Environmental contributions to value computations and population dynamics in ventromedial prefrontal cortex

A DISSERTATION SUBMITTED TO THE FACULTY OF THE UNIVERSITY OF MINNESOTA BY

Priyanka Mehta

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Adviser: Benjamin Hayden
June 2021

Acknowledgments

First, I would like to acknowledge the VIPs (Very Important Primates) who participated in the tasks described below: Joker, Timon, Vader, and Snoopy, without whom none of this would be possible. Their cooperation, intelligence, and heart will be forever appreciated.

Next, I would like to acknowledge members (current and *former)* of the Hayden lab: *Habiba Azab*, Tyler Cash-Padgett, *Becket Ebitz, Benjamin Eisenreich, Hannah Lee*, *Giuliana LoConte, Marc Mancarella, Meghan Castagno Pesce, Maya Wang*, and *Seng Bum Michael Yoo* for their contributions mentally, physically, or both to any and all work I conducted within the lab. This also includes the primates with whom I did not directly work but whose company I still cherished: Batman, Calvin, Hobbes, Kirk, Pumbaa, Spock, and Woodstock.

I thank Giuliana LoConte and Meghan Castagno Pesce in particular for assisting with data collection.

I thank Seng Bum Michael Yoo in particular for mentorship and guidance in the creation of the manuscript that forms Chapter II of this document.

In addition I thank the (current and *former*) members of the Hayden-Zimmerman lab: Indira Conover, *Jen Holmberg*, Brenna Knaebe, Roberto Lopez Cervera, David Maisson, Matthew Maple, Ben Voloh, and Mrunal Zambre, for their ongoing support of the lab.

I extend further gratitude to the (now former) undergraduates who provided valuable insight into many tasks and contributed significantly to all experiments: Kyle Edmonston, Rachel Knoebl, Colin Meyer, Preeta Pavagadhi, Afra Suri, and Sydney Walsh.

I am additionally grateful for the University of Minnesota Graduate Program in Neuroscience entering class of 2017: Chloe Cable, Tyler Cash-Padgett, Adriana Cushnie, Margot DeBaker, Armani Del Franco, Hilary Handler, Haleigh Mulholland, Wenhui Qu, Alexandra Scott, Carrie Sheeler, Bethany Stieve, Nolan Trevino, Brian Trieu, Sarah West, and Nicole Zarate; who all contributed to a particularly supportive and pleasant environment in the program.

Last but not least, I acknowledge my helpful and supportive advisor, Ben Hayden, and my valuable advisory committee: Geoff Ghose (chair), Ben Saunders, Iris Vilares, and Jan Zimmerman.

Table of contents

List of Tables

List of figures

Introduction

On a hot summer day, there's nothing nicer than giving a kid on the roadside a dollar in exchange for an ice cold cup of lemonade. That is, unless you just turned down a fifty-cent cup of lemonade.

From a traditional neuroscience and economics perspective, the disappointment you feel when you pay the full dollar for that cup of lemonade is quantifiable: you've missed out on an offer that was twice as good as the one you ended up with. Such perspectives assume humans are rational in their decisions (Simon, 1979; Tversky & Kahenman, 1981; Heracleous, 1994; Lee 2006), and as such, the value of items with monetary value is just that - their monetary value. Calculations involving such items are a mere matter of mathematics. However, our brains did not evolve in a world of mathematical equations, or rather, any form of modern currency at all. How, then, can we expect our brains to contain built-in currency-computing engines? The early primate brain had access to no lemonade stands; no dollar values were attributed to its food sources. Instead, early humans then (and other primates now), obtained food a slightly different way: they foraged.

Foraging theory is an approach to studying decision making where we assume that because the brain evolved in a naturalistic environment, it is built to make decisions in naturalistic situations (Stephens & Krebs, 1986; Calhoun & Hayden, 2015; Mobbs et. al. 2018). We as neuroscientists should study decision making through the foraging lens: acknowledging that the environment makes a significant impact on our evaluation of food items, and that we can find strong signals of this process in our brain activity. Specifically, the rate of reward consumption, a core aspect of the foraging environment, is particularly fundamental to the neuroeconomic decision making process. Signatures of behaviorally relevant reward rate computations are central to the processing of the ventromedial prefrontal cortex (vmPFC).

In examining some of the ways in which neuroeconomic decision making is studied, we can see how environmental considerations are often ignored by design, and how to incorporate them in experiments. Take the universal task amongst primates, human and non-human, of selecting a beverage. Traditional decision-making experiments follow a particular format: have the subject choose between two simultaneously presented options. For example, an experimenter might ask a subject, "Would you like a glass of lemonade for one dollar, or a glass of lemonade for fifty cents?" This is not the only form such a question can take; there can also be two qualitatively different items pitted against each other (e.g. "Would you like lemonade or cola?"), or the items can vary along several dimensions (e.g. "Would you like a large glass of

lemonade for one dollar, or a small glass of cola for fifty cents?") In all cases, the neuroscientist wishes to uncover how the choice is made.

One popular approach to answering this question is the idea that our brains translate options into a "common currency" - a way to compare different modalities along the same scale by combining multiple attributes of an item into a single value (Christopoulos & Schrater, 2015; Levy & Glimcher, 2012; Padoa-Schioppa, 2011). Take for example the decision between a glass of lemonade and a glass of cola. Both beverages may vary in taste, volume, temperature, price, and more - all factors that you, the beverage-purchaser, must take into account in order to select one or the other. According to the common currency approach, lemonade and cola are fundamentally distinct substances that cannot be compared outright. Instead, you must translate their various traits into a common "space of goods." (Padoa-Schioppa, 2011). In this way, a medium-sized, lukewarm, moderately sweet, cheap cup of lemonade may be converted in-brain to a value of 5 *beverage bucks*, while a large-sized, ice-cold, overly-sweet, overpriced cup of cola may be converted in-brain into 4 *beverage bucks*. The internal computations involved in determining the beverages' value in beverage bucks can vary infinitely from person to person, as long as the computation happens independently for each item (Padoa-Schioppa, 2011). Now that the values for the items have been determined in a common currency, the brain can compare the two values and make the decision: in this case, the lemonade.

Neural correlates of the common currency explanation for decision making have been discovered largely within the prefrontal cortex (PFC). Schioppa and Assad (2006) tested the approach by instructing subjects to compare several different juice flavors and rank them against each other. Once the subjects came up with a "price" for each item, the experimenters looked for brain activity correlating with the subjects' subjective prices. Indeed, they found such brain activity in the orbitofrontal cortex (OFC). Similar signals have also been discovered in the ventromedial prefrontal cortex (Levy & Glimcher, 2012; Padoa-Schioppa, 2011; Chib et al, 2009; Kable & Glimcher, 2009) - a region directly adjacent to the OFC in primates (Passingham & Wise, 2012). The vmPFC in particular has generated considerable interest from neuroeconomics researchers due to its strong response to rewarding food items (Strait et al., 2014; Lim et al, 2011), monetary gain (Vassena et al, 2014; Kolling et al, 2012), and emotional value (Winecoff et al, 2013). In addition, there is evidence that vmPFC encodes value independent of modality (Batra et al, 2013; Metereau & Dreher, 2015).

A value signal independent of modality is a point in favor of the approach that items of varying characteristics are translated into a common currency in order to be compared. However, studying value in this way directs attention away from the fact that value itself is not

one thing; rather, value representations in the brain change constantly depending on numerous factors, such as other people around you, and the quality of the surrounding environment (Stephens, 1989).

Unfortunately, it is easy to eliminate most of these various components of value in a lab setting. When you are given a tube that squirts controlled amounts of lemonade into your mouth, the variables of time, effort, cost, and even temperature can disappear from the equation. Even if the usage of a mental currency fits the neural data obtained under these circumstances, the meaningfulness of this result becomes questionable. An excellent solution to this problem lies in foraging theory.

Using foraging theory to study neuroeconomic decision making solves one of the major problems with the traditional neuroeconomics task: it provides an environmental context for the decision. The common currency is an interesting way to study how value is computed, but is almost always studied in the context of a forced-choice task with two simultaneously presented options (Levy & Glimcher, 2012). As a result, a huge number of computations that go into decision-making are stripped away. When we decide on what beverage to drink, we almost never are faced with the choice described above - one glass of lemonade for a particular price, or one glass of cola for a different price. Instead, these decisions fall inside a much larger context. We often have more than two options - we are often even allowed to select both options, or go back and change our minds. Our thought processes are affected not only by the immediate details of each beverage (which are also usually limited to one or two variables anyway in a particular study, as opposed to the many listed above; Landreth & Bickle, 2008), but also our global environments: the temperature outside, our current financial health, our current physical health, our thirst level, what our friends are drinking, how much time we have, and so on, and so on. This is where foraging theory excels - by adding the world back into tasks, we can glimpse some of these environmental variables at play inside our brains. And there is indeed behavioral evidence that such factors matter - In fact, humans are more optimal when doing a foraging task than a standard economic choice task (Carter et al, 2015).

The first step to creating a foraging neuroeconomics task is to make the decision-making naturalistic. That is, rather than a forced-choice between two options, a serial choice mechanism may be better. (Hayden et al, 2011; Constantino & Daw, 2015). Moreover, a central fixture of foraging theory is the trade-off between costs and rewards (Mobbs et al, 2018). Foraging behavior is often an effort in maximizing the amount of reward obtained per unit time. (Mobbs et al, 2018; Constantino & Daw, 2015), or the *reward rate*. Thus, a foraging task should

have some mechanism by which the forager must wrestle with time costs in addition to receiving rewards.

Another central feature of foraging is the patch. In nature, food is not evenly distributed Instead, it occurs in clumps which foraging theory refers to as *patches* (Stephens & Krebs, 1986; Mobbs et al, 2018). Foragers in patches are always facing a choice: stay in the patch and eat, or leave the patch and find a new one (Mobbs et al 2018; Hayden et al, 2011t). This decision hinges upon the current value of the patch - how much quality food there is (Mobbs et al , 2018), and the cost of going to the next patch (Constantino & Daw, 2015; Hayden et al, 2011). A foraging task, then, acknowledges patches - food does not appear regularly every few seconds, but rather takes time to find, and may have some areas of poor quality and others of rich quality.

Finally, in foraging there are several costs associated with finding food. One of these costs is *search time* - the time an animal spends looking for food.Another is *handling time* - the amount of time it takes to make the food accessible to you once you have it, (e.g. cracking open a nut or wrangling prey). Finally, there is *travel time* - the amount of time spent journeying between patches (Stephens & Krebs, 1986; Hayden et al, 2011; Constantino & Daw, 2015). The forager balances all the time costs they face when they decide whether to stay within or leave a patch. If a patch is rich in food, the animal may wish to stay in the current patch rather than travel to the next one. However, if a patch is poor, the animal may decide that it is worth the cost of travel time to seek out a better patch (Hayden et al, 2011). Indeed, time spent in patches tends to be inversely proportional to the richness of the patch (Stephens & Krebs, 1986; Arditi & Dacorogna, 1988; Hayden et al, 2011). These costs fit extremely naturally into the kinds of tasks already suggested: one that incorporates patches with time costs, and facilitates the calculation of a reward rate.

Foraging engages a wide variety of brain structures: the ACC, (Kolling et al, 2012; Kolling et al, 2016) dorsal anterior cingulate cortex (Calhoun & Hayden, 2015; Blanchard & Hayden 2014; Mobbs et al, 2018) , posterior parietal cortex (Calhoun & Hayden, 2015) , posterior cingulate cortex (Calhoun & Hayden, 2015; Barack et al, 2017), for a start. Conversely, many foraging papers pit the ventromedial prefrontal cortex (vmPFC) against these “foraging regions” as a region that is only concerned with the kind of value referred to variously as “economic,” “current,” and more concerned with evaluating a set of options than encoding environmental factor (Kolling et al, 2012; Shenhav et al, 2016; Boorman et al, 2013). We shall term this conceptualization of vmPFC’s contribution to decision-making as the encoding of *local* variables - the value of what is essentially right in front of the decision-maker’s face.Indeed, the

vmPFC is of particular interest to value researchers. Many studies have shown the vmPFC to increase activity in response to the value of an offer (Daw et al., 2006; Kable & Glimcher, 2009; Chib et al., 2009; Padoa-Schioppa, 2011; Boorman et al., 2013). Thus, the literature dictates that to examine foraging-related activity, one should look for signals in structures like the ACC, and not the vmPFC. However, we posit here that the vmPFC contributes to the foraging process by encoding not just local value but also environmental variables. What better place to look for the influence of the environment on a value computation than a region whose canonical primary focus is value?

Here, we now have an open question - the specific nature of value encoding in prefrontal cortex in the context of the surrounding environment, a framework that is exceptionally well-suited to addressing this problem - foraging theory - and a choice brain region in which to study it - the vmPFC.

To this end, we designed a novel *diet-search* task, for monkeys to complete while we record neural activity in their vmPFC. Our diet-search task is loosely based on the secretary problem (Chow, et al, 1964; Ferguson, 1989; Seale & Rapoport, 1997), where a hiring manager conducts a series of interviews with an infinite number of potential secretaries. Some secretaries are excellent, others are mediocre, and still others are awful. However, there is no way for the hiring manager to know when the 'best' secretary will come along. Instead, they must decide on their own when to end their search - when they have interviewed the best secretary they are probably going to find.

There is a catch, however. Once the hiring manager rejects a secretary, they can never go back to that secretary. As a result, after each interview they must make a permanent decision - to accept, or to reject. They must make this decision based on what they assume about the quality of the environment: the pool of secretaries. This potential secretary is decent, but are there many better secretaries out there? Or by rejecting this secretary, is our hiring manager losing out on the best there is?

Of course, real life is not always exactly like this. When examining a series of options, we often do have the opportunity to go back to previously rejected options. The cost of this action would be time: either having to go back and find the particular candidate's information, or waiting until many candidates have been interviewed before making a decision. As this is beginning to sound a lot like a foraging problem, let us rephrase the secretary problem as something that would be more familiar to a monkey.

Our monkey is hungry, and desires berries. However, the berries are high up in the leaves of trees - trees that he must climb. Some trees are rich with many berries, while others

are poor with only a few. From the ground, they cannot see how many berries are in each tree. They must take the time to climb them, inspect them to see how many berries there are, and then decide if they want to spend further time rummaging around all the branches and leaves to collect all the berries, or just jump ship and find another tree that will give them more berries for the same amount of rummaging. Just like the hiring manager in the secretary problem must balance the time cost of interviewing another candidate with the value of the current candidate, our monkey must balance the cost of foraging with the value of the current tree. The scenarios are comparable, but now by putting the secretary problem into foraging terms, we can apply other pieces of the foraging literature to the equation.

Our novel diet-search task translates the secretary problem into a foraging problem. We preserve the paradigm of searching through a series of offers until the decision-maker settles on one that is 'good enough.' However, the decision-maker may go back and accept an offer they have previously rejected. We also modify the task structure so that it resembles a patchy foraging task. A monkey is presented with a screen filled with a random array of diamonds. If they wish to investigate a diamond (or, climb a tree to examine its berry capacity), they must pay a *search time* cost of 400 ms of staring directly at the diamond. Doing so reveals the richness of this particular offer (how many berries are actually in this tree). At this point, the monkey has 300 ms more of *handling time* to decide whether they wish to accept this offer (remain in the tree and forage for the berries) or reject the offer (jump out of the tree and look for a better one). The monkey may examine each diamond as many times as they wish, but they cannot end the trial (leave the patch) until they have accepted an offer, after which they wait out a *travel time* of four full seconds and the next trial appears.

We trained two monkeys to perform this diet-search task and while recording neurons in their vmPFC. We performed two distinct sets of analyses on the data obtained. First, we sought out the contribution of local and environmental factors to vmPFC neurons' processing of value, detailed below in Chapter 2. Specifically, we examined the animal's behavioral and neural responses to an offer's current magnitude, the magnitude of the previous offer, the time spent in the current trial, the outcome of the previous trial, and the global reward rate. In all cases, we found relationships between the variable in question and the animal's threshold for accepting an offer, as well as encoding in vmPFC in response to the variable.

Next, we were interested in examining vmPFC neurons as a population in the same context of foraging-based decision making. Studying the activity of individual neurons is of course crucial to neuroscience because neurons are the fundamental unit of computation in the brain. However, studying how populations of neurons communicate during cognition is equally

valuable (Rigotti et al, 2013; Gao & Ganguli, 2015; Gallego et al, 2020). Thus, we performed a population analysis on the same foraging dataset as above. Here, we examined how the population of vmpfc neurons moves through stages of the foraging process. Specifically, we determined that distinct computations involving the reward rate occur during travel time versus foraging.

Finally, we sought to flesh out the contributions of environmental computations to behavior further in our foraging task as well as a different task. We find repeated evidence of the importance of tracking reward rate to the success of foraging behavior, and that environmental variables are relevant not only to foraging tasks but even tasks in the traditional two-option format.

Chapter I: Ventromedial prefrontal cortex encodes environmental variable during search

ABSTRACT

To make efficient foraging decisions, we must combine information about the values of available options with non-value information. Some accounts of ventromedial prefrontal cortex (vmPFC) suggest it has a narrow role limited to evaluating immediately available options. We examined responses of neurons in Area 14 (a putative macaque homologue of human vmPFC) as two male macaques performed a novel foraging search task. Although many neurons encoded the values of immediately available offers, they also independently encoded several others 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.

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 prefrontal cortex (vmPFC). 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 (Kolling et al., 2012; Kolling et al., 2014; Daw et al., 2006; Pearson et al., 2009; 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 (Padoa-Schioppa, 2011; Levy & Glimcher, 2012; Rangel et al., 2008; Chen & Stuphorn, 2015). However, some recent work suggests that action selection cannot be dissociated from evaluation and comparison (Pezzulo & Cisek, 2016; Cisek & Kalaska, 2010; Hayden & 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 (Passingham & Wise, 2012; Wise et al., 1996; Calhoun & Hayden, 2015; Mobbs et al., 2018; Hayden, 2018). An important foraging-related region is the ventromedial prefrontal cortex (vmPFC), which appears to assign values to available options (Kolling et al., 2012; Boorman et al., 2013; Kable & Glimcher, 2009; Boorman et al., 2009; Lim et al., 2011; Wunderlich et al., 2009). This proposed role is consistent with its putative role as derived from neuroeconomic studies (Levy & Glimcher, 2012; Bartra et al., 2013; Rushworth et al., 2011; Strait et al., 2014; Monosov & Hikosaka, 2012). 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, while 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 (Hare et al., 2009; Kable & Glimcher, 2007; 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. Gusnard & Raichle, 2001; Bouret & Richmond, 2010; Monosov & 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 homologue of human vmPFC (Mars et al., 2016). The best name for this region is currently unclear, for several reasons. First, the specific macaque homologue 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 Mars et al., 2016 have made the strongest case so far). Third, the primate ventromedial network is large and heterogeneous (Ongur & 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; Strait et al, 2016; Azab & 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.

