

# **Title: Neuronal activity in the posterior cingulate cortex signals environmental information and predicts behavioral variability during routine foraging**

**Authors:** David L Barack<sup>1,2,3\*</sup> and Michael L Platt<sup>3,4,5</sup>

<sup>1</sup>Departments of Philosophy, Neuroscience, and Economics, Columbia University, New York, NY 10027, USA.

<sup>2</sup>Department of Philosophy, Duke University, Durham, NC 27701, USA.

<sup>3</sup>Duke Institute for Brain Sciences, Duke University, Durham, NC, 27701, USA.

<sup>4</sup>Departments of Neurobiology and Psychology and Neuroscience, Duke University, Durham, NC 27701, USA.

<sup>5</sup>Departments of Neuroscience, Psychology, and Marketing, University of Pennsylvania, Philadelphia, PA 19104, USA.

\* corresponding author: dbarack@gmail.com

## **Abstract**

Animals engage in routine behavior in order 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 diverge from routines remains unexplored. To investigate the representation of environmental information during routine foraging, we recorded the activity of single neurons in posterior cingulate cortex (PCC) in 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 diverged 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 mediates the influence of information on variability in choice behavior.

## Introduction

Imagine you are at a horse race, and there are six horses, with Local Field Potential the underdog, facing 100:1 odds against. When LFP wins, a one dollar bet will pay out \$100. But in addition to the reward received from this bet, learning that out of the six 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 must learn not only the sizes of rewards but also their locations, times, 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 upon receiving information that a highly valued resource has become available (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 inter-reward time is too great (McNamara 1982), such 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 a routine foraging path, a strategy known as traplining (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, Leslie, and Thomson 2008), and helping adapt to changes in competition (Ohashi, Leslie, and Thomson 2013). Many animals trapline, including bats (Racey and Swift 1985), bees (Manning 1956; Janzen 1971), butterflies (Boggs, Smiley, and Gilbert 1981), hummingbirds (Gill 1988), and an array of primates including rhesus macaques (Menzel 1973), baboons (Noser and Byrne 2010), vervet monkeys (Cramer and Gallistel 1997), and humans (Hui, Fader, and Bradlow 2009). Wild primates foraging for fruit (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, Fader, and Bradlow 2009) environments all use traplining to minimize total distance traveled and thereby maximize resource intake rates.

Though many primates trapline, information about the state of the environment, such as weather (Janmaat, Byrne, and Zuberbühler 2006), the availability of new foods (Menzel 1991), or possible feeding locations (Hemmi and Menzel 1995; Menzel 1996), can influence choices made while foraging. Such detours result in longer search distances and more variable choices (Noser

and Byrne 2010; Hui, Fader, and Bradlow 2009) but allow animals to identify new resources(Menzel 1991) and engage in novel behaviors(Noser and Byrne 2010). These benefits are consistent with computer simulations that show stochastic traplining yields better long-term returns than pure traplining by uncovering new resources or more efficient routes(Ohashi and Thomson 2005). Thus environmental information may improve foraging efficiency during routine foraging.

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, Gilbert, and Burgess 2011), value representation(Kable and Glimcher 2007; Clithero and Rangel 2014), strategy setting(Wan, Cheng, and Tanaka 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 can induce shifts between options(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 separate out the hedonic value from the informational value of rewards in PCC.

Here we tested the hypothesis that PCC preferentially tracks reward information by recording 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. As they acquired more information about the state of the environment, their trapline foraging behavior was faster and less variable. While foraging, PCC neurons tracked environmental information but not reward and forecast response speed and variability in traplines. 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 task manipulates both reward and information. Reward was manipulated by varying the size of received rewards: on every trial, one of six targets had a large reward, one had a small reward that was half of the size of the large, and the remaining four had zero rewards. Information was manipulated by varying the spatiotemporal pattern of rewarding targets. Given a set of four null rewards, one small, and one large, there are 6! distinct permutations. We made the simplifying assumption that monkeys perceived the pattern of received rewards without distinguishing the different null rewards received. 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  $i$  in the sequence and the total number of possible sequences:

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

where  $|\cdot|$  denotes cardinality,  $P$  is the set of possible permutations, and  $\{P_i\}$  is the set of remaining permutations after the  $i^{\text{th}}$  choice. The amount of information contained in some reward outcome is computed as the difference in the entropy, what has been learned about the current trial's pattern of received reward by receiving the most recent outcome:

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

for the amount of environmental entropy  $H_E$  on the  $i^{\text{th}}$  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 received rewards thus far:

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

for expected information  $E[\Delta H_E]$  for the  $i^{\text{th}}$  choice, possible outcomes  $[\Delta H_E]_i$  for the remaining permutations  $P_{\text{remaining}}$ , and where  $|\cdot|$  again denotes cardinality. Thus, 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 thus 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  $n$ :

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

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.

Given the design of the task, some patterns of returns can have the same amount of expected information for the same choice number but distinct expected rewards (see supplementary information). This partial decorrelation of reward and information occurs as a result of the way expected reward and expected information are computed. Expected information is a function of the number of possible patterns to be excluded by the next outcome, whereas expected reward is a function of the identity of the pattern, and in particular, the amount of reward remaining to be harvested. The same amount of information can be expected from obtaining distinct outcomes because the number of patterns excluded by the different outcomes is identical. But the expected rewards can differ because there are distinct outcomes possible. Hence, our task partially deconfounds information and reward.

## Behavioral Methods

For our behavioral entropy measures, we again used the standard definition of entropy. For behavioral entropy, 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 to 2, 2 to 3, 3 to 4, and 4 to 5 (5 to 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 choice number) 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 (choice number, information outcome, information expectation, reward outcome, and reward expectation), we computed the choicewise behavioral entropy ( $H_B$ ) for that combination as

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

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

Diverge choices were defined as choices that diverged from the daily dominant pattern (DDP). Determining the DDP 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,

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

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  $\bar{D}_{i,i'}$ . The daily dominant pattern corresponded to the pattern with the minimum  $\bar{D}_{i,i'}$ .

Response times, defined as the time from end of saccade for the last decision to end of saccade for the current decision, and behavioral entropy were both regressed against a number of variables and their interactions as covariates with multilinear regression. Covariates included choice number in trial, expected information, expected reward, reward outcome from the previous choice, information outcome from the previous choice, and all 2-way interactions. See Table S2



for full results of the response time regression and Table S6 for full results of the behavioral entropy regressions.

### Neural Methods

Two male rhesus macaque 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 IACUC approved protocols. The animals were on isoflourane 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 traplining task, followed by recordings from BA 23/31 in PCC. MR images were used to locate the relevant anatomical areas and place electrodes. Acute recordings were performed over many months. Approximately one fifth of the recordings were done using FHC (FHC, Inc., Bangor, ME) single contact electrodes and four fifths performed using Plexon (Plexon, Inc., Dallas, TX) 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 non-linearities(Dayan and Abbott 2001), which can be captured using a generalized linear model(Aljadeff et al. 2016). We used a generalized linear model (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 pre-saccade period and the 250 ms hold fixation period to register a choice, and more focally the 250 ms pre-saccade epoch itself. Covariates included choice number in the trial, expected information, expected reward, information outcome from the last choice, reward outcome from the last choice, and all 2-way interactions.

In addition to this GLM, we confirmed our model fits in two ways for each neuron: first, we plotted the residuals against the covariates, to check for higher-order structure, and second, 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 (see supplemental methods). 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 (see supplemental results).

Perievent time histograms (PETHs) were created by binning spikes in 10 ms bins, time-locked to the event of interest. For the anticipation epoch, PETHs were centered on the end of the choice saccade. For the last informative outcome analysis, PETHs were time-locked to the time of last informative feedback, from two seconds before to two seconds after in 50 ms time bins. PETHs were smoothed with a Gaussian and a 20 ms kernel. To analyze encoding of the last informative feedback, a log-linear GLM regression was then 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.

Step sizes, step size probabilities, and choicewise entropies were linearly regressed against the firing rates during the anticipation epoch, when actions were executed, as reflected in the step size. To assess choicewise entropy encoding before and after receipt of the last bit of information, we calculated the number of neurons that significantly encoded choicewise entropy by choice before the receipt of this information to the number after. For the population response, we first separated trials by mean choicewise entropy across all choices. Next, we compared the normalized average population response for high average choicewise entropy trials to low average entropy during the two seconds before the receipt of the last information. Then we ran the same analysis on the normalized average response during the two seconds following receipt of this information, and reported the results of these two analyses below.

## Results

### Trapline Foraging in a Simulated Environment

To explore the effects of information on divergence from routines, two monkeys (*M. mulatta*) solved a simple traveling salesman problem. In this task, monkeys visually foraged through a set of six targets arranged in a circular array, 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 the identity of the rewarding targets from trial-to-trial, reward and information were partially decorrelated. Reward was manipulated by varying the size of received rewards, with a small, 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, determining the remaining uncertainty for the current trial (see methods). Monkeys gathered rewards and thereby reduced uncertainty about the current trial's patterns, thereby gathering information about the environment. The amount of information provided by a specific outcome during a trial was calculated as the difference in the entropy before receiving the outcome compared to after. The expected information is a function of the number of possible patterns excluded by the next outcome, whereas expected reward is a function of the remaining reward to be harvested. 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 returns have been eliminated so far. Distinct possible reward outcomes may offer the same information, and so our task partially de-confounds information and reward. Given the structure of the task, expected reward and expected information are partially decorrelated (linear regression,  $R^2 = 0.13$ ).

As monkeys progressed through the array, forecasts of information still available about the foraging environment influenced the speed of responses (all behavioral analyses:  $n = 145,080$  choices, 24,180 trials). Though monkeys' choice response times, the time between target acquisitions, were more sluggish over the course of a trial (Fig. 1C, top panel; linear regression; Both monkeys:  $\beta = 0.0134 \pm 0.0004$ ,  $p < 1 \times 10^{-197}$ ; Monkey L:  $\beta = 0.0148 \pm 0.0005$ ,  $p < 1 \times 10^{-159}$ ; Monkey R:  $\beta = 0.0100 \pm 0.0007$ ,  $p < 1 \times 10^{-39}$ ), the presence of remaining environmental information defined an information boundary generally separating speedy from sluggish choice behavior (Student's t-test on responses when environmental information remained vs. after; Both monkeys:  $t(145,078) = -37.88$ ,  $p < 1 \times 10^{-311}$ ; Monkey L:  $t(103,912) = -38.49$ ,  $p < 1 \times 10^{-321}$ ; Monkey R:  $t(41,164) = -8.42$ ,  $p < 1 \times 10^{-16}$ ; see supplement and Fig. S1 for full response time regression results). Response times were significantly longer when all information about the array was exhausted than when information remained even after controlling for the number of remaining eye movements (Fig. 1C, bottom panel, response times as function of expected information).

