Neuronal activity in the posterior cingulate cortex signals environmental information and predicts behavioral variability during trapline foraging

1 **Title: Neuronal activity in the posterior cingulate cortex signals environmental**

2 **information and predicts behavioral variability during trapline foraging**

3 Abbreviated Title: Posterior cingulate signals behavioral variability

4 Authors: David L Barack[1*] and Michael L Platt[2]

5 [1]Departments of Philosophy, Neuroscience, and Economics, Columbia University, New York,

6 NY 10027, USA.

7 [2]Departments of Neuroscience, Psychology, and Marketing, University of Pennsylvania,

8 Philadelphia, PA 19104, USA.

9 * corresponding author: dbarack@gmail.com

10 Number of pages: 38

11 Number of figures: 3

12 Number of tables: 1

13 Number of words in abstract: 167

14 Number of words in significance statement: 120 words

15 Number of words in introduction: 949

16 Number of words in discussion: 1220

25

26 Abstract

27 Animals engage in routine behavior in order to efficiently navigate their environments. This

28 routine behavior may be influenced by the state of the environment, such as the location and size

29 of rewards. The neural circuits tracking environmental information and how that information

30 impacts decisions to deviate from routines remains unexplored. To investigate the representation

31 of environmental information during routine foraging, we recorded the activity of single neurons

32 in posterior cingulate cortex (PCC) in two male monkeys searching through an array of targets in

33 which the location of rewards was unknown. Outside the laboratory, people and animals solve

34 such traveling salesman problems by following routine traplines that connect nearest-neighbor

35 locations. In our task, monkeys also deployed traplining routines, but as the environment became

36 better known, they deviate from them despite the reduction in foraging efficiency. While

37 foraging, PCC neurons tracked environmental information but not reward and predicted

38 variability in the pattern of choices. Together, these findings suggest that PCC may mediate the

39 influence of information on variability in choice behavior.

40

41 Significance statement

42   Many animals seek information to better guide their decisions and update behavioral routines. In

43   our study, subjects visually searched through a set of targets on every trial to gather two rewards.

44   Greater amounts of information about the distribution of rewards predicted less variability in

45   choice patterns, whereas smaller amounts predicted greater variability. We recorded from the

46   posterior cingulate cortex, an area implicated in the coding of reward and uncertainty, and

47   discovered that these neurons signaled the expected information about the distribution of rewards

48   instead of signaling expected rewards. The activity in these cells also predicted the amount of

49   variability in choice behavior. These findings suggest that the posterior cingulate helps direct the

50   search for information in order to augment routines.

51

52   **Introduction**

53   Imagine you are at a horse race, and there are six horses, with Local Field Potential the

54   underdog, facing 100:1 odds against. When LFP wins, a one dollar bet will pay out $100. But in

55   addition to the reward received from this bet, learning that out of the six horses, LFP is the

56   winner reduces your uncertainty about the outcome. Hence, LFP crossing the finish line first

57   yields both reward and information.

58   Similar problems are often faced by organisms in their environment. Animals are adept at

59   learning not only the sizes of rewards but also their locations, timing, or other properties. For

60   example, hummingbirds will adapt their nectar foraging in response to unexpected changes in

61   reward timing (Garrison and Gass 1999). Similarly, monkeys will adapt their foraging routines

62   upon receiving information that a highly valued resource has become available (Menzel 1991).

63   In general, animals can make better decisions by tracking such reward information. Perhaps once

64 a reward has been received, it no longer pays to wait for more because the resource is exhausted

65 or the time between rewards is too great (McNamara 1982), as occurs for some foraging animals.

66 Or, perhaps receiving a reward also resolves any remaining uncertainty about an environment

67 (Stephens and Krebs 1986). Keeping track of reward information independent of reward size

68 thus serves as an important input into animals' decision processes.

69   We designed an experiment to probe this oft-neglected informational aspect of reward-

70 based decision making. Our experiment is based on the behavior of animals that exploit

71 renewable resources by following an efficient foraging path, a strategy known as traplining

72 (Freeman 1968, Berger-Tal and Bar-David 2015). Trapline foraging has a number of benefits,

73 including reducing the variance of a harvest and thereby attenuating risk (Possingham 1989),

74 efficiently capitalizing on periodically renewing resources (Possingham 1989, Bell 1990, Ohashi,

75 Leslie et al. 2008), and helping adapt to changes in competition (Ohashi, Leslie et al. 2013).

76 Many animals trapline, including bats (Racey and Swift 1985), bees (Manning 1956, Janzen

77 1971), butterflies (Boggs, Smiley et al. 1981), hummingbirds (Gill 1988), and an array of

78 primates including rhesus macaques (Menzel 1973), baboons (Noser and Byrne 2010), vervet

79 monkeys (Cramer and Gallistel 1997), and humans (Hui, Fader et al. 2009). Wild primates

80 foraging for fruit (Menzel 1973, Noser and Byrne 2010), captive primates searching for hidden

81 foods (Gallistel and Cramer 1996, Desrochers, Jin et al. 2010), and humans moving through

82 simulated (MacGregor and Chu 2011) and real (Hui, Fader et al. 2009) environments all use

83 traplining to minimize total distance traveled and thereby maximize resource intake rates.

84   Though many primates trapline, information about the state of the environment, such as

85 weather(Janmaat, Byrne et al. 2006), the availability of new foods(Menzel 1991), or possible

4

86 feeding locations(Hemmi and Menzel 1995, Menzel 1996), can influence choices made while

87 foraging. Such detours result in longer search distances and more variable choices(Hui, Fader et

88 al. 2009, Noser and Byrne 2010) but allow animals to identify new resources(Menzel 1991) and

89 engage in novel behaviors(Noser and Byrne 2010). These benefits are consistent with computer

90 simulations that show traplining with variation in routes yields better long-term returns than

91 traplining without variation by uncovering new resources or more efficient routes (Ohashi and

92 Thomson 2005). In this way, environmental information may improve foraging efficiency during

93 routine foraging over the longer term.

94   The neural mechanisms that track, update, and regulate the impact of environmental

95 information on decision making remain unknown. Neuroimaging studies have revealed that the

96 posterior cingulate cortex (PCC) is activated by a wide range of cognitive phenomena that

97 involve rewards, including prospection(Benoit, Gilbert et al. 2011), value representation(Kable

98 and Glimcher 2007, Clithero and Rangel 2014), strategy setting(Wan, Cheng et al. 2015), and

99 cognitive control(Leech, Kamourieh et al. 2011). Intracranial recordings in monkeys have found

100 that PCC neurons signal reinforcement learning strategies(Pearson, Hayden et al. 2009), respond

101 to novel stimuli during conditional visuomotor learning(Heilbronner and Platt 2013), represent

102 value(McCoy, Crowley et al. 2003), risk(McCoy and Platt 2005), and task switches (Hayden and

103 Platt 2010), and stimulation there can induce shifts away from a default option (Hayden, Nair et

104 al. 2008). Together, these observations suggest that the PCC mediates the effect of

105 environmental information on variability in routine behavior. However, no studies to date have

106 attempted to disentangle hedonic value from the informational value of rewards in PCC.

5

107   Previously, we reported that in our traplining task neurons in PCC increased their firing

108   rates during choices prior to decisions to diverge from the typical trapline, the most common

109   circular pattern of choices (Barack, Chang et al. 2017). We reported decisions to diverge from

110   typical traplines were driven by the salience of the pattern of total rewards during foraging. PCC

111   neuron firing rates predicted decisions to diverge from typical traplines and signaled the

112   interaction between foraging decision salience, reward, and time. Finally, these cells displayed a

113   large transient increase in activity prior to decisions to diverge that was especially marked in low

114   reward rate environments.

115   Here, we explore how information influenced decisions to deviate from traplines (circular

116   patterns of choices) and test the hypothesis that PCC tracks reward information. We recorded the

117   activity of PCC neurons in monkeys foraging through an array of targets in which environmental

118   information, operationalized as the pattern of rewards, was partially decorrelated from reward

119   size. Monkeys developed traplines in which they moved directly between nearest neighbor

120   targets in a circle. When they expected more information about the state of the environment,

121   their trapline foraging behavior was less variable. While foraging, PCC neurons tracked

122   environmental information but not reward and forecast variability in choice patterns. These

123   findings support our hypothesis that PCC mediates the use of information about the state of the

124   environment to regulate adherence to routines in behavior and cognition.