EXPERIMENTAL 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 two 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 magnetic resonance imaging with the aid of a Brainsight system (Rogue Research Inc.). 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 roughly the same time of day, between 10 am and 3 pm. 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 a region of interest lying within the coronal planes situated between 42 and 31 mm rostral to interaural plane, the horizontal planes situated between 0 to 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 Siemens 3T MAGNETOM Trio Tim 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 & Co., impedance range 0.7 to 5.5 MU) until waveforms of between one and four neurons were isolated. Individual action potentials were isolated on a Plexon system (Plexon, Inc.). 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 over 500 trials (this was our recording target each day).

An infrared eye-monitoring camera system (SR Research) sampled eye position at 1,000 Hz , and computer running Matlab (Mathworks) with Psychtoolbox and Eyelink Toolbox controlled the task presentation. Visual stimuli were colored diamomds 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 (**Figure 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 & Hayden, 2015; Blanchard et al., 2015), intertemporal

choice tasks (Hayden, 2016), two types of gambling tasks (Azab & Hayden, 2017; Strait et al. 2014), attentional tasks (similar to those in Hayden and Gallant, 2013), and two types of reward-based decision tasks (Sleezer, Castagno, & Hayden, 2016; Wang & 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 non-overlapping 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 percent 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 uL for subject T and 23 uL for subject J). Reward values for each offer were generated randomly for each offer using a uniform continuous distribution ranging from 0%-100% of the maximum possible reward value for the individual subject (that is, 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 inital 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 to unmask it and an additional 300 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 second inter-trial interval.

Because the experiment was conducted partially using one computer and partially another, the plot of diamond locations (**Figure 1B**) 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 (such as 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 (that is, 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* preceeding 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

Peri-Stimulus 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 standard deviation 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; Strait et al., 2015; Azab & Hayden, 2017). All comparisons of firing rates over particular epochs were conducted two-samples 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 to 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 'non-monotonic.' According to our classification system, the direction and monotonicity of value tuning are not mutually exclusive, so of the non-monotonic neurons, 8 were also classified as positively tuned, and 5 of these neurons were also classified as negatively tuned.

Population Analyses

We used beta correlation analyses to assess whether neurons encoded particular offer values (Blanchard et al., 2015; Azab & 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 which requires subjects to fixate for 300 ms in order 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 (figure 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 beta correlation analyses to assess whether neurons represented offer acceptance and rejection using similar coding formats (cf. Azab & Hayden, 2017). We found the regression coefficients (that is, the beta 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 beta 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 pre-decisional; 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 beta correlation analysis explained above. For the second analysis, we assume a model in which neurons are purely post-decisional; 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 or greater standard deviations away from the mean. This number was determined prior to performing the analyses. In **Figure 6A**, 2 points were excluded. In **Figure 6D,** 2 points were excluded, in **Figure 6E,** 4 points were excluded, in **Figure 6H**, 4 points were excluded, and in **Figure 7B**, 2 data points were excluded. The vast majority of data points were not outliers, and were within 1 standard deviation of the mean.

In **Figure 6H** (the effect of previous offer value on the strength of neural encoding of the current offer)**,** the error bars of thirty-seven 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 49% or higher. 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 6H** 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 larger than 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, bottom-right). We then performed a one-way analysis of variance (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 beta 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, c. The equation is:

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

Assuming Poisson firing of the neurons, the log likelihood of the Von Mises fit was calculated, and compared to 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 BIC_vonmises > BIC_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 criteria was met:

$$N_{spikes} > alpha/(N^2occ * r^2)$$

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 $alpha$ is a scaling parameter set to be 10000.

Cells are said to be significantly tuned to position when the mutual information exceeds the 95th percentile of the shuffled data.

RESULTS

Macaques are efficient foragers in a computerized search task

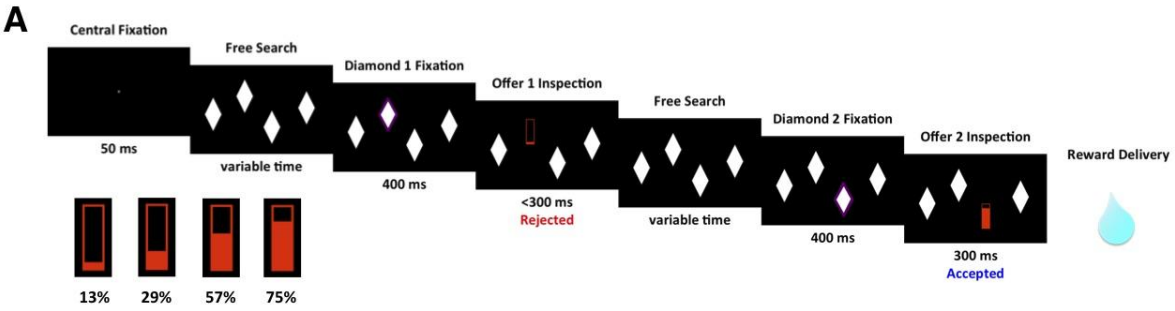
We examined behavior of two macaque subjects performing our foraging search task (subject J: 23,826 trials; subject T: 20,509 trials). On each trial, the subject was presented with an array of visually identical offers randomly positioned on the screen (**Figure 1**, **Methods**). Subjects could inspect 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 re-inspection. Offers were riskless and varied in size continuously from 0 to 230 uL (subject J) or 0 to 200 uL (subject T) fluid reward. Values noted below are in units of percent and refer to proportions of the largest amount for that animal. Results here are combined across the two subjects; all major results were replicated for each animal individually (see **Table 1**) with a few minor exceptions.

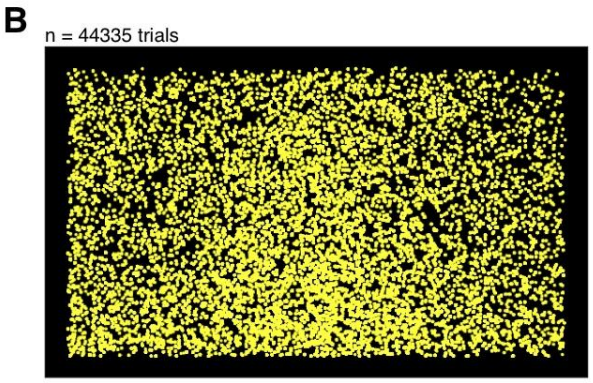
Each trial had either 4 or 7 offers; the two trial types were randomly interleaved and occurred equally often. On average, subjects inspected 1.872 offers on trials with 4-offer arrays and 2.074 offers on trials with 7-offer arrays (for the frequency with which subjects viewed different numbers of offers 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 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 (**Figure 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 (**Figure 2B**). (That is, we assumed subjects' reaction times were unavoidable and not strategic, an approach we developed in Hayden et al., 2011; Blanchard & 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 uL) for subject J and 60.60 % (121 uL) for subject T. For the 7-option case, the optimal threshold was 63.78% (147 uL) for subject J and 60.65 % (121 uL) for subject T. The observed thresholds were significantly lower than the optimal ones in all four 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-offer cases, subjects stopped (i.e. selected an option) at a lower threshold than optimal. Nonetheless, based on these numbers, subjects harvested ~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.

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 more ms results in reward delivery; breaking fixation before 300 ms is up returns the subject to free search. Below: 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 (see* ***Methods*** *for details).* ***(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.*


A
Central Fixation
50 ms
Free Search
variable time
Diamond 1 Fixation
400 ms
Offer 1 Inspection
<300 ms
Rejected
Free Search
variable time
Diamond 2 Fixation
400 ms
Offer 2 Inspection
300 ms
Accepted
Reward Delivery
13%
29%
57%
75%


B
n = 44335 trials


C

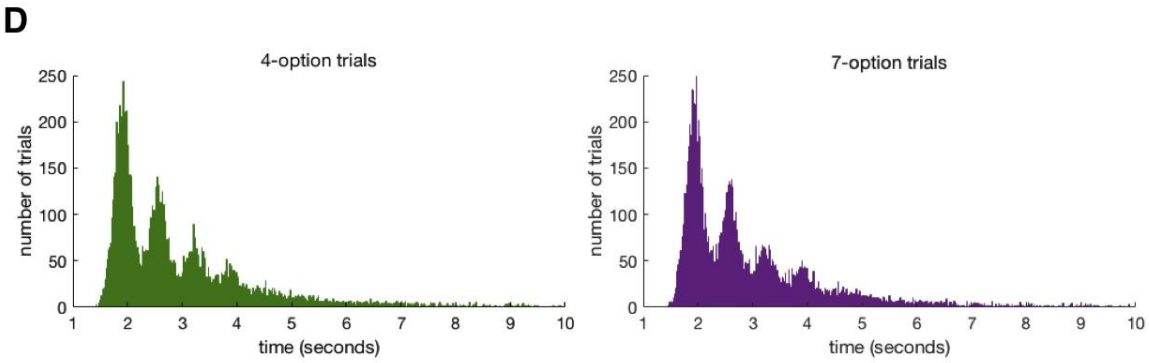
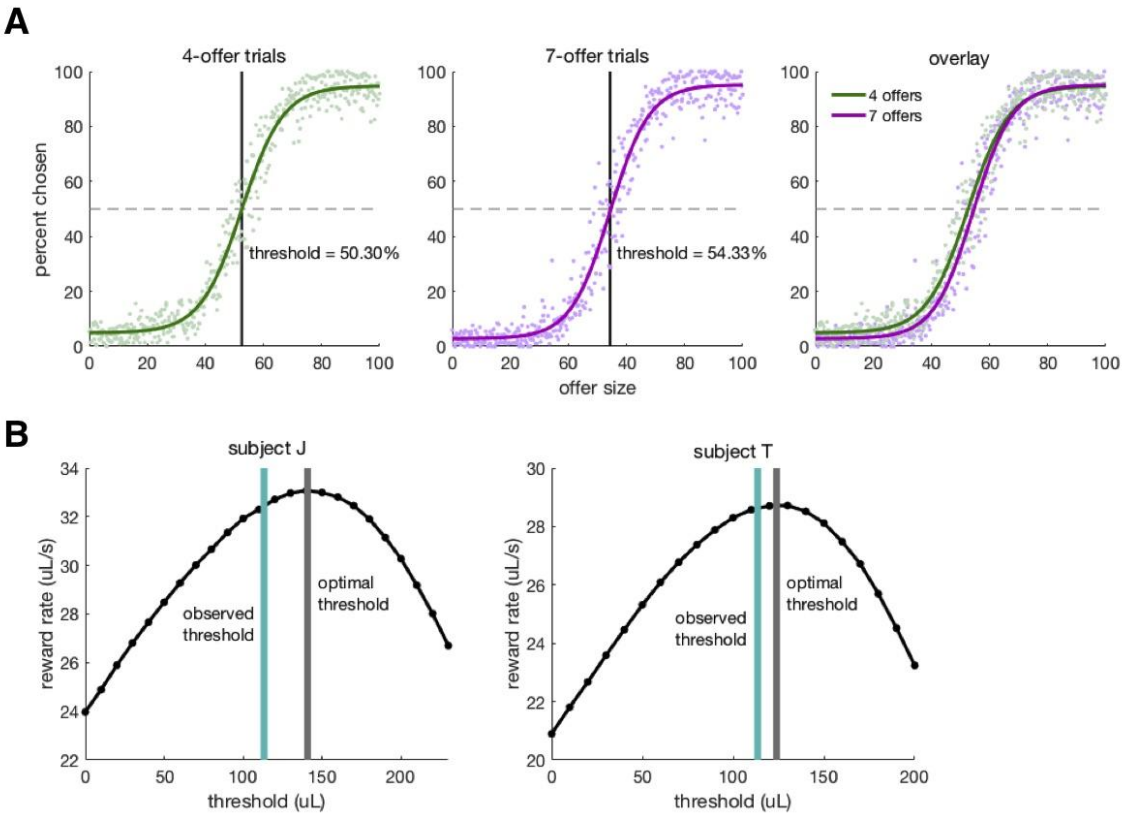

D
4-option trials
7-option trials
number of trials
time (seconds)
0
50
100
150
200
250
1
2
3
4
5
6
7
8
9
10

***Figure 2: Subjects' behavior**. **(A)** Likelihood of accepting a given option was a roughly 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 percent 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 over 100 simulations of 1000-trial task sessions for each threshold between 0 and each subject's reward maximum (J: 230 uL; T: 200 uL) in steps of 10 uL. Gray bars indicate the threshold that yielded the highest reward rate in the simulation. Teal bars indicate 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.*



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; Strait et al., 2016; Azab & Hayden, 2017; Azab & Hayden, 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/sec for high values vs. 11.78 spikes/sec 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 **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; there we reported 16.03% of cells encoded current value, chi-square = 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", "non-monotonic", and "none" (see **Figure 4**; note that some non-monotonic cells can also have an overall positive or negative trend, see **Methods**). We found that 11.48% of cells (n=14/122) demonstrated a significantly positive relationship between firing rate and current offer value, 11.48% (n=14/122) a significantly negative relationship, and 40.16% (n=49/122) cells a non-monotonic relationship (not that these purely non-monotonic neurons are not part of the 33.61% of neurons that linearly encode value). Additionally, we found that 6.56% of cells (n=8/122) demonstrated a non-monotonic relationship with an overall significantly positive trend, and 4.10% of cells (n=5/122) demonstrated a non-monotonic relationship with an overall significantly negative trend. Finally, 26.23% of cells (n=32/122) had no discernable relationship between firing rate and offer value.

Figure 3: Encoding of value in vmPFC. (A) *Peri-stimulus time histogram (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 indicates standard error of firing rate at each time point (10 ms bins). Solid line indicates mean firing and is smoothed over 20 bins. Gray shading indicates the analysis epoch. Vertical lines indicate the beginning of diamond fixation (400ms) 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. (Note that x-axis is different from panel A). Proportions are calculated with a 200 ms sliding boxcar. The peak at ~6 seconds 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 indicates analysis epoch.*

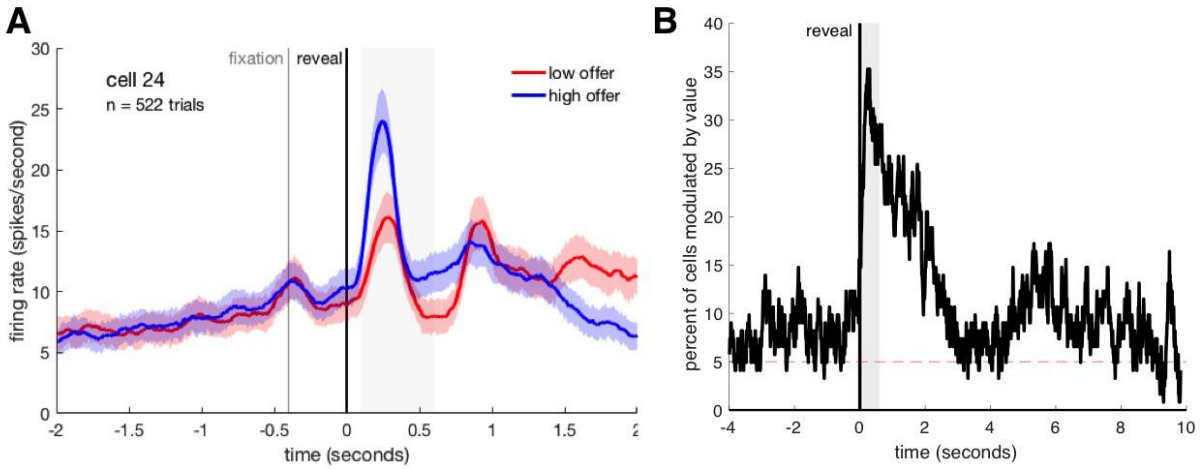
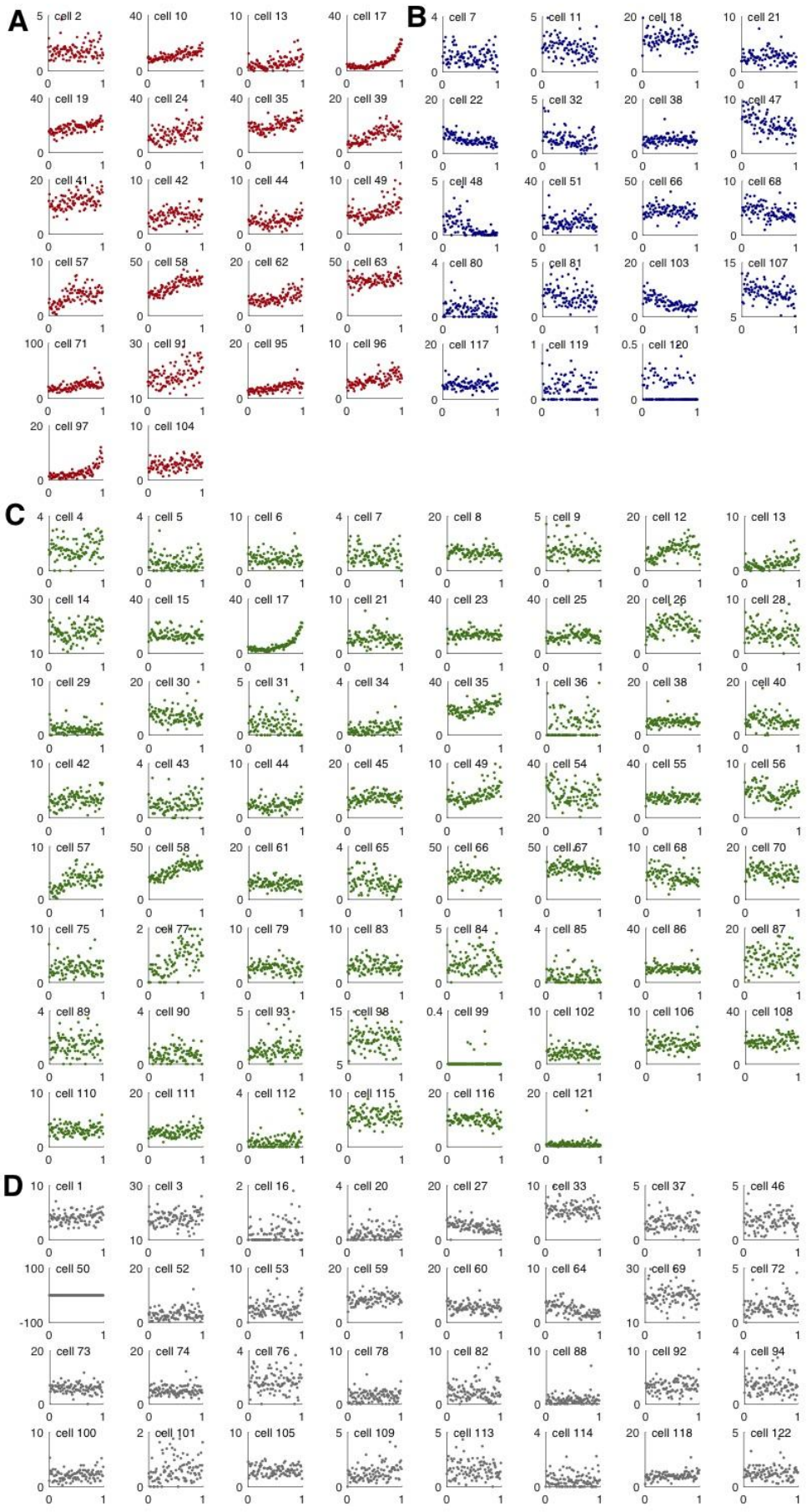


