

Ventromedial Prefrontal Cortex Tracks Multiple Environmental Variables during Search

Priyanka S. Mehta,¹ Jiaxin Cindy Tu,¹ Giuliana A. LoConte,¹ Meghan C. Pesce,² and Benjamin Y. Hayden¹

¹Department of Neuroscience and Center for Magnetic Resonance Research, University of Minnesota, Minneapolis, Minnesota 55455, and ²Department of Brain and Cognitive Sciences University of Rochester, Rochester, New York 14611

To make efficient foraging decisions, we must combine information about the values of available options with nonvalue information. Some accounts of ventromedial PFC (vmPFC) suggest that it has a narrow role limited to evaluating immediately available options. We examined responses of neurons in area 14 (a putative macaque homolog of human vmPFC) as 2 male macaques performed a novel foraging search task. Although many neurons encoded the values of immediately available offers, they also independently encoded several other variables that influence choice, but that are conceptually distinct from offer value. These variables include average reward rate, number of offers viewed per trial, previous offer values, previous outcome sizes, and the locations of the currently attended offer. We conclude that, rather than serving as specialized economic value center, vmPFC plays a broad role in integrating relevant environmental information to drive foraging decisions.

Key words: electrophysiology; foraging; neuroeconomics; search; value; vmPFC

Significance Statement

Decision makers must often choose whether to take an immediately available option or continue to search for a better one. We hypothesized that this process, which is integral to foraging theory, leaves neural signatures in the brain region ventromedial PFC. Subjects performed a novel foraging task in which they searched through differently valued options and attempted to balance their reward threshold with various time costs. We found that neurons not only encode the values of immediately available offers, but multiplexed these with environmental variables, including reward rate, number of offers viewed, previous offer values, and spatial information. We conclude that vmPFC plays a rich role in encoding and integrating multiple foraging-related variables during economic decisions.

Introduction

Foragers typically encounter prey stochastically and must decide whether to pursue them or continue searching for superior ones (Stephens and Krebs, 1986). Deciding whether to pursue a prey or bypass it requires identifying its value (i.e., the value of the immediately available option), but also contemplating the broader context of other variables that influence the expected value of moving on. Such variables include the risk and cost of pursuit, state variables (such as hunger) that change the subjective reward value, information that predicts the delay to and quality of the next reward encountered, and the availability of information which can improve later decisions (Daw et al., 2006;

Pearson et al., 2009; Kolling et al., 2012, 2014; Hills et al., 2015). Moreover, the decision-maker needs to not just decide, but act. Many approaches to modeling choice presume that action selection is divorced from core evaluative processes (Rangel et al., 2008; Padoa-Schioppa, 2011; Levy and Glimcher, 2012; Chen and Stuphorn, 2015). However, some recent work suggests that action selection cannot be dissociated from evaluation and comparison (Cisek and Kalaska, 2010; Pezzulo and Cisek, 2016; Hayden and Moreno-Bote, 2018). We will use the term “environmental variables” to refer to parameters that are conceptually distinct from the value of the encountered prey but that must be combined with it to determine choice and action.

Because the need to forage has been a driving force in the evolution of brain structure and function, understanding the neural basis of foraging decisions is an important goal (Wise et al., 1996; Passingham and Wise, 2012; Calhoun and Hayden, 2015; Hayden, 2018; Mobbs et al., 2018). An important foraging-related region is the ventromedial PFC (vmPFC), which appears to assign values to available options (Boorman et al., 2009, 2013; Kable and Glimcher, 2009; Wunderlich et al., 2009; Lim et al., 2011; Kolling et al., 2012). This proposed role is consistent with

Received Sept. 13, 2018; revised April 16, 2019; accepted April 18, 2019.

Author contributions: P.S.M. and B.Y.H. designed research; P.S.M., G.A.L., and M.C.P. performed research; P.S.M. and J.C.T. analyzed data; P.S.M. and B.Y.H. wrote the first draft of the paper; P.S.M. and B.Y.H. edited the paper; P.S.M. and B.Y.H. wrote the paper.

We thank Marc Mancarella, Shannon Cahalan, Michelle Ficalora, and Marcelina Martynek for technical assistance.

The authors declare no competing financial interests.

Correspondence should be addressed to Priyanka S. Mehta at mehta233@umn.edu.

<https://doi.org/10.1523/JNEUROSCI.2365-18.2019>

Copyright © 2019 the authors

its putative role as derived from neuroeconomic studies (Rushworth et al., 2011; Levy and Glimcher, 2012; Monosov and Hikosaka, 2012; Bartra et al., 2013; Strait et al., 2014). Specifically, an emerging consensus holds that vmPFC is critical and specialized for signaling the values of immediately available offers.

However, the status of vmPFC is contentious (Delgado et al., 2016). It is not clear whether it plays a broader role in choice beyond computing immediate offer values. Indeed, direct tests indicate that human vmPFC hemodynamic activity correlates with economic decision-making variables, whereas dACC, but not vmPFC, encodes environmental variables, such as the cost and value of search (Kolling et al., 2012). This immediate value view is also consistent with self-control literature showing that vmPFC encodes immediate values of offers (Kable and Glimcher, 2007; Hare et al., 2009; Smith et al., 2010).

Other evidence suggests a more integrative role for vmPFC: that its responses track multiple state and cognitive variables that are conceptually distinct from immediately available value (e.g., Gussard and Raichle, 2001; Bouret and Richmond, 2010; Monosov and Hikosaka, 2012). Likewise, there is at least some evidence that vmPFC carries a modest amount of motor-related spatial information in simple binary choice tasks (Strait et al., 2016). We therefore hypothesized that vmPFC plays a role in guiding foraging decisions that goes beyond computing the values of immediately available offers by encoding a variety of additional environmental variables.

We examined responses of neurons in area 14, a presumed macaque homolog of human vmPFC (Mars et al., 2016). The best name for this region is currently unclear, for several reasons. First, the specific macaque homolog of human vmPFC is not fully known, nor is the microstructure of the vmPFC in humans fully elucidated. Most importantly, our recording site, within area 14, probably does not cover the entirety of the region homologous to human vmPFC, which may also include at least parts of areas 10 and 32. Second, this homology to primates has not yet been universally adopted, although we believe that Mars et al. (2016) have made the strongest case so far. Third, the primate ventromedial network is large and heterogeneous (Ongur and Price, 2000). Nonetheless, area 14 is likely a central player in the region homologous to human vmPFC (Mars et al., 2016). We will continue to use the term vmPFC, as we have in our past studies of this region (Strait et al., 2014, 2016; Azab and Hayden, 2017), with the understanding that a more precise functional name is wanting, and that we may only really be examining one subregion of the area homologous to human vmPFC.

On each trial of our novel search task, macaques were presented with an array of obscured offers randomly positioned on a monitor. Revealing an offer's value required a fixed time cost, after which subjects could pay an additional time cost to obtain the reward. Although neurons encoded offer values, these were multiplexed with information, including the value of the current and previous offers, reward rate, and spatial information. These results indicate that encoding immediate offer value is not a specialized function of vmPFC, but is just one of several variables that it tracks, and suggest it plays a rich role in influencing foraging decisions.

Materials and Methods

All procedures were designed and conducted in compliance with the Public Health Service's *Guide for the care and use of animals* and approved by the University Committee on Animal Resources at the University of Rochester. Subjects were 2 male rhesus macaques (*Macaca mulatta*: Subject J, age 10 years; Subject T, age 5 years). We used a small titanium

prosthesis to maintain head position. Training consisted of habituating animals to laboratory conditions and then to perform oculomotor tasks for liquid reward. A Cilux recording chamber (Crist Instruments) was placed over the vmPFC. Position was verified by MRI with the aid of a Brainsight system (Rogue Research). After all procedures, animals received appropriate analgesics and antibiotics. Throughout all sessions, the chamber was kept sterile with regular washes and sealed with sterile caps.

Recording site. All recordings were performed at approximately the same time of day, between 10:00 A.M. and 3:00 P.M. We approached vmPFC through a standard recording grid (Crist Instruments). We defined vmPFC according to the Paxinos atlas (Paxinos et al., 2000). Roughly, we recorded from an ROI lying within the coronal planes situated between 42 and 31 mm rostral to interaural plane, the horizontal planes situated between 0 and 7 mm from the brain's ventral surface, and the sagittal planes between 0 and 7 mm from medial wall. This region falls within the boundaries of area 14 according to the atlas. We used our Brainsight system to confirm recording location before each session with structural magnetic resonance images taken before the experiment. Neuroimaging was performed at the Rochester Center for Brain Imaging, on a 3T MAGNETOM Trio Tim (Siemens) using 0.5 mm voxels. To further confirm recording locations, we listened for characteristic sounds of white and gray matter during recording. These matched the loci indicated by the Brainsight system in all cases.

Electrophysiological techniques, eye tracking, and reward delivery. All methods used have been described previously and are summarized here (Strait et al., 2014). A microdrive (NAN Instruments) was used to lower single electrodes (Frederick Haer, impedance range 0.7–5.5 M Ω) until waveforms of between one and four neurons were isolated. Individual action potentials were isolated on a Plexon system. We only selected neurons based on their isolation quality, never based on task-related response properties. All collected neurons for which we managed to obtain at least 300 trials were analyzed. In practice, 86% of neurons had >500 trials (this was our recording target each day).

An infrared eye-monitoring camera system (SR Research) sampled eye position at 1000 Hz, and computer running MATLAB (The MathWorks) with Psychtoolbox and Eyelink Toolbox controlled the task presentation. Visual stimuli were colored diamonds and rectangles on a computer monitor placed 60 cm from the animal and centered on its eyes. We used a standard solenoid valve to control the duration of juice delivery, and established and confirmed the relationship between solenoid open time and juice volume before, during, and after recording.

Experimental design. Subjects performed a diet-search task (see Fig. 1) that is a conceptual extension of previous foraging tasks we have developed. Previous training history for these subjects included two types of foraging tasks (Blanchard and Hayden, 2015; Blanchard et al., 2015), intertemporal choice tasks (Hayden, 2016), two types of gambling tasks (Strait et al., 2014; Azab and Hayden, 2017), attentional tasks (similar to those in Hayden and Gallant, 2013), and two types of reward-based decision tasks (Sleezer et al., 2016; Wang and Hayden, 2017).

To begin each trial, the animal fixated on a central dot (50 ms), after which either four or seven white diamond shapes appeared in randomly selected nonoverlapping positions on the screen. The number of offers per trial, either 4 or 7, was determined randomly on each trial (50% of trials for each of the two numbers). Each diamond was 200 pixels tall and 100 pixels wide. Continuous fixation on one diamond for 400 ms caused it to disappear and reveal a reward offer. Offers were orange bars, partially filled in to indicate the value of the riskless offered reward. The percentage of the offer bar that was filled in corresponded to the offer value in terms of percentage of the maximum value possible per offer (e.g., an offer bar that was 10% orange and 90% black would indicate an offer worth 10% of the maximum value; 20 μ l for Subject T and 23 μ l for Subject J). Reward values for each offer were generated randomly for each offer using a uniform continuous distribution ranging from 0% to 100% of the maximum possible reward value for the individual subject (i.e., continuously varying rewards were generated, but in practice continuity was limited by pixel size). The 200 pixel height of the bars allows 200 steps between 0 and the maximum reward.

