not reward and forecast variability in choice patterns. These findings support our hypothesis that PCC mediates the use of information about the state of the environment to regulate adherence to routines in behavior and cognition.

## Materials and Methods

### Task analysis

Our experiment required monkeys to select each target in a set of six targets to harvest the rewards. In every trial in our experiment, two fixed rewards (large and small) were assigned to one of six locations in the environment in a pseudorandom fashion (see Fig. 1B). Trials began with monkeys fixating a central cross for a variable amount of time, ranging from 0.5 to 1 s. After fixation offset, six targets arranged in a circle appeared. The same locations were used from trial to trial, and monkeys were free to select the targets in any order. To make a choice, monkeys had to fixate their gaze on a target for 250 ms. In order to advance to the next trial, monkeys had to choose each option, even after they had already harvested the reward available on that trial. Assuming the cost of making a saccade is a monotonic, positive-definite function of distance between targets, the most efficient solution to our task is to minimize saccade times between targets by searching in a circular pattern. This is referred to as a trapline, and sequences of choices that are noncircular are deviations from traplines.

Uncertainty about the current trial's pattern of received rewards is reduced over the course of the trial as the monkey proceeds through all of the targets. This reduction in uncertainty is quantifiable by examining how many possible patterns of rewards are excluded given the rewards revealed by previous choices. For a subset of patterns, the very same information outcome can be delivered by distinct rewards, serving to partially decorrelate and hence de-confound reward and information outcomes. Furthermore, expected reward and expected information, defined as the average amount of information contained in the next outcome given the pattern of rewards received so far, are also partially decorrelated (Table 1).

Given a set of six rewards (four zero, one small, and one large), there are 6! distinct permutations. We made the simplifying assumption that monkeys did not distinguish between the different zero rewards. This assumption reduces the number of distinct patterns from 720 to 30.

Different patterns correspond to different series of received reward. The environmental entropy $H_E$ contained in receiving some reward (zero, small, or large) depends on the choice number (CN) $i$ in the sequence and the total number of possible sequences as follows:

$$H_E = -\log_2 \frac{(|\{P_i\}|)}{(|P|)}$$

where $|\cdot|$ denotes cardinality, $P$ is the set of possible permutations, and $\{P_i\}$ is the set of remaining permutations after the $i^{th}$ choice. The amount of information contained in some reward outcome is computed as the difference
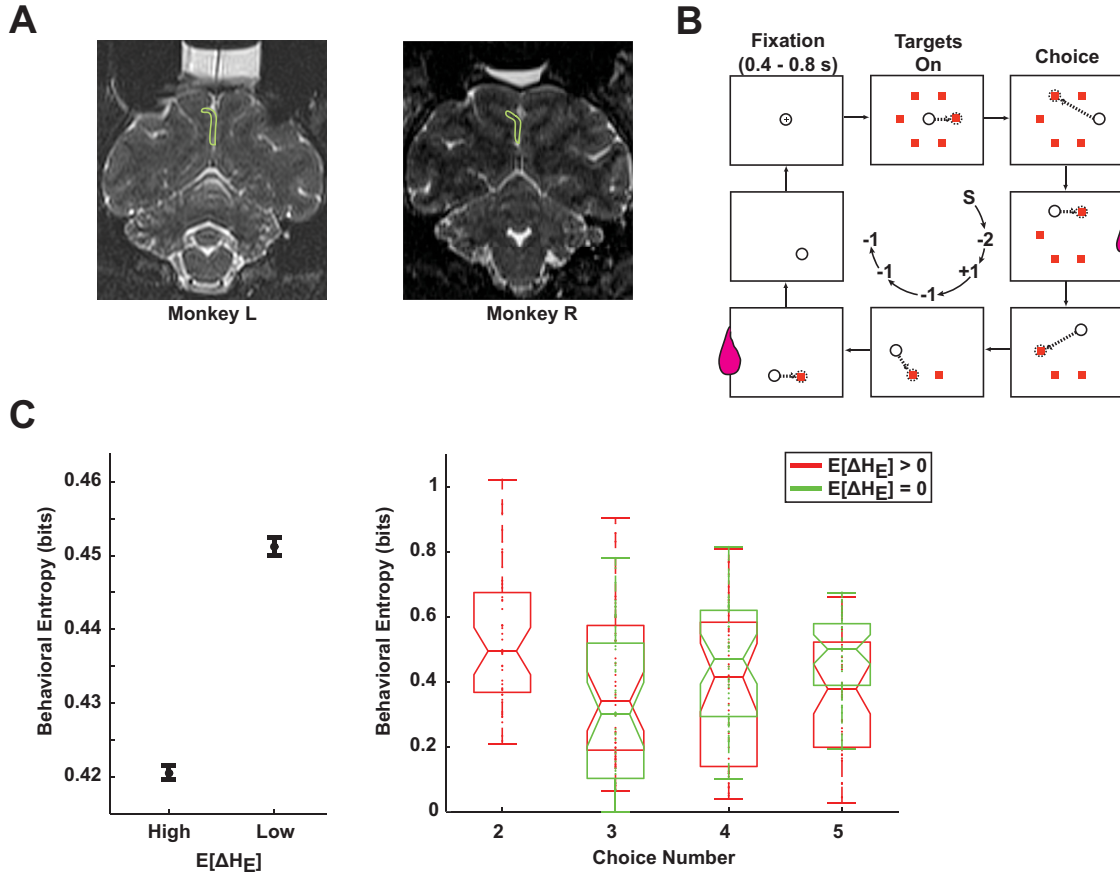
**Figure 1.** Monkeys spontaneously trapline, efficiently choosing targets in a circle, when foraging in a circular array but deviate from these routines as the environment becomes better known. **A**, Recording location in the PCC. Left, Monkey L. Right, Monkey R. **B**, Traplining task, sample trial sequence. Trials began with monkeys fixating a central cross for a variable amount of time. After fixation offset, six targets appeared in the same locations across trials. Monkeys then chose targets in any order. To register a choice, monkeys fixated targets for 250 ms. Only two rewards, one small and one large, were available on every trial, and the identity of the rewarded targets changed in a pseudorandom fashion from trial to trial. In order to advance to the next trial, monkeys had to select every target. Open circle represents simulated eye position. Dashed arrow indicates direction of impending saccade. Dashed circle represents impending saccade endpoint. Small juice drop represents small reward. Large juice drop represents large reward. Central semi-circle represents step size, the clockwise or counterclockwise distance between subsequently chosen targets. S, Start; −2, two targets counterclockwise; −1, one target counterclockwise; +1, one target clockwise. **C**, Left, Mean ± SEM BE for high expected environmental information ($\Delta H_E$) choices ($E[\Delta H_E] >$ mean($E[\Delta H_E]$)) compared with low expected information choices ($E[\Delta H_E] \leq$ mean($E[\Delta H_E]$)) across all sessions and choices. Right, Boxplot by CN2-CN5 across sessions before receipt of last informative outcome ($E[\Delta H_E] > 0$; red points) and after ($E[\Delta H_E] = 0$; green points). Top and bottom of box are interquartile (25%-75%) range of session means. Notch indicates 95% CI for median session. Nonoverlapping notches indicate significantly different medians at $\alpha = 0.05$. Each point is a session mean. CN2 always possesses some expected information; hence, no green box or points. $n = 145{,}524$ choices, 24,254 trials.