125

126   **Materials and Methods**

127   Task Analysis

6

128   Our experiment required monkeys to select each target in a set of six targets to harvest

129   the rewards. In every trial in our experiment, two fixed rewards (large and small) were assigned

130   to one of six locations in the environment in a pseudorandom fashion (Fig. 1B). Trials began

131   with monkeys fixating a central cross for a variable amount of time, ranging from 0.5 – 1 sec.

132   After fixation offset, six targets arranged in a circle appeared. The same locations were used

133   from trial to trial, and monkeys were free to select the targets in any order. To make a choice,

134   monkeys had to fixate their gaze on a target for 250 ms. In order to advance to the next trial,

135   monkeys had to choose each option, even after they'd already harvested the reward available on

136   that trial. Assuming the cost of making a saccade is a monotonic, positive-definite function of

137   distance between targets, the most efficient solution to our task is to minimize saccade times

138   between targets by searching in a circular pattern. This is referred to as a trapline, and sequences

139   of choices that are non-circular are deviations from traplines.

140   Uncertainty about the current trial's pattern of received rewards is reduced over the

141   course of the trial as the monkey proceeds through all of the targets. This reduction in

142   uncertainty is quantifiable by examining how many possible patterns of rewards are excluded

143   given the rewards revealed by previous choices. For a subset of patterns, the very same

144   information outcome can be delivered by distinct rewards, serving to partially decorrelate and

145   hence de-confound reward and information outcomes. Furthermore, expected reward and

146   expected information, defined as the average amount of information contained in the next

147   outcome given the pattern of rewards received so far, are also partially decorrelated (Table 1).

148   Given a set of six rewards (four zero, one small, and one large), there are 6! distinct

149   permutations. We made the simplifying assumption that monkeys did not distinguish between

7

150   the different zero rewards. This assumption reduces the number of distinct patterns from 720 to

151   30.

152

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

Pattern #	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 $EI_i = \frac{\sum_{P_{remaining}} I_i}{(|P_{remaining}|)}$
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

154

155   Each column in the table (except for the leftmost) contains six columns, each corresponding to a choice number during the trial.

156   Total number of choices on every trial = 6. Key: |·| = cardinality of · ; R = reward; n = choice number; i = choice number in trial.

9

157   Different patterns correspond to different series of received reward. The environmental

158 entropy H_E contained in receiving some reward (zero, small, or large) depends on the choice

159 number i in the sequence and the total number of possible sequences:

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

160 where $|\cdot|$ denotes cardinality, P is the set of possible permutations, and $\{P_i\}$ is the set of

161 remaining permutations after the ith choice. The amount of information contained in some reward

162 outcome is computed as the difference in the entropy, what has been learned about the current

163 trial's pattern of received reward by receiving the most recent outcome:

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

164 for the amount of environmental entropy H_E on the ith outcome. Expected information can then

165 be computed as the mean amount of information to be gained by making the next choice, the

166 weighted average over all possible next information outcomes given the pattern of rewards

167 received:

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

168 for expected information $E[\Delta H_E]$ for the ith choice, possible outcomes $[\Delta H_E]_i$ for the remaining

169 permutations $P_{remaining}$, and where $|\cdot|$ again denotes cardinality. As the animal proceeds through

170 the trial, the amount of expected information varies as a function of how many possible patterns

171 of returns have been eliminated so far. Expected reward ER is computed simply as the amount of

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

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

10

174 If the animal harvests all of the reward near the beginning of a trial, the expected reward will be

175 zero. However, if the animal does not harvest the rewards until the end of a trial, the expected

176 reward will increase across the duration of the trial.

177 The linear correlation coefficients between the different task variables (information,

178 expected information, reward, expected reward, etc.) can be computed empirically from the total

179 experienced reward outcomes and information outcomes, and from the total experienced reward

180 expectations and information expectations, derived from the trials the monkeys actually

181 experienced. For the anticipation epoch, this includes expected information and expected reward

182 ($R^2 = 0.1324$), expected information and previous choice information outcome ($R^2 = 0.0292$),

183 expected information and previous choice reward outcome ($R^2 = 0.1348$), expected reward and

184 previous choice information outcome ($R^2 = 2.6458 \times 10^{-06}$), expected reward and previous choice

185 reward outcomes ($R^2 = 0.1082$), previous choice information outcome and previous choice

186 reward outcome ($R^2 = 0.5971$). For the outcome epoch, this includes current choice information

187 outcome and current choice reward outcome ($R^2 = 0.3935$).

188

189 Experimental Design and Statistical Analysis: Behavior

190 In our experiments, two male rhesus macaques performed the task described above on

191 custom software using psychtoolbox {Brainard} and MATLAB (Mathworks, Natick, MA). All

192 statistical comparisons were performed using custom software in MATLAB. Significance was

193 Bonferroni corrected for multiple comparisons, and significance assessed at $p < 0.05$.

194 For our behavioral entropy measures, we again used the standard definition of entropy.

195 Step size was defined as the number of positions clockwise or counter-clockwise of the target

11

196 that the monkey chose in relation to the previous choice's target. For behavioral entropy, the

197 probability of a particular step size was computed for each step size by counting the number of

198 trials with that step size and dividing by the total number of trials. Action step sizes (from -2 to

199 3) and action step size probabilities (probability of taking an action of a given size) were

200 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

201 sizes were calculated on each choice by determining how many targets around clockwise

202 (positive) or counterclockwise (negative) the next choice was from the previous choice; already

203 selected targets were not included in this calculation. Step size probabilities were calculated by

204 holding fixed all of the covariates for a particular choice (information outcome from previous

205 choice, information expectation for next choice, reward outcome from previous choice, reward

206 expectation for next choice, and choice number) and counting the frequencies for each step size

207 and dividing by the total number of trials with that set of covariates. For each unique

208 combination of covariates (choice number, information outcome, information expectation,

209 reward outcome, and reward expectation), we computed the choicewise behavioral entropy (H_B)

210 for that combination as

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

211 for probability of each step size p_s. Finally, a multilinear regression correlated these behavioral

212 entropy scores with the covariates.

213   To analyze neural coding of expectations, we had to remove diverge choices, defined as

214 choices that diverged from the daily dominant pattern. Determining the daily dominant pattern

215 relied on assessing the similarity between pairs of trials, for every possible pair on a given day,

216 by computing the pair's Hamming score(Hamming 1950). To compute the similarity between

217 two trials, each trial's pattern of choices by target number is first coded as a digit string (e.g., 1 –

218 2 – 4 – 5 – 6 – 3). The Hamming distance $D_{i,i'}$ between two strings i, i' of equal length is equal to

219 the sum of the number of differences d between each entry in the string,

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

220 for strings x, y of length n. We computed $D_{i,i'}$ for every pair of trials, and then, for each unique

221 pattern of choices, computed the average Hamming distance $\overline{D}_{i,i'}$. The daily dominant pattern

222 corresponded to the pattern with the minimum $\overline{D}_{i,i'}$ and corresponded to a circular pattern for

223 both monkeys (see Barack et al. 2017). Since the daily dominant pattern was circular, we refer to

224 these as the monkeys' typical traplines.

225 Behavioral entropy was regressed against a number of variables and their interactions

226 using multilinear regression. Covariates included choice number in trial, expected information,

227 expected reward, reward outcome from the previous choice, information outcome from the

228 previous choice, and all 2-way interactions.

229

230 Experimental Design and Statistical Analysis: Neural

231 All neural data were analyzed on custom software in MATLAB. For all tests,

232 significance was Bonferroni corrected for multiple comparisons and assessed at $p < 0.05$.

233 Both monkeys were trained to orient to visual targets for liquid rewards before

234 undergoing surgical procedures to implant a head-restraint post (Crist Instruments) and receive a

235 craniotomy and recording chamber (Crist Instruments) permitting access to PCC. All surgeries

13

236   were done in accordance with Duke University IACUC approved protocols. The animals were on

237   isoflourane during surgery, received analgesics and prophylactic antibiotics after the surgery, and

238   were permitted a month to heal before any recordings were performed. After recovery, both

239   animals were trained on the trapliner task, followed by recordings from BA 23/31 in PCC. MR

240   images were used to locate the relevant anatomical areas and place electrodes. Acute recordings

241   were performed over many sessions. Approximately one fifth of the recordings were done using

242   FHC (FHC, Inc., Bangor, ME) single contact electrodes and four fifths performed using Plexon

243   (Plexon, Inc., Dallas, TX) 8-contact axial array U-probes in monkey L. No statistically

244   significant differences in the proportion of task relevant cells were detected between the

245   populations recorded with the two types of electrodes (χ^2, p > 0.5). All recordings in monkey R

246   were done using the U-probes. Recordings were performed using Plexon neural recording

247   systems. All single contact units were sorted online and then re-sorted offline with Plexon offline

248   sorter. All axial units were sorted offline with Plexon offline sorter.