The influence of information on response times suggested a similar influence may hold for the pattern of choices that monkeys made, resulting in trial-to-trial changes in this pattern. Though monkeys usually chose the 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; Fig. 1C; see methods), they occasionally diverged from this routine. We found that the informativeness of outcomes influenced this variability in the monkeys' patterns of choices as measured by behavioral entropy, the expected degree of divergence. First, monkeys' choices during a trial were egocentrically coded by its step size, the number of targets to the left or right 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. Anticipation of more informative choice outcomes significantly reduced the entropy of the monkeys' choices (Student's t-test; Both monkeys:  $t(96,718) = -19.25$ ,  $p < 1 \times 10^{-81}$ ; Monkey L:  $t(69,274) = -3.24$ ,  $p < 0.005$ ; Monkey R:  $t(27,442) = -23.99$ ,  $p < 1 \times 10^{-125}$ ; Fig. 1E, left panel; see supplement for full behavioral entropy regression results). While still harvesting information about the current trial, monkeys diverged less from routine traplines, but afterward they diverged more, becoming more variable in their choices (Student's t-test on choice numbers (CN) 4 or 5; 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. 1E, right panel). Hence, monkeys diverged less while choices were still informative 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 two monkeys (Fig. 1A; monkey L = 84 neurons; monkey R = 40 neurons). In order to control for behavioral confounds, all choices where monkeys diverged from traplines were excluded from the analyses in this section (those neural findings are reported in Barack, Chang, and Platt 2017).

Overall, we found that during the anticipation epoch (500 ms encompassing a 250 ms pre-choice 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 choice number in the trial (for example, choice number two (CN<sub>2</sub>): Student's t-test,  $p <$

0.0001; firing rate for 0.72 bits =  $23.57 \pm 1.33$  spikes/sec, firing rate for 1.37 bits =  $29.79 \pm 0.85$  spikes/sec). However, after controlling for choice number 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 =  $23.73 \pm 2.17$  spikes/sec, firing rate for 0.4 expected reward =  $23.43 \pm 1.64$  spikes/sec; Fig. 2A, second row from bottom, left panel). The tuning curves for this same cell collapsed across all choice numbers for both expected information and expected reward illustrate the strong sensitivity to large amounts of information (Fig. 2B).

In our population of 124 neurons, significantly more cells were tuned to information than reward when controlling for choice number in trial. A generalized linear model (GLM) regression revealed that during the anticipation epoch, 36 (29%) of 124 neurons (Monkey L: 26 (31%) of 84 neurons; Monkey R: 10 (25%) of 40 neurons) signaled the interaction of choice number 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 choice number and expected reward (all results,  $p < 0.05$ , Bonferroni corrected; Both monkeys:  $\chi^2 = 38.9138$ ,  $p < 1 \times 10^{-9}$ ; Monkey L:  $\chi^2 = 27.5808$ ,  $p < 5 \times 10^{-7}$ ; Monkey R:  $\chi^2 = 11.4286$ ,  $p < 0.001$ ; see methods; see supplement for results of full regression and individual monkey results). A direct test for signaling of expected reward, by comparing the average firing rates for different amounts of expected reward for the same choice number and expected information, revealed that only about 10% of neurons signaled expected reward, except on the last choice when all information had been received (Fig. 2C). In contrast, about 20% of neurons signaled expected information (Fig. 2C).

PCC Neurons Index Response Speed and Variability

We have previously established that PCC neurons signal decisions to diverge from traplines during our task (Barack, Chang, and Platt 2017). However, the extent to which these cells track speed and variability of responses during the task remains to be explored. We next examined the extent to which firing rates in PCC neurons predicted response times across all trials. A high proportion of neurons predicted response times in our task during the pre-saccade epoch, the 250 ms before the end of the choice saccade (both monkeys, 54 (44%) of 124 neurons significantly predicted response times, linear regression,  $p < 0.05$ ; Monkey L, 46 (55%) of 84 neurons; Monkey R, 14 (35%) of 40 neurons). In order to initially assess the possibility that expected information influenced response times through the PCC, a mediation analysis was run on the group of 9 cells (Monkey L: 9; Monkey R: 2; all variables were z-scored including the firing rates and a variable for cell identity was included in the analysis) that significantly encoded the interaction of expected information and choice number and that significantly predicted response times during that epoch. This analysis compared the direct effect of the interaction of expected information and choice number on saccade response time to the indirect, mediated effect of this variable on response time through PCC neurons (Hayes 2013). Mediation analysis revealed that both paths were small but significant: the direct pathway  $CN \times EI \rightarrow$  response times ( $p < 1 \times 10^{-6}$ ; estimated effect =  $-0.0121 \pm 0.0024$ ) as well as the indirect pathway  $CN \times EI \rightarrow$  firing rates  $\rightarrow$  response times ( $p < 5 \times 10^{-10}$ ; estimated effect =  $-0.0014 \pm 0.00022$ ) both significantly sped up responses. Hence, initial indications suggest that PCC partially mediates the influence of expected information on response times.

While PCC neurons predict response speed, we next asked whether these neurons index the degree to which behavior is variable, operationalized as behavioral entropy (see methods).



During the anticipation epoch, behavioral entropy varied significantly with firing rate for 70 (56%) neurons (log-linear regression of behavioral entropy against firing rate,  $p < 0.05$ ; results by choice number: 46 neurons (37%) for CN1, 27 neurons (22%) for CN2, 31 neurons (25%) for CN3, and 41 neurons (33%) for CN4). An example cell was more active for high entropy choices compared to low (Student's t-test,  $p < 1 \times 10^{-32}$ ; Fig. 3A). Across the population, higher firing rates predicted greater behavioral entropy (Student's t-test on mean normalized firing rates during anticipation epoch,  $p < 1 \times 10^{-5}$ ; Fig. 3B). Higher firing rates predicted greater behavioral entropy (BE) in the majority of both the significant subpopulation of 70 cells (GLM,  $p < 0.05$ ,  $\beta_{BE} > 0$  in 51 cells,  $\beta_{BE} \leq 0$  in 19 cells; mean  $\beta_{BE} = 0.0023 \pm 0.0011$  bits<sub>BE</sub>/spike, Student's t-test against  $\beta_{BE} = 0$ ,  $t(69) = 2.08$ ,  $p < 0.05$ ) and the whole population (124 neurons;  $\beta_{BE} > 0$  in 83 cells,  $\beta_{BE} \leq 0$  in 41 cells; mean  $\beta_{BE} = 0.0058 \pm 0.0020$  bits<sub>BE</sub>/spike, Student's t-test,  $t(123) = 2.85$ ,  $p < 0.01$ ). The percent modulation of behavioral entropy per 1 Hz increase in firing rate can be calculated by dividing the mean regression coefficient for the population by the mean behavioral entropy. In the subpopulation of significant neurons, these mean regression coefficients represent an increase of 2.02% in the average behavioral entropy for every additional spike, and across the whole population, of 1.49% for every additional spike.

Monkeys' behavior could be characterized as occupying two states along a continuum from routine to divergent. In the more routine state they made faster, less variable choices, while in the more divergent state they made slower, more variable choices. Recall that a boundary defined by the receipt of the last bit of information, when the pattern of rewards on a given trial becomes fully resolved, generally separated periods of routine from divergent behavior. We next investigated whether PCC neurons signaled this information boundary. After a follow-up regression of each

trial's binned spike counts against the time in the trial and the time of last informative outcome, we found that 98 (79%) of 124 neurons differentiated these two states (GLM, effect of interaction,  $p < 0.05$ ). During a four second epoch centered on the time of the last informative choice outcome, an example cell fired less before that outcome than after (Student's test,  $p < 1 \times 10^{-56}$ ; Fig. 3C). The population of cells also signaled the transition from routine to divergent behavior, firing more after the last bit of information was received (Student's t-test,  $p < 0.005$ ; Fig. 3D).

Finally, behavioral entropy signals and information boundary signals were multiplexed in the PCC population. Significantly fewer cells ( $\chi^2$ ,  $p < 1 \times 10^{-10}$ ) predicted behavioral entropy after receiving all information (24 (19%) of 124 neurons) than before (74 (60%) neurons). A comparison of PCC population responses on choices with high to low behavioral entropy revealed significant differences before receipt of the last informative outcome (Student's t-test,  $p < 1 \times 10^{-4}$ ) but not after (Student's t-test,  $p > 0.5$ ; Fig. S2), with greater modulation for high entropy compared to low.

## Discussion

In this study, we show that environmental information influences responses during routine behavior and that the posterior cingulate cortex (PCC) signals this information and predicts response times and behavioral variability. Despite the fact that in our task monkeys could not utilize environmental information to increase their chance of reward, the receipt of environmental information and the exhaustion of uncertainty impacted behavioral routines. Monkeys' responses were faster and less variable when there was more information to be gathered, but surprisingly slowed and became more variable once the environment became fully known. This pattern of slow responses after resolving all environmental uncertainty departs from the reward rate maximizing

strategy of moving in a circle. While monkeys traplined, neurons in PCC robustly signaled information expectations but not reward expectations. These neurons also predicted the speed of responses as monkeys traplined, with neural activity in the subset of the population that signaled information expectations and predicted response times helping to mediate the influence of informativeness on response. PCC also neurons differentiate the degree of behavioral variability before compared to after all information was received about the pattern of rewards, with increasing activity following receipt of the last informative outcome and decreased representations of behavioral variability. In sum, our experimental findings suggest that PCC tracks the state of the environment in order to influence routine behavior.

Monkeys generally chose targets in the same pattern regardless of which targets had been most recently rewarded or informative, 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 of in groups of wild foraging primates(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(Menzel 1973; Gallistel and Cramer 1996; Cramer and Gallistel 1997; MacDonald and Wilkie 1990). 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 cortico-cortical connectivity(Heilbronner and Haber 2014) and elevated resting state and off-task metabolic activity(Buckner, Andrews-Hanna, and Schacter 2008), is at the heart of the default mode network (DMN)(Buckner, Andrews-Hanna, and Schacter 2008). The DMN is a cortex-spanning network implicated in divergent cognition including imagination(Schacter et al. 2012), creativity(Kühn et al. 2014), and narration(Wise and Braga 2014). Though 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 routine 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, Hayden, and Platt 2011) may in fact 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 in fact 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, Gilbert, and Burgess 2011), decision making(Kable and Glimcher 2007), and creativity(Beaty et al. 2015). Interestingly, our

monkeys adopted a behavioral strategy of slower responses once the state of the environment was fully known. These slower responses may be adaptive insofar as they permit re-evaluation of strategies and potential divergences from routines(Barack, Chang, and Platt 2017).