**Table 1. Equations for expected reward, entropy, information, and expected information for the reward sequence**

| Pattern no. | Permutation $p$ | Expected reward $ER_i = \frac{1}{n}\sum_i^n R_i$ | | Entropy $H_i = -log_2 \frac{(|\{P_i\}|)}{(|P|)}$ | Information $I_i = H_i - H_{i-1}$ | Expected information $EI_i = \frac{\sum_{P_{remaining}} I_i}{\left(|P_{remaining}|\right)}$ |
|---|---|---|---|---|---|---|
| 1 | 0 0 0 0 1 2 | 0.5 0.6 0.75 1.0 1.5 2.0 | | 0.5850 1.3219 3.3219 3.9069 4.9069 4.9069 | 0.5850 0.7370 1 1.5850 1 0 | 1.2516 0.7370 1 1.5850 1 0 |
| 2 | 0 0 0 0 2 1 | 0.5 0.6 0.75 1.0 1.5 1.0 | | 0.5850 1.3219 3.3219 3.9069 4.9069 4.9069 | 0.5850 0.7370 1 1.5850 1 0 | 1.2516 0.7370 1 1.5850 1 0 |
| 3 | 0 0 0 1 0 2 | 0.5 0.6 0.75 1.0 1.0 2.0 | | 0.5850 1.3219 3.3219 3.9069 4.9069 4.9069 | 0.5850 0.7370 1 1.5850 1 0 | 1.2516 0.7370 1 1.5850 1 0 |
| ⋮ | ⋮ | ⋮ | | ⋮ | ⋮ | ⋮ |
| 30 | 2 1 0 0 0 0 | 0.5 0.2 0 0 0 0 | | 2.5850 4.9069 4.9069 4.9069 4.9069 4.9069 | 2.5850 2.3219 0 0 0 0 | 1.2516 2.3219 0 0 0 0 |

Each column in the table (except for the leftmost) contains six columns, each corresponding to a CN during the trial. Total number of choices on every trial = 6. |·| represents cardinality of ·; $n = $ CN; $i = $ CN in trial. R, Reward.

in the entropy, what has been learned about the current trial's pattern of received reward by receiving the most recent outcome as follows:

$$\Delta H_E = H_i - H_{i-1}$$

for the amount of environmental entropy $H_E$ on the $i^{th}$ outcome. Expected information can then be computed as the mean amount of information to be gained by making the next choice, the weighted average over all possible next information outcomes given the pattern of rewards received as follows:

$$E[\Delta H_E]_i = \frac{\sum_{P_{remaining}} [\Delta H_E]_i}{\left(|P_{remaining}|\right)}$$

for expected information $E[\Delta H_E]$ for the $i^{th}$ choice, possible outcomes $[\Delta H_E]_i$ for the remaining permutations $P_{remaining}$, and where |·| again denotes cardinality. As the animal proceeds through the trial, the amount of expected information varies as a function of how many possible patterns of returns have been eliminated so far. Expected reward $ER$ is computed simply as the amount of remaining

reward to be harvested on trial $i$ divided by the number of remaining targets $n$ as follows:

$$ER_i = \frac{1}{n}\sum_i^n R_i$$

If the animal harvests all of the reward near the beginning of a trial, the expected reward will be zero. However, if the animal does not harvest the rewards until the end of a trial, the expected reward will increase across the duration of the trial.

The linear correlation coefficients between the different task variables (information, expected information, reward, expected reward, etc.) can be computed empirically from the total experienced reward outcomes and information outcomes, and from the total experienced reward expectations and information expectations, derived from the trials the monkeys actually experienced. For the anticipation epoch, this includes expected information and expected reward ($R^2 = 0.1324$), expected information and previous choice information outcome ($R^2 = 0.0292$), expected information and previous choice reward outcome ($R^2 = 0.1348$), expected reward and previous choice information outcome ($R^2 = 2.6458 \times 10^{-06}$), expected reward and previous choice reward outcomes ($R^2 = 0.1082$), and previous choice information outcome and previous choice reward outcome ($R^2 = 0.5971$). For the outcome epoch, this includes current choice information outcome and current choice reward outcome ($R^2 = 0.3935$).

*Experimental design and statistical analysis*

*Behavior.* In our experiments, 2 male rhesus macaques performed the task described above on custom software using Psychtoolbox (Brainard) and MATLAB (The MathWorks). All statistical comparisons were performed using custom software in MATLAB. Significance was Bonferroni-corrected for multiple comparisons, and significance assessed at $p < 0.05$.

For our behavioral entropy (BE) measures, we again used the standard definition of entropy. Step size was defined as the number of positions clockwise or counterclockwise of the target that the monkey chose in relation to the previous choice's target. For BE, the probability of a particular step size was computed for each step size by counting the number of trials with that step size and dividing by the total number of trials. Action step sizes (from −2 to 3) and action step size probabilities (probability of taking an action of a given size) were calculated for choices 1-2, 2-3, 3-4, and 4-5 (5-6 had a constant update of 1). Step sizes were calculated on each choice by determining how many targets around clockwise (positive) or counterclockwise (negative) the next choice was from the previous choice; already selected targets were not included in this calculation. Step size probabilities were calculated by holding fixed all of the covariates for a particular choice (information outcome from previous choice, information expectation for next choice, reward outcome from previous choice, reward expectation for next choice, and CN) and counting the frequencies for each step size and dividing by the total number of trials with that set of covariates. For each unique combination of covariates (CN, information outcome, information expectation, reward outcome, and reward expectation), we computed the choicewise BE ($H_B$) for that combination as follows:

$$H_B = -\sum_s p_s \log_2 p_s$$

for probability of each step size $p_s$. Finally, a multilinear regression correlated these BE scores with the covariates.

To analyze neural coding of expectations, we had to remove diverge choices, defined as choices that diverged from the daily dominant pattern. Determining the daily dominant pattern relied on assessing the similarity between pairs of trials, for every possible pair on a given day, by computing the pair's Hamming score (Hamming, 1950). To compute the similarity between two trials, each trial's pattern of choices by target number is first coded as a digit string (e.g., 1-2-4-5-6-3). The Hamming

distance $D_{i,i'}$ between two strings $i$, $i'$ of equal length is equal to the sum of the number of differences $d$ between each entry in the string, as follows:

$$D_{i,i'} = \sum_n d(x_n, y_n)$$

for strings $x$, $y$ of length $n$. We computed $D_{i,i'}$ for every pair of trials, and then, for each unique pattern of choices, computed the average Hamming distance $-D_{i,i'}$. The daily dominant pattern corresponded to the pattern with the minimum $-D_{i,i'}$ and corresponded to a circular pattern for both monkeys (see Barack et al., 2017). Since the daily dominant pattern was circular, we refer to these as the monkeys' typical traplines.