249   Neural responses often show non-linearities (Dayan and Abbott 2001), which can be

250   captured using a generalized linear model (Aljadeff, Lansdell et al. 2016). We used a generalized

251   linear model (GLM) with a log-linear link function and Poisson distributed noise estimated from

252   the data to analyze our neuronal recordings, effectively modeling neuronal responses as an

253   exponential function of a linear combination of the input variables. We analyzed the neural data

254   in two epochs: a 500 ms anticipation epoch, encompassing a 250 ms pre-saccade period and the

255   250 ms hold fixation period to register a choice, as well as the 250 ms pre-saccade epoch itself.

256   Covariates included choice number in the trial, expected information, expected reward,

14

257 information outcome from the last choice, reward outcome from the last choice, and all 2-way

258 interactions.

259   In addition to this GLM, we confirmed our model fits in two ways for each neuron: first,

260 we plotted the residuals against the covariates, to check for higher-order structure, and second,

261 we used elastic net regression, to check that our significant covariates were selected by the best-

262 fit elastic net model (Zou and Hastie 2005). Plotting residuals revealed no significant higher-

263 order structure. Furthermore, elastic net regression confirmed our original GLM results. None of

264 the significant covariates identified by the original GLM received a coefficient of 0 from the

265 elastic net regression, and the sizes of the significant coefficients identified by the original GLM

266 were very close to the sizes of the coefficients computed by the elastic net regression.

267   Perievent time histograms (PETHs) were created by binning spikes time-locked to the

268 event of interest. For the anticipation epoch, PETHs were centered on the end of the choice

269 saccade and spikes binned in 10 ms bins. PETHs were smoothed with a Gaussian kernel with 0

270 mean and 5σ width where $\sigma = 20$ ms (i.e., two samples).

271   To analyze encoding of the information or reward boundary, a log-linear GLM regression

272 was run on vectors of binned spike counts time-locked to the start of the trial, with time in

273 window, time of last informative feedback (a binary covariate encoding whether or not the

274 current time bin was before or after the last informative feedback), and their interaction as

275 covariates. Neuronal spikes were sorted into 50 ms bins starting with trial onset and ending with

276 the time of the last outcome in a trial across the duration of the trial. This activity was regressed

277 against time in trial (coded by the bin number, starting with 1 and ending with the number of 50

278 ms bins for the trial), whether or not the last informative outcome had been received (coded as a

279 0, for before, or a 1, for after receiving the last outcome), and their 2-way interaction. For plots

280 depicting the boundary, PETHs were time-locked to the time of last informative feedback, spikes

281 from two seconds before to two seconds after sorted into 50 ms time bins, and smoothed with a

282 Gaussian kernel with 0 mean 5σ width where σ = 50 ms (i.e., one sample).

283   The failure to find representations of expected reward reported in the results was

284 confirmed by holding fixed expected information and choice number and directly comparing

285 observed firing rates for those combinations for which there was more than one reward level. For

286 choice number (CN) 2, this resulted in one pair of expected rewards; for CN3, one pair; for CN4,

287 one pair; for CN5, one triple; and for CN6, one triple. The observed firing rates for the pairs

288 were compared using Student's t-test and for the triples using ANOVA. A neuron that showed a

289 significant difference in those comparisons was included in the count for that choice number and

290 so could appear as significant for more than one choice (depicted in Fig. 2C).

291   Step sizes, step size probabilities, and choicewise behavioral entropies were linearly

292 regressed against the firing rates during the anticipation epoch, when actions were made. To

293 assess whether neurons showed differences in tonic firing rates for high compared to low

294 behavioral entropies, we fit Gaussians with constant offsets to the mean PETH firing rate and

295 examined the confidence interval for the constant offsets for each. The constant offset for high

296 and low behavioral entropy were considered significantly different if the 95% confidence

297 intervals derived from those fits did not overlap. To assess choicewise entropy encoding before

298 and after receipt of the last bit of information, we used a GLM with log-linear link function and

299 Poisson distributed noise to calculate the number of neurons that significantly encoded

300 choicewise behavioral entropy before the receipt of this information to compare to the number

16

301 after. Covariates included behavioral entropy, choice number in trial, a binary variable with $0 =$

302 before boundary and $1 =$ after boundary, and all 2-way interaction. For the population response,

303 we first separated trials by mean choicewise behavioral entropy across all choices. Next, the

304 normalized average population response for high average choicewise entropy trials was

305 compared to low average entropy during the two seconds before the receipt of the last

306 information using Student's t-test. Then we ran the same analysis on the normalized average

307 response during the two seconds following receipt of this information. We report the results of

308 these two analyses below.

309

310 **Results**

311 Trapline Foraging in a Simulated Environment

312   To explore the effects of information on deviation from routines, two monkeys (*M.*

313 *mulatta*) solved a simple traveling salesman problem. In this trapliner task, monkeys visually

314 foraged through a set of six targets arranged in a circle, only moving on to the next trial after

315 sampling every target (Fig. 1B). On each trial, two of the targets were baited, one with a large

316 reward and one with a small reward, with the identity of the baited targets varying from trial to

317 trial. While foraging, monkeys gathered both rewards, herein defined by the amount of juice

318 obtained, and information, herein defined as the reduction in uncertainty about the location of

319 remaining rewards.

320   By varying which target was rewarded from trial-to-trial, reward and information were

321 partially decorrelated. Reward was manipulated by varying the size of received rewards, with

322 one small, one large, and four zero rewards available on every trial. Information was

17

323 manipulated by varying the spatiotemporal pattern of rewarding targets. Different patterns

324 correspond to different series of received rewards. Based on the series of rewards received up to

325 a particular choice in the trial, some subset of the set of possible sequences remained, and the

326 size of this subset determines the remaining uncertainty for the current trial (see methods). Over

327 the course of a trial, the set of possible patterns shrinks, reducing uncertainty about the current

328 trial's pattern and determining the information gathered about the environment. These

329 differences in reward and information outcomes in turn determine reward and information

330 expectations. The expected reward for each target is the total remaining reward to harvest

331 divided by the number of remaining targets. In contrast, the expected information is the mean

332 amount of information to be gained by making the next choice. As the animal proceeds through

333 the trial, the amount of expected information varies as a function of how many possible patterns

334 of rewards have been eliminated so far. Distinct possible reward outcomes may offer the same

335 information, and so our task partially decorrelates information and reward (linear regression on

336 expected reward and expected information, $R^2 = 0.13$).

337 Information may influence the pattern of choices that monkeys made, resulting in trial-to-

338 trial changes in this pattern (behavioral data are the same as first reported in Barack, Chang et al.

339 2017). On a majority of trials, monkeys chose targets in the same order (the daily dominant

340 pattern, DDP; Monkey R: same DDP across all 14 sessions; Monkey L: same DDP across 24 of

341 30 sessions; across all sessions, 0.4665 ± 0.0317 proportion of trials diverged from the DDP; see

342 methods). More generally, monkeys usually chose the targets in a circle (proportion of trials in

343 average session with circular patterns of choices: Monkey L: 0.6134 ± 0.0418; Monkey R:

344 0.7113 ± 0.0208). However, they occasionally deviated from their circular routine. This

18

345   variability can be measured by finding the behavioral entropy over the distribution of choice

346   probabilities for targets. First, each choice during a trial was egocentrically coded by its step size,

347   the number of targets clockwise or counter-clockwise from the current trial's previously chosen

348   target (Fig. 1B). The probability of a particular step size was computed by counting the number

349   of trials with that step size and dividing by the total number of trials (see methods). Behavioral

350   entropy, the entropy computed over that distribution, significantly predicts adherence to both

351   typical traplines (DDP: logistic regression; significant ($p < 0.05$) β for 22 of 44 sessions) and

352   circular traplines (logistic regression; significant β ($p < 0.05$) for 33 of 44 sessions). We found

353   that the informativeness of outcomes influenced the variability in the monkeys' patterns of

354   choices as measured by behavioral entropy. Anticipation of more informative choice outcomes

355   significantly reduced the entropy of the monkeys' choices on average (Student's t-test across all

356   choices and sessions comparing behavioral entropy for less than average expected information to

357   greater than average; Both monkeys: $t(96,718) = -19.25$, $p < 1\times10^{-81}$; Monkey L: $t(69,274) = -$

358   3.24, $p < 0.005$; Monkey R: $t(27,442) = -23.99$, $p < 1\times10^{-125}$). To better assess the influence of

359   expected information on behavioral variability, we plotted by session and choice number the

360   mean behavioral entropy for zero expected information and compared it to the mean behavioral

361   entropy for non-zero expected information. Median behavioral entropy across sessions was

362   greater for choice numbers 4 and 5 than choice number 3 for no expected information ($p < 0.05$;

363   Fig. 1C, green boxes and points) and was greater for no expected information in comparison to

364   some expected information for choice number 5 ($p < 0.05$; Fig. 1C, choice number 5, red boxes

365   and points compared to green).