The subject could freely search through the diamonds in any order and could accept any offer. Acceptance led to the end of the trial; rejection led to a return to the initial state (viewing an array of diamonds). To accept the offered reward, the animal had to maintain fixation on the offer for 300 ms, after which the screen would go black and the offered amount of liquid reward would be immediately delivered. Thus, selecting a given offer required 700 ms: 400 ms to unmask it and an additional 300 ms to obtain it. If the subject broke fixation on the reward stimulus at any point between 0 and 300 ms from the initial reveal, the reward stimulus would disappear and the diamond would return in its place (a “rejection” of the offer). The subject could then resume freely inspecting other offers. There was no limit to how many offers a subject could inspect, nor to how many times a subject could inspect a particular offer. The trial only ended (and a liquid reward was only delivered) after the subject accepted an offer. Reward delivery was followed by a 4 s intertrial interval.

Because the experiment was conducted partially using one computer and partially another, the plot of diamond locations (see Fig. 1*B*) was generated by translating the pixel coordinates of diamonds on each screen into proportions of the total pixel length and width of the screen.

We computed behavioral threshold by plotting all encountered offer values (x axis) against the percentage of times that that particular offer size was selected when encountered (y axis), and fitting a sigmoidal curve to the data. The behavioral threshold is the x value (offer size) corresponding to the inflection point of the sigmoidal curve. To compute thresholds for subsets of the data (e.g., for trials for which the previous offer was between 0.1 and 0.2), we only included those trials when fitting the sigmoidal.

Reward rate. We calculated reward rate at a particular offer by finding the average reward obtained over the past n trials. For example, to calculate reward rate over the past 10 trials at trial 42 (Offer 1), we first summed all rewards obtained (i.e., only the offer chosen each trial, not all offers viewed) over trials 31–40, and divided this value by 10 to get a “reward-per-trial” value. We excluded the outcome of the immediately preceding trial because we found evidence that vmPFC encodes this value, and we wanted to be certain that reward rate effects are not simply due to previous outcome encoding. We also excluded all trials with less than n preceding trials; in the example above, we would only begin calculating reward rate at trial 12 (calculating reward rate over trials 1–10).

To assess encoding of a recency-weighted reward rate, we first used the subjects’ behavioral data to perform a logistic regression of the current offer and each of the most recent 10 previous outcomes against choice. This produced 11 regression weights: one for the current offer, and one for each of the previous outcomes stretching back 10 trials. We then fit an exponential decay curve to the regression coefficients for each of the 10 previous outcomes, and used the values predicted by this curve to generate weights for each previous outcome. To compute a weighted-average reward rate for the past 10 trials, we multiplied the outcome of each previous trial by its weight, then took the sum of these 10 values.

Statistical analysis for physiology. Peristimulus time histograms (PSTHs) were constructed by aligning spike rasters to the offer reveal and averaging firing rates across multiple trials. Firing rates were calculated in 10 ms bins but were mostly analyzed in longer (500 ms) epochs. Firing rates were normalized (where indicated) by subtracting the mean and dividing by the SD of the entire neuron’s PSTH. We chose analysis epochs before data collection began, to reduce the likelihood of p-hacking. The “current offer” epoch was defined as the 500 ms epoch beginning 100 ms after the offer reveal, to account for sensory processing time. This epoch was used in previous studies of choice behavior (Strait et al., 2014, 2015; Azab and Hayden, 2017). All comparisons of firing rates over particular epochs were conducted using two-sample t tests. All fractions of modulated neurons were tested for significance using a two-sided binomial test.

Types of value tuning. Positive and negatively tuned neurons were classified by performing a regression of firing rate against value (split into 100 equally sized bins). Cells with a positive or negative regression slope (regardless of significance) were classified as positively and negatively tuned, respectively. To estimate monotonicity, we determined the regression slope of firing rate against offer value for low values (0%–50% of the maximum value) and compared that with the slope for high values

(50%–100% of the maximum value). If the sign changed (i.e., went from positive to negative or negative to positive), we considered the cell’s tuning to be “nonmonotonic.” According to our classification system, the direction and monotonicity of value tuning are not mutually exclusive; so of the nonmonotonic neurons, 8 were also classified as positively tuned, and 5 of these neurons were also classified as negatively tuned.

Population analyses. We used β correlation analyses to assess whether neurons encoded particular offer values (Blanchard et al., 2015; Azab and Hayden, 2017). When regressing firing rate against the value of the current offer over the analysis epoch, offer selection is a confound. Subjects are more likely to choose high offers and less likely to choose low offers. Thus, we regressed out offer selection in all regressions of the value of the current offer against firing rate. Furthermore, fixation time is a confound because it correlates with value. Subjects are more likely to accept high offers, and thus more likely to fixate longer on them, due to the nature of the task that requires subjects to fixate for 300 ms to accept an offer. Rejected offers, which are more often of lower value, come with smaller fixation times, as the subject must break fixation before 300 ms to reject the offer. Thus, we regressed out fixation time as well in all regressions of the value of the current offer against firing rate.

To assess the relationships between population encoding of current offer value and other variables, we regressed each neuron’s firing rate against each variable of interest to obtain a list of n coefficients for each variable: one coefficient per neuron per variable. To compare neurons’ encoding of one variable versus the other, for example, current offer value and previous offer value (see Fig. 5), we performed a Pearson correlation between the coefficient set for current offer value and that for previous offer value. The resulting correlation coefficient provides information about how similarly or differently each variable is encoded by the population.

We used a similar β correlation analysis to assess whether neurons represented offer acceptance and rejection using similar coding formats (compare Azab and Hayden, 2017). We found the regression coefficients (i.e., the β weight) of each neuron’s normalized firing rates and the values of accepted offers, and their normalized firing rates and the values of rejected offers. We combined these β weights into two vectors of the same length as the total number of neurons. Each vector indicates the strength and direction of modulation for each neuron in response to the offer reveal in cases where the monkey is about to accept or about to reject the offer. We call this the population “format.” We compared different formats by finding the Pearson correlation coefficient between them. We then compared these distributions of correlation coefficients to distributions that would be obtained under a chance model. For the first set of analyses, we assume a chance model in which neurons are purely predecisional; they do not differentiate between values according to whether they will later be accepted or rejected. To achieve this, we shuffle trials across these two categories at random. This chance models achieve a permutation of the existing data, which we then use for the same β correlation analysis explained above. For the second analysis, we assume a model in which neurons are purely postdecisional; there is no correlation between the coefficients of accepted and rejected offers. We achieve this by comparing the true regression coefficient to 0.

Exclusion of data. When performing regressions of neural and behavioral factors against the values of previous outcomes and offers, we excluded outlier data points, specifically, those that were ≥ 3 SDs away from the mean. This number was determined before performing the analyses. In Figure 6*A*, 2 points were excluded. In Figure 6*D*, 2 points were excluded. In Figure 6*E*, 4 points were excluded. In Figure 6*F*, 4 points were excluded. In Figure 7*B*, 2 data points were excluded. The vast majority of data points were not outliers, and were within 1 SD of the mean.

In Figure 6*F* (the effect of previous offer value on the strength of neural encoding of the current offer), the error bars of 37 points have been removed due to the fact that they are larger than the height of the plot. This is likely due to the low number of samples used to generate these points. These points all represent situations in which the previous offer value was $\geq 49\%$. In the task, high previous offer values are rare because, for an offer to be a “previous offer” within a trial, the subject must have rejected it, and subjects rarely reject offers that are below their value threshold.

To ensure that the positive relationship we found in Figure 6F is not only due to this large proportion of points (37%) that have large errors, we also performed the regression without including the high-error points (points whose error was >10). We found that the positive relationship holds even with these values eliminated ($\beta = 0.6464$, $p < 0.001$).

Quadrant location of offers. To assess neural encoding of quadrant location of offers, we divided the screen into four equivalently sized quadrants (top-left, top-right, bottom-left, and bottom-right). We then performed a one-way ANOVA of the firing rate during the analysis epoch after each offer against the quadrant that offer was in. To assess neural encoding of whether offers were located on one-half of the screen, we assigned each offer a value indicating either up (1) or down (0), or right (1) or left (0). We then performed a linear regression of the firing rates during the analysis epoch after each reveal against these 1/0 values to produce a β coefficient value.

Tuning curve fitting for angular position tuning. We fit a von Mises function to the spike count in the 400 ms diamond viewing period and polar angle (θ) relationship, with parameters a , k , θ , and c . The equation is as follows:

$$\text{spikes} = a * e^{(k * \cos(x - \theta + 1) - 1) / e^{2 * k - 1}} + c$$

Assuming Poisson firing of the neurons, the log likelihood of the von Mises fit was calculated, and compared with that of the mean firing rate uniform tuning curve, and the Bayesian Information Criteria (BIC) difference was calculated. Cells were considered tuned to polar angle of the diamond viewed if $\text{BIC}_{\text{vonmises}} > \text{BIC}_{\text{uniform}}$.

In addition, we also directly measured the mutual information between spatial firing rate and polar angle for each neuron using the method from Skaggs et al. (1993), following an adaptive smoothing method from Skaggs et al. (1996). Briefly, the data were first binned into 100×1 vector of angle bins covering the whole 360 degrees of the field, and then the firing rate at each point in this vector was calculated by expanding a circle around the point until the following criterion was met:

$$N_{\text{spikes}} < \alpha / (N^2_{\text{occ}} * r^2)$$

where N_{occ} is the number of occupancy samples, N_{spikes} is the number of spikes emitted within the circle, r is the radius of the circle in bins, and α is a scaling parameter set to be 10,000. Cells are said to be significantly tuned to position when the mutual information exceeds the 95th percentile of the shuffled data.

Data and code availability statement. All data and code used during data collection and analysis can be made available upon reasonable request to the corresponding author.

Results

Macaques are efficient foragers in a computerized search task

We examined behavior of 2 macaque subjects performing our foraging search task (Subject J: 23,826 trials; Subject T: 20,509 trials). On each trial, the subject was presented with an array of visually identical offers randomly positioned on the screen (Fig. 1; see Materials and Methods). Subjects could inspect offers in any order to reveal their values and, following inspection of an offer, could select or reject it. Rejected offers always remained available for return and reinspection. Offers were riskless and varied in size continuously from 0 to 230 μl (Subject J) or 0 to 200 μl (Subject T) fluid reward. Values noted below are in units of percentage and refer to proportions of the largest amount for that animal. Results here are combined across the 2 subjects; all major results were replicated for each animal individually (Table 1) with a few minor exceptions.

Each trial had either 4 or 7 offers; the two trial types were randomly interleaved and occurred equally often. On average, subjects inspected 1.872 offers on trials with 4 offer arrays and 2.074 offers on trials with 7 offer arrays (for the frequency with which subjects viewed different numbers of offers per trial, see Table 2). A strategy of strictly choosing the first offer revealed

would have obtained 49.62% per trial. Obtained rewards were significantly better than those obtained following this basic strategy ($p < 0.001$, independent-samples t test). Specifically, the average reward obtained on 4 offer trials was 68.72% and on 7 offer trials was 73.46%.

Subjects' behavior approximated a threshold strategy (Fig. 2A). Subjects' measured thresholds were 50.30% in the 4 option case and 54.33% in the 7 option case. To determine optimal thresholds, we simulated performance using a set of response times for each task component drawn from observed behavior (Fig. 2B). That is, we assumed that subjects' reaction times were unavoidable and not strategic, an approach we developed previously (Hayden et al., 2011; Blanchard and Hayden, 2014). Because evaluation of subjects' strategies depends on each subject's idiosyncratic behavior, we provide individual subject data here. For the 4 option case, the optimal threshold was 60.96% (i.e., 140 μl) for Subject J and 60.60% (121 μl) for Subject T. For the 7 option case, the optimal threshold was 63.78% (147 μl) for Subject J and 60.65% (121 μl) for Subject T. The observed thresholds were significantly lower than the optimal ones in all 4 cases ($p < 0.04$ for the 4 option case for Subject T; $p < 0.01$ for the other cases). Thus, in both the 4 and 7 option cases, subjects stopped (i.e., selected an option) at a lower threshold than optimal. Nonetheless, based on these numbers, subjects harvested $\sim 95\%$ of the intake of the optimal chooser with the same reaction time constraints. Together, these results indicate that subjects adopted nearly rate-maximizing thresholds.

vmPFC neurons signal immediate offer value

We recorded responses of 122 neurons in area 14 ($n = 70$ in Subject J and $n = 52$ in Subject T). We collected an average of 523 trials per neuron. We primarily focused on an analysis epoch of 100–600 ms after the time at which the offer was revealed (“reveal”). We chose this epoch before data collection began because we used it in our previous studies of this and similar regions. Using an *a priori* defined epoch reduces the chance of p-hacking (Strait et al., 2014, 2016; Azab and Hayden, 2017, 2018).