BE was regressed against a number of variables and their interactions using multilinear regression. Covariates included CN in trial, expected information, expected reward, reward outcome from the previous choice, information outcome from the previous choice, and all two-way interactions.

*Neural.* All neural data were analyzed on custom software in MATLAB. For all tests, significance was Bonferroni-corrected for multiple comparisons and assessed at $p < 0.05$.

Both monkeys were trained to orient to visual targets for liquid rewards before undergoing surgical procedures to implant a head-restraint post (Crist Instruments) and receive a craniotomy and recording chamber (Crist Instruments) permitting access to PCC. All surgeries were done in accordance with Duke University Institutional Animal Care and Use Committee approved protocols. The animals were on isoflurane during surgery, received analgesics and prophylactic antibiotics after the surgery, and were permitted a month to heal before any recordings were performed. After recovery, both animals were trained on the trapliner task, followed by recordings from BA 23/31 in PCC. MR images were used to locate the relevant anatomic areas and place electrodes. Acute recordings were performed over many sessions. Approximately one-fifth of the recordings were done using FHC single-contact electrodes and four-fifths performed using Plexon 8-contact axial array U probes in Monkey L. No statistically significant differences in the proportion of task-relevant cells were detected between the populations recorded with the two types of electrodes ($\chi^2$, $p > 0.5$). All recordings in Monkey R were done using the U probes. Recordings were performed using Plexon neural recording systems. All single-contact units were sorted online and then re-sorted offline with Plexon offline sorter. All axial units were sorted offline with Plexon offline sorter.

Neural responses often show nonlinearities (Dayan and Abbott, 2001), which can be captured using a GLM (Aljadeff et al., 2016). We used a GLM with a log-linear link function and Poisson-distributed noise estimated from the data to analyze our neuronal recordings, effectively modeling neuronal responses as an exponential function of a linear combination of the input variables. We analyzed the neural data in two epochs: a 500 ms anticipation epoch, encompassing a 250 ms presaccade period and the 250 ms hold fixation period to register a choice, as well as the 250 ms presaccade epoch itself. Covariates included CN in the trial, expected information, expected reward, information outcome from the last choice, reward outcome from the last choice, and all two-way interactions.

In addition to this GLM, we confirmed our model fits in two ways for each neuron: (1) we plotted the residuals against the covariates, to check for higher-order structure; and (2) we used elastic net regression, to check that our significant covariates were selected by the best-fit elastic net model (Zou and Hastie, 2005). Plotting residuals revealed no significant higher-order structure. Furthermore, elastic net regression confirmed our original GLM results. None of the significant covariates identified by the original GLM received a coefficient of 0 from the elastic net regression, and the sizes of the significant coefficients identified by the original GLM were very close to the sizes of the coefficients computed by the elastic net regression.

Perievent time histograms (PETHs) were created by binning spikes time-locked to the event of interest. For the anticipation epoch, PETHs

were centered on the end of the choice saccade and spikes binned in 10 ms bins. PETHs were smoothed with a Gaussian kernel with 0 mean and 5$\sigma$ width where $\sigma = 20$ ms (i.e., two samples).

To analyze encoding of the information or reward boundary, a log-linear GLM regression was run on vectors of binned spike counts time-locked to the start of the trial, with time in window, time of last informative feedback (a binary covariate encoding whether or not the current time bin was before or after the last informative feedback), and their interaction as covariates. Neuronal spikes were sorted into 50 ms bins starting with trial onset and ending with the time of the last outcome in a trial across the duration of the trial. This activity was regressed against time in trial (coded by the bin number, starting with 1 and ending with the number of 50 ms bins for the trial), whether or not the last informative outcome had been received (coded as a 0, for before, or a 1, for after receiving the last outcome), and their two-way interaction. For plots depicting the boundary, PETHs were time-locked to the time of last informative feedback, spikes from 2 s before to 2 s after sorted into 50 ms time bins, and smoothed with a Gaussian kernel with 0 mean 5$\sigma$ width where $\sigma = 50$ ms (i.e., one sample).

The failure to find representations of expected reward reported in Results was confirmed by holding fixed expected information and CN and directly comparing observed firing rates for those combinations for which there was more than one reward level. For CN2, this resulted in one pair of expected rewards; for CN3, one pair; for CN4, one pair; for CN5, one triple; and for CN6, one triple. The observed firing rates for the pairs were compared using Student's $t$ test and for the triples using ANOVA. A neuron that showed a significant difference in those comparisons was included in the count for that CN and so could appear as significant for more than one choice (see Fig. 2C).

Step sizes, step size probabilities, and choicewise behavioral entropies were linearly regressed against the firing rates during the anticipation epoch, when actions were made. To assess whether neurons showed differences in tonic firing rates for high compared with low behavioral entropies, we fit Gaussians with constant offsets to the mean PETH firing rate and examined the confidence interval for the constant offsets for each. The constant offset for high and low BE were considered significantly different if the 95% CIs derived from those fits did not overlap. To assess choicewise entropy encoding before and after receipt of the last bit of information, we used a GLM with log-linear link function and Poisson-distributed noise to calculate the number of neurons that significantly encoded choicewise BE before the receipt of this information to compare with the number after. Covariates included BE, CN in trial, a binary variable with 0 = before boundary and 1 = after boundary, and all two-way interactions. For the population response, we first separated trials by mean choicewise BE across all choices. Next, the normalized average population response for high average choicewise entropy trials was compared with low average entropy during the 2 s before the receipt of the last information using Student's $t$ test. Then we ran the same analysis on the normalized average response during the 2 s following receipt of this information. We report the results of these two analyses below.

## Results
### Trapline foraging in a simulated environment
To explore the effects of information on deviation from routines, 2 monkeys (*Macaca mulatta*) solved a simple traveling salesman problem. In this trapliner task, monkeys visually foraged through a set of six targets arranged in a circle, only moving on to the next trial after sampling every target (Fig. 1B). On each trial, two of the targets were baited, one with a large reward and one with a small reward, with the identity of the baited targets varying from trial to trial. While foraging, monkeys gathered both rewards, herein defined by the amount of juice obtained, and information, herein defined as the reduction in uncertainty about the location of remaining rewards.

By varying which target was rewarded from trial to trial, reward and information were partially decorrelated. Reward was manipulated by varying the size of received rewards, with one

small, one large, and four zero rewards available on every trial. Information was manipulated by varying the spatiotemporal pattern of rewarding targets. Different patterns correspond to different series of received rewards. Based on the series of rewards received up to a particular choice in the trial, some subset of the set of possible sequences remained, and the size of this subset determines the remaining uncertainty for the current trial (see Materials and Methods). Over the course of a trial, the set of possible patterns shrinks, reducing uncertainty about the current trial's pattern and determining the information gathered about the environment. These differences in reward and information outcomes in turn determine reward and information expectations. The expected reward for each target is the total remaining reward to harvest divided by the number of remaining targets. In contrast, the expected information is the mean amount of information to be gained by making the next choice. As the animal proceeds through the trial, the amount of expected information varies as a function of how many possible patterns of rewards have been eliminated so far. Distinct possible reward outcomes may offer the same information, and so our task partially decorrelates information and reward (linear regression on expected reward and expected information, $R^2 = 0.13$).