Figure 4: Classifications of individual neurons' value tuning. ***(A)*** *Neurons classified as "positively tuned" to value.* ***(B)*** *Neurons classified as "negatively tuned" to value.* ***(C)*** *Neurons classified as "non-monotonically tuned" value.* ***(D)*** *Neurons that do not fit into one of the above classifications.*


A
cell 2
cell 10
cell 13
cell 17
cell 19
cell 24
cell 35
cell 39
cell 41
cell 42
cell 44
cell 49
cell 57
cell 58
cell 62
cell 63
cell 71
cell 91
cell 95
cell 96
cell 97
cell 104
B
cell 7
cell 11
cell 18
cell 21
cell 22
cell 32
cell 38
cell 47
cell 48
cell 51
cell 66
cell 68
cell 80
cell 81
cell 103
cell 107
cell 117
cell 119
cell 120
C
cell 4
cell 5
cell 6
cell 7
cell 8
cell 9
cell 12
cell 13
cell 14
cell 15
cell 17
cell 21
cell 23
cell 25
cell 26
cell 28
cell 29
cell 30
cell 31
cell 34
cell 35
cell 36
cell 38
cell 40
cell 42
cell 43
cell 44
cell 45
cell 49
cell 54
cell 55
cell 56
cell 57
cell 58
cell 61
cell 65
cell 66
cell 67
cell 68
cell 70
cell 75
cell 77
cell 79
cell 83
cell 84
cell 85
cell 86
cell 87
cell 89
cell 90
cell 93
cell 98
cell 99
cell 102
cell 106
cell 108
cell 110
cell 111
cell 112
cell 115
cell 116
cell 121
D
cell 1
cell 3
cell 16
cell 20
cell 27
cell 33
cell 37
cell 46
cell 50
cell 52
cell 53
cell 59
cell 60
cell 64
cell 69
cell 72
cell 73
cell 74
cell 76
cell 78
cell 82
cell 88
cell 92
cell 94
cell 100
cell 101
cell 105
cell 109
cell 113
cell 114
cell 118
cell 122

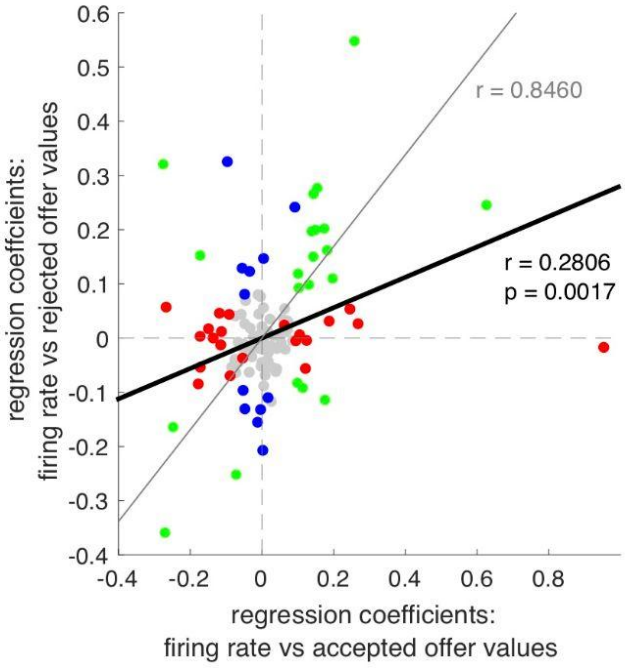
Encoding of categorical choice

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 vs. 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 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 **Methods** and Azab & Hayden, 2017). On the other hand if the code in vmPFC is pre-decisional, 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 would neurons 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 (**Figure 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 pre-decisional 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 **Methods**, $p<0.001$). These results suggest that vmPFC encoding of value is neither *purely* post-decisional nor *purely* pre-decisional; it has properties of both.

***Figure 5: Population encoding for accepted versus rejected offe**rs. 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 (grey line, r = 0.8460).*



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 **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 **Methods**). For both subjects, threshold on the current trial increased with the size of the previous trial's outcome ($\beta = 0.0602$, $p < 0.001$, **Figure 6A**; significantly in one 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=28/122) changed during the choice epoch depending on previous outcome value (linear regression, alpha=0.05). This proportion is significantly greater than expected by chance (i.e. 5%, $p < 0.001$, two-sided binomial test). This analysis was performed across all trials (no exclusions).

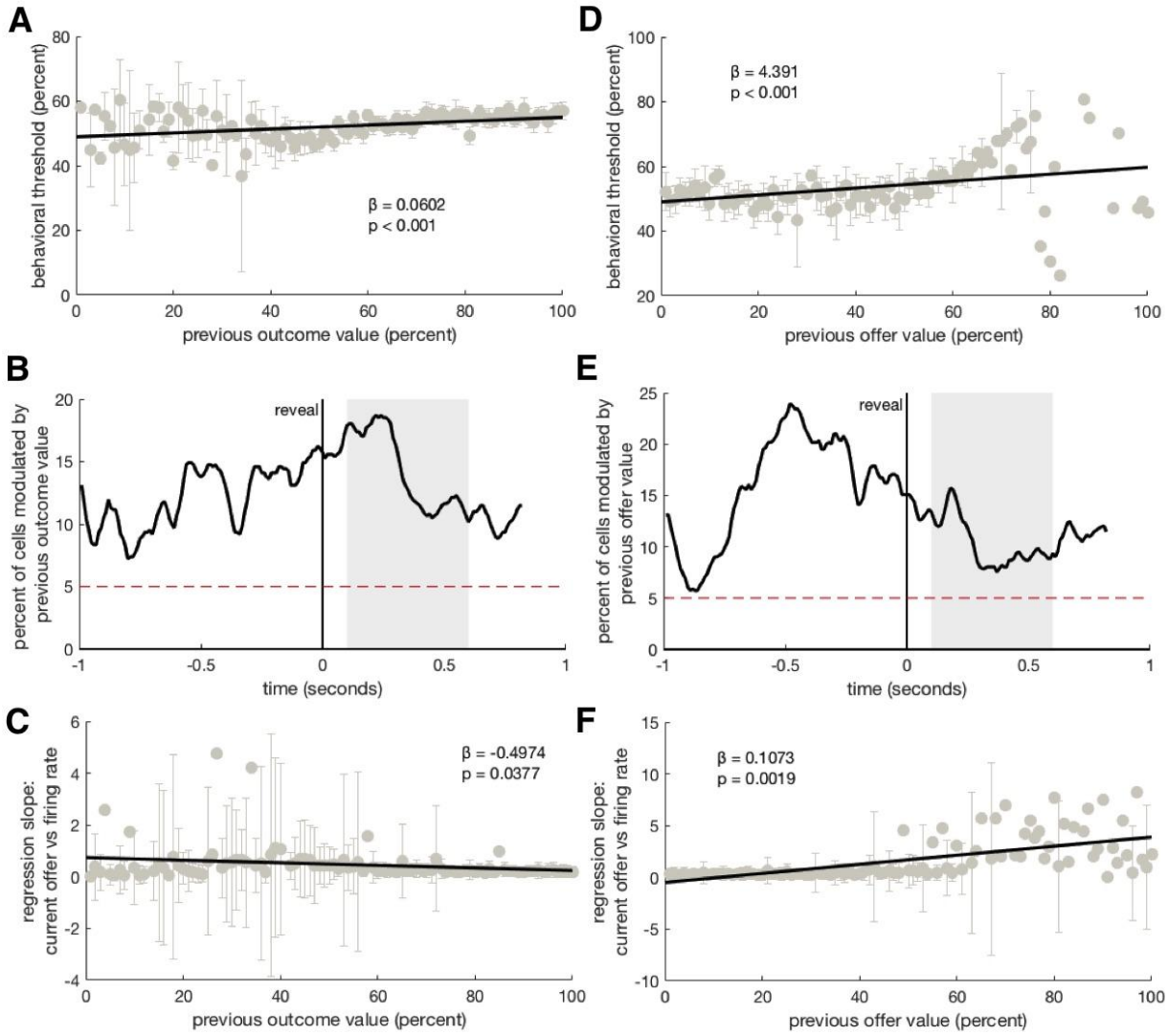
There is some evidence that previous outcome changes the strength of value coding as well. Specifically, we regressed the firing rate against offer value for different previous outcome sizes and found that the previous trial's outcome negatively affected the value encoding of the current offer ($\beta = -0.4974$, $p = 0.0377$), **Figure 6D**; significant in one 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 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 ($\beta = 0.1073$, $p = 0.0019$, **Figure 6E**; significant in one 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=20/122) of neurons (**Figure 6F,** see **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 ($\beta = 4.391$, $p < 0.001$).

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 represent the standard error of the sigmoidal curve fit used to produce each threshold (error bars of four 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 represent the standard deviation of the population of absolute coefficients averaged to produce each point. Error bars from eight 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 sixteen 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 thirty-seven points have been removed (see* ***Methods****).*

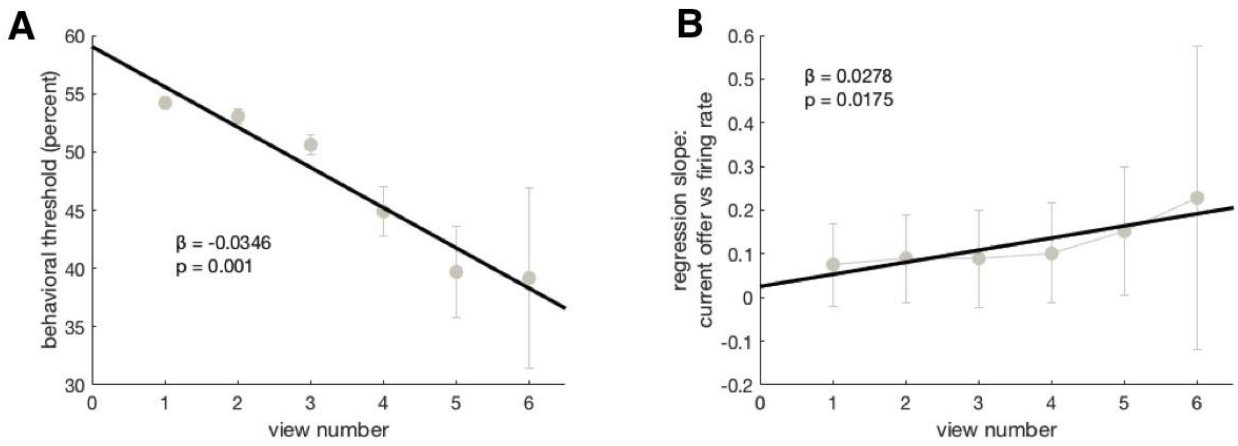


View number

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 less than 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 (β = -0.0346, p=0.001, **Figure 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=40/122) significantly encoded the within-trial view number of an offer. (Note that 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 (β = 0.0278, p=0.0175, **Figure 7B**).

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.*



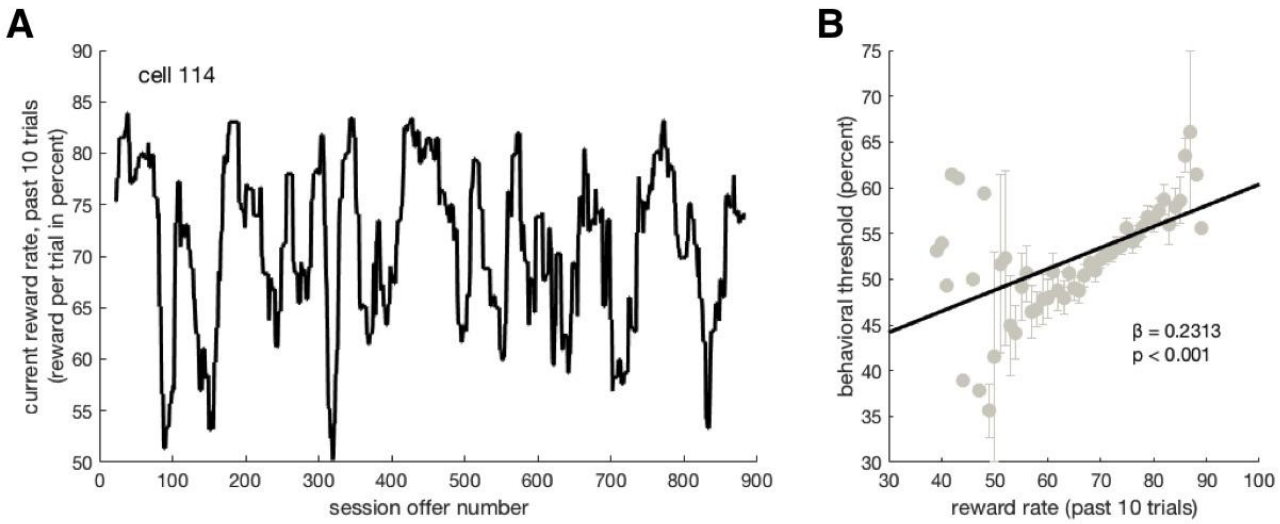
Reward rate

The brain may keep track of recent reward rate – the values of the last several outcomes over time (see **Methods**). Reward rate fluctuates over the course of a behavioral session (**Figure 8A).** Foragers may see this measure as a proxy for environmental richness. It should not, 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, β = 0.2313, p<0.001, **Figure 8B**). We found that 32.79% (n=40/122) 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 & 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 (**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.

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 eleven 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 ($p < 0.001$).*



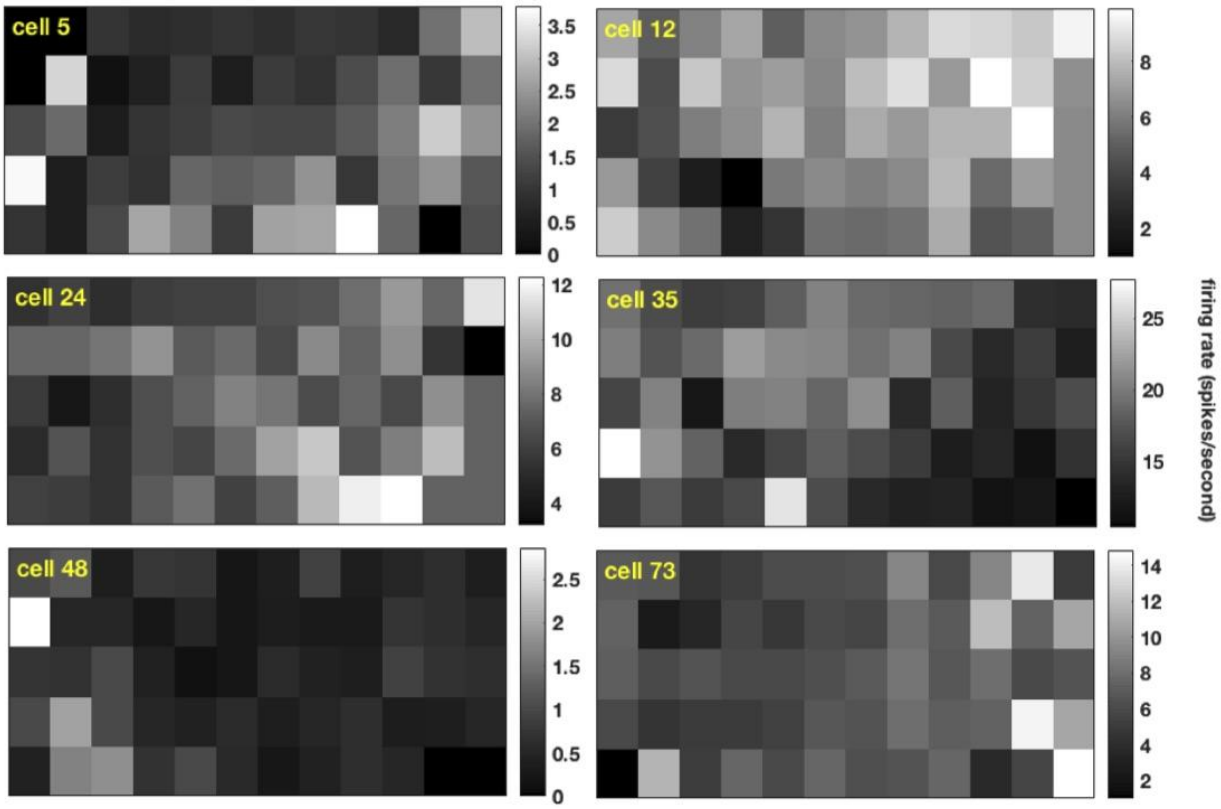
Spatial positions of current offers

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=27/122) showed quadrant spatial selectivity for current gaze position ($p<0.05$, one-way ANOVA, see **Methods**). This proportion is greater than would be expected by chance (that is, 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 information. This proportion is also greater than chance ($p<0.01$, two-sided binomial test). Likewise, 13.93% of neurons had selectivity for up-vs-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 ($r = 0.2444$, confidence interval [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 ($r = 0.1938$, confidence interval [0.0167 0.3592] $p = 0.0323$).

Figure 9: Six example neurons that demonstrate spatial selectivity. *Grayscale in heatmaps indicate mean firing rate (spikes/second) 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 prior to 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 due to the fact that no offers were centered in those bins (or the tops or bottoms of the images would appear cut off). 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.*



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 threshold 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 threshold 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, suggests that the spatial selectivity we see is not a by-product 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 if cells adjusted the strength of their value encoding (changed the regression coefficient of firing rate versus offer value) when those offers were located in cells' preferred spatial locations. We divided the screen into 84 rectangular segments (a 7x12 grid, as in **Figure 9**). For each segment 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=15/122) 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.