Figure 3A shows the responses of an example neuron in this task aligned to the reveal, separated by high and low value offers. Responses of this neuron depended on the value of the presented offer (16.52 spikes/s for high values vs 11.78 spikes/s for low values, $p < 0.001$, independent-samples t test). To examine population encoding of offer value, we performed a regression of firing rate against current offer value; we regressed out time spent fixating on each offer and whether or not that offer was chosen (see Materials and Methods). Figure 3B shows the proportion of neurons in the population significantly encoding offer values. Responses of 33.61% of neurons depended linearly on the offer value over the focal analysis epoch. This proportion is significantly higher than the proportion found in our previous study (Strait et al., 2014) where we reported 16.03% of cells encoded current value ($\chi^2 = 11.68$, $p < 0.001$). This difference indicates that value in this task was more effective at driving responses than in our previous task.

Individual cells' responses to value took a variety of forms: some linear and some nonlinear. We classified the relationship between firing rate and value for each cell into one of four categories: “positive,” “negative,” “nonmonotonic,” and “none” (see Fig. 4; some nonmonotonic cells can also have an overall positive or negative trend; see Materials and Methods). We found that 11.48% of cells ($n = 14$ of 122) demonstrated a significantly positive relationship between firing rate and current offer value, 11.48% ($n = 14$ of 122) a significantly negative relationship, and

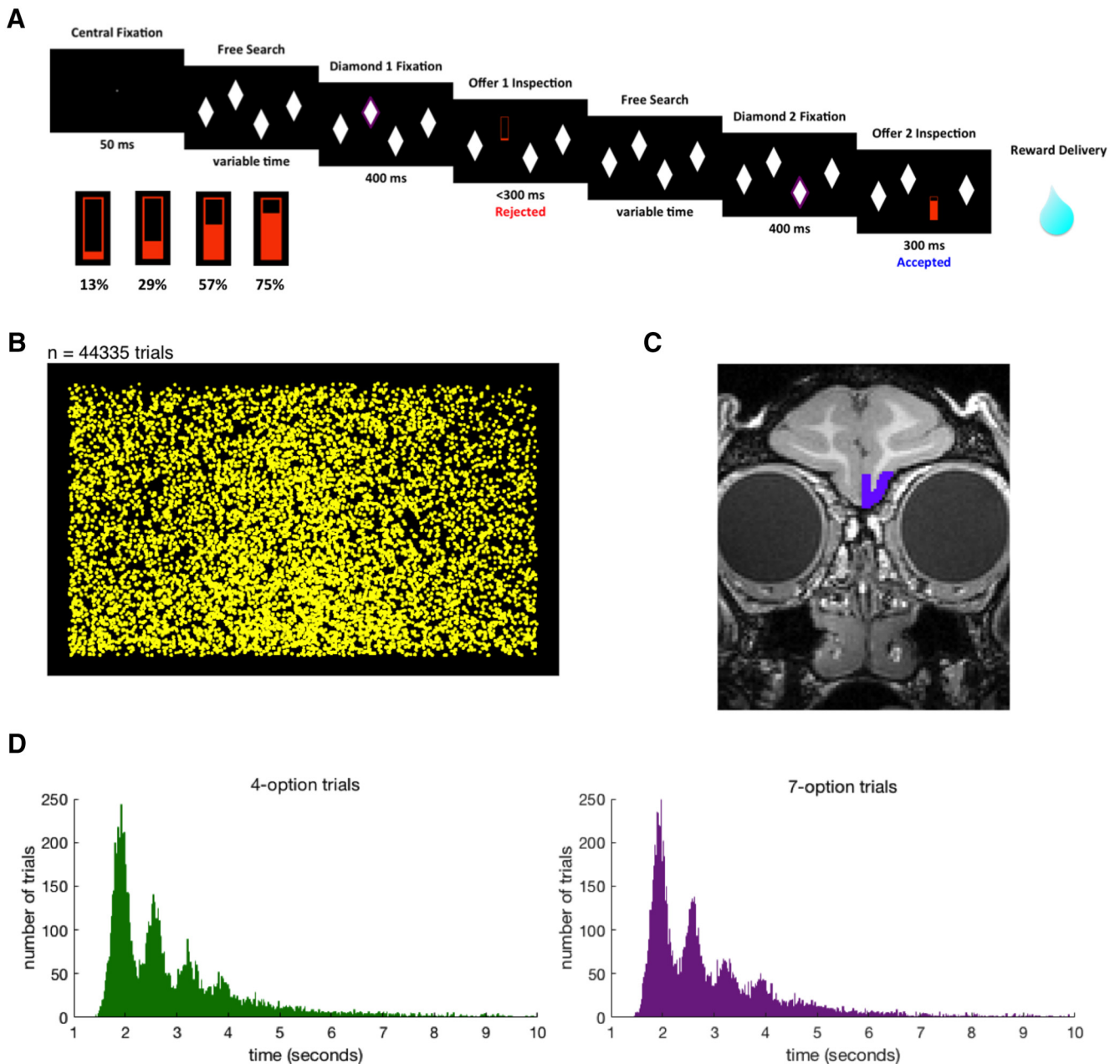


Figure 1. Our novel diet-search task. *A*, Schematic of task. Subjects freely search through a display of diamonds on a computer monitor. They can pay a 400 ms time cost to inspect (reveal) the offer values hidden behind any diamond. Maintaining fixation on a revealed offer for 300 ms more results in reward delivery; breaking fixation before 300 ms up returns the subject to free search. Bottom, Example offer sizes. Percentage indicates percentage of subject's maximum value per offer. *B*, Plot of all diamond positions for all trials superimposed on image of screen (for details, see Materials and Methods). *C*, Illustration of recording position. An example coronal slice of a structural MRI of Subject T with area 14 (vmPFC) superimposed in purple. Recording sites were distributed evenly throughout this region. *D*, Trial durations for 4 option and 7 option trials. Scalloped patterns reflect choices of each sequentially fixated option.

40.16% ($n = 49$ of 122) cells a nonmonotonic relationship (these purely nonmonotonic neurons are not part of the 33.61% of neurons that linearly encode value). Additionally, we found that 6.56% of cells ($n = 8$ of 122) demonstrated a nonmonotonic relationship with an overall significantly positive trend, and 4.10% of cells ($n = 5$ of 122) demonstrated a nonmonotonic relationship with an overall significantly negative trend. Finally, 26.23% of cells ($n = 32$ of 122) had no discernible relationship between firing rate and offer value.

Encoding of categorical choice

It is possible that encoding of pure value of an offer is confounded by encoding of categorical choice: that is, a cell might fire more in

response to a high offer because the offer is high, or because the subject is likely to choose that offer. If the neuron signals choose versus reject, but is otherwise indifferent to value, this will lead to a spurious correlation. For this reason, in the analyses above, we regressed out choice in all our regressions where offer value is a regressor.

To gain more insight into this process, we performed an additional analysis. If vmPFC were simply encoding categorically whether an offer is accepted or rejected (e.g., if its signal is “post-decisional,” after the subject makes the behavioral decision to accept or reject), there would be no reason to expect any similarity in how the values of accepted and rejected offers are encoded. That is, the regression coefficients from a regression of accepted

Table 1. Results of various analyses broken down by individual subject

| | Subject J | Subject T | Combined |
|--|-------------------|-------------------|--------------------|
| Macaques are efficient foragers in a computerized search task | | | |
| No. of trials | 23,826 | 20,509 | 44,335 |
| Offers viewed per 4 offer trial | 1.803 | 1.953 | 1.8722 |
| Offers viewed per 7 offer trial | 1.9996 | 2.1605 | 2.0742 |
| Average reward obtained on 4 offer trials | 69.46% | 67.87% | 68.72% |
| Average reward obtained on 7 offer trials | 73.20% | 73.75% | 73.46% |
| Behavioral threshold across all trials | 49.26% | 56.77% | 52.56% |
| Behavioral threshold for 4 offer trials | 46.77% | 55.06% | 50.30% |
| Behavioral threshold for 7 offer trials | 50.99% | 58.29% | 54.33% |
| vmPFC neurons signal values of offers | | | |
| Percent of neurons significantly encoding current offer value | 40.00% | 25.00% | 33.61% |
| vmPFC neurons encode environmental variables previous outcome | | | |
| Beta correlation coefficient of behavioral threshold against previous outcome value (<i>p</i> value) | 0.0372 (0.2030) | 0.1292 (<0.001) | 0.0602 (<0.001) |
| Percent of neurons significantly encoding previous outcome value (<i>p</i> value) | 27.14% (<0.001) | 17.31% (0.0010) | 22.95% (<0.001) |
| Beta correlation coefficient of firing rate against previous outcome value (<i>p</i> value) | −0.1415 (0.6039) | 0.7718 (0.0042) | −0.4974 (0.0377) |
| Previous offer | | | |
| Beta correlation coefficient of behavioral threshold against previous offer value (<i>p</i> value) | 0.1243 (0.1202) | 0.1515 (<0.001) | 0.1073 (0.0019) |
| Percent of neurons significantly encoding previous offer value (<i>p</i> value) | 24.29 (<0.001) | 5.77 (1.000) | 16.39 (<0.001) |
| Beta correlation coefficient of firing rate against previous offer value (<i>p</i> value) | 7.334 (<0.001) | 0.2565 (0.0736) | 4.391 (<0.001) |
| View number | | | |
| Beta correlation coefficient of behavioral threshold against view number (<i>p</i> value) | −0.04849 (0.0063) | −0.0267 (<0.001) | −0.0346 (<0.001) |
| Percent of neurons significantly encoding view number (<i>p</i> value) | 42.86% (<0.001) | 19.23 (<0.001) | 32.79 (<0.001) |
| Beta correlation coefficient of firing rate against view number (<i>p</i> value) | 0.02553 (0.00895) | 0.03073 (0.03272) | 0.02775 (0.017549) |
| Reward rate | | | |
| Beta correlation coefficient of behavioral threshold against reward rate of past 10 trials (<i>p</i> value) | 0.4708 (<0.001) | 0.3133 (<0.001) | 0.2361 (<0.001) |
| Percent of neurons significantly encoding reward rate of past 10 trials (<i>p</i> value) | 34.29 (<0.001) | 30.77 (<0.001) | 0.3279 (<0.001) |
| Spatial positions of offers | | | |
| Percentage of neurons encoding current offer quadrant position | 30.00% | 11.54% | 22.13% |
| Percentage of neurons encoding current offer left-right position | 18.57% | 19.23% | 18.85% |
| Percentage of neurons encoding current offer up-down | 22.86% | 1.92% | 13.93% |

Table 2. Percentage of trials in which subjects viewed a total of different numbers (*n*) of offers^a

| <i>n</i> | % |
|----------|--------|
| 1 | 47.03 |
| 2 | 25.42 |
| 3 | 13.73 |
| 4 | 7.209 |
| 5 | 3.263 |
| 6 | 1.46 |
| 7 | 0.7394 |
| 8 | 0.4 |
| 9 | 0.2281 |
| 10+ | 0.5082 |

^aIf *n* offers have been viewed, a revisit to a previously viewed offer counts as a new “view” (e.g., *n* + 1).

offer values against firing rate could in principle be totally uncorrelated with the regression coefficients from a regression of rejected offer values against firing rate (see Materials and Methods) (Azab and Hayden, 2017). On the other hand, if the code in vmPFC is predecisional, then neurons would necessarily have to use the same code for soon-to-be-accepted and soon-to-be-rejected offers (if the code were different, the neurons would have access to the outcome of the choice). Thus, qualitative changes in ensemble tuning functions for accepted and rejected offers can provide insight into the status of vmPFC relative to choice processes.