Information may influence the pattern of choices that monkeys made, resulting in trial-to-trial changes in this pattern (behavioral data are the same as first reported in Barack et al., 2017). On a majority of trials, monkeys chose targets in the same order (the daily dominant pattern [DDP]; Monkey R: same DDP across all 14 sessions; Monkey L: same DDP across 24 of 30 sessions; across all sessions, $0.4665 \pm 0.0317$ proportion of trials diverged from the DDP; see Materials and Methods). More generally, monkeys usually chose the targets in a circle (proportion of trials in average session with circular patterns of choices: Monkey L: $0.6134 \pm 0.0418$; Monkey R: $0.7113 \pm 0.0208$). However, they occasionally deviated from their circular routine. This variability can be measured by finding the BE over the distribution of choice probabilities for targets. First, each choice during a trial was egocentrically coded by its step size, the number of targets clockwise or counterclockwise from the current trial's previously chosen target (Fig. 1B). The probability of a particular step size was computed by counting the number of trials with that step size and dividing by the total number of trials (see Materials and Methods). BE, the entropy computed over that distribution, significantly predicts adherence to both typical traplines (DDP: logistic regression; significant [$p < 0.05$] $\beta$ for 22 of 44 sessions) and circular traplines (logistic regression; significant $\beta$ [$p < 0.05$] for 33 of 44 sessions). We found that the informativeness of outcomes influenced the variability in the monkeys' patterns of choices as measured by BE. Anticipation of more informative choice outcomes significantly reduced the entropy of the monkeys' choices on average (Student's $t$ test across all choices and sessions comparing BE for less than average expected information to greater than average; both monkeys: $t_{(96,718)} = -19.25$, $p < 1 \times 10^{-81}$; Monkey L: $t_{(69,274)} = -3.24$, $p < 0.005$; Monkey R: $t_{(27,442)} = -23.99$, $p < 1 \times 10^{-125}$). To better assess the influence of expected information on behavioral variability, we plotted by session and CN the mean BE for zero expected information and compared it with the mean BE for non-zero expected information. Median BE across sessions was greater for CN4 and CN5 than CN3 for no expected information ($p < 0.05$; Fig. 1C, green boxes and points) and was greater for no expected information compared with some expected information for CN5 ($p < 0.05$; Fig. 1C, CN5, red boxes and points compared with green).
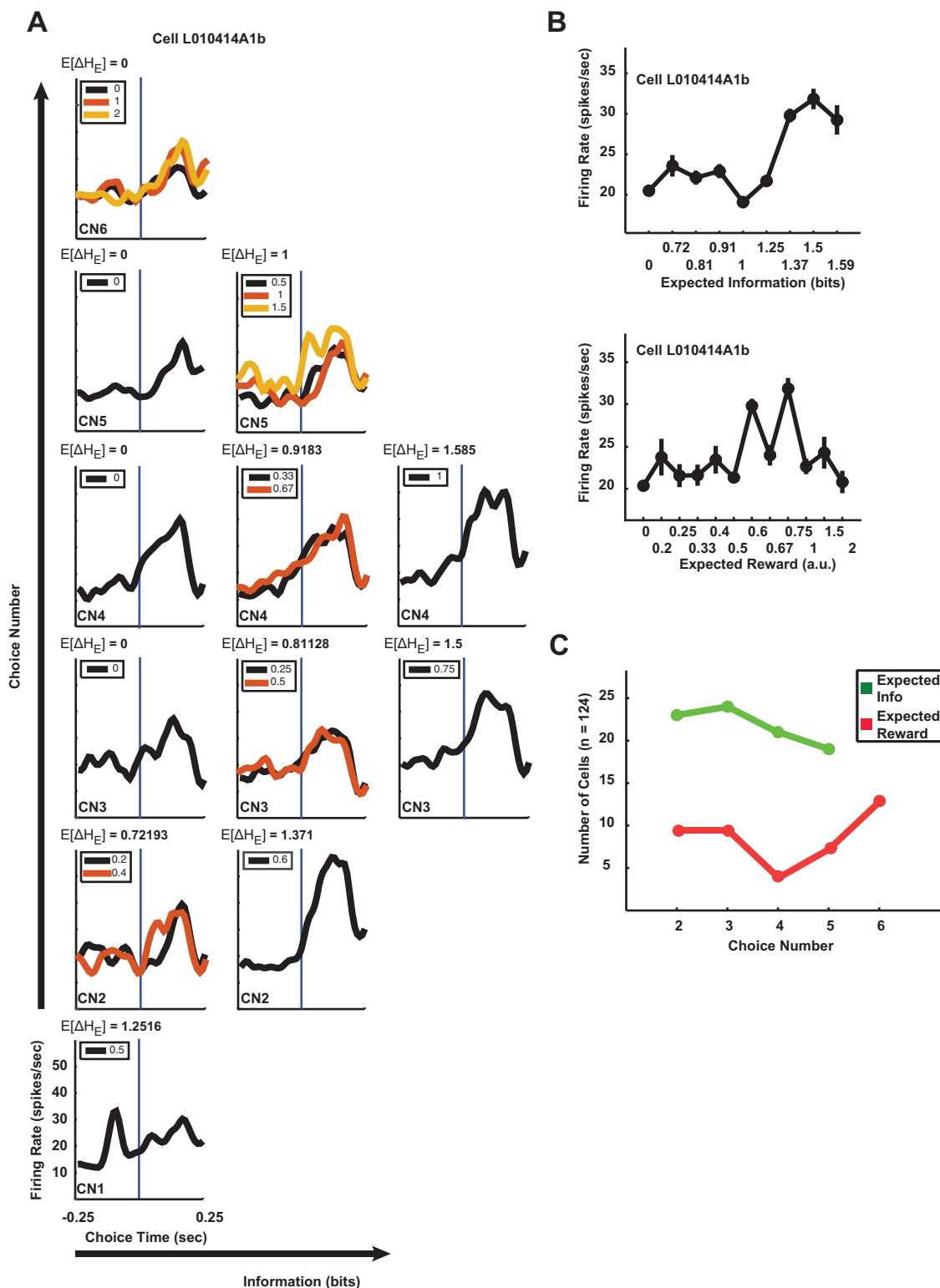
**Figure 2.** PCC neurons preferentially encode environmental information over reward. **A**, Firing rate of sample neuron encoding expected information but not expected reward across all CN1-CN6, plotted separately by expected information ($E[\Delta H_E]$). Legends indicate expected reward(s) for each plot. Blue line indicates end of saccade. **B**, Tuning curves for expected information (top) and expected reward (bottom), collapsed across CNs for better visibility, for the cell plotted in **A**. This example cell showed elevated firing rates for higher amounts of information. The elevated firing rates for certain amounts of expected reward correspond to choices with high expected information with only a single level of expected reward. **C**, Number of cells encoding expected reward by choice (red) for constant expected information, and number of cells encoding expected information by choice (green). Neurons were included in the expected reward counts if Student $t$ tests (CN2-CN4) or ANOVA (CN5-CN6) indicated a significant difference in firing rates ($p < 0.05$, uncorrected for multiple comparisons to allow the weakest criteria for inclusion, and the same cell could appear for more than one CN).