Radial measure of tuning for gaze position

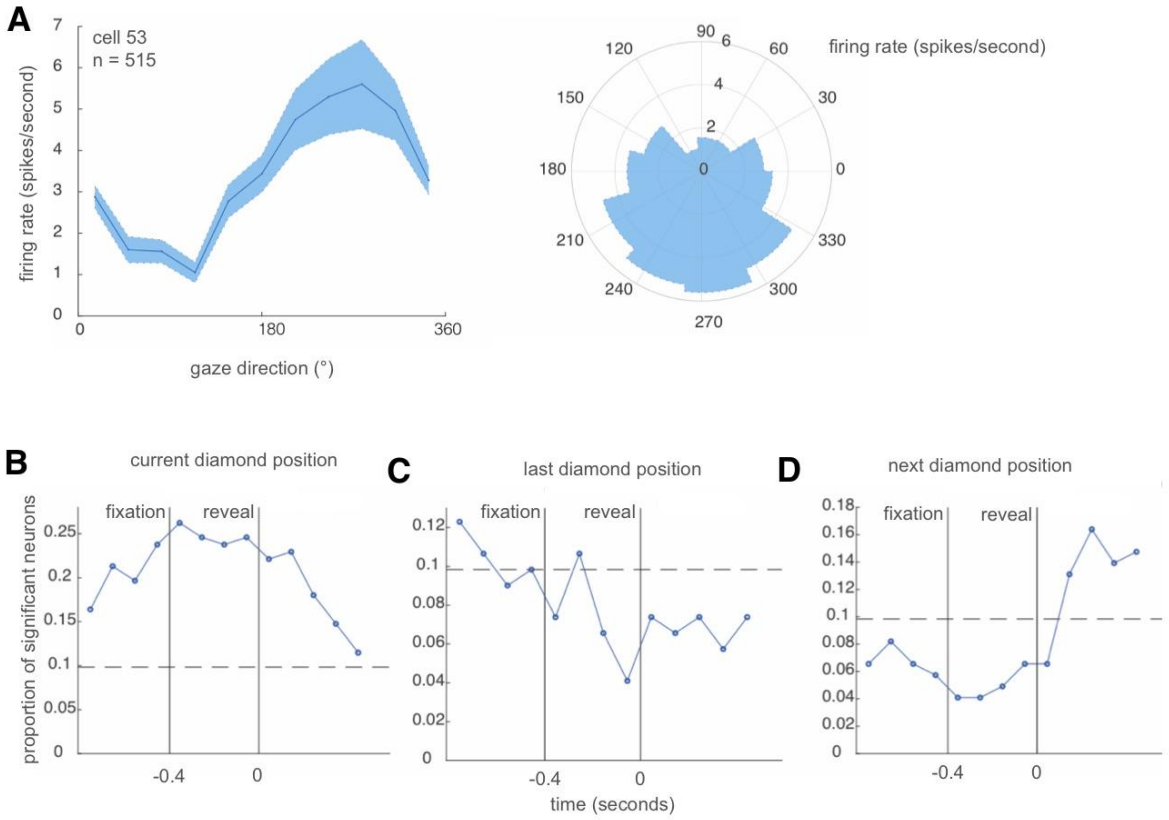
To study spatial selectivity for gaze position in vmPFC with greater precision, we next examined radial tuning functions. We measured each neuron's selectivity for position of current fixation relative to the center of the computer monitor (i.e. straight ahead). We fit Von Mises (circular Gaussian) distributions and compared them to 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 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/second) 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/second, $t = 3.3304$, $p< 0.001$). 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, 0.2327, see **Methods**; BIC decrease = 23). In the population of neurons, a large proportion (34.5%, n=42/122) 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 = 44/122) 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; Skaggs et al., 1996; see **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 (**Figure 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 (**Figure 10C**) and future (**Figure 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, and that selectivity for the future position begins immediately after the present fixation ends.

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.*



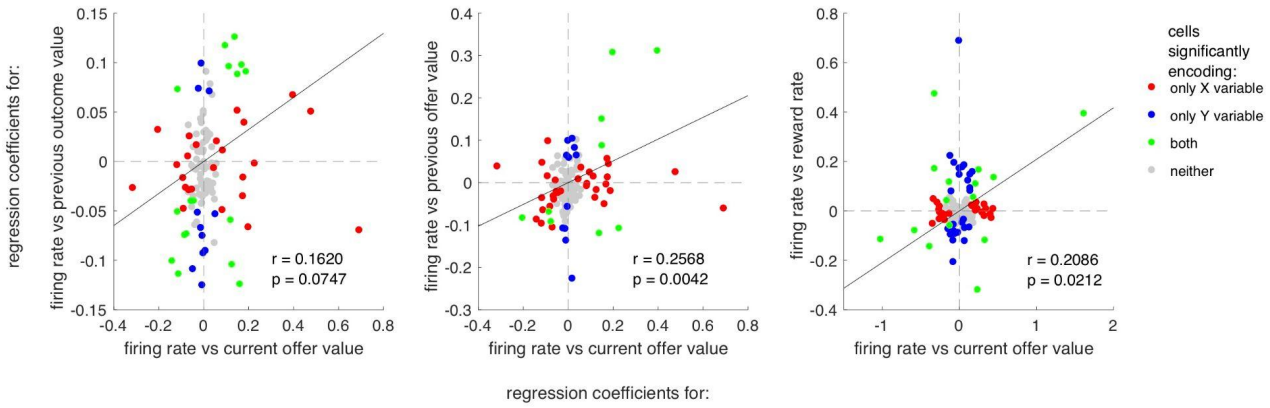
Encoding for environmental variables is weakly correlated with value encoding

We next examined how these codes related to each other. In our sample, 71.3% of cells (n=87/122) significantly encoded at least one variable of the 5 variables we examine (offer value, previous trial outcome, previous offer, view number, reward rate; note that this analysis uses only one analysis epoch to reduce the chance of false positives). From this larger population, 28.69% of cells (n=35/122), encode exactly two variables, 17.21% of cells (n=21/122) encode exactly three variables, 6.56% of cells (n=8/122) encode exactly 4 variables, and 1.64% of cells (n=2/122) 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 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 (**Figure 11**). However, current offer value encoding and previous offer value encoding are correlated at r = 0.2568 (p =

0.004, **Figure 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 (figure 11). This value is significant at p = 0.0211, indicating that cells may also encode offer value and reward rate in a similar way.

Figure 11: Population encoding for current offer value, compared to 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.*



DISCUSSION

We examined neural responses in a putative primate homologue 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 searching 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 employ it consistently. Subjects' behavior was well described by just such a thresholding procedure: average thresholds were approximately reward-maximizing (with a bias towards over-accepting, cf. Blanchard and Hayden, 2014). Neurons strongly encoded values of offers, thus confirming its value-related role (Levy & Glimcher, 2012; Bartra et al., 2013; Clithero & Rangel, 2013; Rushworth et al., 2011). 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 as 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 non-value 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 (Padoa-Schioppa, 2011; Levy & Glimcher, 2012; Rangel et al., 2008). 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 (Bartra, 2013; Rushworth et al., 2011). However, other research casts doubt on the existence of a pure value signal (e.g. Yoo & Hayden, 2018; Rushworth et al., 2012; Smith et al., 2010; Wallis & Kennerley, 2010; Cisek and Kalaska, 2010; Vlaev et al., 2011). 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 & Preston, 2015). Even value variables in vmPFC may not be integrated into a coherent integrated set (Smith et al., 2010; Watson & Platt, 2012; Sescousse et al., 2010).

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 (Kable & Glimcher, 2009; Levy & Glimcher, 2012; Delgado et al., 2016; Plassmann et al., 2007; Chib et al., 2009). This signal incorporates all information that is relevant to choice, but expressed as a single scalar quantity; vmPFC outputs are, by this view, *post-evaluative*. While we cannot speak to whether this 'unitary value signal' exists within vmPFC, we can speak to the post-evaluative nature of the signals within vmPFC. Our data suggests 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 post-evaluative. 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 & Assad, 2006; Padoa-Schioppa, 2011; Rangel & Hare, 2010). The presence of clear spatial information here, then, may be somewhat surprising. One previous study, from our lab, 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 & 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 *economnic salience map*, that is, 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 (Pearson et al., 2014; Passingham & Wise, 2012; Hayden, 2018; Hayden, 2016). 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 (Hayden, 2018; Rich & Wallis, 2016). 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.

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

subject	J	T	Combined
Macaques are efficient foragers in a computerized search task			
number 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 (p value)	0.0372 (0.2030)	0.1292 (<0.001)	0.0602 (<0.001)
percent of neurons significantly encoding previous outcome value (p value)	27.14% (<0.001)	17.31% (0.0010)	22.95% (<0.001)
beta correlation coefficient of firing rate against previous outcome value (p 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 (p value)	0.1243 (0.1202)	0.1515 (<0.001)	0.1073 (0.0019)
percent of neurons significantly encoding previous offer value (p value)	24.29 (<0.001)	5.77 (1.000)	16.39 (<0.001)

beta correlation coefficient of firing rate against previous offer value (p value)	7.334 (<0.001)	0.2565 (0.0736)	4.391 (<0.001)
View Number			
beta correlation coefficient of behavioral threshold against view number (p value)	-0.04849 (0.0063)	-0.0267 (<0.001)	-0.0346 (<0.001)
percent of neurons significantly encoding view number (p value)	42.86% (<0.001)	19.23 (<0.001)	32.79 (<0.001)
beta correlation coefficient of firing rate against view number (p 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 (p value)	0.4708 (<0.001)	0.3133 (<0.001)	0.2361 (<0.001)
percent of neurons significantly encoding reward rate of past 10 trials (p 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. If *n* offers have been viewed, a re-visit to a previously viewed offer counts as a new "view" (eg *n+1*).

n	percentage of all trials in which *n* offers were viewed
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

Chapter II: Signatures of processing complexity during global cognitive states in ventromedial prefrontal cortex

ABSTRACT

Cognitive neuroscience almost exclusively studies behavior during tasks and ignores the unstructured inter-trial interval (ITI). However, it is unlikely that the ITI is simply an idling or paused mode; instead, it likely elicits globally focused cognitive processes, in which attention is disengaged from the task at hand. To gain insight into the computational underpinnings of globally focused cognition, we recorded from neurons in a core decision-making region, Area 14 of ventromedial prefrontal cortex (vmPFC), as macaques performed a foraging search task with long inter-trial intervals (ITIs). We find that while the ITI is associated with reduced firing, it is associated with increased discriminability of recent reward rate and with upcoming search strategy. ITI activity is also associated with increased ensemble dimensionality and fast subspace reorganization, presumed markers of processing complexity. These results demonstrate the flexible nature of mnemonic processing and provide support for the idea that the brain makes use of ostensible downtime to engage in complex processing.

INTRODUCTION

Our minds can readily switch between states that are stimulus-focused and oriented narrowly on the task at hand (***local states***) and stimulus-independent states that are relatively undirected and more attuned to the world around us (***global states***, Raichle, 2010; Raichle, 2015; Mason et al., 2007; Smallwood et al., 2013). During local states, many of the task-irrelevant things on our mind fall away so that we can selectively process the most important and urgent information. When our thoughts move away from the task, our minds then fill with a variety of stimulus-independent thoughts, including recollections of the recent past, as well as background and off-task cognition (Smallwood et al., 2013; Baird et al., 2013; Konishi et al., 2015).

Globally oriented cognition is poorly understood. We know a whole lot more about what goes on in the brain during brief periods of focused performance than during the often much longer times between. However, far from being a mental equivalent of doing nothing, globally oriented cognition is likely to be rich and complex, and presumably extends both backwards and forwards in time (Northoff et. al, 2010, Mazoyer et al., 2001; Zou et al., 2013; Cole et al., 2016). Indeed, even if global cognition is not directly related to the specific details of the most recent

trial, it is associated with several functions, such as reactivating past sequences of action, consolidation of the systems, guiding navigation in future goal-direction movement, detecting subtle patterns, and building cognitive maps (O´lafsdo´ttir et al., 2018; Raichle, 2010). One theme that unites these functions is that they have mnemonic relevance. Specifically, they involve reactivation of memories of the recent and distant past.

Foraging theory may be an unusual lens through which to study globally oriented cognition during decision making tasks, but it is a promising one. In foraging theory, decision making is studied under the assumption that animal brains evolved to make decisions in a naturalistic environment (Stephens & Krebs, 1986; Pearson et al., 2014). As such, economic choice tasks can be framed as food search tasks. Typically, these tasks ask the animal to split time travelling between, and foraging within, *patches*- areas where there are multiple opportunities to accept or reject a food offer (Hayden et al., 2011; Calhoun & Hayden, 2015). While many decision variables are widely studied in such tasks - the richness of the patch, the travel time between patches, the value of each food offer - these variables are locally oriented. That is, they pertain almost completely to the animal's present. Global variables, on the other hand, are also critical to foraging theory; however, prefrontal cortical encoding of these variables has much yet to be examined (Kolling et al., 2012; Hayden, 2018; Adrian et. al, 2019; Gaba & Apps, 2020).

One primary example of a global variable in foraging theory is the average rate of energy consumption over time, or the *reward rate* (McNamara & Houston, 1985; Constantino & Daw, 2015) Examining neural encoding of reward rate is valuable for studying globally oriented cognition because thinking about reward rate requires thinking about the global environment: it concerns the past (the animal must average past rewards), the present (the animal must update its knowledge of reward rate with the current reward value), and the future (the animal may use reward rate information to inform future decisions, Mehta et al, 2019). Furthermore, reward rate is fundamentally a representation of the richness of the environment, and environmental richness is a key influential factor in foraging decisions (Cisek & Pastor-Bernier, 2014; Krebs et al., 1978). Taking all this into account, we studied neural activity as animals completed a task that allowed us to study both local variables in the form of distributions of offer values, and global variables such as reward rate over time.

We examined ensemble firing rates in one structure, the ventromedial prefrontal cortex (vmPFC), as macaques performed a computerized foraging task. The task, like most laboratory tasks, is naturally divided into a task period and an inter-trial interval (ITI). Our analyses therefore centered on comparing neural activity during the task to activity during the ITI. .. To

gain insight into the role of vmPFC in global cognition, we focused on *adjustment*, a rapid learning that depends on past outcomes (Hayden et al., 2008; Hayden et al., 2009a; Hayden et al., 2011; Blanchard et al., 2015). We hypothesized that the ITI may be a critical period for performing adjustment-related computations.

We find three striking differences between these two task phases. First, information about the rate of recent reward intake (over the past 10 trials) is more readily distinguished in the ITI than during the trial. and that this information closely predicts subsequent adjustment. Second, neural ensemble dimensionality is substantially greater during the ITI. Third, the *change rate* of the subspace reorganization is significantly enhanced during the ITI. Together, these results highlight the inherent richness of neuronal processing during the ITI, and during global cognition more generally, and, in particular, suggest that this processing privileges mnemonically relevant information.

RESULTS

Macaques performed a foraging search task in which they inspected a series of hidden offers and selected one of them to obtain a liquid reward (**Figure 12A**, Mehta et al., 2019). The reward was deterministic but its value was randomized by offer and not predictable before the inspection. Search in this task is self-terminating, so the task has features in common with the secretary problem and other stopping problems (Ferguson, 1989). Each trial was separated by a four-second inter-trial interval (ITI). Detailed behavior in the task is summarized in our previous paper and is not relevant to our hypotheses and not repeated here (Mehta et al., 2019).

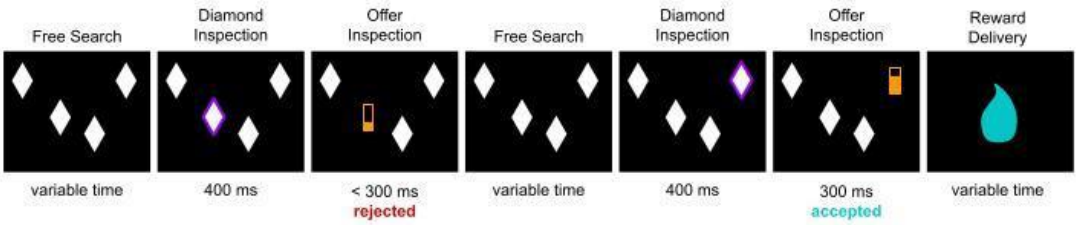
While our subjects performed the task, we recorded responses of 122 single neurons in area 14 of vmPFC. We found that individual neurons encode several foraging-related variables in this task, including the value of the current offer being inspected, the value of the most recent offer inspected, the outcome of the previous trial, the local reward rate (defined as the average reward obtained over the past ten trials, excluding the most recent trial), and the location of the attended offer (Mehta et al., 2019). The key feature of our search task that enables this analysis is the unusually long ITI (4 seconds), which allowed us to study responses temporally dissociated (>1 second) from both the past trial and the upcoming one.

As such, we focused our analyses on two epochs. Because we conducted our principal component analysis over windows of 500 ms, we will describe activity over 500 ms epochs. Our *ITI epoch* is centered on the midpoint of the ITI - exactly two seconds after it begins and two seconds before it ends. This means our window of analysis spans 250 ms before this point, and

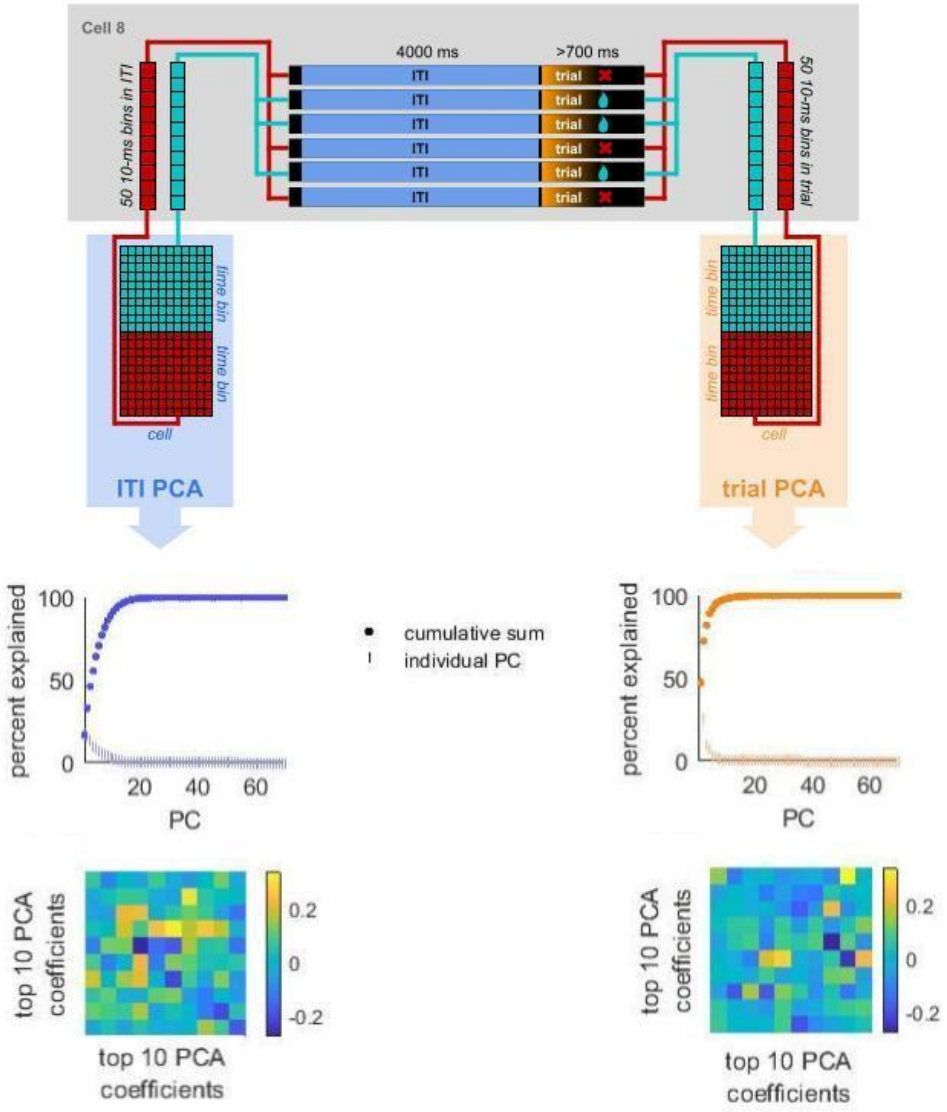
250 ms after. Our *trial epoch* is centered on the midpoint of the 1-second time period after the reveal of the first offer of each trial. This window of analysis spans from 250 ms after the offer reveal to 750 ms after the offer reveal. Offer values are represented as percentages of the maximum possible mL amount per offer for each subject.