19

366   The presence of information or reward left to collect on a trial also drove choice

367   variability. While still harvesting information and reward about the current trial, monkeys'

368   choices were less variable, but afterward they became more variable in their choices (Student's t-

369   test on choice numbers (CN) 4 or 5; Both monkeys: t(48,358) = -125.98, p ~ 0; Monkey L:

370   t(34,636) = -96.32, p ~ 0; Monkey R: t(13,720) = -71.79, p ~ 0; results also significant for each

371   CN separately; Fig. 1C, right panel). Hence, monkeys deviated less while choices were still

372   informative or rewarding and more thereafter.

373

374   Environmental Information Signaling by Posterior Cingulate Neurons

375   We next probed PCC activity during the trapliner task to examine information and reward

376   signaling from 124 cells in two monkeys (Fig. 1A; monkey L = 84 neurons; monkey R = 40

377   neurons; neural data are the same as first reported in Barack, Chang et al. 2017). In order to

378   control for previously uncovered neural effects, all choices where monkeys diverged from

379   typical traplines were excluded from the analyses in this section (those neural findings are

380   reported in Barack, Chang et al. 2017).

381   During the anticipation epoch (500 ms encompassing a 250 ms pre-choice period and a

382   250 ms hold fixation period), neurons in PCC preferentially signaled information expectations

383   over reward expectations. An example cell (Fig. 2A) showed a phasic increase in firing rate

384   during the anticipation epoch when expected information was higher for the same choice number

385   in the trial (for example, choice number two (CN_2): Student's t-test, p < 0.0001, t(283) = -

386   4.3056; firing rate for 0.72 bits = 22.51 ± 1.46 spikes/sec, firing rate for 1.37 bits = 29.84 ± 0.95

387   spikes/sec). However, after controlling for choice number in the trial and expected information,

20

388   the same neuron did not differentiate between different amounts of expected reward (Student's t-

389   test, p > 0.9; firing rate for 0.2 expected reward = 22.23 ± 2.35 spikes/sec, firing rate for 0.4

390   expected reward = 22.76 ± 1.83 spikes/sec; Fig. 2A, second row from bottom, left panel). The

391   tuning curves for this same cell collapsed across all choice numbers for both expected

392   information and expected reward illustrate the strong sensitivity to larger amounts of information

393   (Fig. 2B).

394   In our population of 124 neurons, significantly more cells were tuned to information than

395   reward when controlling for choice number in trial. A generalized linear model (GLM)

396   regression revealed that during the anticipation epoch, 35 (28%) of 124 neurons (Monkey L: 25

397   (30%) of 84 neurons; Monkey R: 10 (25%) of 40 neurons) signaled the interaction of choice

398   number and expected information, but only 1 (~1%) of 124 neurons (Monkey L: 1 (~1%) of 84

399   neurons; Monkey R: 0 (0%) of 40 neurons) signaled the interaction of choice number and

400   expected reward (all results, p < 0.05, Bonferroni corrected; see methods for full list of

401   covariates in the GLM). A further test for signaling of expected reward compares the average

402   firing rates for different amounts of expected reward for the same choice number and expected

403   information. This test revealed that only about 10% of neurons signaled expected reward, except

404   on the last choice when all information had been received (Fig. 2C). In contrast, about 20% of

405   neurons signaled expected information (Fig. 2C). These proportions were not significantly

406   different when all circular traplines were included (expected information X choice number, $\chi^2 >$

407   0.24; expected reward X choice number, $\chi^2 > 0.17$).

408

409   PCC Neurons Index Response Variability

21

410   We have previously established that PCC neurons signal decisions to diverge from

411   typical traplines during our task (Barack, Chang et al. 2017). However, the extent to which these

412   cells track variability of responses during the task remains to be explored. We examined whether

413   PCC neurons index the degree of behavioral variability, operationalized as behavioral entropy

414   (BE; see methods; all trials, including divergences from typical traplines, are included in the

415   following analyses). During the pre-saccade epoch, behavioral entropy varied significantly with

416   firing rate for 48 (39%) of 124 neurons (linear regression of behavioral entropy against firing

417   rate, $p < 0.05$; Monkey L: 37 (44%) of 84 neurons, Monkey R: 11 (28%) of 40 neurons). An

418   example cell was more active for high entropy choices compared to low (linear regression, $\beta_{BE} =$

419   0.0229 ± 0.0026 bits$_{BE}$/spike, $p < 5 \times 10^{-18}$; Fig. 3A). Across the population, higher firing rates

420   predicted greater behavioral entropy (124 neurons; Student's t-test on mean normalized firing

421   rates during pre-saccade epoch, $t(123) = 2.7363$, $p < 0.01$; $\beta_{BE} > 0$ in 80 cells, $\beta_{BE} \leq 0$ in 44 cells;

422   mean $\beta_{BE} = 0.0025 \pm 0.0011$ bits$_{BE}$/spike, Student's t-test against h_0: mean $\beta_{BE} = 0$, $t(123) =$

423   2.3268, $p < 0.05$; Fig. 3B). In addition, in our population of 124 cells, 46 (37%) exhibited

424   significantly different ($p < 0.05$) tonic firing rates for high behavioral entropy compared to low

425   behavioral entropy choices during the anticipation epoch (Monkey L: 35/84 (42%); Monkey R:

426   11/40 (28%)).

427   We next investigated whether PCC neurons signaled the boundary defined by the receipt

428   of the last information or reward, when the pattern of rewards on a given trial becomes fully

429   resolved. Note that this can occur before the last reward is delivered if the last reward is received

430   on the last choice in a trial. A regression of each trial's binned spike counts against the time in

431   the trial and the time of last informative outcome revealed that 84 (68%) of 124 neurons

22

432 differentiated these two states (GLM, effect of interaction, p < 0.05; see methods; monkey L: 61

433 of 84 neurons, 73%; monkey R: 23 of 40 neurons, 58%). During a four second epoch centered on

434 the time of the last informative choice outcome, an example cell fired less before that outcome

435 than after (Student's t-test, $p < 1x10^{-56}$; Fig. 3C). The population of cells also fire more after this

436 boundary (Student's t-test, p < 0.005; Fig. 3D).

437 Finally, behavioral entropy signals and boundary signals were combined in the PCC

438 population. While the time of last information can be partly disambiguated from time of last

439 reward, this occurs only on the last choice when a single target remains. Since behavioral

440 entropy is a measure of response variability, it requires more than one target, which is not

441 available on the last choice. As a result, combined signals of behavioral entropy and the

442 boundary could reflect the end of either information gathering or reward harvesting.

443 Significantly fewer cells (χ^2, $p < 1x10^{-10}$) predicted behavioral entropy after receiving all

444 information or reward (24 (19%) of 124 neurons) than before (74 (60%) neurons). PCC

445 population responses on choices with high behavioral entropy compared to low entropy revealed

446 significant differences before receipt of the last informative or rewarding outcome (Student's t-

447 test, $p < 1x10^{-4}$) but not after (Student's t-test, p > 0.5), with greater modulation for high entropy

448 compared to low.