The presence of information or reward left to collect on a trial also drove choice variability. While still harvesting information and reward about the current trial, monkeys' choices were less variable, but afterward they became more variable in their choices (Student's $t$ test on CN4 or CN5; both monkeys: $t_{(48,358)} = -125.98$, $p \sim 0$; Monkey L: $t_{(34,636)} = -96.32$, $p \sim 0$; Monkey R: $t_{(13,720)} = -71.79$, $p \sim 0$; results also significant for each CN separately; Fig. 1C, right). Hence, monkeys deviated less while choices were still informative or rewarding and more thereafter.

### Environmental information signaling by posterior cingulate neurons

We next probed PCC activity during the trapliner task to examine information and reward signaling from 124 cells in 2 monkeys (Fig. 1A; Monkey L = 84 neurons; Monkey R = 40 neurons; neural data are the same as first reported in Barack et al., 2017). In order to control for previously uncovered neural effects, all choices where monkeys diverged from typical traplines were excluded from the analyses in this section (those neural findings are reported in Barack et al., 2017).

During the anticipation epoch (500 ms encompassing a 250 ms prechoice period and a 250 ms hold fixation period), neurons in PCC preferentially signaled information expectations over reward expectations. An example cell (Fig. 2A) showed a phasic increase in firing rate during the anticipation epoch when expected information was higher for the same CN in the trial (for example, CN2: Student's $t$ test, $p < 0.0001$, $t_{(283)} = -4.3056$; firing rate for 0.72 bits = 22.51 ± 1.46 spikes/s, firing rate for 1.37 bits = 29.84 ± 0.95 spikes/s). However, after controlling for CN in the trial and expected information, the same neuron did not differentiate between different amounts of expected reward (Student's $t$ test, $p > 0.9$; firing rate for 0.2 expected reward = 22.23 ± 2.35 spikes/s, firing rate for 0.4 expected reward = 22.76 ± 1.83 spikes/s; Fig. 2A, second row from bottom, left). The tuning curves for this same cell collapsed across all CNs for both expected information and expected reward illustrate the strong sensitivity to larger amounts of information (Fig. 2B).

In our population of 124 neurons, significantly more cells were tuned to information than reward when controlling for CN in trial. A GLM regression revealed that, during the anticipation epoch, 35 (28%) of 124 neurons (Monkey L: 25 [30%] of 84 neurons; Monkey R: 10 [25%] of 40 neurons) signaled the interaction of CN and expected information, but only 1 (~1%) of 124 neurons (Monkey L: 1 [~1%] of 84 neurons; Monkey R: 0 [0%] of 40 neurons) signaled the interaction of CN and expected reward (all results, $p < 0.05$, Bonferroni-corrected; for full list of covariates in the GLM, see Materials and Methods). A further test for signaling of expected reward compares the average firing rates for different amounts of expected reward for the same CN and expected information. This test revealed that only ~10% of neurons signaled expected reward, except on the last choice when all information had been received (Fig. 2C). In contrast, ~20% of neurons signaled expected information (Fig. 2C). These proportions were not significantly different when all circular traplines were included (expected information × CN, $\chi^2 > 0.24$; expected reward × CN, $\chi^2 > 0.17$).

### PCC neurons index response variability

We have previously established that PCC neurons signal decisions to diverge from typical traplines during our task (Barack et al., 2017). However, the extent to which these cells track variability of responses during the task remains to be explored. We examined whether PCC neurons index the degree of behavioral

variability, operationalized as BE (see Materials and Methods; all trials, including divergences from typical traplines, are included in the following analyses). During the presaccade epoch, BE varied significantly with firing rate for 48 (39%) of 124 neurons (linear regression of BE against firing rate, $p < 0.05$; Monkey L: 37 [44%] of 84 neurons, Monkey R: 11 [28%] of 40 neurons). An example cell was more active for high entropy choices compared with low (linear regression, $\beta_{BE} = 0.0229 \pm 0.0026$ bits$_{BE}$/spike, $p < 5 \times 10^{-18}$; Fig. 3A). Across the population, higher firing rates predicted greater BE (124 neurons; Student's $t$ test on mean normalized firing rates during presaccade epoch, $t_{(123)} = 2.7363$, $p < 0.01$; $\beta_{BE} > 0$ in 80 cells, $\beta_{BE} \leq 0$ in 44 cells; mean $\beta_{BE} = 0.0025 \pm 0.0011$ bits$_{BE}$/spike, Student's $t$ test against $h_0$: mean $\beta_{BE} = 0$, $t_{(123)} = 2.3268$, $p < 0.05$; Fig. 3B). In addition, in our population of 124 cells, 46 (37%) exhibited significantly different ($p < 0.05$) tonic firing rates for high BE compared with low BE choices during the anticipation epoch (Monkey L: 35 of 84 [42%]; Monkey R: 11 of 40 [28%]).

We next investigated whether PCC neurons signaled the boundary defined by the receipt of the last information or reward, when the pattern of rewards on a given trial becomes fully resolved. This can occur before the last reward is delivered if the last reward is received on the last choice in a trial. A regression of each trial's binned spike counts against the time in the trial and the time of last informative outcome revealed that 84 (68%) of 124 neurons differentiated these two states (GLM, effect of interaction, $p < 0.05$; see Materials and Methods; Monkey L: 61 of 84 neurons, 73%; Monkey R: 23 of 40 neurons, 58%). During a 4 s epoch centered on the time of the last informative choice outcome, an example cell fired less before that outcome than after (Student's $t$ test, $p < 1 \times 10^{-56}$; Fig. 3C). The population of cells also fire more after this boundary (Student's $t$ test, $p < 0.005$; Fig. 3D).

Finally, BE signals and boundary signals were combined in the PCC population. While the time of last information can be partly disambiguated from time of last reward, this occurs only on the last choice when a single target remains. Since BE is a measure of response variability, it requires more than one target, which is not available on the last choice. As a result, combined signals of BE and the boundary could reflect the end of either information gathering or reward harvesting. Significantly fewer cells ($\chi^2$, $p < 1 \times 10^{-10}$) predicted BE after receiving all information or reward (24 [19%] of 124 neurons) than before (74 [60%] neurons). PCC population responses on choices with high BE compared with low entropy revealed significant differences before receipt of the last informative or rewarding outcome (Student's $t$ test, $p < 1 \times 10^{-4}$) but not after (Student's $t$ test, $p > 0.5$), with greater modulation for high entropy compared with low.