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 tradeoff between engaging well-rehearsed routines and deviating in search of new, potentially better solutions(Gottlieb et al. 2013; Andrews-Hanna, Smallwood, and Spreng 2014; Beaty et al. 2015). For example, creativity involves diverging from typical patterns of thought, such as occurs in generating ideas(Benedek et al. 2014) or crafting novel concepts(Guilford 1959; Barron 1955). 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 in locating novel resources(Ohashi and Thomson 2005) or discovering 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 provides 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 towards 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, either 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 evidence for making 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 valuationsal computations on the basis of information(Manohar and Husain 2013). 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 sum, harvesting information and behavioral speed and variability were both signaled by PCC neurons, suggesting a central role for PCC in determining how information drives exploration and possibly prospective cognition. Monkeys were sensitive to the amount of uncertainty remaining in the environment, with faster and more reliable patterns of choices while information remained and slower and more variable patterns after environmental uncertainty had been resolved. 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 divergence from routines.

**Acknowledgments:** This work was supported by the National Eye Institute of the National Institutes of Health (R01 EY013496 to M.L.P.) and an Incubator Award from the Duke Institute for Brain Sciences. Correspondence and requests for materials should be addressed to [dbarack@gmail.com](mailto:dbarack@gmail.com), or to the Center for Science and Society, Fayerweather Hall 511, Columbia University, 1180 Amsterdam Ave., New York, NY 10027.

**Contributions:** D.L.B. designed the experiment, D.L.B. collected and analyzed the data, D.L.B. and M.L.P. prepared and revised the manuscript.

## References

- Aljadeff, Johnatan, Benjamin J Lansdell, Adrienne L Fairhall, and David Kleinfeld. 2016. 'Analysis of neuronal spike trains, deconstructed', *Neuron*, 91: 221-59.
- Andrews-Hanna, Jessica R, Jonathan Smallwood, and R Nathan Spreng. 2014. 'The default network and self-generated thought: component processes, dynamic control, and clinical relevance', *Ann N Y Acad Sci*, 1316: 29-52.
- Barack, David L., Steve W. C. Chang, and Michael L. Platt. 2017. 'Posterior Cingulate Neurons Dynamically Signal Decisions to Disengage during Foraging', *Neuron*, 96: 339-47.e5.
- Barron, Frank. 1955. 'The disposition toward originality', *The Journal of Abnormal and Social Psychology*, 51: 478.
- Beaty, Roger E, Mathias Benedek, Scott Barry Kaufman, and Paul J Silvia. 2015. 'Default and executive network coupling supports creative idea production', *Scientific reports*, 5.
- Bell, William J. 1990. 'When to return to a resource patch.' in, *Searching Behaviour* (Springer).
- Benedek, Mathias, Emanuel Jauk, Andreas Fink, Karl Koschutnig, Gernot Reishofer, Franz Ebner, and Aljoscha C Neubauer. 2014. 'To create or to recall? Neural mechanisms underlying the generation of creative new ideas', *Neuroimage*, 88: 125-33.

- 1 Benoit, Roland G., Sam J. Gilbert, and Paul W. Burgess. 2011. 'A Neural Mechanism Mediating  
2 the Impact of Episodic Prospection on Farsighted Decisions', *The Journal of Neuroscience*,  
3 31: 6771-79.
- 4 Berger-Tal, Oded, and Shirli Bar-David. 2015. 'Recursive movement patterns: review and  
5 synthesis across species', *Ecosphere*, 6: art149.
- 6 Boggs, Carol L, John T Smiley, and Lawrence E Gilbert. 1981. 'Patterns of pollen exploitation by  
7 Heliconius butterflies', *Oecologia*, 48: 284-89.
- 8 Buckner, R. L., J. R. Andrews-Hanna, and D. L. Schacter. 2008. 'The brain's default network:  
9 anatomy, function, and relevance to disease', *Ann N Y Acad Sci*, 1124: 1-38.
- 10 Clithero, John A, and Antonio Rangel. 2014. 'Informatic parcellation of the network involved in  
11 the computation of subjective value', *Social Cognitive and Affective Neuroscience*, 9: 1289-  
12 302.
- 13 Cramer, Audrey E., and C. R. Gallistel. 1997. 'Vervet monkeys as travelling salesmen', *Nature*,  
14 387: 464.
- 15 Cunningham, Elena, and Charles Janson. 2007. 'Integrating information about location and value  
16 of resources by white-faced saki monkeys (*Pithecia pithecia*)', *Animal cognition*, 10: 293-  
17 304.
- 18 Dayan, P., and L.F. Abbott. 2001. *Theoretical Neuroscience* (MIT Press: Cambridge,  
19 Massachusetts).
- 20 Desrochers, Theresa M., Dezhe Z. Jin, Noah D. Goodman, and Ann M. Graybiel. 2010. 'Optimal  
21 habits can develop spontaneously through sensitivity to local cost', *Proceedings of the*  
22 *National Academy of Sciences*, 107: 20512-17.
- 23 Di Fiore, Anthony, and Scott A. Suarez. 2007. 'Route-based travel and shared routes in sympatric  
24 spider and woolly monkeys: cognitive and evolutionary implications', *Animal cognition*,  
25 10: 317-29.
- 26 Fu, Wai-Tat, and Peter Pirolli. 2007. 'SNIF-ACT: A cognitive model of user navigation on the  
27 World Wide Web', *Human-Computer Interaction*, 22: 355-412.
- 28 Gallistel, C. R., and Audrey E. Cramer. 1996. 'Computations on metric maps in mammals: getting  
29 oriented and choosing a multi-destination route', *The Journal of Experimental Biology*,  
30 199: 211-17.
- 31 Garber, Paul A. 1988. 'Foraging decisions during nectar feeding by tamarin monkeys (*Saguinus*  
32 *mystax* and *Saguinus fuscicollis*, Callitrichidae, Primates) in Amazonian Peru', *Biotropica*:  
33 100-06.
- 34 Garrison, Jennifer SE, and Clifton Lee Gass. 1999. 'Response of a traplining hummingbird to  
35 changes in nectar availability', *Behavioral Ecology*, 10: 714-25.
- 36 Gill, Frank B. 1988. 'Trapline foraging by hermit hummingbirds: competition for an undefended,  
37 renewable resource', *Ecology*: 1933-42.
- 38 Gottlieb, Jacqueline, Pierre-Yves Oudeyer, Manuel Lopes, and Adrien Baranes. 2013.  
39 'Information-seeking, curiosity, and attention: computational and neural mechanisms',  
40 *Trends in Cognitive Sciences*, 17: 585-93.
- 41 Guilford, Joy Paul. 1959. 'Traits of creativity', *Creativity and its cultivation*, 10: 141-61.
- 42 Hamming, Richard W. 1950. 'Error detecting and error correcting codes', *Bell System technical*  
43 *journal*, 29: 147-60.