We find that the correlation between regression coefficients of accepted and rejected offers is 0.2806 (Fig. 5). This value is significantly different from 0 (*p* = 0.002). Therefore, there is some degree of similarity in how accepted and rejected offers are encoded, consistent with a predecisional role for vmPFC. We can also use a cross-validation procedure to estimate the correlation

that would be observed with perfectly correlated variables, given the noise properties of our dataset (Azab and Hayden, 2017). This analysis shows that 0.2806 is lower than the ceiling estimate of perfect correlation (i.e., randomly reshuffled value), which was 0.8460 (see Materials and Methods, *p* < 0.001). These results suggest that vmPFC encoding of value is neither purely postdecisional nor purely predecisional; it has properties of both.

vmPFC neurons encode other environmental variables

We next investigated encoding of environmental variables, parameters that are distinct from the immediate offer value, but that influence the choice made. All results presented below exclude outliers (see Materials and Methods).

Previous outcome

In search, as in many other contexts, the value of the background is determined by the environmental richness. We estimated our subjects’ accept-reject thresholds based on their aggregate behavior (see Materials and Methods). For both subjects, threshold on the current trial increased with the size of the previous trial’s outcome (β = 0.0602, *p* < 0.001; Fig. 6A; significantly in 1 subject, positive trend in the other). This suboptimal behavior is consistent with subjects having an erroneous belief that recent good outcomes reflect changing environmental richness (as with a hot hand fallacy) (Blanchard et al., 2014).

Neurons in vmPFC encoded the value of the previous trial’s outcome. Figure 6B shows the proportion of neurons in the population encoding previous outcome values. Responses of 22.95% of neurons (*n* = 28 of 122) changed during the choice epoch depending on previous outcome value (linear regression, α = 0.05). This proportion is significantly greater than expected by

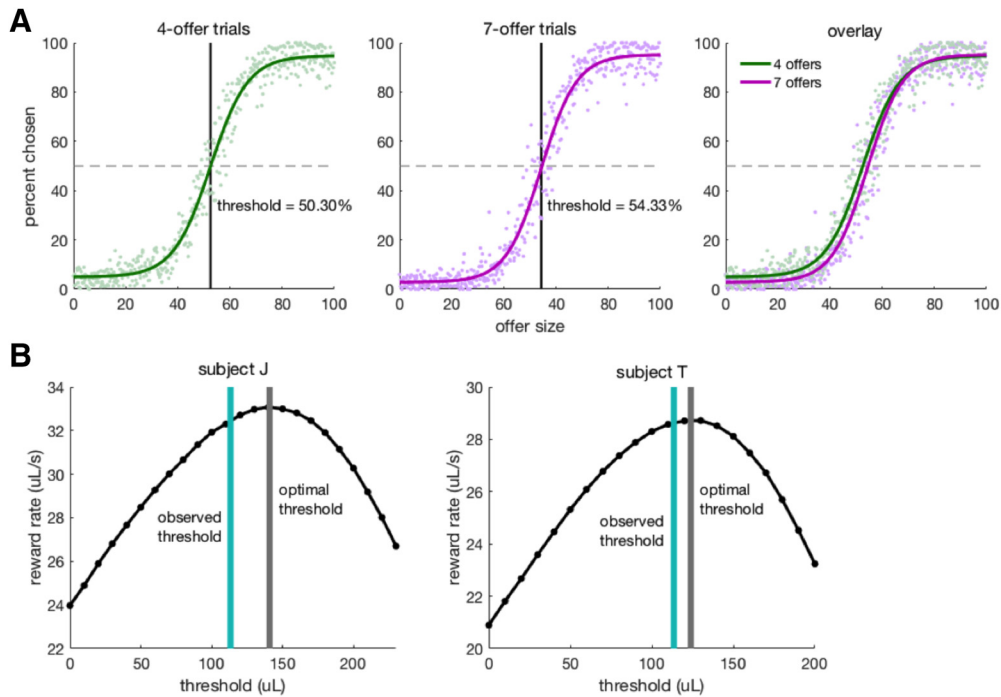


Figure 2. Subjects' behavior. **A**, Likelihood of accepting a given option was an approximately sigmoidal function of its value. Observed threshold was slightly higher on 7 option trials than on 4 option trials. Right, Overlay of the two plots. Points represent the percentage of times a particular value was selected out of the times it was encountered. Curves are sigmoidal best fits to these points. **B**, Optimal behavioral threshold results from computer simulation of task. Black line indicates the average reward rate earned >100 simulations of 1000 trial task sessions for each threshold between 0 and each subject's reward maximum (Subject J: 230 μ L; Subject T: 200 μ L) in steps of 10 μ L. Gray bars represent the threshold that yielded the highest reward rate in the simulation. Teal bars represent the subjects' observed behavioral threshold. For both subjects, observed threshold was significantly lower than optimal, although in both cases, due to the asymmetry of the curve, provided nearly the same intake as the optimal threshold.

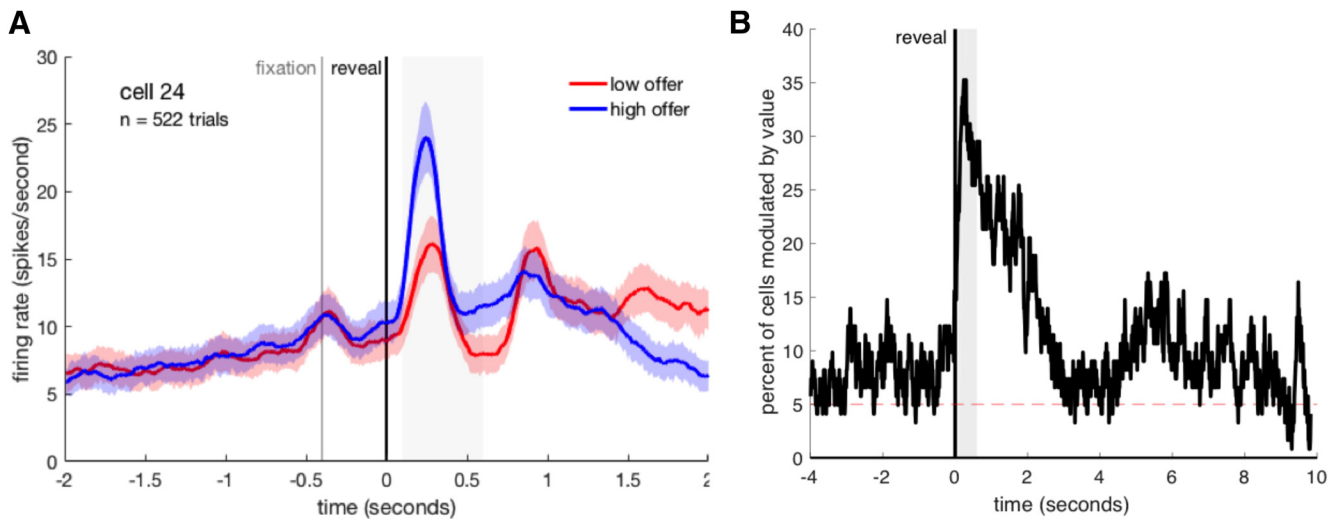


Figure 3. Encoding of value in vmPFC. **A**, PSTH showing the firing rate of an example neuron aligned to offers and separated by high and low offer values. For this cell, the firing rate is higher when the offer is high (>50% of maximum value per offer). Colored shading represents SE of firing rate at each time point (10 ms bins). Solid line indicates mean firing and is smoothed >20 bins. Gray shading represents the analysis epoch. Vertical lines indicate the beginning of diamond fixation (400 ms) and the subsequent reveal of the offer. **B**, Proportion of neurons in our entire sample whose firing rate responses were affected by reward value as a function of time relative to offer period. The x axis is different from **A**. Proportions are calculated with a 200 ms sliding boxcar. The peak at ~6 s after offer reveal corresponds to the approximate beginning of the following trial, during which we see a reactivation of encoding of the current offer (consistent with our finding that neurons encode the value of the outcome of the previous trial; see Results). Gray shading represents analysis epoch.

chance (i.e., 5%, $p < 0.001$, two-sided binomial test). This analysis was performed across all trials (no exclusions).

There is some evidence that previous outcome changes the strength of value coding as well. Specifically, we regressed the firing rate against offer value for different previous outcome sizes and found that the previous trial's outcome negatively affected the value

encoding of the current offer ($\beta = -0.4974$, $p = 0.0377$; Fig. 6D; significant in 1 subject, with a negative trend in the other; see Table 1).

Previous offer

In heterogeneous environments, the value of the previous offers so far encountered within a trial provides information about the

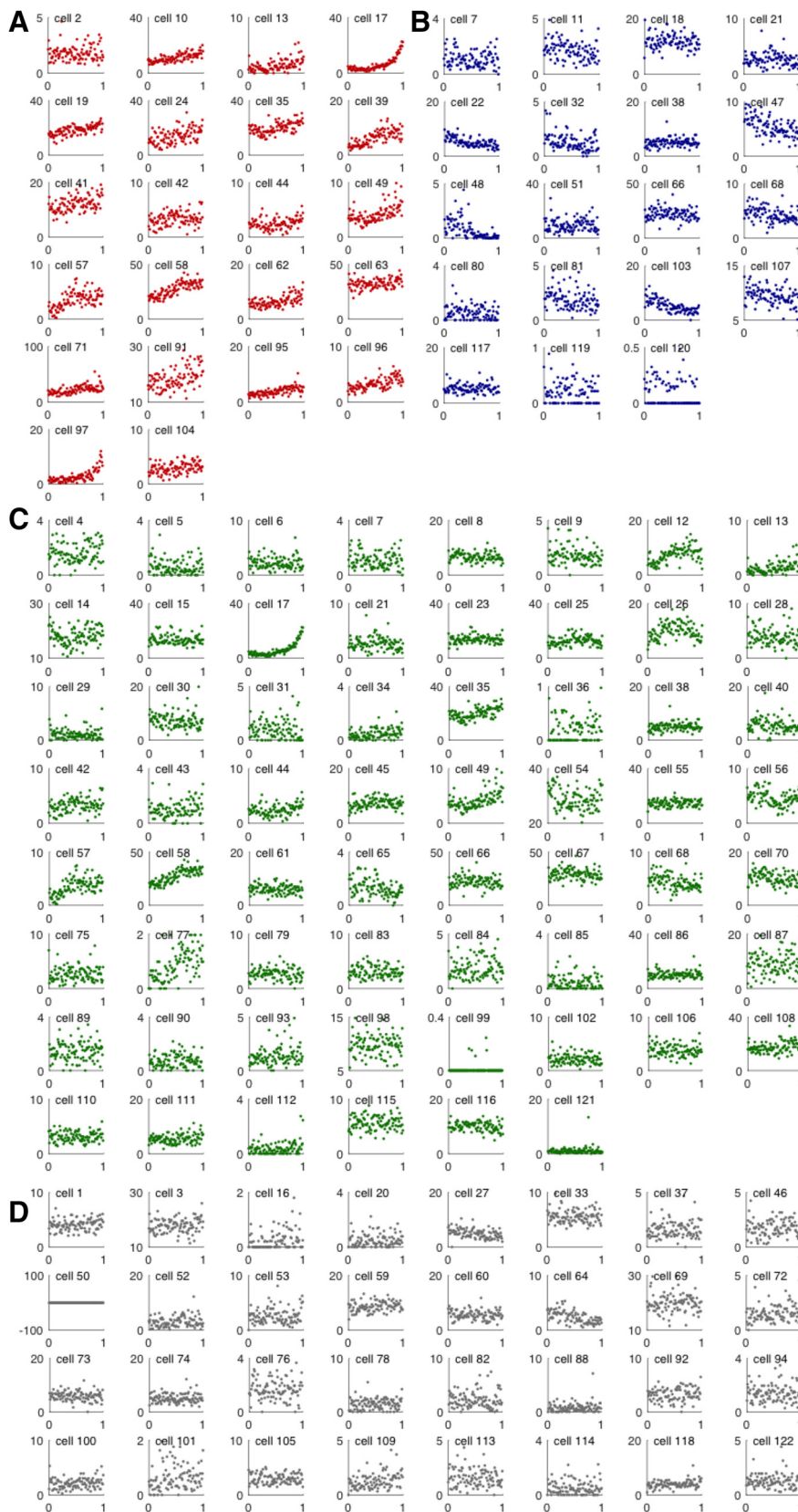


Figure 4. Classifications of individual neurons' value tuning. **A**, Neurons classified as “positively tuned” to value. **B**, Neurons classified as “negatively tuned” to value. **C**, Neurons classified as “nonmonotonically tuned” value. **D**, Neurons that do not fit into one of the above classifications.