## Discussion

In this study, we show that environmental information influences responses during routine behavior and that firing rates of PCC neurons carry this information and predict behavioral variability. Despite the fact that, in our task, monkeys could not use environmental information to increase their chance of reward, the receipt of environmental information and the exhaustion of uncertainty impacted behavioral routines. Monkeys' responses were less variable when there was more information to be gathered, but became more variable once the environment became fully known. This pattern of variable responses after resolving all environmental uncertainty departs from the reward rate
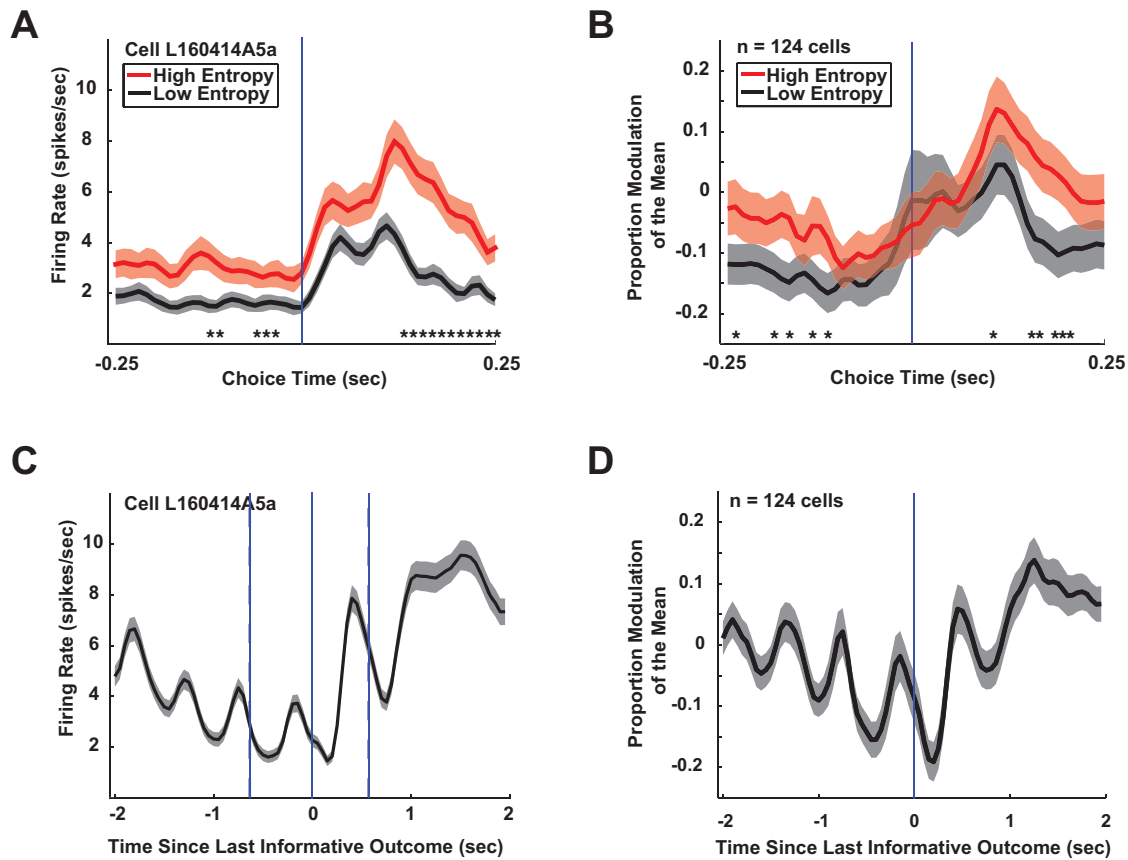
**Figure 3.** Neurons in PCC forecast deviations in behavior. *A*, Sample neuron encoding BE during the anticipation epoch. This cell was more active for high entropy choices than for low entropy choices. *B*, Population encoding of BE. The population was more active for high entropy choices than low. *C*, Sample cell encoding the end of information gathering. This cell had higher firing rates after the last informative or rewarding outcome compared with before. Thin blue lines with dashed lines very close on either side indicate average time of choice before (left) or after (right) last informative or rewarding choice ± 1 SEM. *D*, The population also encoded this boundary, with higher firing rates after the last informative or rewarding outcome compared with before. *B*, *D*, n = 124 cells. *A*, *B*, Blue line indicates end of saccade. *C*, *D*, Middle blue line indicates time of outcome. *A-D*, Shading represents ± 1 SEM. *$p < 0.05$ for that 10 ms time bin.

maximizing strategy of selecting targets in a circle to minimize saccade lengths. While monkeys traplined, neurons in PCC robustly signaled information expectations, but not reward expectations, and predicted the variability in the patterns of choices. Finally, PCC neurons differentiate the degree of behavioral variability before all information or reward was received about the pattern of rewards compared with after, with an increase in activity following receipt of the last informative outcome and concomitant decreases in forecasting behavioral variability. In sum, our experimental findings suggest that PCC tracks the state of the environment to influence routine behavior.

Monkeys often chose targets in the same pattern, consistent with previous findings of repetitive stereotyped foraging in wild primate groups (Noser and Byrne, 2007). They also generally moved in a circle, visiting the next nearest neighbor after the current target, likewise consistent with previous findings in groups of wild foraging primates (E. W. Menzel, 1973; Garber, 1988; Janson, 1998). These foraging choices almost always result in straight line routes (Janson, 1998; Pochron, 2001; Cunningham and Janson, 2007; Valero and Byrne, 2007) or a series of straight lines (Di Fiore and Suarez, 2007; Noser and Byrne, 2007). Experiments on captive primates have also observed nearest neighbor or near optimal path finding (E. W. Menzel, 1973; MacDonald and Wilkie, 1990; Gallistel and Cramer, 1996; Cramer and Gallistel, 1997). Our monkeys' choices are also consistent with human behavior on traveling salesman problems,

wherein next nearest neighbor paths are usually chosen for low numbers of points (Hirtle and Gärling, 1992; MacGregor and Ormerod, 1996; MacGregor and Chu, 2011).

The PCC, a posterior midline cortical region with extensive corticocortical connectivity (Heilbronner and Haber, 2014) and elevated resting state and off-task metabolic activity (Buckner et al., 2008), is at the heart of the default mode network (Buckner et al., 2008). The default mode network is a cortex-spanning network implicated in exploratory cognition, including imagination (Schacter et al., 2012), creativity (Kühn et al., 2014), and narration (Wise and Braga, 2014). Although implicated in a range of cognitive functions, activity in PCC may be unified by a set of computations related to harvesting information from the environment to regulate behavior. Signals in PCC that carry information about environmental decision variables, such as value (McCoy et al., 2003), risk (McCoy and Platt, 2005), and decision salience (Heilbronner et al., 2011), may indeed reflect the tracking of information returns from the immediate environment. For example, in a two-alternative forced-choice task, neurons in PCC preferentially signaled the resolution of a risky choice with a variable reward over the value of choosing a safe choice with a guaranteed reward (McCoy and Platt, 2005). Such signals may reflect the information associated with the resolution of uncertainty regarding the risky option. PCC neurons also signal reward-based exploration (Pearson et al., 2009), and microstimulation in PCC can shift monkeys from a preferred option to one

they rarely choose (Hayden et al., 2008). Both of these functions may reflect signaling of environmental information as well; for example, the signaling of exploratory choices may reflect the information from an increase in the number of recent sources of reward (Pearson et al., 2009). Evidence from neuroimaging studies in humans similarly reveals PCC activation in a wide range of cognitive processes related to adaptive cognition, including imagination (Benoit et al., 2011), decision-making (Kable and Glimcher, 2007), and creativity (Beaty et al., 2015).