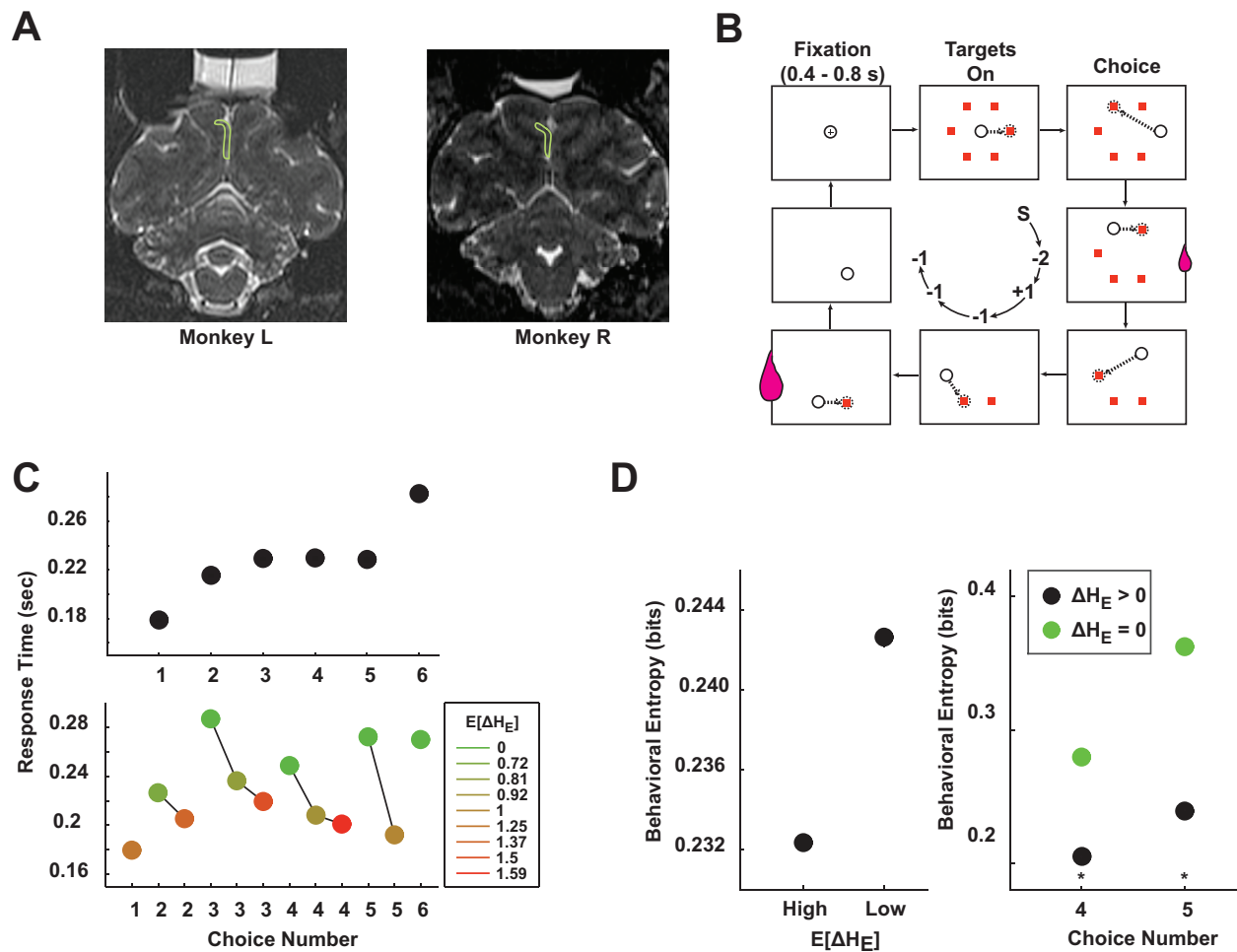


- 1 Hayden, B. Y., A. C. Nair, A. N. McCoy, and M. L. Platt. 2008. 'Posterior cingulate cortex  
2 mediates outcome-contingent allocation of behavior', *Neuron*, 60: 19-25.
- 3 Hayden, B. Y., and M. L. Platt. 2010. 'Neurons in anterior cingulate cortex multiplex information  
4 about reward and action', *Journal of Neuroscience*, 30: 3339-46.
- 5 Hayes, Andrew F. 2013. *Introduction to Mediation, Moderation, and Conditional Process  
6 Analysis: A Regression-Based Approach*. (The Guilford Press: New York, NY).
- 7 Heilbronner, Sarah R., and Michael L Platt. 2013. 'Causal Evidence of Performance Monitoring by  
8 Neurons in Posterior Cingulate Cortex during Learning', *Neuron*, 80: 1384-91.
- 9 Heilbronner, Sarah R., and Suzanne N. Haber. 2014. 'Frontal Cortical and Subcortical Projections  
10 Provide a Basis for Segmenting the Cingulum Bundle: Implications for Neuroimaging and  
11 Psychiatric Disorders', *The Journal of Neuroscience*, 34: 10041-54.
- 12 Heilbronner, Sarah R., Benjamin Y. Hayden, and Michael L. Platt. 2011. 'Decision salience signals  
13 in posterior cingulate cortex', *Front Neurosci*, 5.
- 14 Hemmi, Jan M, and Charles R Menzel. 1995. 'Foraging strategies of long-tailed macaques, *Macaca  
15 fascicularis*: directional extrapolation', *Animal Behaviour*, 49: 457-64.
- 16 Hirtle, Stephen C., and Tommy Gärling. 1992. 'Heuristic rules for sequential spatial decisions',  
17 *Geoforum*, 23: 227-38.
- 18 Hui, Sam K., Peter S. Fader, and Eric T. Bradlow. 2009. 'Research note-The traveling salesman  
19 goes shopping: The systematic deviations of grocery paths from TSP optimality',  
20 *Marketing Science*, 28: 566-72.
- 21 Janmaat, Karline R. L., Richard W. Byrne, and Klaus Zuberbühler. 2006. 'Evidence for a spatial  
22 memory of fruiting states of rainforest trees in wild mangabeys', *Animal Behaviour*, 72:  
23 797-807.
- 24 Janson, Charles H. 1998. 'Experimental evidence for spatial memory in foraging wild capuchin  
25 monkeys, *Cebus apella*', *Animal Behaviour*, 55: 1229-43.
- 26 Janzen, DH. 1971. 'Euglossine bees as long-distance pollinators of tropical plants', *Science*, 171:  
27 203-05.
- 28 Johnson, Adam, Zachary Varberg, James Benhardus, Anthony Maahs, and Paul Schrater. 2012.  
29 'The hippocampus and exploration: dynamically evolving behavior and neural  
30 representations', *Frontiers in human neuroscience*, 6.
- 31 Kable, J. W., and P. W. Glimcher. 2007. 'The neural correlates of subjective value during  
32 intertemporal choice', *Nat Neurosci*, 10: 1625-33.
- 33 Kühn, Simone, Simone M. Ritter, Barbara C. N. Müller, Rick B. Baaren, Marcel Brass, and Ap  
34 Dijksterhuis. 2014. 'The importance of the default mode network in creativity—a structural  
35 MRI study', *The Journal of Creative Behavior*, 48: 152-63.
- 36 Leech, Robert, Salwa Kamourieh, Christian F Beckmann, and David J Sharp. 2011. 'Fractionating  
37 the default mode network: distinct contributions of the ventral and dorsal posterior  
38 cingulate cortex to cognitive control', *The Journal of Neuroscience*, 31: 3217-24.
- 39 MacDonald, Suzanne E., and Donald M. Wilkie. 1990. 'Yellow-nosed monkeys'(*Cercopithecus  
40 ascanius whitesidei*) spatial memory in a simulated foraging environment', *Journal of  
41 Comparative Psychology*, 104: 382.
- 42 MacGregor, James N., and Yun Chu. 2011. 'Human performance on the traveling salesman and  
43 related problems: A review', *The Journal of Problem Solving*, 3: 2.

- 1 MacGregor, James N., and Tom Ormerod. 1996. 'Human performance on the traveling salesman  
2 problem', *Perception & Psychophysics*, 58: 527-39.
- 3 Manning, Aubrey. 1956. 'Some aspects of the foraging behaviour of bumble-bees', *Behaviour*, 9:  
4 164-200.
- 5 Manohar, Sanjay G, and Masud Husain. 2013. 'Attention as foraging for information and value',  
6 *Frontiers in human neuroscience*, 7: 711.
- 7 McCoy, A. N., J. C. Crowley, G. Haghighian, H. L. Dean, and M. L. Platt. 2003. 'Saccade reward  
8 signals in posterior cingulate cortex', *Neuron*, 40: 1031-40.
- 9 McCoy, A. N., and M. L. Platt. 2005. 'Risk-sensitive neurons in macaque posterior cingulate  
10 cortex', *Nature neuroscience*, 8: 1220-27.
- 11 McNamara, John. 1982. 'Optimal patch use in a stochastic environment', *Theoretical Population  
12 Biology*, 21: 269-88.
- 13 Menzel, Charles. 1996. 'Structure-guided foraging in long-tailed macaques', *American Journal of  
14 Primatology*, 38: 117-32.
- 15 Menzel, Charles R. 1991. 'Cognitive aspects of foraging in Japanese monkeys', *Animal Behaviour*,  
16 41: 397-402.
- 17 Menzel, Emil W. 1973. 'Chimpanzee spatial memory organization', *Science*, 182: 943-45.
- 18 Miller, George. 1983. 'Informavores.' in F. Machlup and U. Mansfield (eds.), *The study of  
19 information: Interdisciplinary messages* (Wiley-Interscience).
- 20 Noser, Rahel, and Richard W. Byrne. 2007. 'Travel routes and planning of visits to out-of-sight  
21 resources in wild chacma baboons, *Papio ursinus*', *Animal Behaviour*, 73: 257-66.
- 22 ———. 2010. 'How do wild baboons (*Papio ursinus*) plan their routes? Travel among multiple  
23 high-quality food sources with inter-group competition', *Animal cognition*, 13: 145-55.
- 24 Ohashi, Kazuharu, Alison Leslie, and James D Thomson. 2008. 'Trapline foraging by bumble bees:  
25 V. Effects of experience and priority on competitive performance', *Behavioral Ecology*,  
26 19: 936-48.
- 27 ———. 2013. 'Trapline foraging by bumble bees: VII. Adjustments for foraging success following  
28 competitor removal', *Behavioral Ecology*, 24: 768-78.
- 29 Ohashi, Kazuharu, and James D. Thomson. 2005. 'Efficient harvesting of renewing resources',  
30 *Behavioral Ecology*, 16: 592-605.
- 31 Pearson, J. M., B. Y. Hayden, S. Raghavachari, and M. L. Platt. 2009. 'Neurons in posterior  
32 cingulate cortex signal exploratory decisions in a dynamic multioption choice task', *Curr  
33 Biol*, 19: 1532-7.
- 34 Pirolli, Peter, and Stuart Card. 1999. 'Information foraging', *Psychological review*, 106: 643.
- 35 Pirolli, Peter L. T. 2007. *Information foraging theory: Adaptive interaction with information*  
36 (Oxford University Press).
- 37 Pochron, Sharon T. 2001. 'Can concurrent speed and directness of travel indicate purposeful  
38 encounter in the yellow baboons (*Papio hamadryas cynocephalus*) of Ruaha National Park,  
39 Tanzania?', *International Journal of Primatology*, 22: 773-85.
- 40 Possingham, Hugh P. 1989. 'The distribution and abundance of resources encountered by a  
41 forager', *American Naturalist*: 42-60.
- 42 Racey, PA, and SM Swift. 1985. 'Feeding ecology of *Pipistrellus pipistrellus* (Chiroptera:  
43 Vespertilionidae) during pregnancy and lactation. I. Foraging behaviour', *The Journal of  
44 Animal Ecology*: 205-15.

- 1 Schacter, Daniel L, Donna Rose Addis, Demis Hassabis, Victoria C Martin, R. Nathan Spreng,  
2 and Karl K Szpunar. 2012. 'The Future of Memory: Remembering, Imagining, and the  
3 Brain', *Neuron*, 76: 677-94.
- 4 Schmidhuber, Jürgen. 1991. "Curious model-building control systems." In *1991 IEEE*  
5 *International Joint Conference on Neural Networks.*, 1458-63. IEEE.
- 6 Stephens, D.W., and J.R. Krebs. 1986. *Foraging Theory* (Princeton University Press: Princeton,  
7 NJ).
- 8 Sutton, Richard S., and Andrew G. Barto. 1998. *Reinforcement learning : an introduction* (MIT  
9 Press: Cambridge, Mass.).
- 10 Valero, Alejandra, and Richard W. Byrne. 2007. 'Spider monkey ranging patterns in Mexican  
11 subtropical forest: do travel routes reflect planning?', *Animal cognition*, 10: 305-15.
- 12 Wan, Xiaohong, Kang Cheng, and Keiji Tanaka. 2015. 'Neural encoding of opposing strategy  
13 values in anterior and posterior cingulate cortex', *Nature neuroscience*, 18: 752-59.
- 14 Wise, Richard J. S., and Rodrigo M. Braga. 2014. 'Default mode network: the seat of literary  
15 creativity?', *Trends in Cognitive Sciences*, 18: 116-17.
- 16 Zou, Hui, and Trevor Hastie. 2005. 'Regularization and variable selection via the elastic net',  
17 *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67: 301-20.
- 18