Figure 12. Our search task and core PCA analysis. (A) *Schematic of a typical trial. The subject inspects an offer, rejects it, then inspects another offer and selects it, leading to a liquid reward.* ***(B)*** *Schematic of our core PCA analysis. For each cell, neural activity during a 500 ms epoch during the ITI and a 500 ms epoch during the first offer of each trial was averaged together in two separate groups - trials where the subject accepted the first offer, and those in which they rejected the first offer (upper panel). These averaged time slices from all cells were put together into an m x n matrix, where m is double the number of time bins in one epoch, and n is the number of cells (middle panel). We performed a PCA on this matrix for the ITI and a separate PCA for the trial. Here, we display some data from these PCAs for subject J (lower panels): the percent of variance explained by all 70 PCs and the coefficients resulting from the PCA.*

A


Free Search
Diamond Inspection
Offer Inspection
Free Search
Diamond Inspection
Offer Inspection
Reward Delivery
variable time
400 ms
< 300 ms
rejected
variable time
400 ms
300 ms
accepted
variable time

B


Cell 8
4000 ms
>700 ms
50 10-ms bins in ITI
50 10-ms bins in trial
ITI
trial
time bin
cell
ITI PCA
trial PCA
percent explained
100
50
0
20
40
60
PC
cumulative sum
individual PC
top 10 PCA coefficients
0.2
0
-0.2

Mnemonic information is more decodable during the ITI than during the task

Our central hypothesis is that firing patterns during the inter-trial interval are neither functionally irrelevant nor otiose, but instead have a rich functional repertoire that differs from that of the task period. Based on existing literature, we hypothesized that functionally meaningful activity during the ITI would be associated with an expressly mnemonic role. In this task, rapid learning takes place on the basis of recent outcomes (Mehta et al., 2019). This adjustment may reflect a strategic response to evidence that environmental richness has changed (cf. Stephens & Krebs, 1986). In that study, we showed that *recent reward rate* specifically, the average reward obtained over the past ten trials, influences the subject's behavioral threshold on the current trial (β = 0.2313, p< 0.001, linear regression of combined data across subjects between local reward rate and behavioral threshold). Note that this effect is not simply due to single-trial adjustment (cf. Hayden et al., 2008; Sugrue et al., 2004). We know this because our analysis excludes the most recent trial. Thus all the events that drive the response to reward rate as we define it come from 10-70 seconds before the choice is expressed.

The recent reward rate is robustly encoded by responses of vmPFC neurons. Indeed, during the 500 ms trial epoch, 22.13% of neurons in vmPFC encoded the average reward rate on the past ten trials (18.57% and 26.92% in subjects J and T, respectively). (As above, this and the next analysis excludes the most recent trial to eliminate bias by possible hysteretic effects). A similar proportion of neurons (23.77%) encode the same information during the 500 ms ITI epoch (J: 28.57%; T:17.31%). All four of these proportions are substantially greater than chance (binomial test, p<0.0001 in all cases).

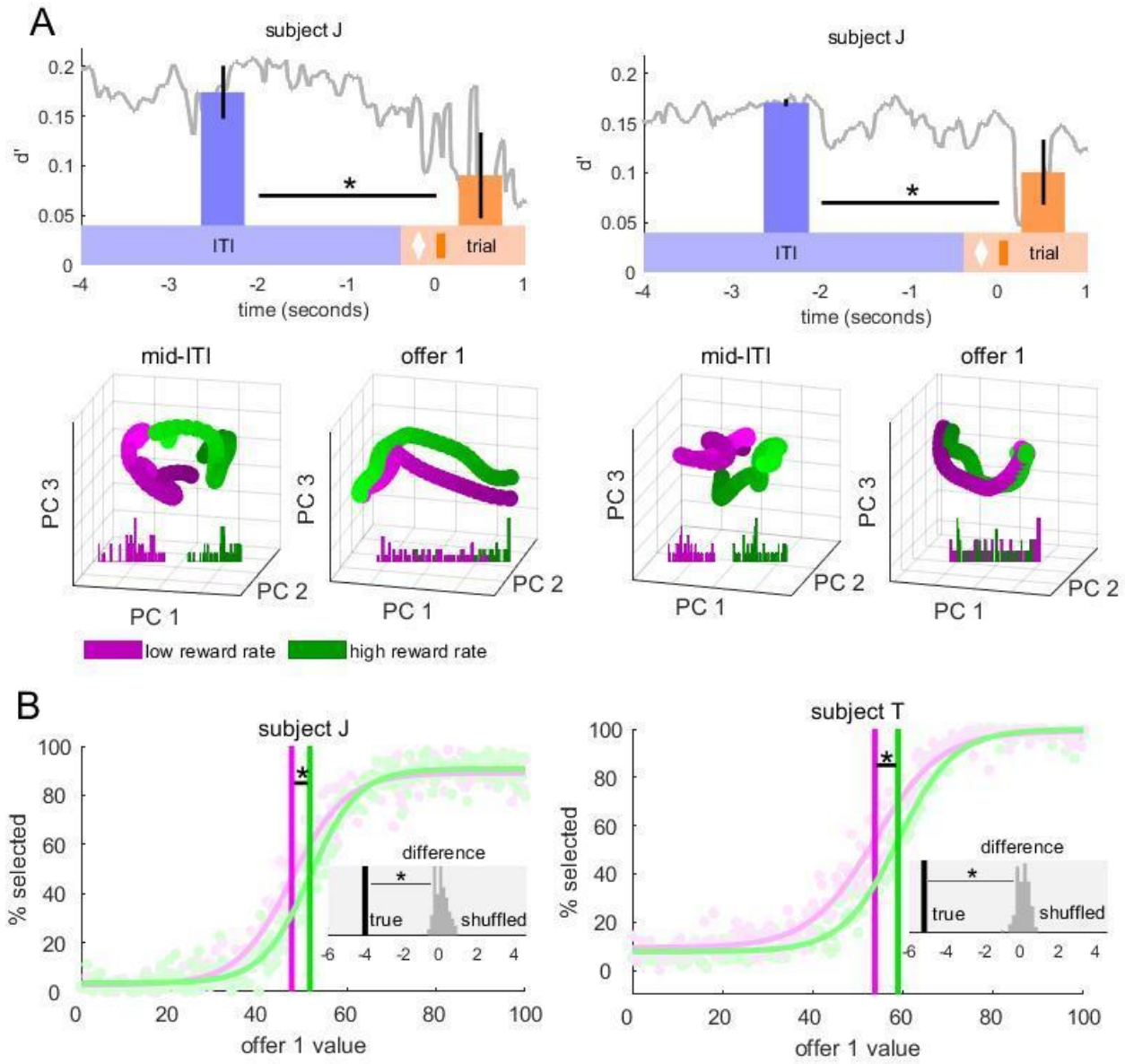
We next hypothesized that, if ITI was associated with mnemonic function, then the recent reward rate would be more *informationally accessible* - that is, encoded in a way that could be simply and efficiently decoded from a downstream computation - during the ITI than during the task period. To test this hypothesis, we first devised a measure of informational accessibility based closely on a technique previously devised by Cohen and Maunsell (2010) and extended by Keemink and Machens (2019). We used a dimensionality reduction approach; specifically, one that uses Principal Component Analysis (PCA). We first projected our condition-averaged peri-stimulus time histograms (which represent time binned vectors of firing) into a PCA space. That is, we averaged all the peri-stimulus time histograms in an epoch across trials within each condition (in this case, trials where the recent reward rate exceeded the mean reward rate (*high reward rate* trials) and trials where the recent reward rate was lower than the mean reward rate (*low reward rate* trials).

In this way, we mathematically reduced our set of firing rates down to a set of *principal components* that efficiently describe the population activity as it varies with reward rate. We focused on the three principal components that describe the largest amount of neural activity for each group of trials (high-reward-rate trials and low-reward-rate trials). This allowed us to graphically represent the population activity: each principal component became an axis, with each point on the resulting three-dimensional plot representing a projection of the neural population (**Figure 13A**). We compared the two sets of points - the neural population during low-reward-rate trials versus the neural population during high-reward-rate trials.

We were interested in the question of whether information is more separable (i.e., decodable) during the ITI than it is during the trial epoch (Cohen & Maunsell, 2010). To determine the separability between high-reward-rate and low-reward-rate trials, we calculated the summed Euclidean distance (*d'*) between the cloud of high-reward-rate points and that of low-reward-rate points. A larger distance means a greater difference in the way the vmPFC population represents high vs. low reward rate. In other words, a downstream brain region with access to this information would have an easier time determining whether reward rate was high or low with a high *d'.* We found that high and low reward rate trials were significantly more separable during the ITI than during the trial in both subjects (subject J: ITI epoch *d'* 6.97, trial epoch *d'* 0.63; $p<0.001$, paired t-test; subject T: ITI epoch *d'* 5.17, trial epoch *d'* 0.61, $p < 0.001$, paired t-test; **Figure 13A, Methods**).

If the reward rate is actively calculated during the ITI, then it should influence choice behavior during foraging. We previously reported that the behavioral threshold (the amount of reward the subject accepts 50% of the time) does change systematically with the reward rate (Mehta et al., 2019). Here we demonstrate a significant difference between behavioral threshold over the same categories we found separation between during the ITI: high and low reward rate (subject J: low reward rate threshold 47.41%, high reward rate threshold 51.50%, $p< 0.01$, 100-repetition bootstrap; subject T: low reward rate threshold 53.73%, high reward rate threshold 58.89%; $p<0.01$,100-repetition bootstrap; **Figure 13B, Methods**). This illustrates a potential use for distinguishing high and low reward rate before the trial starts: adjusting behavioral strategy in accordance with the reward rate.

Figure 13. Discriminability of reward rate is higher at ITI than task-on period, and subjects behaviorally exploit reward rate information. (A) *Upper panels: Discriminability of high versus low reward rate in top 3 PCs over time. Discriminability (d') at the midpoint of the ITI epoch is significantly higher than d' at the midpoint of the trial epoch in both subjects. Error bars represent standard deviation. Lower panels: Projection of principal components for high (green) and low (magenta) reward rates into a 3D space, where each point represents the top three principal components resulting from a PCA conducted on a 500 ms period centered around one time point. There are 50 time points for each condition, describing the 500 ms epoch centered on the midpoint of the ITI and trial epochs. Histograms depict the distances of each point from the plane that separates the two groups of points, as determined by linear discriminant analysis.* ***(B)*** *Subjects' behavioral threshold for accepting an offer changes as recent reward rate changes: behavior threshold is higher for high reward rates than low reward rates for both subjects. Insets: the difference between the high reward rate threshold and low reward rate thresholds is significantly greater than the difference resulting from shuffled data.*



Dimensionality of neural subspace is greatest during the inter-trial-interval

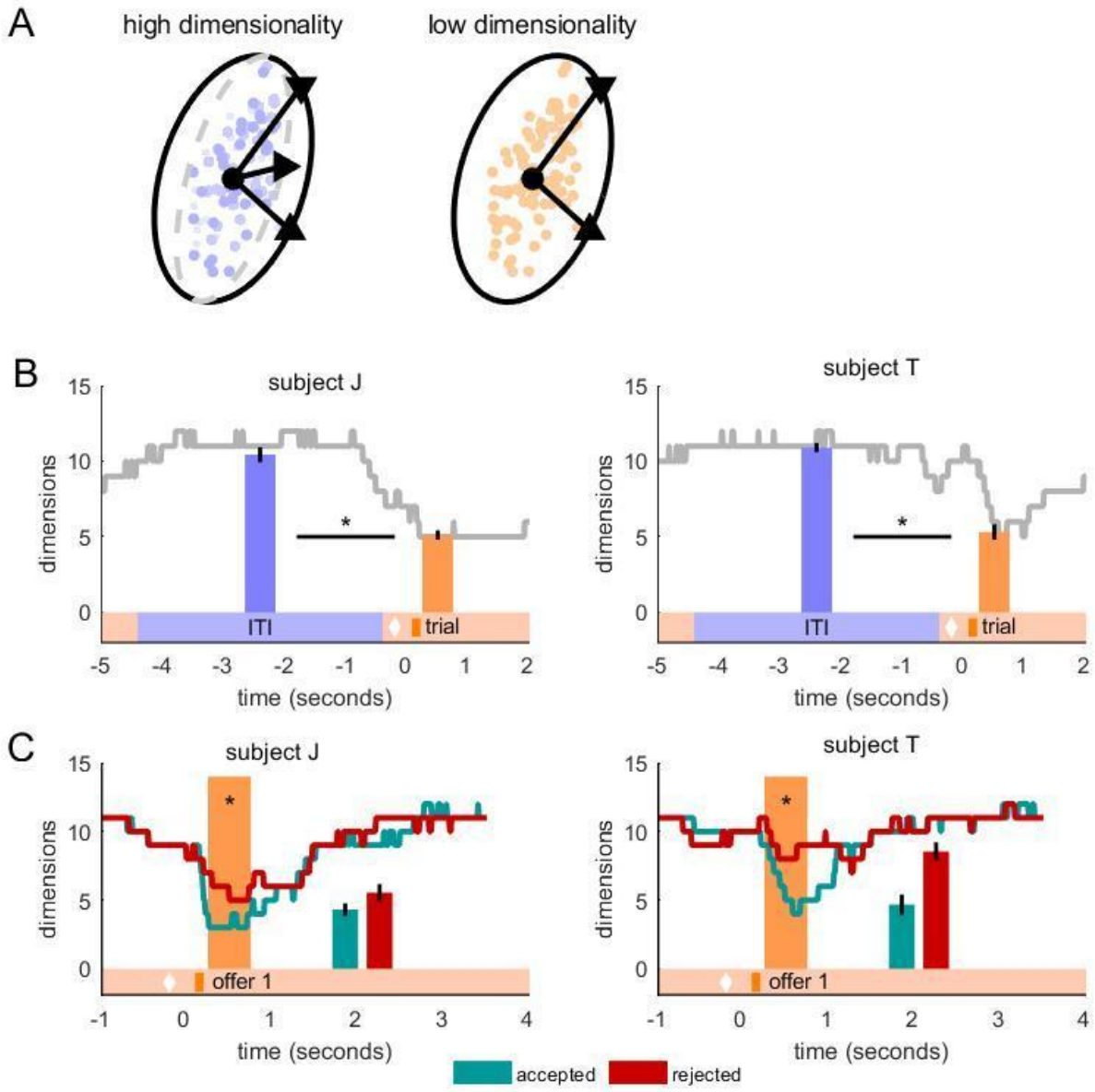
We next examined the relationship between ITI and trial epoch ensemble **dimensionality**. Ensembles of neurons have constrained firing rate patterns; that is, single neurons have a limited range of firing rates they can take, and this range is determined in part by the responses of other neurons. That range, collectively, defines the neural manifold or subspace of the population (Gallego et al., 2017; Fusi et al., 2016). Dimensionality, then, refers to the number of the axes within the larger space spanned by the subspace. It is a useful measure because it provides a potential proxy for the complexity of processing going on within

an area. Among other things, greater dimensionality also allows for a more ready separation of response patterns, thus allowing for more abstract processing and learning (Fusi et al., 2016).

There are many ways to quantify the dimensionality of a subspace. We used a previously developed method (Machens et al., 2010; Rouse & Scheiber, 2018). Specifically, we defined dimensionality as the number of principal components it takes to explain 90% of the variance in the neural activity at any particular moment (**Figure 14A**). We found that dimensionality is substantially greater during the ITI epoch than during the trial epoch in both subjects (subject J: 10.43 dimensions to 5.11 dimensions; $p< 0.001$, paired-t-test; subject T: 10.90 dimensions to 5.32 dimensions, $p<0.001$, paired t-test). (**Figure 14B)**. To compute these values and test significance, we conducted a bootstrap in which we left out 25% of our sample each time we computed this dimensionality value (using the remaining 75%) and averaged these values.

To gain deeper insight into these processes, we next asked whether this change in dimensionality is limited to the ITI vs task periods. We hypothesized that if it reflects variations in focus, we should be able to see changes as focus changes within the task itself (cf. Hayden et al., 2009). For example, post-reward periods should have greater task focus than search periods because they involve greater local focus (on consumption), while search periods involve relatively more global focus (because they involve a search-like contemplation of multiple options). Our data are consistent with this hypothesis. We compared the dimensionality of the trial epoch between accepted and rejected offers using the same method specified in Figure 3B. Viewing an accepted offer (which led to a reward consumption period) resulted in a significantly lower dimensionality than viewing a rejected offer (which continued the trial); the effect was observed in both subjects (**Figure 14C**, subject J: accepted offers 4.31 dimensions, rejected offers 5.54 dimensions, $p < 0.001$, unpaired T-test; subject T: accepted offers 4.67 dimensions, rejected offers 8.54 dimensions, $p<0.001$, unpaired T-test).