Uncovering the neural circuits that underlie variability in foraging behavior may provide insight into more complex cognitive functions. A fundamental feature of what we call prospective cognition, thoughts about times, places, and objects beyond the here and now, involves consideration of different ways the world might turn out. Various types of prospective cognition, including imagination, exploration, and creativity, impose a trade-off between engaging well-rehearsed routines and deviating in search of new, potentially better solutions (Gottlieb et al., 2013; Andrews-Hanna et al., 2014; Beaty et al., 2015). For example, creativity involves diverging from usual patterns of thought, such as occurs in generating ideas (Benedek et al., 2014) or crafting novel concepts (Barron, 1955; Guilford, 1959). During creative episodes, the PCC shows increased activity during idea generation (Benedek et al., 2014) and higher connectivity with control networks during idea evaluation (Beaty et al., 2015), perhaps reflecting imagined, anticipated, or predicted variation in the environment. Exploration similarly involves diverging from the familiar, such as to locate novel resources (Ohashi and Thomson, 2005) or discover shorter paths (Sutton and Barto, 1998) between known locations. Such prospective cognition requires diverging from routine thought, and the identification of the neural circuits that mediate deviations from motor routines may provide initial insight into the computations and mechanisms of prospective cognition. The discovery that the PCC preferentially signals the state of the environment and predicts behavioral variability relative to that state is a first step toward understanding these circuits.

The reinforcement learning literature is replete with models where exploration is driven by the search for information (Schmidhuber, 1991; Johnson et al., 2012). These models hypothesize that agents should take actions that maximize the information gleaned from the environment, by reducing uncertainty about the size of offered rewards (Schmidhuber, 1991), the location of rewards in the environment (Johnson et al., 2012), or otherwise maximizing information for subsequent decisions. Furthermore, evidence from initial studies studying information-based exploration shows that humans are avid information-seekers (Miller, 1983; Fu and Pirolli, 2007) and regulate attentional and valuational computations on the basis of information (Manohar and Husain, 2013; Blanchard et al., 2015). In our task, the PCC represented environmental information and tracked when learning about the environment was complete, two variables central to information-based exploration. In particular, the dramatic change in firing rates associated with the end of information gathering suggests that PCC represents the information state of the environment and possibly also the rate of information intake, a central variable in information foraging models (Pirolli and Card, 1999; Fu and Pirolli, 2007; Pirolli, 2007). PCC appears poised to regulate exploration for information.

In conclusion, harvested information and response variability were both signaled by PCC neurons, suggesting a central role for PCC in how information drives exploration and possibly prospective cognition. Monkeys were sensitive to the amount of uncertainty remaining in the environment, with more reliable patterns of choices while information remained and more variable patterns after environmental uncertainty had been resolved and all rewards collected. PCC neurons preferentially tracked this information and predicted the variability in monkeys' behavior. Our findings implicate the PCC in the regulation of foraging behavior, and specifically the information-driven deviation from routines. When at the races, PCC will both track who won and set the stage for changing up your bets.

## References

Aljadeff J, Lansdell BJ, Fairhall AL, Kleinfeld D (2016) Analysis of neuronal spike trains, deconstructed. Neuron 91:221–259.

Andrews-Hanna JR, Smallwood J, Spreng RN (2014) The default network and self-generated thought: component processes, dynamic control, and clinical relevance. Ann NY Acad Sci 1316:29–52.

Barack DL, Chang SW, Platt ML (2017) Posterior cingulate neurons dynamically signal decisions to disengage during foraging. Neuron 96:339–347. e335.

Barron F (1955) The disposition toward originality. J Abnorm Psychol 51:478–485.

Beaty RE, Benedek M, Kaufman SB, Silvia PJ (2015) Default and executive network coupling supports creative idea production. Sci Rep 5:10964.

Bell WJ (1990) When to return to a resource patch. In: Searching behaviour, pp 117–129. New York: Springer.

Benedek M, Jauk E, Fink A, Koschutnig K, Reishofer G, Ebner F, Neubauer AC (2014) To create or to recall? Neural mechanisms underlying the generation of creative new ideas. Neuroimage 88:125–133.

Benoit RG, Gilbert SJ, Burgess PW (2011) A neural mechanism mediating the impact of episodic prospection on farsighted decisions. J Neurosci 31:6771–6779.

Berger-Tal O, Bar-David S (2015) Recursive movement patterns: review and synthesis across species. Ecosphere 6:art149.

Blanchard TC, Hayden BY, Bromberg-Martin ES (2015) Orbitofrontal cortex uses distinct codes for different choice attributes in decisions motivated by curiosity. Neuron 85:602–614.

Boggs CL, Smiley JT, Gilbert LE (1981) Patterns of pollen exploitation by Heliconius butterflies. Oecologia 48:284–289.

Buckner RL, Andrews-Hanna JR, Schacter DL (2008) The brain's default network: anatomy, function, and relevance to disease. Ann NY Acad Sci 1124:1–38.

Clithero JA, Rangel A (2014) Informatic parcellation of the network involved in the computation of subjective value. Soc Cogn Affect Neurosci 9:1289–1302.

Cramer AE, Gallistel CR (1997) Vervet monkeys as travelling salesmen. Nature 387:464.

Cunningham E, Janson C (2007) Integrating information about location and value of resources by white-faced saki monkeys (*Pithecia pithecia*). Anim Cogn 10:293–304.

Dayan P, Abbott LF (2001) Theoretical neuroscience. Cambridge, MA: Massachusetts Institute of Technology.

Desrochers TM, Jin DZ, Goodman ND, Graybiel AM (2010) Optimal habits can develop spontaneously through sensitivity to local cost. Proc Natl Acad Sci USA 107:20512–20517.

Di Fiore A, Suarez SA (2007) Route-based travel and shared routes in sympatric spider and woolly monkeys: cognitive and evolutionary implications. Anim Cogn 10:317–329.

Freeman R (1968) Charles Darwin on the routes of male bumble bees. Bull Br Museum 3:177–189.

Fu WT, Pirolli P (2007) SNIF-ACT: a cognitive model of user navigation on the World Wide Web. Hum Comput Interaction 22:355–412.

Gallistel CR, Cramer AE (1996) Computations on metric maps in mammals: getting oriented and choosing a multi-destination route. J Exp Biol 199:211–217.

Garber PA (1988) Foraging decisions during nectar feeding by tamarin monkeys (*Saguinus mystax* and *Saguinus fuscicollis*, Callitrichidae, Primates) in Amazonian Peru. Biotropica 20:100–106.

Garrison JS, Gass CL (1999) Response of a traplining hummingbird to changes in nectar availability. Behav Ecol 10:714–725.