449

450 **Discussion**

451 In this study, we show that environmental information influences responses during

452 routine behavior and that firing rates of PCC neurons carry this information and predict

453 behavioral variability. Despite the fact that in our task monkeys could not utilize environmental

23

454   information to increase their chance of reward, the receipt of environmental information and the

455   exhaustion of uncertainty impacted behavioral routines. Monkeys' responses were less variable

456   when there was more information to be gathered, but became more variable once the

457   environment became fully known. This pattern of variable responses after resolving all

458   environmental uncertainty departs from the reward rate maximizing strategy of selecting targets

459   in a circle to minimize saccade lengths. While monkeys traplined, neurons in PCC robustly

460   signaled information expectations but not reward expectations and predicted the variability in the

461   patterns of choices. Finally, PCC neurons differentiate the degree of behavioral variability before

462   all information or reward was received about the pattern of rewards compared to after, with an

463   increase in activity following receipt of the last informative outcome and concomitant decreases

464   in forecasting behavioral variability. In sum, our experimental findings suggest that PCC tracks

465   the state of the environment in order to influence routine behavior.

466   Monkeys often chose targets in the same pattern, consistent with previous findings of

467   repetitive stereotyped foraging in wild primate groups (Noser and Byrne 2007). They also

468   generally moved in a circle, visiting the next nearest neighbor after the current target, likewise

469   consistent with previous findings in groups of wild foraging primates (Menzel 1973, Garber

470   1988, Janson 1998). These foraging choices almost always result in straight line routes (Janson

471   1998, Pochron 2001, Cunningham and Janson 2007, Valero and Byrne 2007) or a series of

472   straight lines (Di Fiore and Suarez 2007, Noser and Byrne 2007). Experiments on captive

473   primates have also observed nearest neighbor or near optimal path finding (Menzel 1973,

474   MacDonald and Wilkie 1990, Gallistel and Cramer 1996, Cramer and Gallistel 1997). Our

475   monkeys' choices are also consistent with human behavior on traveling salesman problems,

476   wherein next nearest neighbor paths are usually chosen for low numbers of points (Hirtle and

477   Gärling 1992, MacGregor and Ormerod 1996, MacGregor and Chu 2011).

478   The PCC, a posterior midline cortical region with extensive cortico-cortical connectivity

479   (Heilbronner and Haber 2014) and elevated resting state and off-task metabolic activity

480   (Buckner, Andrews-Hanna et al. 2008), is at the heart of the default mode network (DMN)

481   (Buckner, Andrews-Hanna et al. 2008). The DMN is a cortex-spanning network implicated in

482   exploratory cognition including imagination (Schacter, Addis et al. 2012), creativity (Kühn,

483   Ritter et al. 2014), and narration (Wise and Braga 2014). Though implicated in a range of

484   cognitive functions, activity in PCC may be unified by a set of computations related to

485   harvesting information from the environment to regulate behavior. Signals in PCC that carry

486   information about environmental decision variables such as value (McCoy, Crowley et al. 2003),

487   risk (McCoy and Platt 2005), and decision salience (Heilbronner, Hayden et al. 2011) may in fact

488   reflect the tracking of information returns from the immediate environment. For example, in a

489   two alternative forced choice task, neurons in PCC preferentially signaled the resolution of a

490   risky choice with a variable reward over the value of choosing a safe choice with a guaranteed

491   reward (McCoy and Platt 2005). Such signals may reflect the information associated with the

492   resolution of uncertainty regarding the risky option. PCC neurons also signal reward-based

493   exploration (Pearson, Hayden et al. 2009) and microstimulation in PCC can shift monkeys from

494   a preferred option to one they rarely choose (Hayden, Nair et al. 2008). Both of these functions

495   may reflect signaling of environmental information as well; for example, the signaling of

496   exploratory choices may reflect the information from an increase in the number of recent sources

497   of reward (Pearson, Hayden et al. 2009). Evidence from neuroimaging studies in humans

25

498   similarly reveals PCC activation in a wide range of cognitive processes related to adaptive

499   cognition, including imagination (Benoit, Gilbert et al. 2011), decision making (Kable and

500   Glimcher 2007), and creativity (Beaty, Benedek et al. 2015).

501   Uncovering the neural circuits that underlie variability in foraging behavior may provide

502   insight into more complex cognitive functions. A fundamental feature of what we call

503   prospective cognition, thoughts about times, places, and objects beyond the here and now,

504   involves consideration of different ways the world might turn out. Various types of prospective

505   cognition, including imagination, exploration and creativity, impose a tradeoff between engaging

506   well-rehearsed routines and deviating in search of new, potentially better solutions (Gottlieb,

507   Oudeyer et al. 2013, Andrews-Hanna, Smallwood et al. 2014, Beaty, Benedek et al. 2015). For

508   example, creativity involves diverging from usual patterns of thought, such as occurs in

509   generating ideas (Benedek, Jauk et al. 2014) or crafting novel concepts (Barron 1955, Guilford

510   1959). During creative episodes the PCC shows increased activity during idea generation

511   (Benedek, Jauk et al. 2014) and higher connectivity with control networks during idea evaluation

512   (Beaty, Benedek et al. 2015), perhaps reflecting imagined, anticipated, or predicted variation in

513   the environment. Exploration similarly involves diverging from the familiar, such as to locate

514   novel resources (Ohashi and Thomson 2005) or discover shorter paths (Sutton and Barto 1998)

515   between known locations. Such prospective cognition requires diverging from routine thought,

516   and the identification of the neural circuits that mediate deviations from motor routines may

517   provide initial insight into the computations and mechanisms of prospective cognition. The

518   discovery that the PCC preferentially signals the state of the environment and predicts behavioral

519   variability relative to that state is a first step towards understanding these circuits.

26

520   The reinforcement learning literature is replete with models where exploration is driven

521 by the search for information (Schmidhuber 1991, Johnson, Varberg et al. 2012). These models

522 hypothesize that agents should take actions that maximize the information gleaned from the

523 environment, either by reducing uncertainty about the size of offered rewards (Schmidhuber

524 1991), the location of rewards in the environment (Johnson, Varberg et al. 2012), or otherwise

525 maximizing information for subsequent decisions. Furthermore, evidence from initial studies

526 studying information-based exploration shows that humans are avid information-seekers (Miller

527 1983, Fu and Pirolli 2007) and regulate attentional and valuational computations on the basis of

528 information (Manohar and Husain 2013, Blanchard, Hayden et al. 2015). In our task, the PCC

529 represented environmental information and tracked when learning about the environment was

530 complete, two variables central to information-based exploration. In particular, the dramatic

531 change in firing rates associated with the end of information gathering suggests that PCC

532 represents the information state of the environment and possibly also the rate of information

533 intake, a central variable in information foraging models (Pirolli and Card 1999, Fu and Pirolli

534 2007, Pirolli 2007). PCC appears poised to regulate exploration for information.

535   In sum, harvested information and response variability were both signaled by PCC

536 neurons, suggesting a central role for PCC in how information drives exploration and possibly

537 prospective cognition. Monkeys were sensitive to the amount of uncertainty remaining in the

538 environment, with more reliable patterns of choices while information remained and more

539 variable patterns after environmental uncertainty had been resolved and all rewards collected.

540 PCC neurons preferentially tracked this information and predicted the variability in monkeys'

541 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.

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, J., B. J. Lansdell, A. L. Fairhall and D. Kleinfeld (2016). "Analysis of neuronal spike trains, deconstructed." Neuron **91**(2): 221-259.

Andrews-Hanna, J. R., J. Smallwood and R. N. Spreng (2014). "The default network and self-generated thought: component processes, dynamic control, and clinical relevance." Annals of the New York Academy of Sciences **1316**(1): 29-52.

Barack, D. L., S. W. C. Chang and M. L. Platt (2017). "Posterior cingulate neurons dynamically signal decisions to disengage during foraging." Neuron **96**(2): 339-347.e335.