Figure 14. Subspace dimensionality is higher in the ITI than during the trial. (A) *Schematic representing high and low dimensionality subspaces in abstraction. Axes represent dimensions, while points represent principal components of neural activity. The high dimensionality subspace consists of points in a three-dimensional space, while the low dimensionality subspace consists of points in a two-dimensional space - three dimensions being unnecessary to accurately and completely represent the data.* ***(B)*** *Dimensionality is significantly higher during the ITI epoch than the trial epoch for both subjects. Each point on the gray line represents the dimensionality of subspace over the preceding 500 ms. Blue and orange bars represent the dimensionality values centered on the ITI epoch and trial epoch generated by bootstrapped data. Error bars are standard deviation.* ***(C)*** *Dimensionality is significantly lower for accepted offers (teal) than rejected offers (red) for both subjects. Teal and red lines represent dimensionality at each point in time as in (B). Bars represent dimensionality centered on the midpoint of the trial epoch for accepted and rejected offers, using bootstrapped data. Error bars are standard deviation.*



Reduction in dimensionality is not explained by task-driven firing rate changes

The observed reduction in dimensionality could potentially be explained by some other factor, such as the increase in firing at specific phases of the trial across all neurons. For example, a few neurons dramatically increasing their firing rate at stimulus onset could compress the subspace dimensionality, even though the effect would not be reflective of processing. Indeed, we found that average firing rate significantly increases between the ITI and task (significantly in subject J: mid-ITI normalized mean firing rate -1.03 spikes/second, offer 1

firing rate 1.39 spikes/second, p<0.001, paired t-test; subject T: mid-ITI normalized mean firing rate -0.09, offer 1 normalized firing rate 0.00, p =0.01, paired t-test, data smoothed over 20 bins;

To test for the possibility that firing rate changes are responsible for the dimensionality changes we see, we created control data that had the same mean firing rate as the real data, but scrambled relationships between individual neurons **(Methods).** In both subjects, the shuffled dimensionality is significantly higher than the true dimensionality (subject J: 7.57 dimensions versus 5.35 dimensions, p<0.001; subject T: 8.30 vs 7.33 dimensions, p, = 0.003). This result suggests that while some of the reduction in dimensionality could be due to the trial-driven firing rate transient, a large part of it is independent of that.

We created an additional set of control data that matched the ITI firing rate amplitude range to that of the trial. We found that ITI dimensionality, even with mean firing rate scaled to the trial firing rate, was still greater than trial dimensionality in both subjects (subject J: ITI mean ITI dimensionality 6.30 dimensions, mean trial dimensionality 5.33 dimensions, p<0.001, unpaired t-test; subject T: mean ITI dimensionality 5.66 dimensions, mean trial dimensionality 5.26 dimensions, p<0.001, unpaired t-test).

Reorganization rate of neural subspace reorganization is greatest during the inter-trial-interval

The subspace that a group of neurons' response occupies can change. This change, which can implement functional partition, is often an indicator of changing function roles of the population (Yoo & Hayden, 2020; Elsayed et al., 2016; Jiang et al., 2020; Tang et al., 2020) Here, we refer to the amount of subspace reorganization at each point in time as subspace *reorganization rate* - where a high amount of reorganization from one moment to the next reflects high change rate (and its converse, low stability). . Changes in neural activity patterns during rest states are generally considered to be slow (<0.1 Hz; Cabral et al., 2014; Deco & Jirsa, 2012; de Pasquale & Marzetti, 2019) . That is, techniques ranging from functional magnetic resonance imaging, to magnetoencephalography, to electrophysiology, have revealed that low-frequency patterns characterize neural activity during rest. As a result, one would expect the pattern of changes we find during the ITI to be slow as well. However, there is emerging evidence of rapid computations during rest broadly across the default mode network (Baker et al., 2014; Khanna et al., 2015), and specifically in the motor cortex (Vidaurre et al., 2016). Here we provide evidence of fast-changing neural patterns (significant reorganization of the neural subspace between 500 ms epochs) during rest in the ventromedial prefrontal cortex.

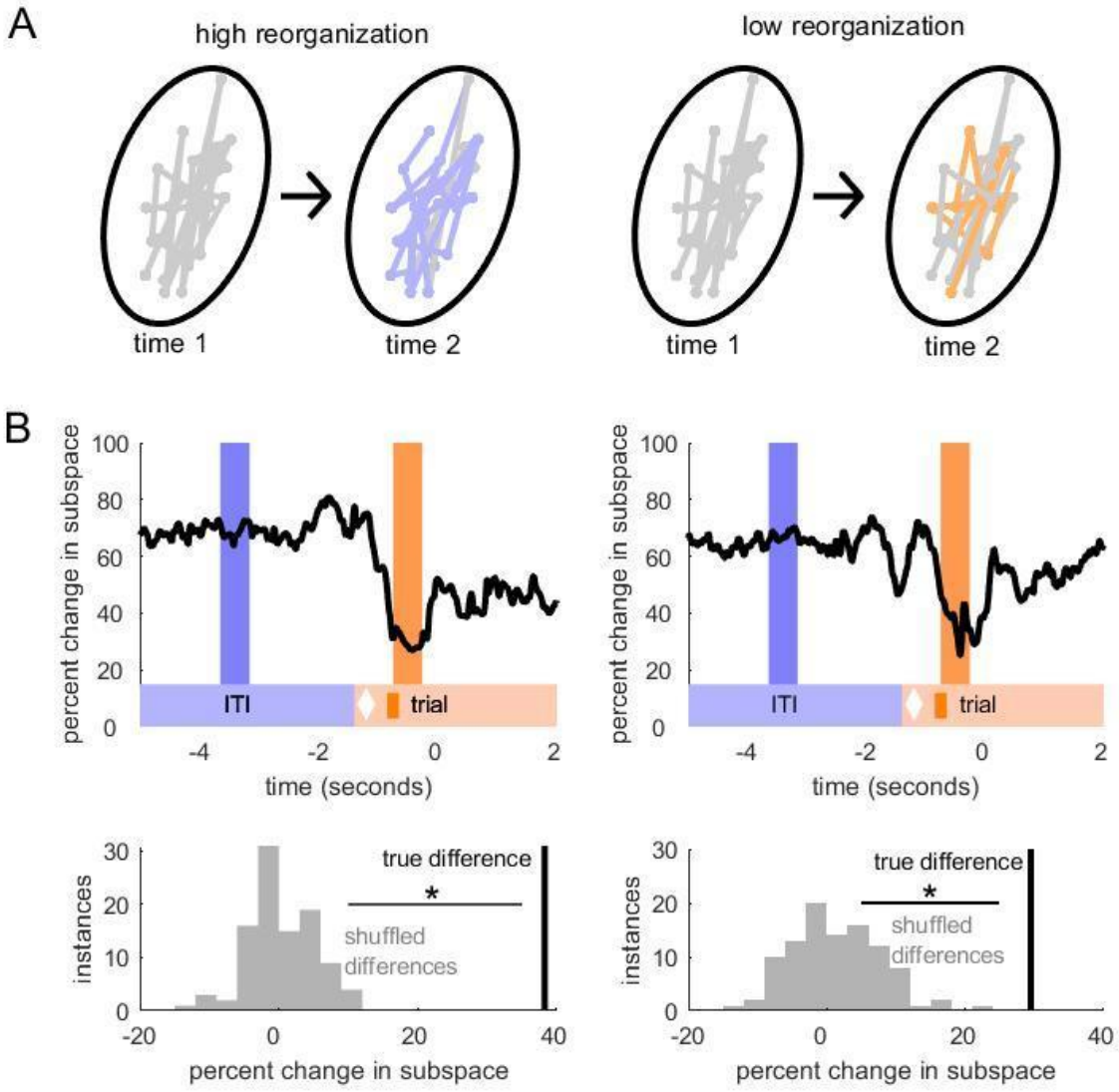
These results point to a changing view of default mode activity as not only slow oscillatory activity patterns, but also fast-changing dynamic activity.

To assess the subspace reorganization rate, we measured the similarity of the eigenvectors between each pair of subspaces before and after every given time point (10 ms intervals) of the task with 500 ms epochs (**Methods**). We considered every timepoint from the beginning of the ITI through two seconds after the reveal of the first offer. For each timepoint, we defined two time periods - 500 ms before, and 500 ms after - and computed the PCA coefficients for both time periods. We then projected the coefficients from the *before* time period into the *after* subspace and determined the percentage of *after* variance that was explained by the *before* coefficients (Elsayed et al., 2016; Yoo & Hayden, 2020; Jiang et al., 2020; Khanna et al., 2019). This overlap, expressed as a percentage, indicates the amount of reorganization of the neural subspace. A reorganization rate of 100% means the *before* subspace coefficients explained none of the *after* subspace (fast reorganization), and thus has completely reorganized , while 0% velocity means the *before* subspace explains all of the *after* subspace, and thus has not reorganized at all.

We found that the subspace reorganization rate is significantly greater during the ITI than during the task for both subjects (**Figure 15B**). Specifically, the average reorganization for sequential epochs was 66.59% in subject J during the ITI epoch but only 28.34% during the trial epoch ($p<0.001$, one-sample t-test between true difference and shuffled differences, see **Methods**). For subject T, these numbers were 64.59% in the ITI epoch and 43.22% during the trial epoch ($p<0.001$, one-sample t-test between true difference and shuffled differences). In other words, vmPFC neuronal subspace occupancy changes much more over time during the ITI than during the task; once the trial begins, however, the neural subspace becomes much more stable.

Finally, we note that the ITI subspace is distinct from the trial subspace. That is, we conducted the same projection analysis on the 500 ms epoch centered around the midpoint of the ITI and the 500 ms epoch centered around the midpoint of the trial epoch, and found that the PC space of the ITI epoch changes 83.07% in subject J and 76.76% in subject T.

Figure 15. Subspace reorganizes much more during ITI than during trial.
(A) *Schematic of high subspace reorganization versus low subspace reorganization in abstraction. Each network of points represents principal components of neural activity at a particular time. Edges between points represent relationships between PCs. In high reorganization, most of the relationships change over time. In low reorganization, a few relationships change over time.* ***(B)*** *Upper panels: The percent change in subspace over the midpoint of the ITI epoch is greater than that over the midpoint of the trial epoch in both subjects. Lower panels: this difference is significantly greater than the distribution of differences generated by shuffling data 100 times.*


A
high reorganization
low reorganization
time 1
time 2
time 1
time 2
B
percent change in subspace
100
80
60
40
20
0
ITI
trial
-4
-2
0
2
time (seconds)
percent change in subspace
100
80
60
40
20
0
ITI
trial
-4
-2
0
2
time (seconds)
instances
30
20
10
0
true difference
*
shuffled
differences
-20
0
20
40
percent change in subspace
instances
30
20
10
0
true difference
*
shuffled
differences
-20
0
20
40
percent change in subspace

DISCUSSION

The ITI, and globally focused cognition more generally, occupies a large amount of our lives and yet has not been studied nearly as much as task-oriented locally focused cognition. Indeed, despite its quotidian nature, it remains strangely inaccessible to many conventional cognitive neuroscientific methods (Raichle, 2006). Nonetheless, studies of foraging behavior have demonstrated that performing computations on global variables is critical to foraging success (Clark & Mangel, 1984; Hayden et al., 2011; Cisek & Pastor-Bernier, 2014;). Here, we propose that some of these crucial computations occur during travel time - the ITI.

Here, we examined the responses of a population of neurons in area 14 of the vmPFC while macaques performed a foraging search task (Mehta et al., 2019). We imposed an unusually long interval between sequential trials (duration of 4 seconds), a design feature that allowed us to compare neuronal ensemble responses during task-on (trial) and task-off (inter-trial interval, ITI) periods while avoiding immediate post-trial and pre-trial effects. We examined encoding of a key piece of mnemonic information in the task - recent reward rate over multiple trials, which we show are attended and learned. During the ITI, that information is encoded in a more linearly separable format: an effect we demonstrate with our discriminability analysis. In other words, information is present during both the task and ITI epochs, but it is more untangled during the ITI (DiCarlo et al., 2012; Yoo & Hayden, 2018; DiCarlo & Cox, 2007). This finding suggests that the brain actively reformats task-relevant information depending on task context and, in particular, that it makes mnemonic information more functionally available during the ITI than during the task itself.

Why would reward rate information be more separable during the ITI than during the task itself? During the trial, the subject must do two things - (1) apply learned information for purposes of guiding decisions and (2) be alert for new incoming information. Both processes may be antagonistic to the kinds of neural processes that are associated with mnemonic processing, including computing and representing the recent average reward value. In other words, explicit reward representation is a guide for behavior but it is conceptually quite distinct from its implementation, and cognition may have a finite capacity. It may be efficient, then, to process it asynchronously during low-demand periods. Indeed, we hypothesize that the more accessible format during the delay reflects the need to process it and develop a decision strategy; that strategy only needs to be implemented during choice, so the information that guides the strategy can be stored in a more latent manner.

We also found that the ITI is associated with potentially more complex processing, as indicated by two complementary measures: subspace dimensionality and subspace change

rate. Dimensionality refers to the volume occupied by an ensemble in a high-dimensional space: a higher dimension is presumed to reflect less redundancy and more complex and multifarious response patterns (Cunningham & Byron, 2014). Subspace change rate, on the other hand is a new measure that we introduce here, but it is one that follows naturally from past work on subspace reorganization - it is, essentially the rate at which reorganization occurs. We reasoned that faster reorganization may imply rapid cognitive processing with multiple changing priorities. Some of the dimensionality and velocity results are likely related to the specific mnemonic function discussed above, but some likely reflect even more complex processing to which we don't have experimental access.

Why would the brain alternate between processing different types of information in different modes? We speculate that the brain is not a strictly parallel processor, and that its capacity is highly restricted (Shenav et al., 2017). Specifically, we conjecture that broad and narrow-focused attention naturally competes and cannot be effectively implemented at the same time. In other words, we propose that information processing is analogous to preemptive multiple processing in computer science. This type of multiple processing operates by prioritizing the task and context switch operation allows the interruption of certain tasks and resumes later. The global information processing can be interrupted by task engagement (local phase) and resumed once the individual occasion of search ends. According to preemptive multiple processing, the operating processor never shuts down but they switch the information being processed.

METHODS

We conducted all procedures 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. These data were previously analyzed in a published study (Mehta et al., 2019). Subjects were two male rhesus macaques (*Macaca mulatta*: subject J age 10 years; subject T age 5 years). To maintain head position, we used a small titanium prosthesis. We trained the subjects first to habituate to laboratory conditions and then to perform oculomotor tasks for liquid reward. To record neural data, we placed a Cilux recording chamber (Crist instruments) over the ventromedial prefrontal cortex (vmPFC) of each animal. We verified the accuracy of this position using magnetic resonance and a Brainsight system (Rogue Research Inc.). We provided the animals with all appropriate analgesics and antibiotics post-procedure, and regularly sterilely cleaned and sealed the implanted chamber.

Recording site

We performed all recordings between morning and midday (10am to 3pm), using a standard recording grid (Crist Instruments) to approach area 14 (vmPFC) as described by the Paxinos atlas (Paxinos et al., 2000). Specifically, we recorded neurons located between 42 and 31 mm rostral to the interaural plane, 0 and 7 mm horizontally from the brain's ventral surface, and 0 to 7 mm sagittal from the medial wall.

Electrophysiological techniques, eye tracking and reward delivery

All methods used have been described previously and are summarized here (Strait et al., 2014). We used a microdrive (NAN instruments) to guide single electrodes (Frederick Haer & Co., impedance range 0.7 to 5.5 MU). Each recording session, we isolated between one and four neurons on a Plexon system (Plexon, Inc.). We used data from all neurons we could isolate for at least 300 consecutive trials, regardless of their behavior during the task. In practice, 86% of neurons had over 500 trials

To obtain eye position data ast 1,000 Hz, we used an infrared eye-monitoring camera system (SR Research). We controlled the task presentation with a computer running Matlab (Mathworks) with Psychtoolbox and Eyelink Toolbox. Our visual stimuli were colored shapes on a computer monitor located 60 cm from the animal at eye level. We controlled liquid reward delivery with a standard solenoid valve.

Experimental Design

Subjects performed a diet-search task (Figure 1) that is described in detail in a previous publication (Mehta et al., 2019). The task is a conceptual extension of previous foraging tasks we have developed. Each trial, subjects were presented with either four or seven white diamonds that appeared in random positions on the screen. 400 ms of continuous fixation on one of these diamonds caused it to disappear and reveal a reward offer underneath. Reward offers were partially-filled-in orange bars that indicated riskless reward amount by the proportion of their area that was filled in. There were 200 possible reward offer sizes per offer. To accept an offer, subjects had to continuously fixate on one for 300 ms.

Subjects freely searched through the diamonds in any order, but had to accept an offer to obtain a reward and end the trial. Rejection of an offer (breaking fixation before 300 ms had elapsed) resulted in a continuation of the trial (free search through the diamonds). Reward was delivered immediately after acceptance of an offer, which was followed by a four-second inter-trial-interval.

Previous training history for these subjects included two types of foraging tasks (Blanchard and Hayden, 2015; Blanchard et al., 2015), complex choice tasks with time elements (Blanchard et al., 2014), two types of gambling tasks (Farashahi et al., 2018; Azab and Hayden, 2017), attentional tasks (similar to those in Hayden and Gallant, 2013), and two types of reward-based decision tasks (Pirrone et al., 2018; Wang and Hayden, 2017).

Statistical Analysis for Physiology

We formed Peri-Stimulus Time Histograms (PSTHs) by aligning spike rasters to each offer reveal of each trial and averaging firing rates across multiple trials. Firing rates were normalized (where indicated) by subtracting the mean and dividing by the standard deviation of the entire neuron's PSTH.

Analysis epochs

Before beginning analysis, we determined two epochs over the course of the trial to correspond to the ITI and the task. We selected the 500 ms period in the middle of the 4-second ITI to represent the *ITI epoch*, and the 500ms in the middle of the 1-second epoch following the reveal of the first offer to represent the *offer epoch*.

Creating the PC spaces

To create a PC-space for the ITI and first offer of each trial, we began with the PSTHs of all the trials, sorted into two groups: trials where the subject accepted the first offer, and trials where the subject rejected the first offer. These PSTHs were Gaussian smoothed over 15 time bins (150 ms), and normalized. (Note that while data was divided into accept and reject data for the dimensionality and orthogonality analyses, it was divided into low and high reward rate datasets for the discriminability analysis).

We then found the mean PSTH for each cell for each of these groups, giving us two matrices per subject - each with cells as rows and time points of the PSTH as columns. In other words, we computed the average behavior of each cell for trials where the subject accepted the first offer, and trials where the subject rejected the first offer. We then appended the *reject* matrix to the *accept* matrix to create a combined matrix.

To obtain PC spaces over the course of the ITI and first offer, we used a sliding boxcar of 50 time bins (500 ms) to perform a PCA on each 500-ms epoch of PSTH data from the beginning of the ITI through 2 seconds following the reveal of the first offer.