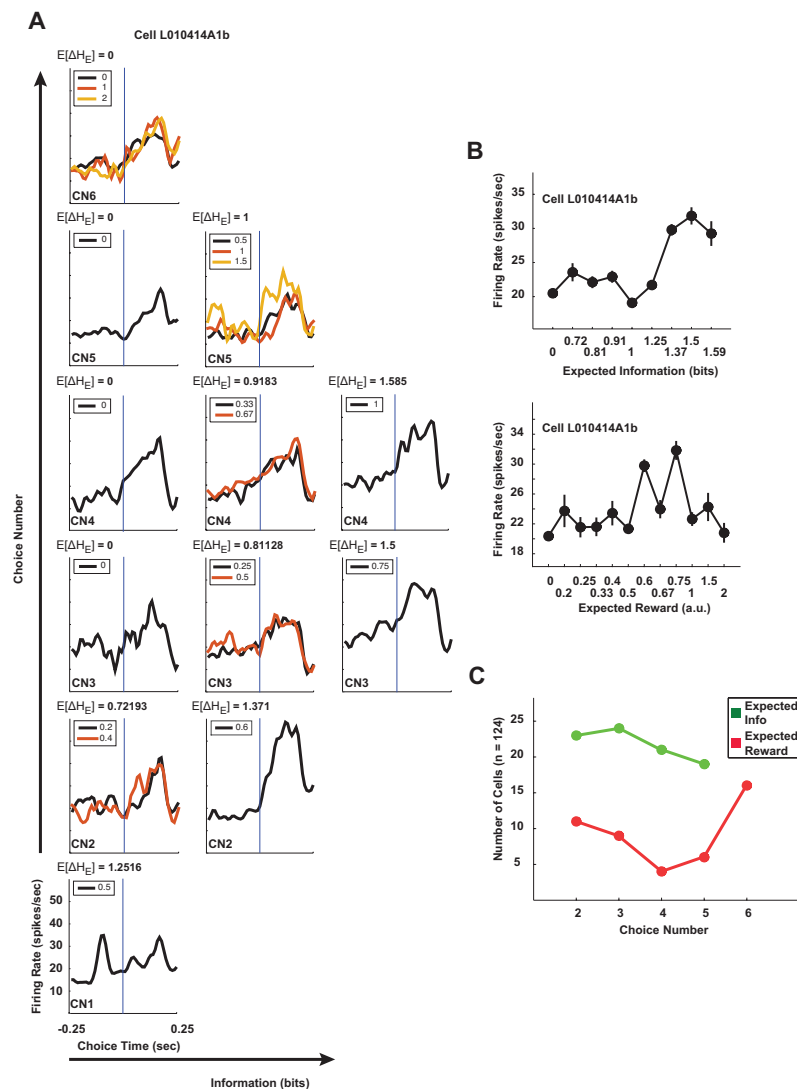
## 19 **Figures**



**Fig. 1.** Monkeys spontaneously trapline when foraging in a circular array but diverge from these efficient routines as the environment becomes better known. **A.** Recording location in the posterior cingulate cortex (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

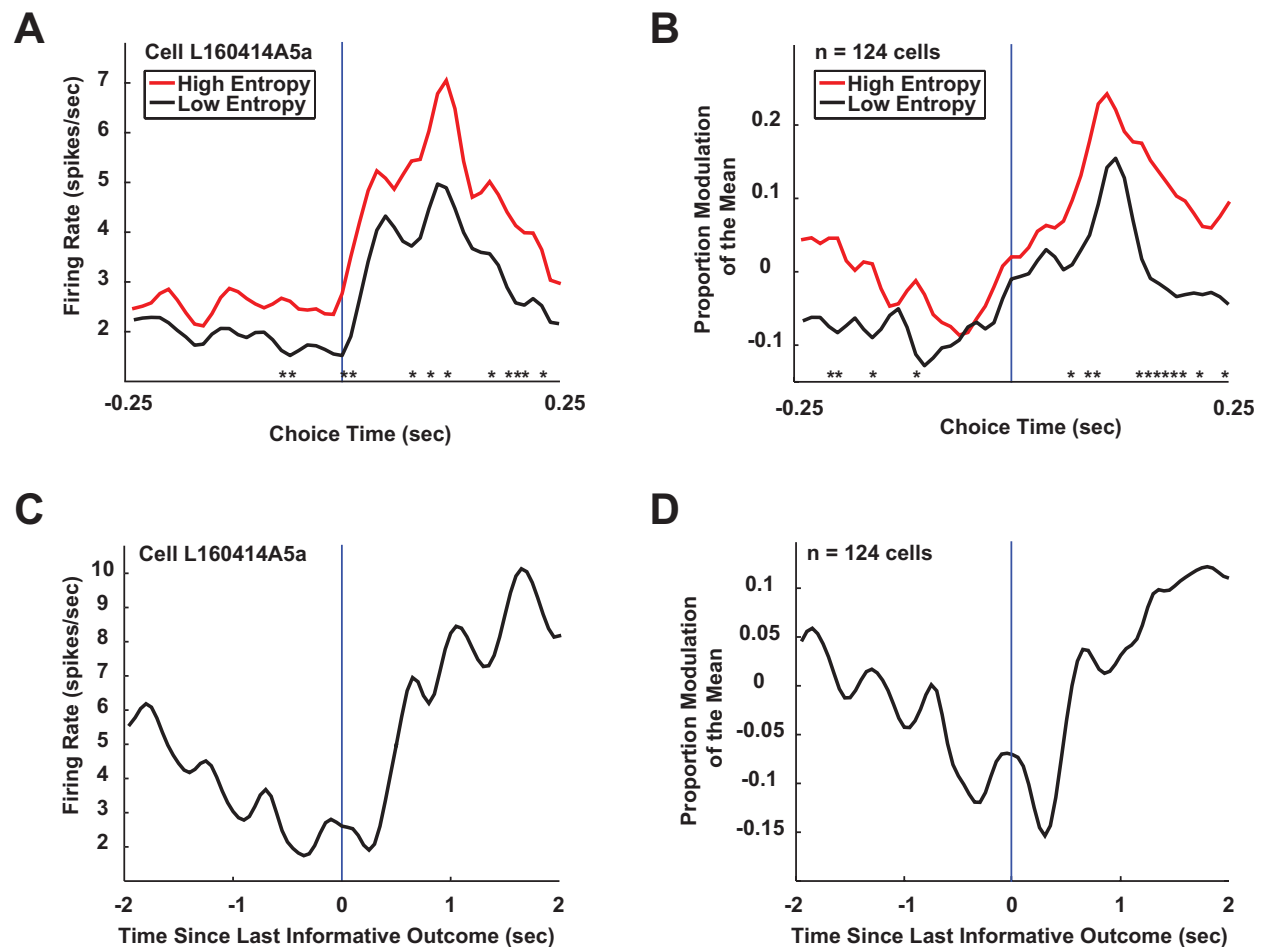
1 select every target. Open circle: simulated eye position; dashed arrow: direction of impending  
2 saccade; dashed circle: impending saccade endpoint; small juice drop: small reward; large juice  
3 drop: large reward. Central circle: step size, the clockwise or counter-clockwise distance between  
4 subsequently chosen targets; ‘S’ = start; ‘-2’ = two targets counter-clockwise; ‘-1’ = one target  
5 counter-clockwise; ‘+1’ = one target clockwise. **C.** Top panel: response times increased throughout  
6 the trial across all six choices; bottom panel: response times were faster for greater expected  
7 information, controlling for choice number in trial. **D.** Left panel: behavioral entropy for high  
8 expected information choices ( $E[\Delta H_E] > \text{mean}(E[\Delta H_E])$ ) compared to low expected information  
9 choices ( $E[\Delta H_E] \leq \text{mean}(E[\Delta H_E])$ ); right panel: behavioral entropy for choice numbers four and  
10 five, before receipt of last informative outcome ( $\Delta H_E > 0$ ; black points) and after ( $\Delta H_E = 0$ ; green

1 points). **All** plots: error bars occluded by points; \* =  $p < 0.05$ , Bonferroni corrected;  $n = 145,080$   
 2 choices, 24,180 trials.



3  
 4 **Fig. 2.** PCC neurons preferentially encode environmental information over reward. **A.** Firing rate  
 5 of sample neuron encoding expected information but not expected reward across all choice  
 6 numbers (CN) one through six, plotted separately by expected information ( $E[\Delta H_E]$ ). Legends  
 7 indicate expected reward(s) for each plot. Blue line = end of saccade. **B.** Tuning curves for

expected information (top panel) and expected reward (bottom panel), collapsed across choice numbers for better visibility, for the cell plotted in **A**. This example cell showed elevated firing rates for higher amounts of information. **C**. Percent of cells encoding expected reward by choice (red) for constant expected information, and percent of cells encoding expected information by choice (green).



**Fig. 3.** Neurons in PCC forecast divergent behavior. **A**. Same neuron encoding behavioral entropy during the anticipation epoch. This cell was more active for high entropy choices than for low entropy choices. **B**. Population encoding of behavioral entropy. 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 outcome compared to before. **D.** The population also encoded this information boundary, with higher firing rates after the last informative outcome compared to before. **B** and **D** plots:  $n = 125$  cells. **All** plots: blue line = end of saccade;  $* = p < 0.05$  for that 10 ms time bin.

## Supplement

To assess the effect of information on behavior independently of reward, we designed a multi-option choice task that partially decorrelated information and reward. Monkeys performed an iterated variant of a Hamiltonian path problem, similar to the uncapacitated traveling purchaser problem, a generalization of the traveling salesman problem (Ramesh 1981; Boctor, Laporte, and Renaud 2003). Hamiltonian path problems are like traveling salesmen problems but the agent is not required to return to the starting location after their trip. In this problem, the agent must efficiently visit a series of markets that may or may not have the desired product, such that in addition to potentially purchasing the product (receiving reward), the agent also learns if the product is available (receiving information). This version of the problem is uncapacitated because the problem assumes that the markets always have sufficient capacity (or supply of a product) to meet the demand. Using this paradigm with varying reward sizes, reward and information can be partially decorrelated.

## Supplemental Methods

### Task Analysis



Our experiment required monkeys to select once each of a set of six targets to harvest the rewards. In every trial in our experiment, two fixed rewards (large and small) were each assigned to one of six locations in the environment in a pseudorandom fashion (Fig. 1B). Trials began with monkeys fixating a central cross for a variable amount of time, ranging from (0.5 – 1 sec). 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'd 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.

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 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 S1).

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 ( $R^2 = 0.1082$ ), 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 ( $R^2 = 0.3935$ ).