565 Barron, F. (1955). "The disposition toward originality." <u>The Journal of Abnormal and Social</u>
566 <u>Psychology</u> **51**(3): 478.
567 Beaty, R. E., M. Benedek, S. B. Kaufman and P. J. Silvia (2015). "Default and executive
568 network coupling supports creative idea production." <u>Scientific reports</u> **5**.
569 Bell, W. J. (1990). When to return to a resource patch. <u>Searching Behaviour</u>, Springer**:** 117-129.
570 Benedek, M., E. Jauk, A. Fink, K. Koschutnig, G. Reishofer, F. Ebner and A. C. Neubauer
571 (2014). "To create or to recall? Neural mechanisms underlying the generation of creative new
572 ideas." <u>NeuroImage</u> **88**: 125-133.
573 Benoit, R. G., S. J. Gilbert and P. W. Burgess (2011). "A Neural Mechanism Mediating the
574 Impact of Episodic Prospection on Farsighted Decisions." <u>The Journal of Neuroscience</u> **31**(18):
575 6771-6779.
576 Berger-Tal, O. and S. Bar-David (2015). "Recursive movement patterns: review and synthesis
577 across species." <u>Ecosphere</u> **6**(9): art149.
578 Blanchard, T. C., B. Y. Hayden and E. S. Bromberg-Martin (2015). "Orbitofrontal cortex uses
579 distinct codes for different choice attributes in decisions motivated by curiosity." <u>Neuron</u> **85**(3):
580 602-614.
581 Boggs, C. L., J. T. Smiley and L. E. Gilbert (1981). "Patterns of pollen exploitation by
582 Heliconius butterflies." <u>Oecologia</u> **48**(2): 284-289.
583 Buckner, R. L., J. R. Andrews-Hanna and D. L. Schacter (2008). "The brain's default network:
584 anatomy, function, and relevance to disease." <u>Ann N Y Acad Sci</u> **1124**: 1-38.
585 Clithero, J. A. and A. Rangel (2014). "Informatic parcellation of the network involved in the
586 computation of subjective value." <u>Social cognitive and affective neuroscience</u> **9**(9): 1289-1302.
587 Cramer, A. E. and C. R. Gallistel (1997). "Vervet monkeys as travelling salesmen." <u>Nature</u> **387**:
588 464.
589 Cunningham, E. and C. Janson (2007). "Integrating information about location and value of
590 resources by white-faced saki monkeys (Pithecia pithecia)." <u>Animal cognition</u> **10**(3): 293-304.
591 Dayan, P. and L. F. Abbott (2001). <u>Theoretical Neuroscience</u>. Cambridge, Massachusetts, MIT
592 Press.
593 Desrochers, T. M., D. Z. Jin, N. D. Goodman and A. M. Graybiel (2010). "Optimal habits can
594 develop spontaneously through sensitivity to local cost." <u>Proceedings of the National Academy</u>
595 <u>of Sciences</u> **107**(47): 20512-20517.
596 Di Fiore, A. and S. A. Suarez (2007). "Route-based travel and shared routes in sympatric spider
597 and woolly monkeys: cognitive and evolutionary implications." <u>Animal cognition</u> **10**(3): 317-
598 329.
599 Freeman, R. (1968). "Charles Darwin on the routes of male humble bees." <u>Bulletin of the British</u>
600 <u>Museum (Natural History) Historical Series</u> **3**(6): 177-189.
601 Fu, W.-T. and P. Pirolli (2007). "SNIF-ACT: A cognitive model of user navigation on the World
602 Wide Web." <u>Human–Computer Interaction</u> **22**(4): 355-412.
603 Gallistel, C. R. and A. E. Cramer (1996). "Computations on metric maps in mammals: getting
604 oriented and choosing a multi-destination route." <u>The Journal of Experimental Biology</u> **199**(1):
605 211-217.

Garber, P. A. (1988). "Foraging decisions during nectar feeding by tamarin monkeys (Saguinus mystax and Saguinus fuscicollis, Callitrichidae, Primates) in Amazonian Peru." Biotropica: 100-106.

Garrison, J. S. and C. L. Gass (1999). "Response of a traplining hummingbird to changes in nectar availability." Behavioral Ecology 10(6): 714-725.

Gill, F. B. (1988). "Trapline foraging by hermit hummingbirds: competition for an undefended, renewable resource." Ecology: 1933-1942.

Gottlieb, J., P.-Y. Oudeyer, M. Lopes and A. Baranes (2013). "Information-seeking, curiosity, and attention: computational and neural mechanisms." Trends in cognitive sciences 17(11): 585-593.

Guilford, J. P. (1959). "Traits of creativity." Creativity and its cultivation 10: 141-161.

Hamming, R. W. (1950). "Error detecting and error correcting codes." Bell System technical journal 29(2): 147-160.

Hayden, B. Y., A. C. Nair, A. N. McCoy and M. L. Platt (2008). "Posterior cingulate cortex mediates outcome-contingent allocation of behavior." Neuron 60(1): 19-25.

Hayden, B. Y. and M. L. Platt (2010). "Neurons in anterior cingulate cortex multiplex information about reward and action." J Neurosci 30(9): 3339-3346.

Heilbronner, S. R. and S. N. Haber (2014). "Frontal Cortical and Subcortical Projections Provide a Basis for Segmenting the Cingulum Bundle: Implications for Neuroimaging and Psychiatric Disorders." The Journal of Neuroscience 34(30): 10041-10054.

Heilbronner, S. R., B. Y. Hayden and M. L. Platt (2011). "Decision salience signals in posterior cingulate cortex." Frontiers in neuroscience 5.

Heilbronner, Sarah R. and Michael L. Platt (2013). "Causal Evidence of Performance Monitoring by Neurons in Posterior Cingulate Cortex during Learning." Neuron 80(6): 1384-1391.

Hemmi, J. M. and C. R. Menzel (1995). "Foraging strategies of long-tailed macaques, Macaca fascicularis: directional extrapolation." Animal behaviour 49(2): 457-464.

Hirtle, S. C. and T. Gärling (1992). "Heuristic rules for sequential spatial decisions." Geoforum 23(2): 227-238.

Hui, S. K., P. S. Fader and E. T. Bradlow (2009). "Research note-The traveling salesman goes shopping: The systematic deviations of grocery paths from TSP optimality." Marketing Science 28(3): 566-572.

Janmaat, K. R. L., R. W. Byrne and K. Zuberbühler (2006). "Evidence for a spatial memory of fruiting states of rainforest trees in wild mangabeys." Animal Behaviour 72(4): 797-807.

Janson, C. H. (1998). "Experimental evidence for spatial memory in foraging wild capuchin monkeys, Cebus apella." Animal Behaviour 55(5): 1229-1243.

Janzen, D. (1971). "Euglossine bees as long-distance pollinators of tropical plants." Science 171(3967): 203-205.

Johnson, A., Z. Varberg, J. Benhardus, A. Maahs and P. Schrater (2012). "The hippocampus and exploration: dynamically evolving behavior and neural representations." Frontiers in human neuroscience 6.

Kable, J. W. and P. W. Glimcher (2007). "The neural correlates of subjective value during intertemporal choice." Nat Neurosci 10(12): 1625-1633.

648 Kühn, S., S. M. Ritter, B. C. N. Müller, R. B. Baaren, M. Brass and A. Dijksterhuis (2014). "The
649 importance of the default mode network in creativity—a structural MRI study." The Journal of
650 Creative Behavior **48**(2): 152-163.

651 Leech, R., S. Kamourieh, C. F. Beckmann and D. J. Sharp (2011). "Fractionating the default
652 mode network: distinct contributions of the ventral and dorsal posterior cingulate cortex to
653 cognitive control." The Journal of Neuroscience **31**(9): 3217-3224.

654 MacDonald, S. E. and D. M. Wilkie (1990). "Yellow-nosed monkeys'(Cercopithecus ascanius
655 whitesidei) spatial memory in a simulated foraging environment." Journal of Comparative
656 Psychology **104**(4): 382.

657 MacGregor, J. N. and Y. Chu (2011). "Human performance on the traveling salesman and related
658 problems: A review." The Journal of Problem Solving **3**(2): 2.

659 MacGregor, J. N. and T. Ormerod (1996). "Human performance on the traveling salesman
660 problem." Perception & Psychophysics **58**(4): 527-539.

661 Manning, A. (1956). "Some aspects of the foraging behaviour of bumble-bees." Behaviour **9**(1):
662 164-200.

663 Manohar, S. G. and M. Husain (2013). "Attention as foraging for information and value."
664 Frontiers in human neuroscience **7**: 711.

665 McCoy, A. N., J. C. Crowley, G. Haghighian, H. L. Dean and M. L. Platt (2003). "Saccade
666 reward signals in posterior cingulate cortex." Neuron **40**(5): 1031-1040.

667 McCoy, A. N. and M. L. Platt (2005). "Risk-sensitive neurons in macaque posterior cingulate
668 cortex." Nature Neuroscience **8**(9): 1220-1227.

669 McNamara, J. (1982). "Optimal patch use in a stochastic environment." Theoretical Population
670 Biology **21**(2): 269-288.