To obtain PC spaces for trials following the acceptance or rejection of an offer, we used a sliding boxcar of 50 time bins to perform a PCA on each 500-ms epoch of PSTH data from 2 seconds before the reveal of each offer to 3 seconds after.

Discriminability analysis

We defined *discriminability* between two task conditions as the separation between representations of those variables in principal component space. To this end, we performed PCAs as described above, using neural data with each trial split into either a low-reward rate (between 0 and 71% for subject J; 0 and 72% for subject T, see Mehta et al. 2019 for calculation of reward rate), or high-reward-rate (71%-100% for subject J; 72% and above for subject T) condition. We then obtained the top three principal components (the principal components that explain the three largest amounts of variance in the neural activity) for each condition and computed the summed euclidean distance between pairs of points in each category. That is, for each of the 50 time points, we computed the euclidean distance between the low-reward-rate point and the high-reward-rate point. We defined the distance between the two categories (*d'*) as the sum of all 50 of these values.

While we visualized the data using only the top three PCs from each PCA, this result holds across many different numbers of PCs. For subject J, discriminability between low and high reward rate is significant through 8 PCs, and for subject T, it is significant through all 52 PCs. Furthermore, discriminability for subject J is significant for all 70 PCs if the length of the analysis epoch is extended from 500 ms to 1 s, due to the fact that the discriminability drop at the trial onset in subject J occurs a little later than that of subject T. However, we conducted the analysis on the same predetermined 500 ms epoch in both subjects in the interest of preventing p-hacking.

Dimensionality analysis

We defined *dimensionality* of an epoch as the number of principal components it takes to explain 90% of the variance in the neural activity during that epoch (Elsayed et al., 2016; Yoo &Hayden, 2020). We used the PCAs run as detailed above on our epochs of interest and obtained the percentage of the variance that each PC explained, sorted from greatest variance explained to least. For each epoch of interest, we found the number of principal components, starting with the one that explained the most variance, that it took to explain a minimum of 90% of the total variance. This value we termed the 'dimensionality.'

It is important to note that dimensionality can only be a whole number: it represents a number of principal components. When we report dimensionality as having decimal values, it is because we have averaged several whole numbers over a period of time. We interpret these decimal values as estimates of dimensionality.

To perform significance tests on dimensionality data, we had to bear in mind that each sequential time point's dimensionality is created from data that overlaps with the next time point. Thus, we could not perform a simple t-test between, for example, the dimensionalities of the time points between the reveal of the first offer and one second following. Instead, we used a bootstrap: we selected the central time point of each analysis epoch (which represents data from 250 ms before that point and 250 ms after) and recalculated it 100 times using 75% of the original data each time. We then performed a t-test between these two groups of 100 points between epochs.

To create control data, we randomly swapped the PSTHs of the *offer 1* epoch between cells. For example, if there were three real cells, Cell A, B and C and three control cells, Cell A', B', and C, Cell A' may have the *offer 1* epoch of Cell C', Cell B' might have the *offer 1* epoch of Cell A, and Cell C would have the *offer 1* epoch of Cell B. Thus, the control data would have the same mean firing rate as the real data. However, suppose that the *combination* of Cells A and B specifically is meaningful to the vmPFC; that is, population information is meaningful. Now, the combination of Cell A and Cell B's firing rate is meaningless, because their PSTHs have been swapped out with those of other cells. Thus, our control data was such that the mean firing rate at every time point of the *offer 1* epoch was the same, but the properties of the data at the population level are meaningless.

To create our set of control data where ITI firing rate amplitude range matched that of the trial, we did the following. First, we determined the minimum and maximum meaned firing rate over the 1-second epoch following the reveal of the first offer for each cell. Then, we scaled that cell's ITI PSTH (4-second epoch) so that it had the same maximum and minimum value, and then proceeded with our dimensionality analysis in the same way as the original.

Reorganization analysis

We defined *reorganization* as the percentage of variance in the neural activity of one epoch that that of the immediately previous epoch fails to explain. In other words, it is the extent to which the principal component space has changed. To estimate the amount of reorganization between subspaces at a moment in time, we did the following (Elsayed et al., 2016; Yoo & Hayden, 2020). We performed PCAs on two adjacent 500 ms epochs: one spanning time *t :*

t+499 (epoch 1), and one spanning time *t+500 : t+999* (epoch 2). To obtain the amount of reorganization between epoch 1 and epoch 2, we projected the input data from epoch 1 into the PC space generated by the PCA on epoch 2. We then calculated the percent of variance in epoch 2 explained by the data from epoch 1, and subtracted it from the percent of variance in epoch 2 explained by the data from epoch 2. This percentage is the amount of reorganization. For example, if the data from epoch 1 had explained 30% of the variance in epoch 2, and the data from epoch 2 (using the same projection method) explains 98% of the variance in epoch 2, then we would say the amount of reorganization is 68%.

To perform significance tests on orthogonality data, we noted that orthogonality values are derived from principal components that describe overlapping epochs of neural data. Thus again, we could not perform a simple t-test between the average orthogonality of different epochs. Instead, we created shuffled data by scrambling the PCA inputs between the two epochs of interest, 100 times. We then calculated the orthogonality of the central time point of the ITI and the offer epoch 100 times with these sets of scramble data. We compared the true difference in orthogonality to the distribution of shuffled differences with a t-test to obtain a p-value.

Chapter III: Further explorations into the contributions of reward rate to behavior, and future directions

Environmental factors including, but not limited to, reward rate are clearly involved in computations of the vmPFC during a foraging task. But a deeper look into behavior can elucidate more on how these computations are being made.

In chapter I, we found a positive correlation between regression coefficients for firing rates in response to the current offer versus that of the previous offer. In other words, we found that neurons that preferred high-value offers at one point in time tended to also prefer higher-value offers at a subsequent point in time. This is evidence against the idea that the brain is actively comparing one offer to the next while foraging; if it were, a neuron that preferred a high value offer 1 would not also prefer a high-value offer 2, as a comparison signal would signal the difference between the two offers. Instead, this result may be a point in favor of the common currency approach - the pure values alone seem to be recorded in vmPFC in the same currency.

While these results mean that consecutive offers are encoded using the same neural code, they do not necessitate that the foragers are using a common currency to compare offers. Instead, they may not be comparing offers at all. One of the core elements of a foraging task is that it is not a forced-choice - the animal is under no obligation to compare *offer 2* to *offer 1*. All they must do is compare offer 1 to *something* - be it something about the environment, or some internal metric.

To this end, we performed some further analysis in our diet-search task. First, we compared the coefficients resulting from a regression of firing rate versus the magnitude of the previous offer (with respect to the current offer) with the coefficients from a regression of firing rate vs the outcome of the previous trial. In both subjects, we found that previous outcome and previous offer are negatively correlated (subject J: $R = -0.32$, $p = 0.0065$; subject T: $R = -0.25$, $p = 0.069$), indicating that it is likely that these values are being compared. Thus when the subject has just rejected an offer, their vmPFC neurons are comparing the outcome of the previous trial to the offer the subject just rejected.

This comparison can be interpreted as a signal of foraging success - if the animal just rejected something lower than the previous outcome, they are behaving in keeping with what they assume about the current environmental richness. However, if they just rejected something higher than what they received on the previous trial, they are behaving a little recklessly. Such information is important for the foraging brain.

The forager might also take into account everything they currently believe about the environment, and then make a decision from that result. In this vein, we performed several more behavioral analyses seeking the influence of reward rate on behavior during the trial. We found several interesting phenomena. First, reward rate is positively correlated with value of the chosen offer (subject J beta = 0.13, $p<0.001$; subject T beta = 0.19, $p <0.001$;) and the number of offers each subject decided to inspect each trial (subject J beta = 1.39, $p < 0.001$; subject T beta = 2.84, $p< 0.001$). Conversely, reward rate has a negative relationship with time needed to make a decision to reject an offer (subject J beta = -0.01, $p = 0.004$; subject T beta = -0.05, $p< 0.001$). This information indicates that reward rate information is being incorporated into the decision at each point to accept or reject each offer. It must be part of the equation. Furthermore, we found that (in all cases but that of encoding of the current offer), the effect of reward rate on these behavioral variables (measured as the mean absolute beta coefficient values) was *stronger* than that of the previous offer or outcome (decision time: subject J, reward rate coefficients = 0.013, previous outcome coefficients = 0.003, $p < 0.001$, subject T reward rate coefficients = 0.05, previous outcome coefficients = 0.02, $p<0.001$; number of offers inspected: subject J reward rate coefficients = 1.13, previous outcome coefficients = 0.32, $p < 0.001$, subject T reward rate coefficients 2.54, previous outcome coefficients 0.64, $p < 0.001$; trial times: subject J reward rate coefficients = 14.14, previous outcome coefficients = 3.55, $p < 0.001$, subject T reward rate coefficients = 4.00, previous outcome coefficients = 2.30, $p < 0.001$).

So far we have discussed at great length the value of the foraging perspective in examining neural activity during a foraging task. The true test of the strength of the idea, however, is in applying it somewhere not explicitly designed for it to apply. Thus, we examined the impact of the environment on behavior in a forced-choice, two-option task.

Our *conflict task* was designed to encourage the subject to compare two offers where one offer was better on one dimension, but the other offer was better on the other (for example, an expensive, but large, glass of lemonade, versus a cheap, but small glass of lemonade). Stimuli for the task were circles that had two distinct signals for value: color and radius. Three different possible colors indicated 3 different values. 200 different circle radii indicated 200 different values. The ‘correct’ response for each trial was the offer that represented the highest combination of color and radius (*integrated value,* a simple multiplication of the two numerical values represented by color and radius).

We looked here for signals that environmental values contribute to the processing of value. We found promising behavioral evidence for our central metric of task performance:

accuracy: the percentage of instances where the subject chose the offer with a higher integrated value. We found our subject's accuracy to be higher when the outcome of the previous trial was higher (low previous outcome accuracy = 75.38%, high previous outcome accuracy = 80.73%, $p = 0.006$), when the total value of the two simultaneous offers was higher (low environmental value accuracy = 70.59%, high environmental value accuracy = 86.35%, $p < 0.001$), and when the reward rate was higher (low reward rate accuracy = 76.05%, high reward rate accuracy = 81.55%, $p < 0.001$). All three of these variables are signals for environmental richness, clearly present even when the task is not structured like a foraging task.

In the future, we hope to be able to access even further insight into the neural evaluation of complex value signals through both monkey and human fMRI. To this end, we have built a mock scanner designed to gradually desensitize monkeys to the environment of magnetic resonance imaging, in which they have been learning to perform a decision-making task for juice rewards. The task in question is a two-option forced-choice task, but with several qualities that make it interesting to study from a foraging perspective.

The task consists of two phases: exposure and choice. During the exposure phase, the subject fixates on a neutral image while offers appear on either side of the image. During the choice phase, the neutral image disappears and the subject may move their eyes to one side or the other to choose that offer.

Offers are displayed as piles of colored sticks, where the number of sticks represents the offer magnitude (1-3 drops of juice), the color of the sticks represents the time cost of selecting that offer (1, 6, or 12 seconds), and the width of the sticks represents the probability of receiving a reward if that offer is chosen (33%, 66%, or 99%). Because each offer simultaneously conveys information about multiple distinct reward attributes, this task will allow us to examine where value signals are disparate according to modality and where they are integrated. Furthermore, manipulating the particularly long inter-trial-interval (6-12 seconds) and the offer time costs can provide us with valuable foraging information in seeing how the subjects' behavior and brain responds to different rates of reward intake.

Even so, the ultimate value of the task lies in its methodology - by performing fMRI with awake, behaving monkeys, we can perform the same exact experiment with humans and directly compare the data. When scientists work with animals, there is very often an underlying assumption that the information gained can somehow, in some way, be interpreted in a human context. Comparing data from the same task, using the same method, in both humans and monkeys would be invaluable information for the scientific community because it would shed light upon what kinds of data can legitimately be cross-interpreted between the two species.

This can lead to less wasteful science, which in turn leads not only to better data but more ethical practices within animal research.

Conclusion

In chapter I, we described how environmental factors like the magnitude of recent rewards and the rate of reward intake influence a forager's brain and behavior. These results make a compelling argument for applying foraging principles to the study of neuroeconomics, but they also demonstrate that the vmPFC is more than just a "local value region" of the brain. Instead, it dynamically adjusts its value-related activity to the surrounding environment.

In chapter II, we demonstrated that high reward-rate environments were more distinguishable from low reward-rate environments during travel time than during foraging, and in Chapter III, we further explored the myriad influences of reward rate and environmental variables on decision making behavior. In addition to bolstering the argument towards more widely using foraging as a tool to study neuroeconomics, we also see that the vmPFC plays a more complex role in value-related cognition than previously considered. Not only does it encode foraging variables as a population in addition to with individual neurons, but it is also actively involved in computations even in a *task-off* situation - when there is nothing explicit for the animal to do. VmPFC can no longer be considered a region primarily tied to the value of local offers if it is engaging in complex, behaviorally-relevant processing when there are no offers to be seen.

While these conclusions are exciting, they do not disprove the idea of a common currency approach to decision-making. Indeed, the purpose of these analyses is not to claim the common currency approach to studying decision making is wrong. Instead we propose that decision scientists forgo considering how offers are compared in the first place, because offer comparison is not the most natural way we make decisions.

The claim that values of various items are translated into a currency to be compared to other values in the brain holds some merit, but our data necessitates a second look into what we consider to be "value comparison" in the first place. Perhaps, our brains are not actually wired to compare pairs of values, but rather to compare potential rewards to current assumptions about the environment. Instead of looking for signatures of vmPFC activity pitting two simultaneously-presented offers against each other, we can think about value in terms of how likely an actor is to choose an offer in a particular circumstance versus another. This likelihood is influenced by a range of factors including its monetary or caloric value, but also the richness of the environment and all the various costs associated with obtaining the reward. VmPFC may not be computing a common currency at all, but rather, the signal we are seeing is a symptom of vmPFC reflecting how likely we are to select an option.

Padoa-Schippa (2011) describes the decision-making process with a table that posits an intermediary step between computing the different attributes and making a decision, where the decision-maker must translate the two options into a common currency. However, evolutionarily speaking, currency did not exist when our brains evolved. Computationally, the step is unnecessary - there are many other ways to arrive at a decision; for example a ramping of activity indicating bias towards a particular direction (Blanchard et al, 2015; Huk & Shadlen, 2005).

It is natural to use a modern analogy to describe the brain. Society has described the brain variously as a series of animate forces, an electrical circuit, a switchboard, and of course a computer (Gentner & Grudin, 1985; Eliasmith, 2003), in lock-step with the evolution of technology. So it is natural to imagine that our brains have taken into account the modern human invention of money. But as mentioned before, animals do not use money. In foraging, whether an animal receives a food item comes down to one central question - to accept it or to reject it. Thus, the decision to accept or reject is the central focus of all our analyses. With this, we hope to shift the narrative of vmPFC's contribution to a choice from *valuing* the option to *influencing* the choice to accept or reject.

We might find the most success moving away from the two-option choice model and towards a model where the decision maker's choice is not one or the other, but a serial search through multiple available options, each one at which they must say "no" or "yes." After all, life usually does not function like the meal choice menu at a wedding, where you must put a tick mark next to exactly one option at every juncture. Even when faced with such a decision, - a menu of food that only has two options - we *still* rarely have exactly two options. Often, we are allowed to purchase both items, and if not, we usually have the option to accept *neither* offer. When we limit ourselves to only looking at how local value is encoded in the brain, we miss out on a huge piece of the puzzle. Lemonade is never just lemonade - it's five or ten or a hundred different pros and cons swirling around in our minds, compressing so fast into a decision to drink it or not that we don't even know we have made a choice. Thinking about choice in this way facilitates not only a better understanding of neuroeconomic decision making, but also anything that involves behavior. Ultimately, the power of the foraging framework is that it takes everything that is generally a pesky impediment to clear science - uncontrollable, outside variables - and turns it into an advantage. Armed with this power, we can become free of many of the constraints in our experiments, and open the doors to new worlds of knowledge.

REFERENCES

Aridan, N., Malecek, N. J., Poldrack, R. A., & Schonberg, T. (2019). Neural correlates of effort-based valuation with prospective choices. Neuroimage, 185, 446-454.

Arditi, R., & Dacorogna, B. (1988). Optimal foraging on arbitrary food distributions and the definition of habitat patches. *The American Naturalist*, *131*(6), 837-846.

Azab, H., & Hayden, B. Y. (2017). Correlates of decisional dynamics in the dorsal anterior cingulate cortex. PLoS biology, 15(11), e2003091.

Baird, B., Smallwood, J., Gorgolewski, K. J., & Margulies, D. S. (2013). Medial and lateral networks in anterior prefrontal cortex support metacognitive ability for memory and perception. *Journal of Neuroscience*, *33*(42), 16657-16665.

Baker, A. P., Brookes, M. J., Rezek, I. A., Smith, S. M., Behrens, T., Smith, P. J. P., & Woolrich, M. (2014). Fast transient networks in spontaneous human brain activity. Elife, 3, e01867.

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage*, *76*, 412-427.

Barack, D. L., Chang, S. W. C., & Platt, M. L. (2017). Posterior cingulate neurons dynamically signal decisions to disengage during foraging. Neuron, 96(2), 339–347.

Bechara, A., Tranel, D., & Damasio, H. (2000). Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. Brain, 123(11), 2189-2202.

Behrens, T. E., Muller, T. H., Whittington, J. C., Mark, S., Baram, A. B., Stachenfeld, K. L., & Kurth-Nelson, Z. (2018). What is a cognitive map? Organizing knowledge for flexible behavior. Neuron, 100(2), 490-509.

Blanchard, T. C., Wilke, A., & Hayden, B. Y. (2014). Hot-hand bias in rhesus monkeys. Journal of Experimental Psychology: Animal Learning and Cognition, 40(3), 280.

Blanchard, T. C., & Hayden, B. Y. (2014). Neurons in dorsal anterior cingulate cortex signal postdecisional variables in a foraging task. *Journal of Neuroscience*, *34*(2), 646-655.

Blanchard, T. C., & Hayden, B. Y. (2015). Monkeys are more patient in a foraging task than in a standard intertemporal choice task. PloS one, 10(2), e0117057.

Blanchard, T. C., Hayden, B. Y., & Bromberg-Martin, E. S. (2015). Orbitofrontal cortex uses distinct codes for different choice attributes in decisions motivated by curiosity. Neuron, 85(3), 602-614.

Blanchard, T. C., Strait, C. E., & Hayden, B. Y. (2015). Ramping ensemble activity in dorsal anterior cingulate neurons during persistent commitment to a decision. *Journal of neurophysiology*, *114*(4), 2439-2449.