## Neural Regressions

To corroborate our generalized linear model (GLM) regression, we used elastic net regression. Elastic net regression is a model fitting technique that balances two different ways of finding model fits: LASSO (least absolute shrinkage and selection operator) regression and ridge regression. These two model fitting techniques correspond to two ways of computing the distance from the model to the data: using L1 norms (Manhattan distance; LASSO) and L2 norms (Euclidean distance; ridge regression). LASSO regression tends to shift to zero the values of one of two correlated coefficients in order to increase model fit, and ridge regression includes a bias term that helps resolve multi-collinearity in the coefficients. The tradeoff between LASSO and ridge regression is accomplished by means of a parameter  $\alpha$ . We used *fmincon*, a constrained search function in MATLAB, to find the maximum likelihood estimate of the best-fit tradeoff parameter on a cell-by-cell, epoch-by-epoch basis (average by epoch: both monkeys: anticipation: , outcome: ; monkey L: anticipation:  $\alpha = 0.9327 \pm 0.0132$ ; monkey R: anticipation:  $\alpha = 0.8963 \pm$

1 0.0239). We then compared these fits to the fits from our GLM in three [four?] ways: i. the identity  
2 of the covariates that were retained in the best-fit elastic net model; ii. the value of the coefficients  
3 computed by the best-fit elastic net model; and iii. the value of just the significant coefficients  
4 computed by the best-fit elastic net model. To confirm the identity of the significant covariates,  
5 we compared the covariates retained by the elastic net regression to the ones deemed significant  
6 by the original GLM. To compare the value of the coefficients computed by the two types of  
7 regression, we computed an elastic net score (e\_score) for each covariate for each cell for each  
8 epoch, equal to the ratio of the elastic net coefficient to the GLM coefficient (i.e.,  $e\_score =$   
9  $\beta_{\text{elastic\_net}} / \beta_{\text{GLM}}$ ). Values around 1 correspond to roughly equivalent computed coefficients; values  
10 greater than 1 correspond to a greater weight placed on the covariate by the elastic net compared  
11 to the GLM; and values less than 1 correspond to a greater weight placed on the covariate by the  
12 GLM compared to the elastic net. Finally, we also computed e\_scores but for the significant  
13 covariates only.

1

2 Table S1: Equations for expected reward, entropy, information, and expected information for the reward sequence.

Patter n #	Permutati on $P$	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 $EI_i = \frac{\sum_{P_{remaining}} I_i}{( P_{remaining} )}$
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

3

- 1 Each column in the table (except for the leftmost) contains six columns, each corresponding to a choice number during the trial.
- 2 Total number of choices on every trial = 6. Key:  $|\cdot|$  = cardinality of  $\cdot$  ; R = reward; n = choice number.

## Additional Behavioral Results: Response Times

Response times grew longer over the course of the trial (linear regression,  $p < 1 \times 10^{-197}$ ; Fig. 1D, top panel), consistent with past findings that saccade response grow more sluggish over the course of a sequence (Zingale and Kowler 1987). Despite this increasing sluggishness, monkeys were faster to select targets when expected information for making the next choice was higher (multilinear regression,  $p < 1 \times 10^{-9}$ ; Fig. 1D, bottom panel) but surprisingly not when the expected reward for the next choice was greater (multilinear regression,  $p > 0.93$ ; Fig. S1B). Thus, in addition to increased variability contributing to decreased reward rates, monkeys responses also reflected the state of their knowledge of the environment, growing slower as the trial progressed and further depressing reward rates.

Since different choice numbers in the sequence possess distinct reward and information expectations, response times were regressed against choice number (CN; from 1 to 6), previous information (I-1; from 0 to 2.585 bits<sub>info</sub>) and reward (R-1; from 0 to 2 units<sub>juice</sub>) outcomes, expected information (EI; from 0 to 1.585 bits<sub>info</sub>) and reward (ER; from 0 to 2 units<sub>juice</sub>), and all 2-way interactions. The regression coefficients, standard errors of the mean, and significance are reported in table S2 and displayed in Figure S1B.

Table S2:

Covariate	Beta	SE	p-value
Choice number (CN)	0.0325	0.0028	0.0000

Expected information (EI)	0.0236	0.0155	0.1295
Expected reward (ER)	-0.0833	0.0406	0.0404
Previous choice information (I-1)	0.0586	0.0109	0.0000
Previous choice reward (R-1)	-0.4772	0.0496	0.0000
CN x EI	-0.0158	0.0025	0.0000
CN x ER	0.0005	0.0068	0.9370
CN x I-1	-0.0241	0.0023	0.0000
CN x R-1	0.0524	0.0062	0.0000
EI x ER	0.0457	0.0087	0.0000
EI x I-1	-0.0066	0.0108	0.5417
EI x R-1	-0.0092	0.0068	0.1756
ER x I-1	0.0363	0.0106	0.0006
ER x R-1	0.0062	0.0046	0.1783
I-1 x R-1	0.1461	0.0136	0.0000
Monkey identity	-0.0460	0.0017	0.0000

Like response times after the end of informative outcomes, response times for choices after the animal had harvested the last reward in a trial were significantly longer than for choices when reward remained to be harvested (Student's t-test,  $t(145,078) = -46.3694$ ,  $p < 10^{-300}$ ; mean RT for reward remaining =  $0.2037 \pm 0.00055$  sec, mean RT for reward finished =  $0.2877 \pm 0.0028$  sec). Interestingly and surprisingly, we did not see an effect of expected rewards on response times.

Naïve regression revealed that response times were faster for higher expected information (linear regression,  $\beta = -0.0524$  sec/bits,  $p < 1 \times 10^{-317}$ ; Fig. S1A, main text) and seemingly for higher expected reward (linear regression,  $\beta = -0.0664$  sec/units juice,  $p < 1 \times 10^{-282}$ ). However, after Bonferroni correction for the number of tests, the only significant covariates that include expected reward also include an information covariate, either expected information (EI) or the previous choice's information outcome (I-1). In addition, response times were faster following more informative ( $\beta_{\text{CNxI-1}} = -0.0241 \pm 0.0023$  sec/bits<sub>info</sub>) but slower following more rewarding ( $\beta_{\text{CNxR-1}} = 0.0524 \pm 0.0062$  sec/units<sub>juice</sub>) previous choice outcomes. There was some monkey-to-monkey variability; the influence of the interaction of choice number and expected information was significant in monkey L but, after Bonferroni correction, not in monkey R (tables S4 and S5).

To corroborate the lack of effect of reward size on response times, we directly compared response times for identical trial histories except for the size of past rewards. This test explicitly controls for all other covariates. To directly assess the effect of expected reward on response time, we compared the response times for choices that were matched for the history of choices on that trial but for the size of the rewards, that is, they were matched for the choice number of the choice, information outcomes received thus far, and choice numbers of the rewards received, but not for the size of the rewards received on those choice numbers. Only one such t-test was significant regardless of Bonferroni correction, for response times during the second choice following receipt of a large or small reward on the first choice. Table S3 lists all the p-values from this set of tests. Each row corresponds to the choice number for which response times were compared. Each column corresponds to the choice number in the trial during which either the large or small reward



was received, thereby selecting a subset of the response times during the corresponding choice number to compare. For example, the first row (Choice number = 2) picks out all response times during the second choice, and the first column (First choice in trial) picks out the response times for which either a large or small reward was received during the first choice in the trial. The first entry (Choice number = 2, First choice in trial) lists the p-value from the t-test comparing the responses during the second choice where the first choice was the large reward to those response times during the second choice where the first choice was the small reward. As can be observed from Table S3, only the first test is significant.

Table S3:

Choice Number	First choice in trial	Second choice	Third choice	Fourth choice	Fifth choice
2	0.0028	0	0	0	0
3	0.9290	0.0665	0	0	0
4	0.9665	0.1059	0.2784	0	0
5	0.5433	0.2474	0.9829	0.1807	0
6	0.8676	0.3773	0.5329	0.0669	0.3958

Individual regression results of response times against choice number, expected information, expected reward, previous behavioral information outcomes, previous choice reward outcome, and all 2-way interactions now follow:

1 Table S4: Monkey L (103,914 choices, 17319 trials), RT regression:

Covariate	Beta	p-value
Choice number (CN)	0.0372	0.0000
Expected information (EI)	0.0255	0.1783
Expected reward (ER)	-0.1040	0.0359
Previous information (I-1)	0.0699	0.0000
Previous reward (R-1)	-0.4660	0.0000
CN x EI	-0.0175	0.0000
CN x ER	-0.0030	0.7213
CN x I-1	-0.0300	0.0000
CN x R-1	0.0515	0.0000
EI x ER	0.0610	0.0000
EI x I-1	-0.0297	0.0246
EI x R-1	0.0025	0.7646
ER x I-1	0.0661	0.0000
ER x R-1	0.0035	0.5433
I-1 x R-1	0.1469	0.0000

2

3 Table S5: Monkey R (41,166 choices, 6861 trials), RT regression:

Covariate	Beta	p-value
Choice number (CN)	0.0205	0.0000
Expected information (EI)	0.0170	0.5211
Expected reward (ER)	-0.0154	0.8228
Previous information (I-1)	0.0304	0.1001
Previous reward (R-1)	-0.4908	0.0000
CN x EI	-0.0107	0.0108

CN x ER	0.0070	0.5437
CN x I-1	-0.0088	0.0257
CN x R-1	0.0532	0.0000
EI x ER	0.0029	0.8431
EI x I-1	0.0512	0.0054
EI x R-1	-0.0370	0.0012
ER x I-1	-0.0405	0.0229
ER x R-1	0.0116	0.1345
I-1 x R-1	0.1392	0.0000

### Model Fits to the Behavior

Model fits to the behavior revealed that monkeys generally chose targets in the same pattern regardless of which targets had been most recently rewarded or informative.

This finding was corroborated by the daily dominant pattern results reported in the main text. Monkey R chose targets in the same pattern across all 14 recording sessions, whereas monkey L chose targets in the same pattern on 24 sessions, with four patterns accounting for the remaining six days (two days each for two of the patterns and one each for the remaining two patterns).