671 Menzel, C. (1996). "Structure-guided foraging in long-tailed macaques." American Journal of
672 Primatology **38**(2): 117-132.

673 Menzel, C. R. (1991). "Cognitive aspects of foraging in Japanese monkeys." Animal Behaviour
674 **41**(3): 397-402.

675 Menzel, E. W. (1973). "Chimpanzee spatial memory organization." Science **182**(4115): 943-945.

676 Miller, G. (1983). Informavores. The study of information: Interdisciplinary messages. F.
677 Machlup and U. Mansfield, Wiley-Interscience**:** 111-113.

678 Noser, R. and R. W. Byrne (2007). "Travel routes and planning of visits to out-of-sight resources
679 in wild chacma baboons, Papio ursinus." Animal Behaviour **73**(2): 257-266.

680 Noser, R. and R. W. Byrne (2010). "How do wild baboons (Papio ursinus) plan their routes?
681 Travel among multiple high-quality food sources with inter-group competition." Animal
682 cognition **13**(1): 145-155.

683 Ohashi, K., A. Leslie and J. D. Thomson (2008). "Trapline foraging by bumble bees: V. Effects
684 of experience and priority on competitive performance." Behavioral Ecology **19**(5): 936-948.

685 Ohashi, K., A. Leslie and J. D. Thomson (2013). "Trapline foraging by bumble bees: VII.
686 Adjustments for foraging success following competitor removal." Behavioral ecology **24**(3):
687 768-778.

688 Ohashi, K. and J. D. Thomson (2005). "Efficient harvesting of renewing resources." Behavioral
689 Ecology **16**(3): 592-605.

31

690 Pearson, J. M., B. Y. Hayden, S. Raghavachari and M. L. Platt (2009). "Neurons in posterior
691 cingulate cortex signal exploratory decisions in a dynamic multioption choice task." Curr Biol
692 **19**(18): 1532-1537.
693 Pirolli, P. and S. Card (1999). "Information foraging." Psychological review **106**(4): 643.
694 Pirolli, P. L. T. (2007). Information foraging theory: Adaptive interaction with information,
695 Oxford University Press.
696 Pochron, S. T. (2001). "Can concurrent speed and directness of travel indicate purposeful
697 encounter in the yellow baboons (Papio hamadryas cynocephalus) of Ruaha National Park,
698 Tanzania?" International Journal of Primatology **22**(5): 773-785.
699 Possingham, H. P. (1989). "The distribution and abundance of resources encountered by a
700 forager." American Naturalist: 42-60.
701 Racey, P. and S. Swift (1985). "Feeding ecology of Pipistrellus pipistrellus (Chiroptera:
702 Vespertilionidae) during pregnancy and lactation. I. Foraging behaviour." The Journal of Animal
703 Ecology: 205-215.
704 Schacter, Daniel L., Donna R. Addis, D. Hassabis, Victoria C. Martin, R. N. Spreng and Karl K.
705 Szpunar (2012). "The Future of Memory: Remembering, Imagining, and the Brain." Neuron
706 **76**(4): 677-694.
707 Schmidhuber, J. (1991). Curious model-building control systems. 1991 IEEE International Joint
708 Conference on Neural Networks., IEEE.
709 Stephens, D. W. and J. R. Krebs (1986). Foraging Theory. Princeton, NJ, Princeton University
710 Press.
711 Sutton, R. S. and A. G. Barto (1998). Reinforcement learning : an introduction. Cambridge,
712 Mass., MIT Press.
713 Valero, A. and R. W. Byrne (2007). "Spider monkey ranging patterns in Mexican subtropical
714 forest: do travel routes reflect planning?" Animal cognition **10**(3): 305-315.
715 Wan, X., K. Cheng and K. Tanaka (2015). "Neural encoding of opposing strategy values in
716 anterior and posterior cingulate cortex." Nature neuroscience **18**(5): 752-759.
717 Wise, R. J. S. and R. M. Braga (2014). "Default mode network: the seat of literary creativity?"
718 Trends in cognitive sciences **18**(3): 116-117.
719 Zou, H. and T. Hastie (2005). "Regularization and variable selection via the elastic net." Journal
720 of the Royal Statistical Society: Series B (Statistical Methodology) **67**(2): 301-320.

721

722 **Figures**

723
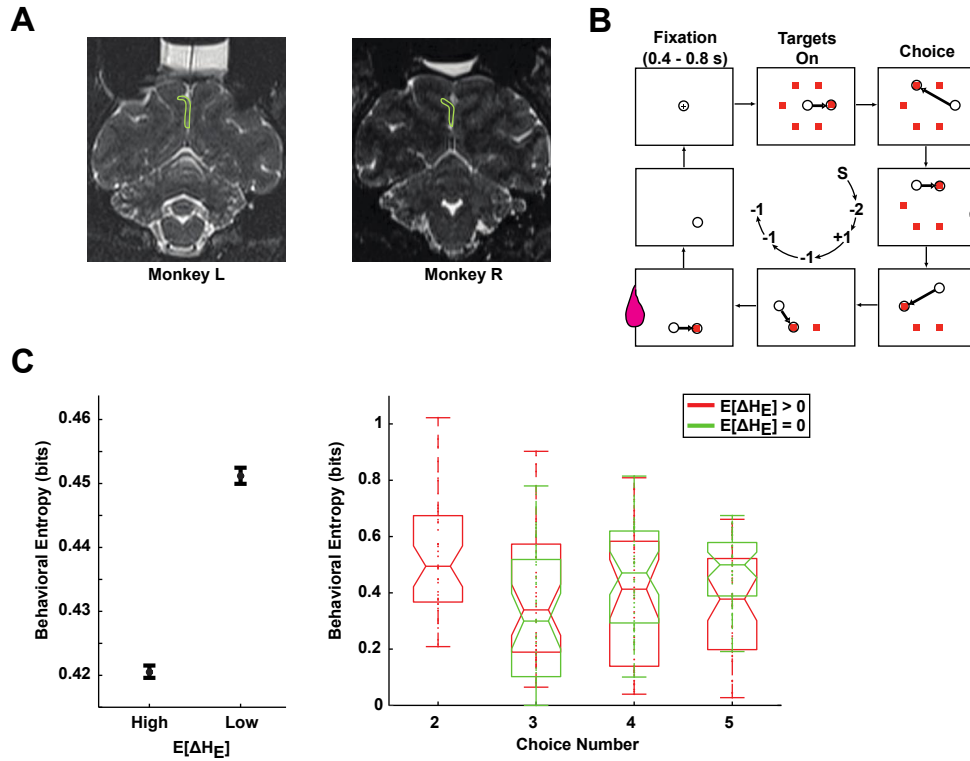
Fig. 1. Monkeys spontaneously trapline, efficiently choosing targets in a circle, when foraging in a circular array but deviate from these routines as the environment becomes better known. **A**. Recording location in the posterior cingulate cortex (PCC). Left: Monkey L. Right: Monkey R. **B**. Traplining 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

33

731   identity of the rewarded targets changed in a pseudorandom fashion from trial to trial. In order to

732   advance to the next trial, monkeys had to select every target. Open circle: simulated eye position;

733   dashed arrow: direction of impending saccade; dashed circle: impending saccade endpoint; small

734   juice drop: small reward; large juice drop: large reward. Central semi-circle: step size, the

735   clockwise or counter-clockwise distance between subsequently chosen targets; 'S' = start; '-2' =

736   two targets counter-clockwise; '-1' = one target counter-clockwise; '+1' = one target clockwise.

737   **C**. Left panel: mean ± s.e.m. behavioral entropy for high expected environmental information

738   (ΔH_E) choices ($E[\Delta H_E] > \text{mean}(E[\Delta H_E])$) compared to low expected information choices

739   ($E[\Delta H_E] \leq \text{mean}(E[\Delta H_E])$) across all sessions and choices; right panel: boxplot by choice number

740   (2-5) across sessions before receipt of last informative outcome ($E[\Delta H_E] > 0$; red points) and

741   after ($E[\Delta H_E] = 0$; green points). Top and bottom of box are interquartile (25% - 75%) range of

742   session means, and notch indicates 95% CI for median session. Non-overlapping notches

743   indicate significantly different medians at $\alpha = 0.05$. Each point is a session mean. Choice number

744   2 always possesses some expected information, hence no green box or points. n = 145,524

745   choices, 24,254 trials.