richness of the current environment. In our task, where offers are generated independently, the optimal strategy is to ignore previous offers. However, we found evidence that subjects incorporate the value of the previous offer into their strategy when considering the current offer. Specifically, we found a positive relationship between the size of previous offers and threshold for the current offer ($\beta = 0.1073$, $p = 0.0019$; Fig. 6E; significant in 1 subject and positive trend in the other).

The brain also encodes the value of the previous offer when evaluating the current offer. To examine population encoding of offer value, we regressed firing rate against previous offer value. Previous offer affected responses of 16.39% ($n = 20$ of 122) of neurons (Fig. 6F; Table 1). This proportion is significantly greater than chance ($p < 0.001$, two-sided binomial test). Larger previous offer increased the neural encoding strength of the current offer ($\beta = 4.391$, $p < 0.001$).

View number

Another environmental variable in the task is the position of the offer in the sequence of available offers (regardless of whether the subject has seen the offer already, as the subject is allowed to freely view each offer multiple times). We performed analyses on view numbers only up to the sixth view because each value beyond 6 constitutes <1% of the data (Table 2). View number changed subjects' behavioral thresholds: for both subjects, threshold for a given offer decreased with the number of views ($\beta = -0.0346$, $p = 0.001$; Fig. 7A). Because rejection occurred with low-quality offers, view number is, in essence, a proxy for poor within-trial environmental richness. Neurons in vmPFC encoded view number: 32.79% of cells ($n = 40$ of 122) significantly encoded the within-trial view number of an offer (this estimate is conservative; we regressed out the value of the current offer as this tends to decrease as view number increases). This proportion is significantly greater than chance ($p < 0.001$, two-sided binomial test). Higher offer view numbers were also associated with increased value coding strength ($\beta = 0.0278$, $p = 0.0175$; Fig. 7B).

Reward rate

The brain may keep track of recent reward rate: the values of the last several outcomes over time (see Materials and Methods). Reward rate fluctuates over the course of a behavioral session (Fig. 8A). Foragers may see this measure as a proxy for environmental richness. It should not,

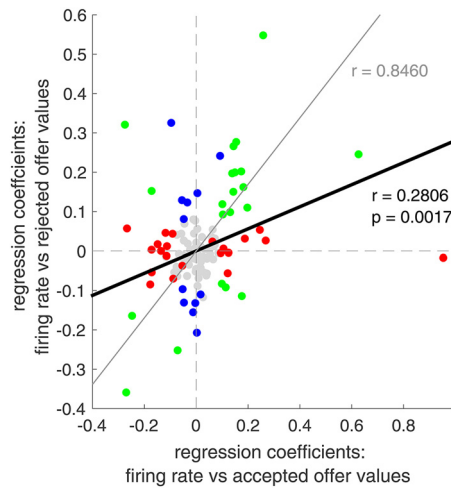


Figure 5. Population encoding for accepted versus rejected offers. Correlation between regression coefficients for accepted offers and rejected offers (black line, $r = 0.2806$) falls between 0 (x axis) and the perfectly correlated value, accounting for noise (gray line, $r = 0.8460$).

normatively, affect subjects' behavior thresholds, but it does. Threshold increased with the reward rate over the past 10 outcomes (after ignoring the most recent outcome and regressing out trial number, $\beta = 0.2313$, $p < 0.001$; Fig. 8*B*). We found that 32.79% ($n = 40$ of 122) of vmPFC neurons significantly encoded the reward rate of the past 10 values encountered. This proportion is significantly greater than chance ($p < 0.001$, two-sided binomial test).

There are other ways of computing reward rate, such as using a recency-weighted average (Constantino and Daw, 2015). Instead of computing reward rate as the average of the last several values, one can compute it as a weighted average where more recent outcomes loom larger than more distant outcomes (see Materials and Methods). Using this method, we found that 29.51% of cells significantly encode recency-weighted reward rate. While this result is not precisely the same as the number computed using the other method (i.e., 32.79%), they are very similar.

Spatial positions of current offers

Foragers make their decisions in the real world. Dealing with spatial information is essential to foraging decisions but is conceptually quite distinct from value. By using positions arrayed randomly across the screen, we were able to measure selectivity for saccadic position in a relatively unbiased manner. This provides information hitherto unavailable in vmPFC. Figure 9 shows heat maps indicating the average firing associated with the gaze position for six example neurons in our dataset. All six show spatial selectivity. A visual inspection of our entire dataset revealed that, to the extent that spatial selectivity exists, it is broad (i.e., associated with hemifields or quadrants, similar to posterior cingulate cortex) (Dean et al., 2004). We therefore assessed spatial selectivity by splitting the task display into four quadrants (top-right, top-left, bottom-left, bottom-right).

In our sample, 22.13% of vmPFC neurons ($n = 27$ of 122) showed quadrant spatial selectivity for current gaze position ($p < 0.05$, one-way ANOVA; see Materials and Methods). This proportion is greater than would be expected by chance (i.e., 5%; $p < 0.01$, two-sided binomial test). Comparing firing rate for fixation on the left- and right-hand side of the screen, we found a similar result: 18.85% of neurons carried left-right gaze position infor-

mation. This proportion is also greater than chance ($p < 0.01$, two-sided binomial test). Likewise, 13.93% of neurons had selectivity for up-versus-down halves of the screen (also greater than chance, $p < 0.01$, two-sided binomial test).

It does not appear that this spatial information came from a separate set of neurons from the neurons that encoded value. Instead, both variables appear to be multiplexed by a single set of neurons. Specifically, strength of encoding of spatial information correlated positively with strength of value encoding. Absolute (unsigned) regression coefficients of neurons encoding whether the current offer was on the right or left half of the screen correlated significantly with absolute regression coefficients of currently viewed value ($r = 0.2444$, CI 0.06975–0.4046, $p < 0.01$). Absolute regression coefficients of neurons encoding whether the current offer was on the top or bottom half of the screen also correlated significantly with absolute regression coefficients of currently viewed value ($r = 0.1938$, CI 0.0167–0.3592, $p = 0.0323$).

Valuation of space

It is conceivable that our subjects intrinsically assign more value to specific spatial positions and that our observed spatial tuning is not spatial at all, but an artifact of the link between value and space. Fortunately, due to the design of our task, any intrinsic value that subjects assign to space is directly measurable in terms of behavioral threshold at each spatial position. Over all behavioral sessions, we find small effects of space on threshold. For example, Subject J's threshold for offers in the lower visual field (48.74%) is lower than threshold for offers in the upper visual field (51.02%). This difference is significant ($p < 0.001$), which is not surprising given our very large dataset. Subject T's threshold for offers in the lower visual field (57.35%) is significantly greater than threshold for offers in the upper visual field (54.35%, also significant, $p < 0.001$). Despite the large dataset sizes, we do not find corresponding effects for left/right choices: Subject J's thresholds for offers in the left visual field (49.82%) and right visual field (49.18%) do not significantly differ ($p = 0.1314$), and Subject T's thresholds for offers in the left visual field (56.70%) and right visual field (56.02%) do not significantly differ ($p = 0.0797$). However, we do still find spatial selectivity for left/right choices. These findings, the lack of value difference on the vertical dimension and the weak in magnitude (albeit significant) value difference on the horizontal dimension, suggest that the spatial selectivity we see is not a byproduct of differential value assignment to spatial positions.

While this result is suggestive, a more sensitive test is a cell-by-cell analysis. To assess whether valuation of offers was correlated with spatial location, we next examined whether value encoding and firing rate show a correlation associated with their corresponding spatial locations. That is, we looked to see whether cells adjusted the strength of their value encoding (changed the regression coefficient of firing rate vs offer value) when those offers were located in cells' preferred spatial locations. We divided the screen into 84 rectangular segments (a 7×12 grid, as in Fig. 9). For each segment, we computed a regression coefficient for offer value for offers located within it, as well as an average firing rate. For each cell, we correlated these two factors. We found that only 12.30% ($n = 15$ of 122) of cells had a significant correlation ($p < 0.05$) between value encoding and firing rate. That is, for most cells, the strength of value encoding was unrelated to spatial tuning.