Gill FB (1988) Trapline foraging by hermit hummingbirds: competition for an undefended, renewable resource. Ecology 69:1933–1942.

Gottlieb J, Oudeyer PY, Lopes M, Baranes A (2013) Information-seeking, curiosity, and attention: computational and neural mechanisms. Trends Cogn Sci 17:585–593.

Guilford JP (1959) Traits of creativity. Creativity Cultivation 10:141–161.

Hamming RW (1950) Error detecting and error correcting codes. Bell Syst Tech J 29:147–160.

Hayden BY, Platt ML (2010) Neurons in 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.

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, Hayden BY, Platt ML (2011) Decision salience signals in posterior cingulate cortex. Front Neurosci 5:55.

Heilbronner SR, Platt ML (2013) Causal evidence of performance monitoring by neurons in posterior cingulate cortex during learning. Neuron 80:1384–1391.

Hemmi JM, Menzel CR (1995) Foraging strategies of long-tailed macaques, *Macaca fascicularis*: directional extrapolation. Anim Behav 49:457–464.

Hirtle SC, Gärling T (1992) Heuristic rules for sequential spatial decisions. Geoforum 23:227–238.

Hui SK, Fader PS, Bradlow ET (2009) Research note. The traveling salesman goes shopping: the systematic deviations of grocery paths from TSP optimality. Marketing Sci 28:566–572.

Janmaat KR, Byrne RW, Zuberbühler K (2006) Evidence for a spatial memory of fruiting states of rainforest trees in wild mangabeys. Anim Behav 72:797–807.

Janson CH (1998) Experimental evidence for spatial memory in foraging wild capuchin monkeys, *Cebus apella*. Anim Behav 55:1229–1243.

Janzen D (1971) Euglossine bees as long-distance pollinators of tropical plants. Science 171:203–205.

Johnson A, Varberg Z, Benhardus J, Maahs A, Schrater P (2012) The hippocampus and exploration: dynamically evolving behavior and neural representations. Front Hum Neurosci 6:216.

Kable JW, Glimcher PW (2007) The neural correlates of subjective value during intertemporal choice. Nat Neurosci 10:1625–1633.

Kühn S, Ritter SM, Müller BC, Baaren RB, Brass M, Dijksterhuis A (2014) The importance of the default mode network in creativity: a structural MRI study. J Creat Behav 48:152–163.

Leech R, Kamourieh S, Beckmann CF, Sharp DJ (2011) Fractionating the default mode network: distinct contributions of the ventral and dorsal posterior cingulate cortex to cognitive control. J Neurosci 31:3217–3224.

MacDonald SE, Wilkie DM (1990) Yellow-nosed monkeys' (*Cercopithecus ascanius whitesidei*) spatial memory in a simulated foraging environment. J Comp Psychol 104:382–387.

MacGregor JN, Chu Y (2011) Human performance on the traveling salesman and related problems: a review. J Problem Solving 3:2.

MacGregor JN, Ormerod T (1996) Human performance on the traveling salesman problem. Percept Psychophys 58:527–539.

Manning A (1956) Some aspects of the foraging behaviour of bumble-bees. Behavior 9:164–200.

Manohar SG, Husain M (2013) Attention as foraging for information and value. Front Hum Neurosci 7:711.

McCoy AN, Crowley JC, Haghighian G, Dean HL, Platt ML (2003) Saccade reward signals in posterior cingulate cortex. Neuron 40:1031–1040.

McCoy AN, Platt ML (2005) Risk-sensitive neurons in macaque posterior cingulate cortex. Nat Neurosci 8:1220–1227.

McNamara J (1982) Optimal patch use in a stochastic environment. Theor Popul Biol 21:269–288.

Menzel CR (1991) Cognitive aspects of foraging in Japanese monkeys. Anim Behav 41:397–402.

Menzel CR (1996) Structure-guided foraging in long-tailed macaques. Am J Primatol 38:117–132.

Menzel EW (1973) Chimpanzee spatial memory organization. Science 182:943–945.

Miller G (1983) Informavores. The study of information: interdisciplinary messages (Machlup F, MansfieldU, eds), pp 111–113. New York: Wiley-Interscience.

Noser R, Byrne RW (2007) Travel routes and planning of visits to out-of-sight resources in wild chacma baboons, *Papio ursinus*. Anim Behav 73:257–266.

Noser R, Byrne RW (2010) How do wild baboons (*Papio ursinus*) plan their routes? Travel among multiple high-quality food sources with intergroup competition. Anim Cogn 13:145–155.

Ohashi K, Thomson JD (2005) Efficient harvesting of renewing resources. Behav Ecol 16:592–605.

Ohashi K, Leslie A, Thomson JD (2008) Trapline foraging by bumble bees: V. Effects of experience and priority on competitive performance. Behav Ecol 19:936–948.

Ohashi K, Leslie A, Thomson JD (2013) Trapline foraging by bumble bees: VII. Adjustments for foraging success following competitor removal. Behav Ecol 24:768–778.

Pearson JM, Hayden BY, Raghavachari S, Platt ML (2009) Neurons in posterior cingulate cortex signal exploratory decisions in a dynamic multioption choice task. Curr Biol 19:1532–1537.

Pirolli P, Card S (1999) Information foraging. Psychol Rev 106:643–675.

Pirolli PL (2007) Information foraging theory: adaptive interaction with information. Oxford: Oxford UP.

Pochron ST (2001) Can concurrent speed and directness of travel indicate purposeful encounter in the yellow baboons (*Papio hamadryas cynocephalus*) of Ruaha National Park, Tanzania? Int J Primatol 22:773–785.

Possingham HP (1989) The distribution and abundance of resources encountered by a forager. Am Naturalist 133:42–60.

Racey P, Swift S (1985) Feeding ecology of *Pipistrellus pipistrellus* (Chiroptera: Vespertilionidae) during pregnancy and lactation: I. Foraging behaviour. J Anim Ecol 54:205–215.

Schacter DL, Addis DR, Hassabis D, Martin VC, Spreng RN, Szpunar KK (2012) The future of memory: remembering, imagining, and the brain. Neuron 76:677–694.

Schmidhuber J (1991) Curious model-building control systems. IEEE International Joint Conference on Neural Networks.

Stephens DW, Krebs JR (1986) Foraging theory. Princeton, NJ: Princeton UP.

Sutton RS, Barto AG (1998) Reinforcement learning: an introduction. Cambridge, MA: Massachusetts Institute of Technology.

Valero A, Byrne RW (2007) Spider monkey ranging patterns in Mexican subtropical forest: do travel routes reflect planning? Anim Cogn 10:305–315.

Wan X, Cheng K, Tanaka K (2015) Neural encoding of opposing strategy values in anterior and posterior cingulate cortex. Nat Neurosci 18:752–759.

Wise RJ, Braga RM (2014) Default mode network: the seat of literary creativity? Trends Cogn Sci 18:116–117.

Zou H, Hastie T (2005) Regularization and variable selection via the elastic net. J R Stat Soc B 67:301–320.