A

Cell L010414A1b

$E[\Delta H_E] = 0$

▬ 0	
▬ 1	
▬ 2	

CN6

$E[\Delta H_E] = 0$ $E[\Delta H_E] = 1$

| ▬ 0 |

| ▬ 0.5 |
| ▬ 1 |
| ▬ 1.5 |

CN5   CN5

$E[\Delta H_E] = 0$ $E[\Delta H_E] = 0.9183$ $E[\Delta H_E] = 1.585$

| ▬ 0 |

| ▬ 0.33 |
| ▬ 0.67 |

| ▬ 1 |

CN4   CN4   CN4

$E[\Delta H_E] = 0$ $E[\Delta H_E] = 0.81128$ $E[\Delta H_E] = 1.5$

| ▬ 0 |

| ▬ 0.25 |
| ▬ 0.5 |

| ▬ 0.75 |

CN3   CN3   CN3

$E[\Delta H_E] = 0.72193$ $E[\Delta H_E] = 1.371$

| ▬ 0.2 |
| ▬ 0.4 |

| ▬ 0.6 |

CN2   CN2

$E[\Delta H_E] = 1.2516$

| ▬ 0.5 |

Firing Rate (spikes/sec)
50
40
30
20
10

CN1

-0.25   0.25
Choice Time (sec)

Choice Number

Information (bits)

B

Firing Rate (spikes/sec)
35
30
25
20

Cell L010414A1b

0   0.72   0.91   1.25   1.5
  0.81   1   1.37 1.59
Expected Information (bits)

Firing Rate (spikes/sec)
35
30
25
20

Cell L010414A1b

0   0.25 0.4   0.6   0.75 1.5
  0.2 0.33 0.5 0.67   1   2
Expected Reward (a.u.)

C

Number of Cells (n = 124)
25
20
15
10
5

| ■ | Expected Info |
| ■ | Expected Reward |

2   3   4   5   6
Choice Number

35

746

747 **Fig. 2.** PCC neurons preferentially encode environmental information over reward. **A**. Firing rate

748 of sample neuron encoding expected information but not expected reward across all choice

749 numbers (CN) one through six, plotted separately by expected information ($E[\Delta H_E]$). Legends

750 indicate expected reward(s) for each plot. Blue line = end of saccade. **B**. Tuning curves for

751 expected information (top panel) and expected reward (bottom panel), collapsed across choice

752 numbers for better visibility, for the cell plotted in **A**. This example cell showed elevated firing

753 rates for higher amounts of information. Note that the elevated firing rates for certain amounts of

754 expected reward correspond to choices with high expected information with only a single level

755 of expected reward. **C**. Number of cells encoding expected reward by choice (red) for constant

756 expected information, and number of cells encoding expected information by choice (green).

757 Neurons were included in the expected reward counts if Student t-tests (CN2 – 4) or ANOVA

758 (CN5 – 6) indicated a significant difference in firing rates ($p < 0.05$, uncorrected for multiple

759 comparisons to allow the weakest criteria for inclusion, and the same cell could appear for more
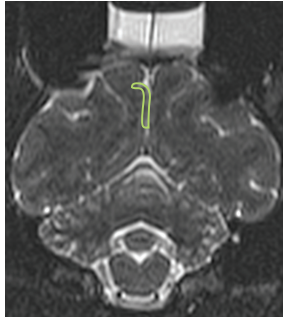
760 than one choice number).

761

762
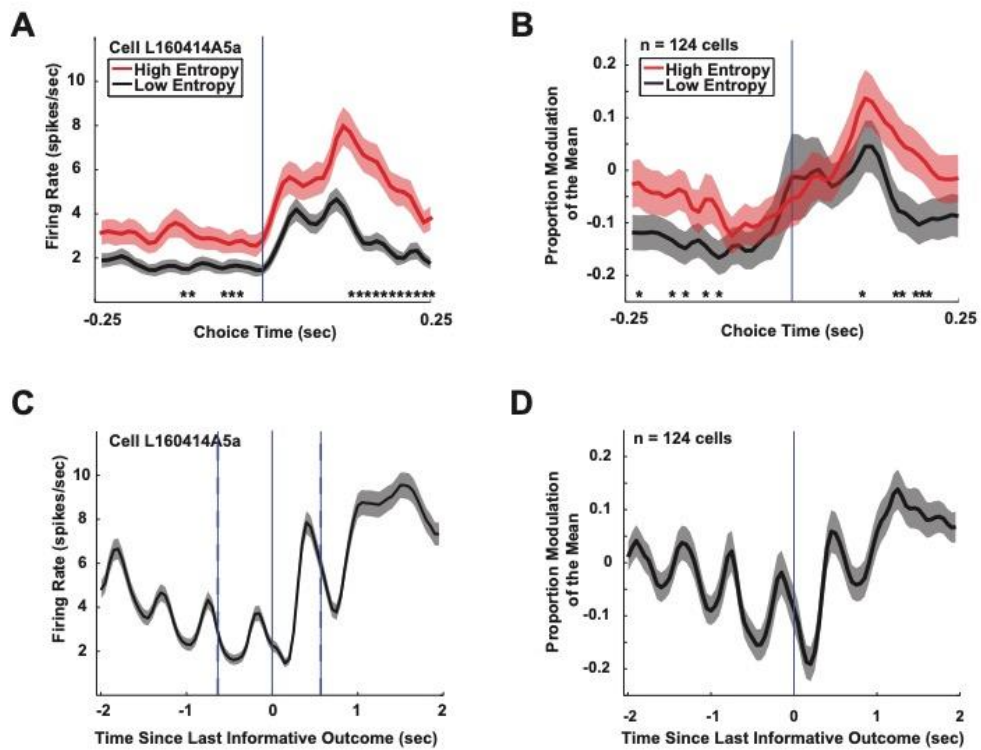
Fig. 3. Neurons in PCC forecast deviations in behavior. **A**. Sample neuron encoding behavioral

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 or rewarding outcome

compared to before. Thin blue lines with dashed lines very close on either side: average time of

choice before (left) or after (right) last informative or rewarding choice ± 1 s.e.m. **D**. The
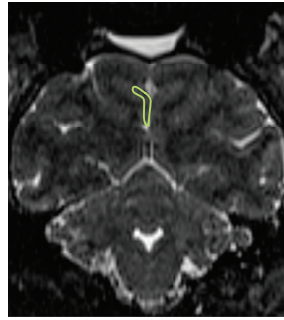
37

770   population also encoded this boundary, with higher firing rates after the last informative or

771   rewarding outcome compared to before. **B** and **D** plots: n = 124 cells. **A** and **B** plots: blue line =

772   end of saccade; **C** and **D** plots: central blue line = time of outcome. **All** plots: shading = ± 1

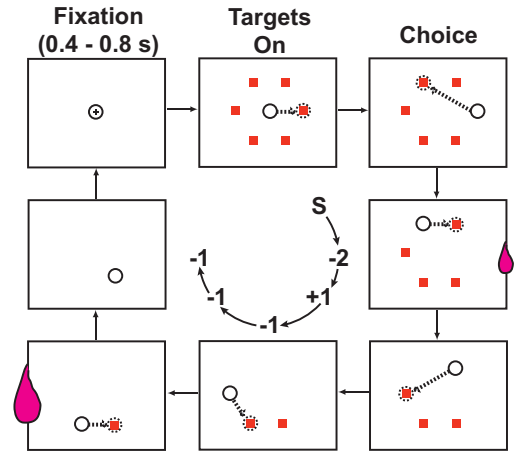773   s.e.m. * = p < 0.05 for that 10 ms time bin.

774

A

Monkey L   Monkey R

B

C

A

Cell L010414A1b

B

Cell L010414A1b

Cell L010414A1b

C

A

Cell L160414A5a

High Entropy
Low Entropy

Firing Rate (spikes/sec)

10
8
6
4
2

** ***   ***********

-0.25   Choice Time (sec)   0.25

B

n = 124 cells

High Entropy
Low Entropy

Proportion Modulation
of the Mean

0.2
0.1
0
-0.1
-0.2

*   * * *   *   ** ***

-0.25   Choice Time (sec)   0.25

C

Cell L160414A5a

Firing Rate (spikes/sec)

10
8
6
4
2

-2   -1   0   1   2
Time Since Last Informative Outcome (sec)

D

n = 124 cells

Proportion Modulation
of the Mean

0.2
0.1
0
-0.1
-0.2

-2   -1   0   1   2
Time Since Last Informative Outcome (sec)