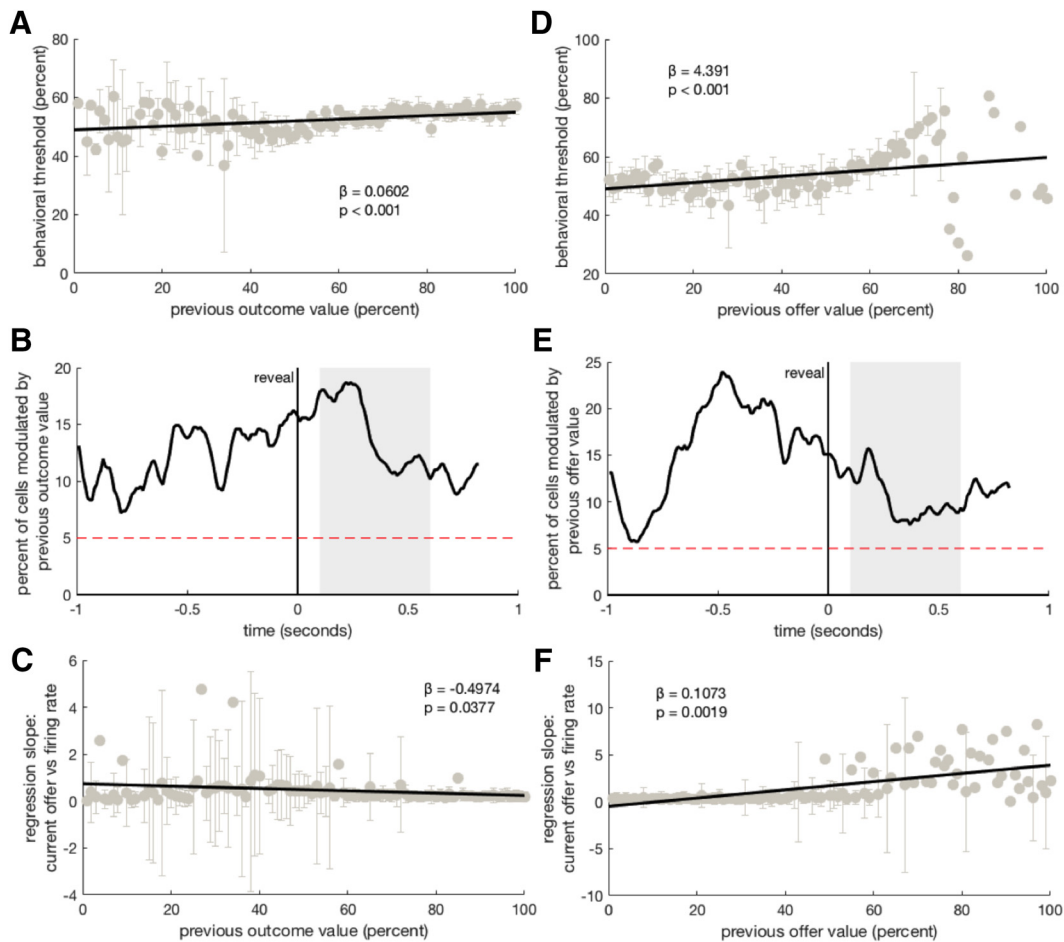


Figure 6. Effects of previous outcome value and previous offer value on behavior and neurons. **A**, Behavioral threshold increases immediately after trials with higher outcomes. Error bars indicate the SE of the sigmoidal curve fit used to produce each threshold. Error bars of 4 points have been removed for visualization purposes because they are larger than the height of the plot. **B**, Proportion of neurons in our entire sample whose firing rate responses were affected by the previous outcome value as a function of time relative to offer period. Proportions are calculated with a 200 ms sliding boxcar. Line is smoothed over 5 data points. **C**, Neural encoding of the current offer decreases with higher previous outcomes. Neural encoding slopes are the β values for the correlation between firing rate and current offer values, calculated for each current offer value in steps of 1%. Error bars indicate the SD of the population of absolute coefficients averaged to produce each point. Error bars from 8 points have been removed because they are larger than the height of the plot. **D**, Behavioral threshold increases after rejecting higher offers within trial. Error bars from 16 points have been removed. **E**, Proportion of neurons in our entire sample whose firing rate responses were affected by previous offer value as a function of time relative to offer period, calculated as in **B**. **F**, Neural encoding strength increases with higher previous offers. Error bars from 37 points have been removed (see Materials and Methods).

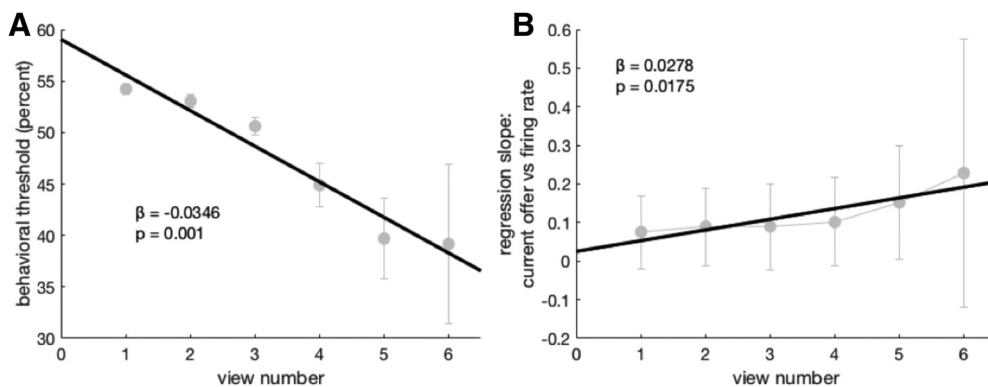


Figure 7. Effects of view number on behavioral threshold and neural encoding. **A**, Behavioral threshold decreases with view number. **B**, Neural encoding strength increases with view number.

Radial measure of tuning for gaze position

To study spatial selectivity for gaze position in vmPFC with greater precision, we next examined radial tuning functions. We measured each neuron’s selectivity for position of current fixation relative to the center of the computer monitor (i.e., straight

ahead). We fit von Mises (circular Gaussian) distributions and compared them with the uniform model (i.e., lack of tuning) (Yoo et al., 2018). We also used a cosine function, which is quantitatively similar to a von Mises, and obtained nearly identical results (data not shown).

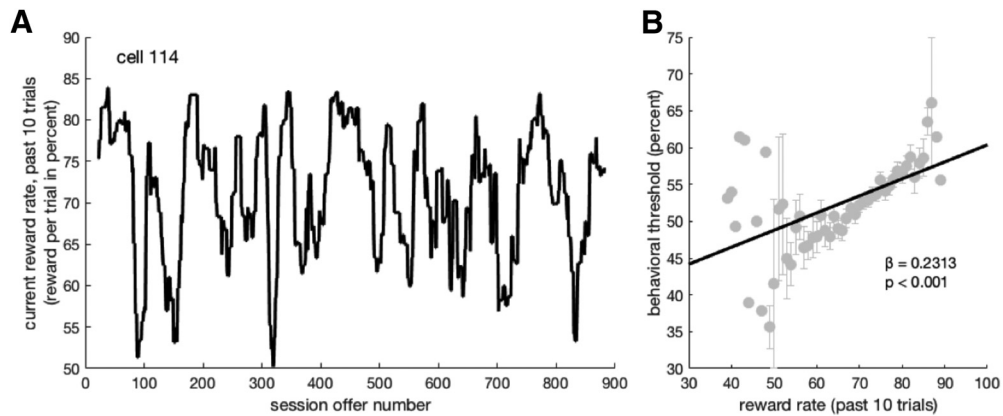


Figure 8. Effect of reward rate on behavioral threshold. **A**, Reward rate fluctuates over the course of a typical behavioral session. **B**, Behavioral threshold increases with reward rate over the past 10 trials. Error bars from 11 points have been removed because they are larger than the height of the plot (these points are at the highest and lowest reward rates, which are the rarest and have the fewest samples). The regression remains significant upon removal of these points ($p < 0.001$).

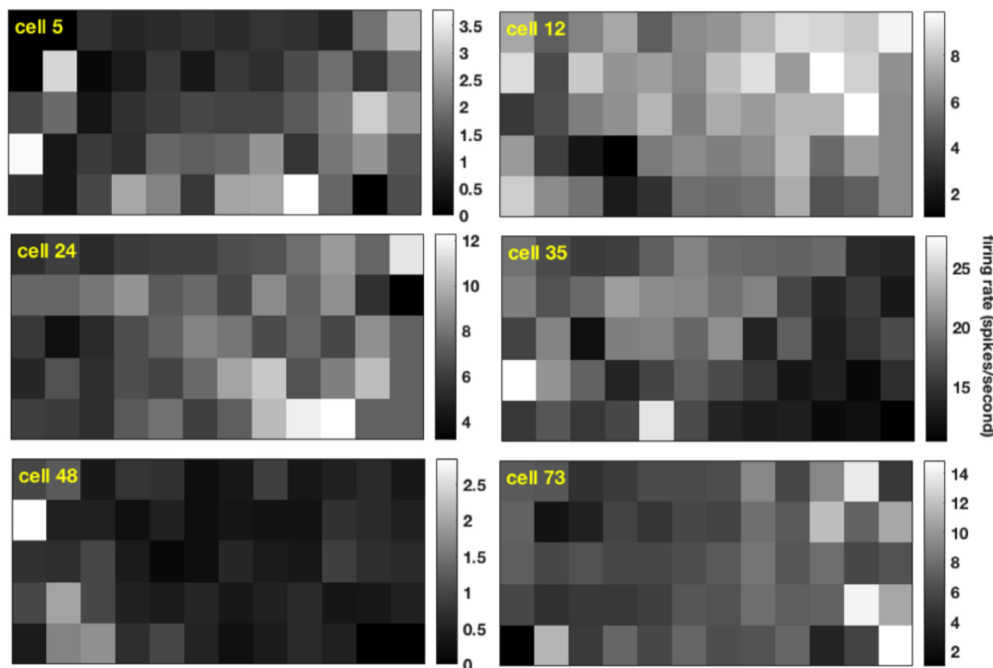


Figure 9. Six example neurons that demonstrate spatial selectivity. Grayscale in heatmaps represents mean firing rate (spikes/s) for each cell when offers were present within each rectangular bin (as determined by the coordinates of the center of the offer stimulus) during the analysis epoch (100–600 ms after offer reveal). Bins were determined before analysis as having a width of one-twelfth of the width of the task screen and a height of one-seventh of the height of the task screen. For display purposes, the first and last bin rows have been eliminated because no offers were centered in those bins (or the tops or bottoms of the images would appear cutoff). Bins in which the subject did not view any offers during that cell's session are assigned a firing rate equivalent to mean of all the other bins.

Figure 10A shows the responses of an example neuron. During the fixation epoch (the 400 ms before the values were revealed), this neuron fired more strongly (5.6 spikes/s) when the subject fixated on targets to the lower field than when subjects fixated on targets to the rest of the field (2.9 spikes/s, $t = 3.3304$, $p < 0.001$). This neuron's firing rate was better fit with a von Mises function than a uniform one, indicating that it is selective for gaze angle (parameters = 0.6789, 64.6084, 1.4878, and 0.2327; see Materials and Methods; BIC decrease = 23). In the population of neurons, a large proportion (34.5%, $n = 42$ of 122) showed sensitivity to gaze position. This proportion is greater than that expected by chance (two-sided binomial test, $p < 0.0001$). The average effect size (measured with BIC decrease) was 30.0 ± 50.5 . A similar proportion of neurons (36%, $n = 44$ of 122) was tuned to gaze

angle of the fixated diamonds using a direct measure of mutual information between the spikes and the angle following adaptive smoothing (method developed by Skaggs et al., 1993, 1996) (see Materials and Methods).

We then repeated this analysis over time relative to target onset. To gain more insight into temporal dynamics of spatial selectivity, we used a shorter time window (sliding window of size 300 ms): sliding width of 100 ms (Fig. 10B). We found that the incidence of gaze angle encoding rises upon fixation and then tapers following the decision to reject an offer. We then expanded this analysis using both past (Fig. 10C) and future (Fig. 10D) gaze position, excluding the last and first viewed diamond in each trial accordingly. These data demonstrate that selectivity for past position is maintained weakly during fixation of the present one,

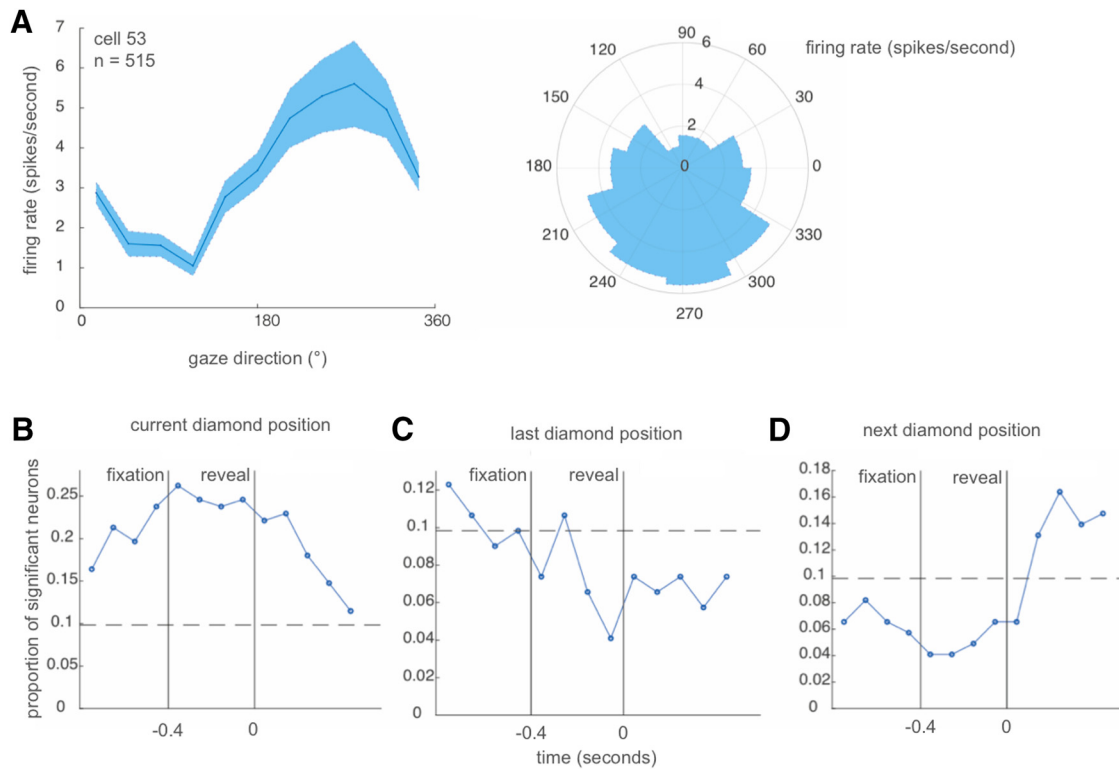


Figure 10. Tuning for the position of the field. **A**, Left, Example neuron tuning curve. Right, Angular probability densities representing firing rates. **B**, The proportion of neurons with significant tuning for the current diamond position during the time course of diamond fixation and bar viewing. **C**, Same as **B**, but for the previous diamond position. **D**, Same as **B**, but for the next diamond position.