### Behavioral Entropy Additional Results

From the distribution of step sizes on each choice number, we calculated the behavioral entropy (BE) (0.0294 to 0.5021 bits<sub>BE</sub>; mean<sub>BE</sub> = 0.1139 ± 0.0170 bits<sub>BE</sub> across all choices) and correlated these values with choice number (CN), the next choice's expected information (EI) and

expected reward (ER), the previous choice's information (I-1) and reward (R-1) outcomes, and all 2-way interactions. After combining both monkeys' data, all covariates significantly influenced behavioral entropy ( $p < 0.05$ , Bonferroni corrected) except for the interaction of expected information and reward from the previous choice (EI x R-1,  $p > 0.6$ ). The effects of previous information outcome (I-1) and expected information (EI) interactions with choice number were both significant across the two monkeys (multilinear regression, all  $p$ 's  $< 0.0000001$ ; both monkeys:  $\beta_{\text{CNxEI}} = -0.1260 \pm 0.0011 \text{ bits}_{\text{BE}}/\text{bits}_{\text{info}}$ ;  $\beta_{\text{CNxI-1}} = -0.0538 \pm 0.0008 \text{ bits}_{\text{BE}}/\text{bits}_{\text{info}}$ ; Fig. S1C) and significant for both monkeys individually (monkey L:  $\beta_{\text{CNxI-1}} = -0.0672 \pm 0.0009 \text{ bits}_{\text{BE}}/\text{bits}_{\text{info}}$ ;  $\beta_{\text{CNxEI}} = -0.1626 \pm 0.0012 \text{ bits}_{\text{BE}}/\text{bits}_{\text{info}}$ ; monkey R:  $\beta_{\text{CNxI-1}} = -0.0181 \pm 0.0017 \text{ bits}_{\text{BE}}/\text{bits}_{\text{info}}$ ;  $\beta_{\text{CNxEI}} = -0.0296 \pm 0.0024 \text{ bits}_{\text{BE}}/\text{bits}_{\text{info}}$ ; Fig. S1C). The effects of the interaction of choice number and reward were significant but inconsistent: later previous reward outcomes resulted in significantly increased variability in the second monkey but not the first, whereas later reward expectations resulted in significantly increased variability in the first monkey but not the second (multilinear regression, all  $p$ 's  $< 0.0000001$  except where indicated; both monkeys:  $\beta_{\text{CNxR-1}} = 0.0235 \pm 0.0030 \text{ bits}_{\text{BE}}/\text{units}_{\text{juice}}$ ,  $\beta_{\text{CNxER}} = 0.0775 \pm 0.0026 \text{ bits}_{\text{BE}}/\text{units}_{\text{juice}}$ ; monkey L:  $\beta_{\text{CNxR-1}} = 0.0014 \pm 0.0033 \text{ bits}_{\text{BE}}/\text{units}_{\text{juice}}$ ,  $p > 0.66$ ,  $\beta_{\text{CNxER}} = 0.0980 \pm 0.0029 \text{ bits}_{\text{BE}}/\text{units}_{\text{juice}}$ ; monkey R:  $\beta_{\text{CNxR-1}} = 0.0734 \pm 0.0063 \text{ bits}_{\text{BE}}/\text{units}_{\text{juice}}$ ,  $\beta_{\text{CNxER}} = 0.0064 \pm 0.0055 \text{ bits}_{\text{BE}}/\text{units}_{\text{juice}}$ ,  $p > 0.24$ ; Fig. S1C). Here we report the full regression results (Table S7).

Table S6: Beta coefficient values, standard errors, and p-values from regression of behavioral entropy against the listed covariates and 2-way interactions.

Covariate	Beta	SE	p-value
Choice number	0.1238	0.0013	0
Previous info outcomes	0.1882	0.0033	0
Expected info	0.4966	0.0053	0
Previous reward outcomes	-0.1367	0.0265	0.0000
Expected reward	-0.4285	0.0116	0.0000
CN x I-1	-0.0538	0.0008	0
CN x EI	-0.1260	0.0011	0
CN x R-1	0.0235	0.0030	0.0000
CN x ER	0.0775	0.0026	0.0000
I-1 x EI	-0.0438	0.0041	0.0000
I-1 x R-1	0.0256	0.0077	0.0009
I-1 x ER	0.0448	0.0039	0.0000
EI x R-1	0.0013	0.0029	0.6608
EI x ER	0.0835	0.0025	0.0000
R-1 x ER	0.0907	0.0034	0.0000
Monkey ID	0.0110	0.0005	0.0000

1

2 We now report the full individual monkey regressions:

3

4 Table S7: Monkey L (103,914 choices, 17319 trials), behavioral entropy regression:

Covariate	Beta	SE	p-value
Choice number	0.1503	0.0014	0

Previous info outcomes	0.2552	0.0037	0
Expected info	0.6727	0.0060	0
Previous reward outcomes	0.1177	0.0299	0.0001
Expected reward	-0.4810	0.0130	0.0000
CN x I-1	-0.0672	0.0009	0
CN x EI	-0.1626	0.0012	0
CN x R-1	0.0014	0.0033	0.6646
CN x ER	0.0980	0.0029	0.0000
I-1 x EI	-0.0659	0.0047	0.0000
I-1 x R-1	-0.0580	0.0087	0.0000
I-1 x ER	0.0518	0.0044	0.0000
EI x R-1	0.0095	0.0033	0.0044
EI x ER	0.0680	0.0029	0.0000
R-1 x ER	0.0847	0.0038	0.0000

1

2 Table S8: Monkey R (41,166 choices, 6861 trials), behavioral entropy regression:

Covariate	Beta	SE	p-value
Choice number	0.0603	0.0027	0.0000
Previous info outcomes	0.0196	0.0070	0.0053
Expected info	0.0450	0.0113	0.0001
Previous reward outcomes	-0.7385	0.0565	0.0000
Expected reward	-0.1875	0.0246	0.0000

CN x I-1	-0.0181	0.0017	0.0000
CN x EI	-0.0296	0.0024	0.0000
CN x R-1	0.0734	0.0063	0.0000
CN x ER	0.0064	0.0055	0.2483
I-1 x EI	0.0031	0.0089	0.7224
I-1 x R-1	0.2299	0.0164	0.0000
I-1 x ER	0.0258	0.0083	0.0019
EI x R-1	-0.0159	0.0062	0.0110
EI x ER	0.1156	0.0054	0.0000
R-1 x ER	0.0993	0.0072	0.0000

1

2 The uninformative outcomes could occur for choices four and five, as some choices occurred when

3 the previous choice's outcome was uninformative, after the current trial's pattern was already

4 revealed. In the text, we reported that behavioral entropy was greater following uninformative than

5 informative choices, collapsing across choices four and five. Both choices four and five

6 individually reflected this finding (CN<sub>4</sub>: BE<sub>info</sub> = 0.2050 ± 0.0003678 bits, BE<sub>no\_info</sub> = 0.2801 ±

7 0.0017 bits, t(24178) = -52.0376, p ~ 0; CN<sub>5</sub>: BE<sub>info</sub> = 0.2402 ± 0.0005666 bits, BE<sub>no\_info</sub> = 0.3627

8 ± 0.0012, t(24178) = -96.3793, p ~ 0). The individual monkey results reflected this difference as

9 well (for monkey L: for choice number four: BE<sub>info</sub> = 0.2907 ± 0.0007 bits, BE<sub>no\_info</sub> = 0.3272 ±

10 0.0032 bits, t(7857) = -12.4444, p < 1x10<sup>-34</sup>; for choice number five: BE<sub>info</sub> = 0.3481 ± 0.001

11 bits, BE<sub>no\_info</sub> = 0.4334 ± 0.0017, t(7857) = -37.2945, p < 1x10<sup>-279</sup>; for monkey R: for choice

12 number four: BE<sub>info</sub> = 0.2889 ± 0.001 bits, BE<sub>no\_info</sub> = 0.4264 ± 0.0034 bits, t(6859) = -35.3207, p

<  $1 \times 10^{-250}$ ; for choice number five:  $BE_{\text{info}} = 0.3005 \pm 0.0012$  bits,  $BE_{\text{no\_info}} = 0.4455 \pm 0.0017$ ,  $t(6859) = -58.4394$ ,  $p \sim 0$ ). Thus, the effects were consistent and strong in both monkeys.

#### Additional Neural Results: Elastic Net Regression

Our three checks—verifying the presence of the significant covariates, comparing the elastic net weights to the original GLM weights for all covariates, and a similar comparison restricted to only the significant covariate coefficients from the original GLM (see supplemental methods)—confirmed our original GLM analysis. First, none of the significant covariates identified by our original GLM for any cell during any epoch for either monkey received a coefficient of 0 from the elastic net regression. This suggests that our regression results were robust despite the correlations in the covariates.

Second, we compared the  $e\_scores$  (see supplemental methods) for the elastic net and original GLM regressions. The average  $e\_score = 1.3747 \pm 0.5120$  (monkey L:  $e\_score = 1.5968 \pm 0.7405$ ; monkey R:  $e\_score = 0.9084 \pm 0.0823$ ), suggesting that the elastic net regression tended to weight the covariates more than the original GLM, though this was clearly driven by monkey L. (Examination of the individual covariate comparisons further revealed that this effect was driven by the immense weight that the elastic net regression put on expected information for monkey L; see table below.) The following table reports the individual covariate  $e\_scores$  (mean  $\pm$  sem). Note that the elastic net regression weighted expected information much more than the original GLM but only for monkey L. All other fits for that epoch were comparatively reasonably close to the original GLM.



1

2 Table S9: Elastic-net regression results for main GLM.

Covariate	Monkey L	Monkey R	Both
Choice number (CN)	$1.1228 \pm 0.2436$	$1.0377 \pm 0.0728$	$1.0953 \pm 0.1664$
Expected reward (ER)	$0.8528 \pm 0.0441$	$0.8867 \pm 0.0337$	$0.8637 \pm 0.0317$
Expected information (EI)	$11.1105 \pm 10.1277$	$0.8904 \pm 0.0505$	$7.8137 \pm 6.8609$
Previous choice info (I-1)	$0.7180 \pm 0.1358$	$0.6169 \pm 0.1955$	$0.6854 \pm 0.1112$
Previous choice reward (R-1)	$0.7588 \pm 0.0303$	$0.7914 \pm 0.0350$	$0.7693 \pm 0.0234$
CN x ER	$0.8339 \pm 0.0982$	$0.8451 \pm 0.0446$	$0.8375 \pm 0.0679$
CN x EI	$0.9729 \pm 0.0595$	$0.7348 \pm 0.1256$	$0.8961 \pm 0.0577$
CN x I-1	$0.9328 \pm 0.0921$	$0.9516 \pm 0.0424$	$0.9389 \pm 0.0637$
CN x R-1	$0.8031 \pm 0.0440$	$0.8104 \pm 0.0303$	$0.8054 \pm 0.0313$
ER x EI	$1.0155 \pm 0.0494$	$0.9859 \pm 0.0530$	$1.0060 \pm 0.0375$
ER x I-1	$0.9935 \pm 0.0135$	$1.0092 \pm 0.0111$	$0.9986 \pm 0.0098$
ER x R-1	$1.0699 \pm 0.0803$	$1.0123 \pm 0.0946$	$1.0513 \pm 0.0622$
EI x I-1	$0.9987 \pm 0.0088$	$1.0003 \pm 0.0300$	$0.9992 \pm 0.0113$
EI x R-1	$0.9921 \pm 0.0097$	$0.9481 \pm 0.0508$	$0.9779 \pm 0.0176$
I-1 x R-1	$0.7766 \pm 0.0703$	$1.1050 \pm 0.3644$	$0.8825 \pm 0.1266$

3 Third, we ran these same statistics, but restricted only to the significant covariates from our  
4 original GLM. The average  $e\_score = 0.9682 \pm 0.0035$  (monkey L:  $e\_score = 0.9694 \pm 0.0039$ ;  
5 monkey R:  $e\_score = 0.9615 \pm 0.0087$ ), a very good fit between the elastic net results and the

original GLM, especially compared to the  $e\_score$  for all covariates, suggesting that the original GLM was a fair representation of the influence of the different covariates.

#### Additional Neural Results: Anticipation Epoch PCC Neuron Responses

In the main text, we reported that PCC neurons preferentially signaled expected information over expected reward. Significantly more neurons encoded expected information than expected reward ( $\chi^2 = 12.86$ ,  $p < 0.0005$ ), and significantly more neurons encoded the interaction of choice number and expected information than the interaction of choice number and expected reward ( $\chi^2 = 23.04$ ,  $p < 0.000005$ ). This preference was maintained until the last choice number in a trial (CN2: 23 (19%) EI, 10 (8%) ER; CN3: 24 (19%) EI, 9 (7%) ER; CN4: 21 (17%) EI, 4 (3%) ER; CN5: 19 (15%) EI, 6 (5%) ER; CN6: 16 (13%) ER; Fig. 2C). Of the cells significantly encoding the interaction of choice number and expected information, 19 (63%) were negatively encoding and 11 (37%) were positively encoding, with a mean regression weight ( $\beta_{CN \times EI}$ ) =  $-0.11 \pm 0.075$  (spikes/sec)/(choice\*bits<sub>info</sub>) (Student's t-test against  $\beta_{CN \times EI} = 0$ ,  $t(29) = -1.52$ ,  $p > 0.1$ ). The mean effect of the interaction on firing rates across the whole population was significant and negative, as 71 (57%) cells showed a negative influence and 53 (43%) a positive influence, with a mean regression weight ( $\beta_{CN \times EI}$ ) =  $-0.056 \pm 0.025$  (spikes/sec)/(choice\*bits<sub>info</sub>) (Student's t-test,  $t(123) = -2.24$ ,  $p < 0.05$ ). As the trial progressed, PCC neurons generally showed lower firing rates for more informative expectations, matching the decrease in subject response times and in choice variability for larger information expectations. The full results of our GLM are reported in table S10.

Table S10: Full numbers of cells encoding all covariates for anticipation epoch, all choices inclusive (including diverge choices).

Covariate	Monkey L (84 neurons)	Monkey R (40 neurons)
Choice number (CN)	15	2
Expected reward (ER)	4	0
Expected information (EI)	18	3
Previous choice info (I-1)	13	3
Previous choice reward (R-1)	17	4
CN x ER	4	0
CN x EI	23	7
CN x I-1	8	3
CN x R-1	11	2
ER x EI	11	2
ER x I-1	15	1
ER x R-1	0	0
EI x I-1	22	5
EI x R-1	15	0
I-1 x R-1	9	2

#### Additional Behavioral Entropy Encoding Results by Monkey

In the anticipation epoch, 51 neurons in monkey L ( $37 \beta's > 0$ ,  $14 \beta's \leq 0$ ; mean  $\beta_{BE} = 0.0077 \pm 0.0045$ ,  $t(50) = 1.70$ ,  $p = 0.095$ ; mean across all 84 neurons  $\beta_{BE} = 0.0061 \pm 0.0028$ ,  $t(83) = 2.15$ ,  $p < 0.05$ ) and 19 neurons in monkey R ( $14 \beta's > 0$ ,  $5 \beta's \leq 0$ ; mean  $\beta_{BE} = 0.0062 \pm 0.0033$ ,

$t(18) = 1.86$ ,  $p = 0.078$ ; mean across all 40 neurons  $\beta_{BE} = 0.0051 \pm 0.0014$ ,  $t(39) = 2.45$ ,  $p < 0.05$ ) significantly predicted behavioral entropy.

#### Additional Encoding Results of Time of Receipt of Last Information

In the main text, we reported that 98 of 124 neurons distinguished before and after receiving the last information. This fraction was similar across the two monkeys: monkey L: 67 neurons (80%); monkey R: 31 neurons (78%).

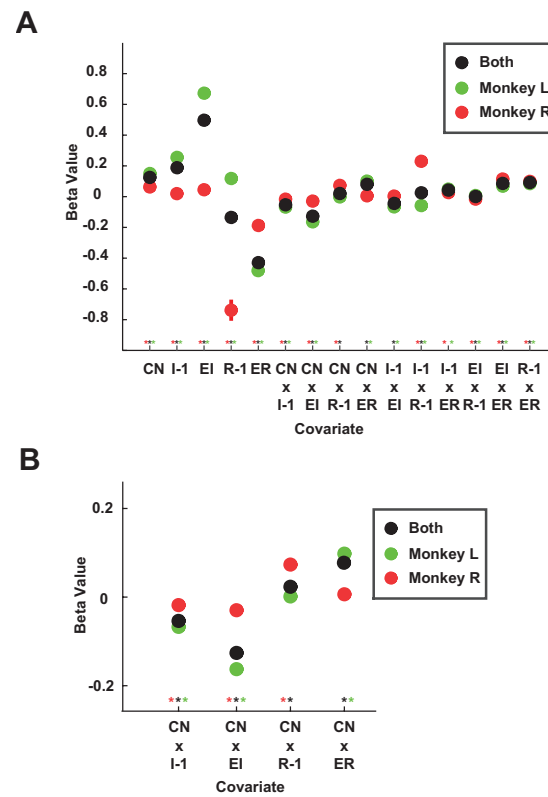
#### Additional Behavioral Entropy Encoding Results Before and After Receiving Last Information

Each trial's binned spike counts were regressed against the time in the trial (TT), time of last informative outcome (TL), and their interaction (TTxTL). Of the cells significantly encoding the different states, 21 (21%) fired less and 77 (79%) fired more after the animal received all information compared to before, with a mean regression coefficient  $\beta_{TTxTL} = 0.017 \pm 0.0030$  spikes/information period (Student's t-test against  $\beta_{TTxTL} = 0$ ,  $t(97) = 5.52$ ,  $p < 5 \times 10^{-7}$ ). Across the entire population, 30 (24%) fired more and 94 (76%) fired less before compared to after, with a mean regression coefficient  $\beta_{TTxTL} = 0.014 \pm 0.0025$  spikes/information period (Student's t-test,  $t(123) = 5.52$ ,  $p < 5 \times 10^{-7}$ ).

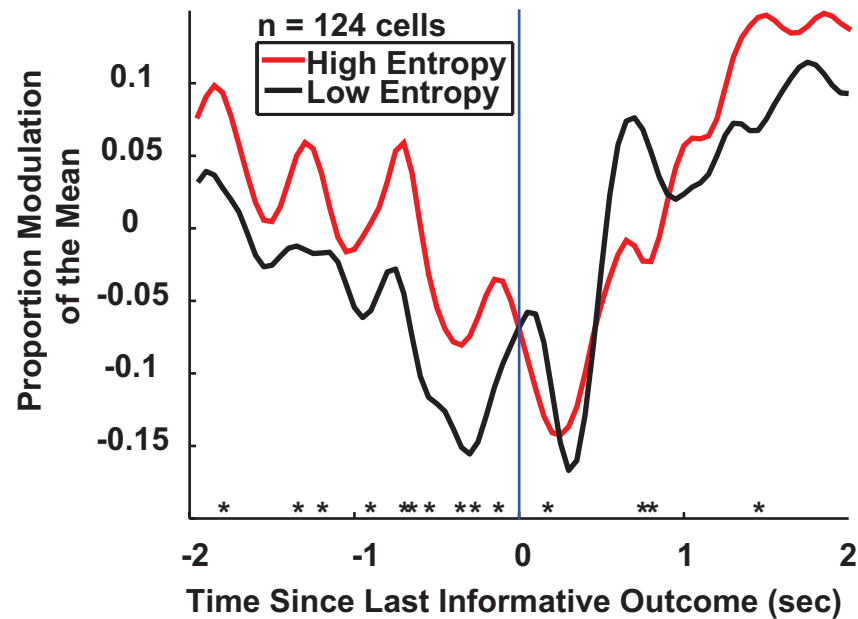
Higher firing in individual neurons tended to predict behavioral entropy both before and after receiving the last bit of information. We next examined the size of the modulation of this choice-by-choice signaling. In our population, 74 (60%) cells predicted behavioral entropy before receiving the last informative outcome (all cells: mean  $\beta_{BE\_before} = 0.0016 \pm 0.00064$  bits/spike,

Student's t-test:  $t(123) = 2.43$ ,  $p < 0.05$ ; significant cells: mean  $\beta_{BE\_before} = 0.0018 \pm 0.00096$  bits/spike, Student's t-test:  $t(73) = 1.83$ ,  $p = 0.071$ ) whereas only 24 (19%) predicted behavioral entropy after (all cells: mean  $\beta_{BE\_after} = 0.0025 \pm 0.0012$  bits/spike, Student's t-test:  $t(123) = 2.05$ ,  $p < 0.05$ ; significant cells: mean  $\beta_{BE\_after} = 0.0044 \pm 0.0022$  bits/spike, Student's t-test:  $t(23) = 2.27$ ,  $p < 0.05$ ; Fig. S3). Though the mean normalized population response failed to differentiate high from low average behavioral entropy trials after receipt of the last information as reported in the text, the choice-by-choice analysis of behavioral entropy revealed consistent and significant positive modulation both before and after. In sum, averaging the modulation of neuronal firing rates showed that PCC neurons showed higher firing for high behavioral entropy both before and after, albeit with many fewer significant cells after.

## Supplemental Figures



**Figure S1.** Coefficient values from response time and behavioral entropy regressions. **A.** Coefficient values from regression of response time against listed covariates. **B.** Coefficient values from regression of behavioral entropy against listed covariates. CN = choice number; EI = expected information; I – 1 = previous choice information outcome; ER = expected reward; R – 1 = previous choice reward outcome. Error bars either occluded by points or  $\pm 1$  s.e.m. \* =  $p < 0.05$ , Bonferroni corrected.



**Figure S2.** The multiplexing of behavioral entropy and time of receipt of last information in the PCC population. The population multiplexed information boundary signaling with behavioral entropy encoding, with higher firing rates for more variable choices before the last informative outcome but not after.  $n = 124$  cells;  $* = p < 0.05$ ; blue line = time of last informative outcome.

## Supplement References

- Boctor, Faye F., Gilbert Laporte, and Jacques Renaud. 2003. 'Heuristics for the traveling purchaser problem', *Computers & Operations Research*, 30: 491-504.
- Ramesh, T. 1981. 'Traveling purchaser problem', *Opsearch*, 18: 78-91.
- Zingale, Carolina M., and Eileen Kowler. 1987. 'Planning sequences of saccades', *Vision Res*, 27: 1327-41.