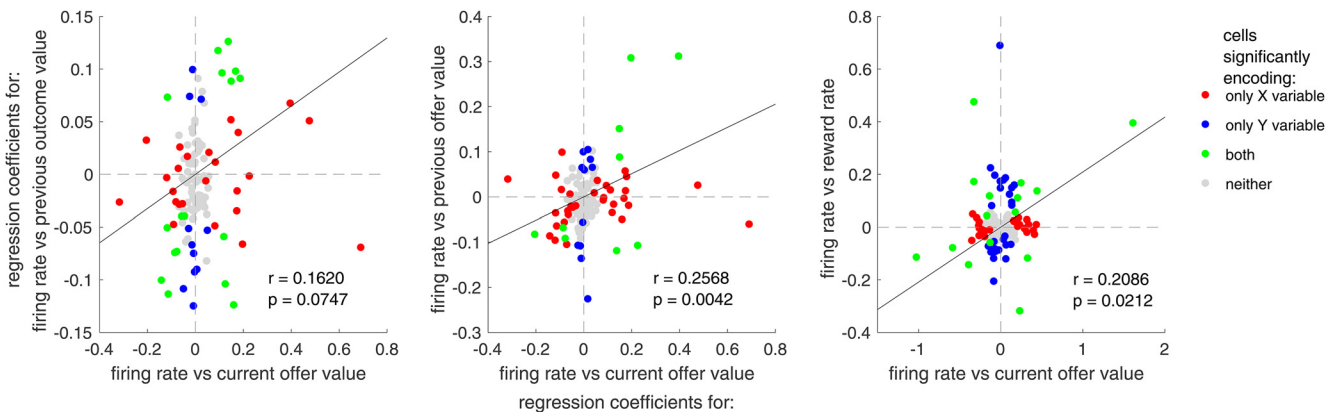


Figure 11. Population encoding for current offer value, compared with population encoding for previous outcome value, previous offer value, and reward rate. Regression coefficients for current offer value are not significantly correlated with those of previous outcome values but are slightly positively correlated with those of previous offer value and reward rates.

and that selectivity for the future position begins immediately after the present fixation ends.

Encoding for environmental variables is weakly correlated with value encoding

We next examined how these codes related to each other. In our sample, 71.3% of cells ($n = 87$ of 122) significantly encoded at least one variable of the five variables we examine (offer value, previous trial outcome, previous offer, view number, and reward rate; this analysis uses only one analysis epoch to reduce the chance of false positives). From this larger population, 28.69% of cells ($n = 35$ of 122) encode exactly two variables, 17.21% of cells ($n = 21$ of 122) encode exactly three variables, 6.56% of cells ($n =$

8 of 122) encode exactly four variables, and 1.64% of cells ($n = 2$ of 122) encode all five.

It is possible that cells whose firing rates have a positive relationship with current offer value have a negative relationship with previous offer value. We assessed the relationship between encoding of variables by correlating regression coefficients of firing rate versus different variables (see Materials and Methods). First, encoding of offer value and encoding of previous outcome are weakly and insignificantly correlated, providing no evidence for a relationship between how cells encode current offer value and previous trial outcome (Fig. 11). However, current offer value encoding and previous offer value encoding are correlated at $r = 0.2568$ ($p = 0.004$; Fig. 11), indicating that cells encode

current offer value and the previous offer value in a slightly similar way. Similarly, regression coefficients of offer value encoding and reward rate encoding are positively correlated at $r = 0.2086$ (Fig. 11). This value is significant at $p = 0.0211$, indicating that cells may also encode offer value and reward rate in a similar way.

Discussion

We examined neural responses in a putative primate homolog of human vmPFC as macaques performed a novel 4 or 7 option search task. A key element of the task was that subjects could stop searching whenever they chose to, meaning that they had to trade off the potential gain of foraging against the time costs of doing so. Unlike some other foraging tasks, the optimal strategy does not involve a variable foreground/background comparison; instead, the best strategy is to learn a single threshold and use it consistently. Subjects' behavior was well described by just such a thresholding procedure: average thresholds were approximately reward-maximizing (with a bias toward overaccepting) (compare Blanchard and Hayden, 2014). Neurons strongly encoded values of offers, thus confirming its value-related role (Rushworth et al., 2011; Levy and Glimcher, 2012; Bartra et al., 2013; Clithero and Rangel, 2014). However, they also encoded two other categories of broader environmental variables: (1) parameters that affected accept-reject decisions that were conceptually distinct from immediate offer value; and (2) information about the spatial positions of offers on the screen.

Our results have important implications about the function of this region. The role of the vmPFC is not limited to encoding the immediately available value of options. It clearly encodes other factors, which we call environmental variables. We found that several of these encoded variables (the size of the previous trial's outcome, the size of the previous offer within a trial, the view number within a trial, and the reward rate over the past 10 trials) correlate with changes in the subjects' primary strategy during the task. That is, subjects alter their behavioral choice thresholds as environmental variables change. These correlations are indicative of the importance of environmental variables in foraging-based decision-making; and because we did not perform causal manipulations, it would be interesting to investigate these relationships in greater detail. Whether these different environmental variables are considered to be different kinds of value, rather than considering them as nonvalue factors, is a separate question also worthy of future investigation.

Furthermore, our results relate to important fundamental debates about the function of vmPFC (Delgado et al., 2016). A major hope in neuroeconomics is to identify a single scalar value signal (Rangel et al., 2008; Padoa-Schioppa, 2011; Levy and Glimcher, 2012). This signal should be one that transcends multiple dimensions along which options can vary, and that can encode values of categorically different types of options. Finding the neural location of this hypothesized signal has been a major goal of neuroeconomics; a large amount of recent work (but by no means all of it) has narrowed in on the vmPFC (Rushworth et al., 2011; Bartra et al., 2013). However, other research casts doubt on the existence of a pure value signal (e.g., Cisek and Kalaska, 2010; Smith et al., 2010; Wallis and Kennerley, 2010; Vlaev et al., 2011; Rushworth et al., 2012; Yoo and Hayden, 2018). Work in this vein highlights the diversity of factors contributing to value-based choices and, in some cases, the evidence that there are multiple competing signals. Our study is consistent with recent work indicating that vmPFC carries cognitive variables other than immediate offer value (Delgado et al., 2016). For example, hemodynamic responses in vmPFC track decisional confidence

(Lebreton et al., 2015). They also appear to track decisional conflict (Shenhav et al., 2017) and memory integration (Schlichting and Preston, 2015). Even value variables in vmPFC may not be integrated into a coherent integrated set (Sescousse et al., 2010; Smith et al., 2010; Watson and Platt, 2012).

However, the level of integration between environmental variables and value within vmPFC is less clear. We found small correlations between neural encoding of current offer value and previous offer value, as well as current offer value and reward rate. These results provide evidence for weak integration of environmental signals with value within vmPFC. On the other hand, we found no correlation between current offer value previous outcome value. Thus, it remains to be seen whether environmental variables remain as disparate signals within vmPFC or are integrated with pure offer value into a single scalar quantity: what we call the "unitary value signal" theory of vmPFC function. In this view, the major, perhaps even sole, role of vmPFC is to carry a domain general value signal that fully incorporates all factors that influence value (Plassmann et al., 2007; Chib et al., 2009; Kable and Glimcher, 2009; Levy and Glimcher, 2012; Delgado et al., 2016). This signal incorporates all information that is relevant to choice but expressed as a single scalar quantity; vmPFC outputs are, by this view, postevaluative. While we cannot speak to whether this "unitary value signal" exists within vmPFC, we can speak to the postevaluative nature of the signals within vmPFC. Our data suggest that vmPFC carries a "mid-decisional" value signal: current offer values that are ultimately accepted versus those that are ultimately rejected are encoded neither perfectly similarly nor completely differently. This places vmPFC somewhere in the middle of the decision process: it is neither completely blind to whether an offer will ultimately be accepted, nor does it fully distinguish between accepted and rejected offers. Thus, we do not view vmPFC as purely postevaluative. Whether this is the case because vmPFC lacks an integrated unitary value signal or not is an interesting question, but not the focus of our current investigation.

The vmPFC is often classified as one of the core value regions, or even as a site of common currency value representation. Such representations are presumed to be amodal and thus not biased by spatial information (Padoa-Schioppa and Assad, 2006; Rangel and Hare, 2010; Padoa-Schioppa, 2011). The presence of clear spatial information here, then, may be somewhat surprising. One previous study, from our laboratory, reported spatial selectivity in vmPFC, although the task was much simpler, tuning was weaker there, and encoding was associated with actions and not gaze position (Strait et al., 2016). A limitation of that study was that spatial location may be confounded with object selectivity (Padoa-Schioppa and Cai, 2011; Yoo et al., 2018). Our present work gets around this problem by using new positions on each trial so that subjects cannot readily link locations to stimuli across trials. We thus confirm our previous finding and extend it to more strongly quantify the effect and to demonstrate coding for gaze position. One speculative interpretation of this result is that vmPFC may serve as an abstract economic salience map (i.e., as a map of potential values in the peripersonal field).

We view foraging as a special type of economic choice that is one for which the brain has evolved and is thus adapted (Passingham and Wise, 2012; Pearson et al., 2014; Hayden, 2016, 2018). Our task is a foraging task in the sense that it is inspired by natural foraging problems; the choices our subjects face are explicitly accept-reject, and the choice to accept or reject an option affects the suite of options available on future decisions. We take the perspective that the animal is seeking to maximize reward per

unit time while also keeping mental effort low. One nearly optimal strategy in such tasks is to choose a threshold and stick to it regardless of recent outcomes if the decision-maker believes the environment is potentially variable, they may adjust their strategy in a lawful manner in response to recent outcomes. This is precisely the pattern we observe: behavior depends on both the value on the trial and the other factors that also alter choice. We have previously argued that laboratory tasks should be structured around the types of decisions animals have evolved to perform rather than types that are mathematically convenient (Pearson et al., 2010; Hayden, 2018). The present results highlight several benefits of doing so. First, we were able to observe stronger and more frequent value-related coding than previous studies of this region; we presume that the naturalness of the task make it more efficient at driving neural responses. Second, by focusing on single accept-reject decisions, we were able to get around ambiguities associated with binary choices, and the difficult-to-measure shifts in attention they bring (Rich and Wallis, 2016; Hayden, 2018). Third, the structure of our task brings with it a natural case of within-trial changes in threshold, which allows us to measure their neural correlates.

References

- Azab H, Hayden BY (2017) Correlates of decisional dynamics in the dorsal anterior cingulate cortex. *PLoS Biol* 15:e2003091.
- Azab H, Hayden BY (2018) Correlates of economic decisions in the dorsal and subgenual anterior cingulate cortices. *Eur J Neurosci* 47:979–993.
- Bartra O, McGuire JT, Kable JW (2013) The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage* 76:412–427.
- Blanchard TC, Hayden BY (2014) Neurons in dorsal anterior cingulate cortex signal postdecisional variables in a foraging task. *J Neurosci* 34:646–655.
- Blanchard TC, Hayden BY (2015) Monkeys are more patient in a foraging task than in a standard intertemporal choice task. *PLoS One* 10:e0117057.
- Blanchard TC, Wilke A, Hayden BY (2014) Hot-hand bias in rhesus monkeys. *J Exp Psychol Anim Learn Cogn* 40:280–286.
- Blanchard TC, Strait CE, Hayden BY (2015) Ramping ensemble activity in dorsal anterior cingulate neurons during persistent commitment to a decision. *J Neurophysiol* 114:2439–2449.
- Boorman ED, Behrens TE, Woolrich MW, Rushworth MF (2009) How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. *Neuron* 62:733–743.
- Boorman ED, Rushworth MF, Behrens TE (2013) Ventromedial prefrontal and anterior cingulate cortex adopt choice and default reference frames during sequential multi-alternative choice. *J Neurosci* 33:2242–2253.
- Bouret S, Richmond BJ (2010) Ventromedial and orbital prefrontal neurons differentially encode internally and externally driven motivational values in monkeys. *J Neurosci* 30:8591–8601.
- Calhoun AJ, Hayden BY (2015) The foraging brain. *Curr Opin Behav Sci* 5:24–31.
- Chen X, Stuphorn V (2015) Sequential selection of economic good and action in medial frontal cortex of macaques during value-based decisions. *Elife* 4:e09418.
- Chib VS, Rangel A, Shimojo S, O'Doherty JP (2009) Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *J Neurosci* 29:12315–12320.
- Cisek P, Kalaska JF (2010) Neural mechanisms for interacting with a world full of action choices. *Annu Rev Neurosci* 33:269–298.
- Clithero JA, Rangel A (2014) Informatic parcellation of the network involved in the computation of subjective value. *Soc Cogn Affect Neurosci* 9:1289–1302.
- Constantino SM, Daw ND (2015) Learning the opportunity cost of time in a patch-foraging task. *Cogn Affect Behav Neurosci* 15:837–853.
- Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ (2006) Cortical substrates for exploratory decisions in humans. *Nature* 441:876–879.
- Dean HL, Crowley JC, Platt ML (2004) Visual and saccade-related activity in macaque posterior cingulate cortex. *J Neurophysiol* 92:3056–3068.
- Delgado MR, Beer JS, Fellows LK, Huettel SA, Platt ML, Quirk GJ, Schiller D (2016) Viewpoints: dialogues on the functional role of the ventromedial prefrontal cortex. *Nat Neurosci* 19:1545–1552.
- Gusnard DA, Raichle ME (2001) Searching for a baseline: functional imaging and the resting human brain. *Nat Rev Neurosci* 2:685–694.
- Hare TA, Camerer CF, Rangel A (2009) Self-control in decision-making involves modulation of the vmPFC valuation system. *Science* 324:646–648.
- Hayden BY (2016) Time discounting and time preference in animals: a critical review. *Psychonom Bull Rev* 23:39–53.
- Hayden BY (2018) Economic choice: the foraging perspective. *Curr Opin Behav Sci* 24:1–6.
- Hayden BY, Gallant JL (2013) Working memory and decision processes in visual area V4. *Front Neurosci* 7:18.
- Hayden BY, Moreno-Bote R (2018) A neuronal theory of sequential economic choice. *Brain Neurosci Adv* 2:2398212818766675.
- Hayden BY, Pearson JM, Platt ML (2011) Neuronal basis of sequential foraging decisions in a patchy environment. *Nat Neurosci* 14:933–939.
- Hills TT, Todd PM, Lazer D, Redish AD, Couzin ID (2015) Exploration versus exploitation in space, mind, and society. *Trends Cogn Sci* 19:46–54.
- Kable JW, Glimcher PW (2007) The neural correlates of subjective value during intertemporal choice. *Nat Neurosci* 10:1625–1633.
- Kable JW, Glimcher PW (2009) The neurobiology of decision: consensus and controversy. *Neuron* 63:733–745.
- Kolling N, Behrens TE, Mars RB, Rushworth MF (2012) Neural mechanisms of foraging. *Science* 336:95–98.
- Kolling N, Wittmann M, Rushworth MF (2014) Multiple neural mechanisms of decision making and their competition under changing risk pressure. *Neuron* 81:1190–1202.
- Lebreton M, Abitbol R, Daunizeau J, Pessiglione M (2015) Automatic integration of confidence in the brain valuation signal. *Nat Neurosci* 18:1159–1167.
- Levy DJ, Glimcher PW (2012) The root of all value: a neural common currency for choice. *Curr Opin Neurobiol* 22:1027–1038.
- Lim SL, O'Doherty JP, Rangel A (2011) The decision value computations in the vmPFC and striatum use a relative value code that is guided by visual attention. *J Neurosci* 31:13214–13223.
- Mars RB, Verhagen L, Gladwin TE, Neubert FX, Sallet J, Rushworth MF (2016) Comparing brains by matching connectivity profiles. *Neurosci Biobehav Rev* 60:90–97.
- Mobbs D, Trimmer PC, Blumstein DT, Dayan P (2018) Foraging for foundations in decision neuroscience: insights from ethology. *Nat Rev Neurosci* 19:419–427.
- Monosov IE, Hikosaka O (2012) Regionally distinct processing of rewards and punishments by the primate ventromedial prefrontal cortex. *J Neurosci* 32:10318–10330.
- Padoa-Schioppa C (2011) Neurobiology of economic choice: a good-based model. *Annu Rev Neurosci* 34:333–359.
- Padoa-Schioppa C, Assad JA (2006) Neurons in the orbitofrontal cortex encode economic value. *Nature* 441:223–226.
- Padoa-Schioppa C, Cai X (2011) The orbitofrontal cortex and the computation of subjective value: consolidated concepts and new perspectives. *Ann N Y Acad Sci* 1239:130–137.
- Passingham RE, Wise SP (2012) The neurobiology of the prefrontal cortex: anatomy, evolution, and the origin of insight, Vol 50. Oxford: Oxford UP.
- Paxinos G, Huang XF, Toga AW (2000) The rhesus monkey brain in stereotaxic coordinates. San Diego: Academic.
- Pearson JM, Hayden BY, Raghavachari S, Platt ML (2009) Neurons in posterior cingulate cortex signal exploratory decisions in a dynamic multi-option choice task. *Curr Biol* 19:1532–1537.
- Pearson JM, Watson KK, Platt ML (2014) Decision making: the neuroethological turn. *Neuron* 82:950–965.
- Pearson JM, Hayden BY, Platt ML (2010) Explicit information reduces discounting behavior in monkeys. *Front Psychol* 1:237.
- Pezzulo G, Cisek P (2016) Navigating the affordance landscape: feedback control as a process model of behavior and cognition. *Trends Cogn Sci* 20:414–424.
- Plassmann H, O'Doherty J, Rangel A (2007) Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J Neurosci* 27:9984–9988.
- Rangel A, Hare T (2010) Neural computations associated with goal-directed choice. *Curr Opin Neurobiol* 20:262–270.

- Rangel A, Camerer C, Montague PR (2008) A framework for studying the neurobiology of value-based decision making. *Nat Rev Neurosci* 9:545–556.
- Rich EL, Wallis JD (2016) Decoding subjective decisions from orbitofrontal cortex. *Nat Neurosci* 19:973–980.
- Rushworth MF, Noonan MP, Boorman ED, Walton ME, Behrens TE (2011) Frontal cortex and reward-guided learning and decision-making. *Neuron* 70:1054–1069.
- Rushworth MF, Kolling N, Sallet J, Mars RB (2012) Valuation and decision-making in frontal cortex: one or many serial or parallel systems? *Curr Opin Neurobiol* 22:946–955.
- Schlichting ML, Preston AR (2015) Memory integration: neural mechanisms and implications for behavior. *Curr Opin Behav Sci* 1:1–8.
- Sescousse G, Redouté J, Dreher JC (2010) The architecture of reward value coding in the human orbitofrontal cortex. *J Neurosci* 30:13095–13104.
- Shenhav A, Musslick S, Lieder F, Kool W, Griffiths TL, Cohen JD, Botvinick MM (2017) Toward a rational and mechanistic account of mental effort. *Annu Rev Neurosci* 40:99–124.
- Skaggs WE, McNaughton BL, Gothard KM (1993) An information-theoretic approach to deciphering the hippocampal code. In: *Advances in neural information processing systems*, pp 1030–1037. Neural Information Processing Systems Conference, San Francisco.
- Skaggs WE, McNaughton BL, Wilson MA, Barnes CA (1996) Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences. *Hippocampus* 6:149–172.
- Sleezer BJ, Castagno MD, Hayden BY (2016) Rule encoding in orbitofrontal cortex and striatum guides selection. *J Neurosci* 36:11223–11237.
- Smith DV, Hayden BY, Truong TK, Song AW, Platt ML, Huettel SA (2010) Distinct value signals in anterior and posterior ventromedial prefrontal cortex. *J Neurosci* 30:2490–2495.
- Stephens DW, Krebs JR (1986) *Foraging theory*. Princeton, NJ: Princeton UP.
- Strait CE, Blanchard TC, Hayden BY (2014) Reward value comparison via mutual inhibition in ventromedial prefrontal cortex. *Neuron* 82:1357–1366.
- Strait CE, Sleezer BJ, Blanchard TC, Azab H, Castagno MD, Hayden BY (2016) Neuronal selectivity for spatial positions of offers and choices in five reward regions. *J Neurophysiol* 115:1098–1111.
- Strait CE, Sleezer BJ, Hayden BY (2015) Signatures of value comparison in ventral striatum neurons. *PLoS Biol* 13:e1002173.
- Vlaev I, Chater N, Stewart N, Brown GD (2011) Does the brain calculate value? *Trends Cogn Sci* 15:546–554.
- Wallis JD, Kennerley SW (2010) Heterogeneous reward signals in prefrontal cortex. *Curr Opin Neurobiol* 20:191–198.
- Wang MZ, Hayden BY (2017) Reactivation of associative structure specific outcome responses during prospective evaluation in reward-based choices. *Nat Commun* 8:15821.
- Watson KK, Platt ML (2012) Social signals in primate orbitofrontal cortex. *Curr Biol* 22:2268–2273.
- Wise SP, Murray EA, Gerfen CR (1996) The frontal cortex-basal ganglia system in primates. *Crit Rev Neurobiol* 10:3.
- Wunderlich K, Rangel A, O'Doherty JP (2009) Neural computations underlying action-based decision making in the human brain. *Proc Natl Acad Sci U S A* 106:17199–17204.
- Yoo SB, Hayden BY (2018) Economic choice as an untangling of options into actions. *Neuron* 99:434–447.
- Yoo SB, Sleezer BJ, Hayden BY (2018) Robust encoding of spatial information in orbitofrontal cortex and striatum. *J Cogn Neurosci* 30:898–913.