Boly, M., Balteau, E., Schnakers, C., Degueldre, C., Moonen, G., Luxen, A., ... & Laureys, S. (2007). Baseline brain activity fluctuations predict somatosensory perception in humans. *Proceedings of the National Academy of Sciences*, *104*(29), 12187-12192.

Boorman, E. D., Behrens, T. E., Woolrich, M. W., & Rushworth, M. F. (2009). How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. Neuron, 62(5), 733-743.

Boorman, E. D., Rushworth, M. F., & Behrens, T. E. (2013). Ventromedial prefrontal and anterior cingulate cortex adopt choice and default reference frames during sequential multi-alternative choice. *Journal of Neuroscience*, *33*(6), 2242-2253.

Bouret, S., & Richmond, B. J. (2010). Ventromedial and orbital prefrontal neurons differentially encode internally and externally driven motivational values in monkeys. Journal of Neuroscience, 30(25), 8591-8601.

Cabral, J., Kringelbach, M. L., & Deco, G. (2014). Exploring the network dynamics underlying brain activity during rest. Progress in neurobiology, 114, 102-131.

Calhoun, A. J., & Hayden, B. Y. (2015). The foraging brain. *Current Opinion in Behavioral Sciences*, *5*, 24-31.

Carter, E. C., Pedersen, E. J., & McCullough, M. E. (2015). Reassessing intertemporal choice: human decision-making is more optimal in a foraging task than in a self-control task. *Frontiers in psychology*, *6*, 95.

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, V. S., Rangel, A., Shimojo, S., & O'Doherty, J. P. (2009). Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *Journal of Neuroscience*, *29*(39), 12315-12320.

Christopoulos, V., & Schrater, P. R. (2015). Dynamic integration of value information into a common probability currency as a theory for flexible decision making. *PLoS Comput Biol*, *11*(9), e1004402.

Chow, Y. S., Moriguti, S., Robbins, H., & Samuels, S. M. (1964). Optimal selection based on relative rank (the “secretary problem”). *Israel Journal of mathematics*, *2*(2), 81-90.

Cisek, P., & Pastor-Bernier, A. (2014). On the challenges and mechanisms of embodied decisions. Philosophical Transactions of the Royal Society B: Biological Sciences, 369(1655), 20130479.

Cisek, P., & Kalaska, J. F. (2010). Neural mechanisms for interacting with a world full of action choices. Annual review of neuroscience, 33, 269-298.

Clark, C. W., & Mangel, M. (1984). Foraging and flocking strategies: information in an uncertain environment. The American Naturalist, 123(5), 626-641.

Clithero, J. A., & Rangel, A. (2013). Informatic parcellation of the network involved in the computation of subjective value. Social cognitive and affective neuroscience, 9(9), 1289-1302.

Cohen, M. R., & Maunsell, J. H. (2010). A neuronal population measure of attention predicts behavioral performance on individual trials. *Journal of Neuroscience*, *30*(45), 15241-15253.

Cole, M. W., Ito, T., Bassett, D. S., & Schultz, D. H. (2016). Activity flow over resting-state networks shapes cognitive task activations. *Nature neuroscience*, *19*(12), 1718-1726.

Constantino, S. M., & Daw, N. D. (2015). Learning the opportunity cost of time in a patch-foraging task. *Cognitive, Affective, & Behavioral Neuroscience*, *15*(4), 837-853.

Cunningham, J. P., & Byron, M. Y. (2014). Dimensionality reduction for large-scale neural recordings. Nature neuroscience, 17(11), 1500-1509.

Daw, N. D., O'doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*(7095), 876-879.

Daselaar, S. M., Prince, S. E., & Cabeza, R. (2004). When less means more: deactivations during encoding that predict subsequent memory. *Neuroimage*, *23*(3), 921-927.

Dean, H. L., Crowley, J. C., & Platt, M. L. (2004). Visual and saccade-related activity in macaque posterior cingulate cortex. Journal of neurophysiology, 92(5), 3056-3068.

Deco, G., & Jirsa, V. K. (2012). Ongoing cortical activity at rest: criticality, multistability, and ghost attractors. Journal of Neuroscience, 32(10), 3366-3375.

Delgado, M. R., Beer, J. S., Fellows, L. K., Huettel, S. A., Platt, M. L., Quirk, G. J., & Schiller, D. (2016). Viewpoints: dialogues on the functional role of the ventromedial prefrontal cortex.

de Pasquale, F., & Marzetti, L. (2019). Temporal and spectral signatures of the default mode network. Magnetoencephalography: From Signals to Dynamic Cortical Networks, 571-603.

DiCarlo, J. J., & Cox, D. D. (2007). Untangling invariant object recognition. Trends in cognitive sciences, 11(8), 333-341.

DiCarlo, J. J., Zoccolan, D., & Rust, N. C. (2012). How does the brain solve visual object recognition?. Neuron, 73(3), 415-434.

Eliasmith, C. (2003). Moving beyond metaphors: Understanding the mind for what it is. *The Journal of philosophy*, *100*(10), 493-520.

Elsayed, G. F., Lara, A. H., Kaufman, M. T., Churchland, M. M., & Cunningham, J. P. (2016). Reorganization between preparatory and movement population responses in motor cortex. *Nature communications*, *7*(1), 1-15.

Farashahi, S., Azab, H., Hayden, B., & Soltani, A. (2018). On the flexibility of basic risk attitudes in monkeys. Journal of Neuroscience, 38(18), 4383-4398.

Ferguson, T. S. (1989). Who solved the secretary problem?. *Statistical science*, *4*(3), 282-289.

FitzGerald, T. H., Seymour, B., & Dolan, R. J. (2009). The role of human orbitofrontal cortex in value comparison for incommensurable objects. Journal of Neuroscience, 29(26), 8388-8395.

Fusi, S., Miller, E. K., & Rigotti, M. (2016). Why neurons mix: high dimensionality for higher cognition. *Current opinion in neurobiology*, *37*, 66-74.

Gaba, A. S., & Apps, M. A. (2020). Foraging Optimally in Social Neuroscience: Computations and Methodological considerations. Social cognitive and affective neuroscience.

Gallego, J. A., Perich, M. G., Chowdhury, R. H., Solla, S. A., & Miller, L. E. (2020). Long-term stability of cortical population dynamics underlying consistent behavior. *Nature neuroscience*, *23*(2), 260-270.

Gallego, J. A., Perich, M. G., Miller, L. E., & Solla, S. A. (2017). Neural manifolds for the control of movement. *Neuron*, *94*(5), 978-984.

Gao, P., & Ganguli, S. (2015). On simplicity and complexity in the brave new world of large-scale neuroscience. *Current opinion in neurobiology*, *32*, 148-155.

Gentner, D., & Grudin, J. (1985). The evolution of mental metaphors in psychology: A 90-year retrospective. *American Psychologist*, *40*(2), 181.

Gusnard, D. A., & Raichle, M. E. (2001). Searching for a baseline: functional imaging and the resting human brain. Nature Reviews Neuroscience, 2(10), 685.

Hare, T. A., Camerer, C. F., & Rangel, A. (2009). Self-control in decision-making involves modulation of the vmPFC valuation system. Science

Hayden, B. Y. (2016). Time discounting and time preference in animals: a critical review. Psychonomic bulletin & review, 23(1), 39-53.

Hayden, B. Y. (2018). Economic choice: the foraging perspective. Current Opinion in Behavioral Sciences, 24, 1-6.

Hayden, B. Y., & Moreno-Bote, R. (2018). A neuronal theory of sequential economic choice. Brain and Neuroscience Advances, 2, 2398212818766675.

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

Hayden, B. Y., Pearson, J. M., & Platt, M. L. (2009). Fictive reward signals in the anterior cingulate cortex. *science*, *324*(5929), 948-950.

Hayden, B. Y., Smith, D. V., & Platt, M. L. (2009). Electrophysiological correlates of default-mode processing in macaque posterior cingulate cortex. *Proceedings of the National Academy of Sciences*, *106*(14), 5948-5953.

Hayden, B. Y., Pearson, J. M., & Platt, M. L. (2011). Neuronal basis of sequential foraging decisions in a patchy environment. *Nature neuroscience*, *14*(7), 933.

Hayden, B., & Gallant, J. (2013). Working memory and decision processes in visual area v4. Frontiers in neuroscience, 7, 18.

Heilbronner, S. R., & Hayden, B. Y. (2016). Dorsal anterior cingulate cortex: a bottom-up view. Annual review of neuroscience, 39, 149-170.

Heilbronner, S. R., & Hayden, B. Y. (2016). The description-experience gap in risky choice in nonhuman primates. Psychonomic bulletin & review, 23(2), 593-600.

Heilbronner, S. R. (2017). Modeling risky decision-making in nonhuman animals: shared core features. Current opinion in behavioral sciences, 16, 23-29.

Heracleous, L. T. (1994). Rational decision making: myth or reality?. *Management development review*.

Hills, T. T., Todd, P. M., Lazer, D., Redish, A. D., Couzin, I. D., & Cognitive Search Research Group. (2015). Exploration versus exploitation in space, mind, and society. Trends in cognitive sciences, 19(1), 46-54.

Huk, A. C., & Shadlen, M. N. (2005). Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. *Journal of Neuroscience*, *25*(45), 10420-10436.

Jiang, X., Saggar, H., Ryu, S. I., Shenoy, K. V., & Kao, J. C. (2020). Structure in Neural Activity during Observed and Executed Movements Is Shared at the Neural Population Level, Not in Single Neurons. *Cell reports*, *32*(6), 108006.

Kable, J. W., & Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. Nature neuroscience, 10(12), 1625.

Kable, J. W., & Glimcher, P. W. (2009). The neurobiology of decision: consensus and controversy. *Neuron*, *63*(6), 733-745.

Kao, T. C., Sadabadi, M. S., & Hennequin, G. (2020). Anticipatory control of movement in a thalamo-cortical circuit model. *bioRxiv*.

Keemink, S. W., & Machens, C. K. (2019). Decoding and encoding (de) mixed population responses. *Current Opinion in Neurobiology*, *58*, 112-121.

Khanna, A., Pascual-Leone, A., Michel, C. M., & Farzan, F. (2015). Microstates in resting-state EEG: current status and future directions. Neuroscience & Biobehavioral Reviews, 49, 105-113.

Khanna, S. B., Snyder, A. C., & Smith, M. A. (2019). Distinct sources of variability affect eye movement preparation. *Journal of Neuroscience*, *39*(23), 4511-4526.

Kolling, N., Behrens, T. E., Mars, R. B., & Rushworth, M. F. (2012). Neural mechanisms of foraging. *Science*, *336*(6077), 95-98.

Kolling, N., Wittmann, M., & Rushworth, M. F. (2014). Multiple neural mechanisms of decision making and their competition under changing risk pressure. Neuron, 81(5), 1190-1202.

Kolling, N., Behrens, T. E., Wittmann, M. K., & Rushworth, M. F. (2016). Multiple signals in anterior cingulate cortex. *Current opinion in neurobiology*, *37*, 36-43.

Konishi, Mahiko, Donald George McLaren, Haakon Engen, and Jonathan Smallwood. "Shaped by the past: the default mode network supports cognition that is independent of immediate perceptual input." *PloS one* 10, no. 6 (2015): e0132209.

Krebs, J. R., Kacelnik, A., & Taylor, P. (1978). Test of optimal sampling by foraging great tits. *Nature*, *275*(5675), 27-31.

Landreth, A., & Bickle, J. (2008). Neuroeconomics, neurophysiology and the common currency hypothesis. *Economics and Philosophy*, *24*(3), 419.

Lebreton, M., Abitbol, R., Daunizeau, J., & Pessiglione, M. (2015). Automatic integration of confidence in the brain valuation signal. Nature neuroscience, 18(8), 1159.

Lee, D. (2006). Neural basis of quasi-rational decision making. *Current opinion in neurobiology*, *16*(2), 191-198.

Levy, D. J., & Glimcher, P. W. (2012). The root of all value: a neural common currency for choice. *Current opinion in neurobiology*, *22*(6), 1027-1038.

Lim, S. L., O'Doherty, J. P., & Rangel, A. (2011). The decision value computations in the vmPFC and striatum use a relative value code that is guided by visual attention. *Journal of Neuroscience*, *31*(37), 13214-13223.

Machens, C. K., Romo, R., & Brody, C. D. (2010). Functional, but not anatomical, separation of “what” and “when” in the prefrontal cortex. *Journal of Neuroscience*, *30*(1), 350-360.

Mars, R. B., Verhagen, L., Gladwin, T. E., Neubert, F. X., Sallet, J., & Rushworth, M. F. (2016). Comparing brains by matching connectivity profiles. Neuroscience & Biobehavioral Reviews, 60, 90-97.

Mason, M. F., Norton, M. I., Van Horn, J. D., Wegner, D. M., Grafton, S. T., & Macrae, C. N. (2007). Wandering minds: the default network and stimulus-independent thought. *Science*, *315*(5810), 393-395.

Mazoyer, B., Zago, L., Mellet, E., Bricogne, S., Etard, O., Houdé, O., ... & Tzourio-Mazoyer, N. (2001). Cortical networks for working memory and executive functions sustain the conscious resting state in man. *Brain research bulletin*, *54*(3), 287-298.

Mehta, P. S., Tu, J. C., LoConte, G. A., Pesce, M. C., & Hayden, B. Y. (2019). Ventromedial prefrontal cortex tracks multiple environmental variables during search. *Journal of Neuroscience*, *39*(27), 5336-5350.

Metereau, E., & Dreher, J. C. (2015). The medial orbitofrontal cortex encodes a general unsigned value signal during anticipation of both appetitive and aversive events. *Cortex*, *63*, 42-54.

Mobbs, D., Trimmer, P. C., Blumstein, D. T., & Dayan, P. (2018). Foraging for foundations in decision neuroscience: insights from ethology. *Nature Reviews Neuroscience*, *19*(7), 419-427.

Monosov, I. E., & Hikosaka, O. (2012). Regionally distinct processing of rewards and punishments by the primate ventromedial prefrontal cortex. Journal of Neuroscience, 32(30), 10318-10330.

Öngür, D., & Price, J. L. (2000). The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. Cerebral cortex, 10(3), 206-219.

Northoff, G., Qin, P., & Nakao, T. (2010). Rest-stimulus interaction in the brain: a review. *Trends in neurosciences*, *33*(6), 277-284.

Padoa-Schioppa, C. (2011). Neurobiology of economic choice: a good-based model. *Annual review of neuroscience*, *34*, 333-359.

Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature*, *441*(7090), 223-226.

Padoa‐Schioppa, C., & Cai, X. (2011). The orbitofrontal cortex and the computation of subjective value: consolidated concepts and new perspectives. Annals of the New York Academy of Sciences, 1239(1), 130-137.

Passingham, R. E., & Wise, S. P. (2012). *The neurobiology of the prefrontal cortex: anatomy, evolution, and the origin of insight* (No. 50). Oxford University Press.

Paxinos, G., Huang, X. F., & Toga, A. W. (2000). The rhesus monkey brain in stereotaxic coordinates.

Pearson, J. M., Watson, K. K., & Platt, M. L. (2014). Decision making: the neuroethological turn. Neuron, 82(5), 950-965.

Pearson, J. M., Hayden, B. Y., Raghavachari, S., & Platt, M. L. (2009). Neurons in posterior cingulate cortex signal exploratory decisions in a dynamic multioption choice task. Current biology, 19(18), 1532-1537.

Pearson, J., Hayden, B., & Platt, M. (2010). Explicit information reduces discounting behavior in monkeys. Frontiers in Psychology, 1, 237.

Pezzulo, G., & Cisek, P. (2016). Navigating the affordance landscape: feedback control as a process model of behavior and cognition. Trends in cognitive sciences, 20(6), 414-424.

Pirrone, A., Azab, H., Hayden, B. Y., Stafford, T., & Marshall, J. A. (2018). Evidence for the speed–value trade-off: Human and monkey decision making is magnitude sensitive. Decision, 5(2), 129.

Plassmann, H., O'Doherty, J., & Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. Journal of neuroscience, 27(37), 9984-9988.

Raichle, M. E. (2006). The brain's dark energy. Science-New York Then Washington-, 314(5803), 1249.

Raichle, M. E. (2010). Two views of brain function. *Trends in cognitive sciences*, *14*(4), 180-190.

Raichle, M. E. (2015). The brain's default mode network. *Annual review of neuroscience*, *38*, 433-447.

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. Nature reviews neuroscience, 9(7), 545.

Rangel, A., & Hare, T. (2010). Neural computations associated with goal-directed choice. Current opinion in neurobiology, 20(2), 262-270.

Rich, E. L., & Wallis, J. D. (2016). Decoding subjective decisions from orbitofrontal cortex. Nature neuroscience, 19(7), 973.

Rigotti, M., Barak, O., Warden, M. R., Wang, X. J., Daw, N. D., Miller, E. K., & Fusi, S. (2013). The importance of mixed selectivity in complex cognitive tasks. *Nature*, *497*(7451), 585-590.

Rouse, A. G., & Schieber, M. H. (2018). Condition-dependent neural dimensions progressively shift during reach to grasp. *Cell reports*, *25*(11), 3158-3168.

Rushworth, M. F., Noonan, M. P., Boorman, E. D., Walton, M. E., & Behrens, T. E. (2011). Frontal cortex and reward-guided learning and decision-making. Neuron, 70(6), 1054-1069.

Rushworth, M. F., Kolling, N., Sallet, J., & Mars, R. B. (2012). Valuation and decision-making in frontal cortex: one or many serial or parallel systems?. Current opinion in neurobiology, 22(6), 946-955.

Schlichting, M. L., & Preston, A. R. (2015). Memory integration: neural mechanisms and implications for behavior. Current opinion in behavioral sciences, 1, 1-8.

Seale, D. A., & Rapoport, A. (1997). Sequential decision making with relative ranks: An experimental investigation of the" secretary problem". *Organizational behavior and human decision processes*, *69*(3), 221-236.

Sescousse, G., Redouté, J., & Dreher, J. C. (2010). The architecture of reward value coding in the human orbitofrontal cortex. Journal of neuroscience, 30(39), 13095-13104.

Shenhav, A., Straccia, M. A., Botvinick, M. M., & Cohen, J. D. (2016). Dorsal anterior cingulate and ventromedial prefrontal cortex have inverse roles in both foraging and economic choice. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(6), 1127-1139.

Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., & Botvinick, M. M. (2017). Toward a rational and mechanistic account of mental effort. *Annual review of neuroscience*, *40*, 99-124.

Simon, H. A. (1979). Rational decision making in business organizations. *The American economic review*, *69*(4), 493-513.

Skaggs, W. E., McNaughton, B. L., & Gothard, K. M. (1993). An information-theoretic approach to deciphering the hippocampal code. In Advances in neural information processing systems (pp. 1030-1037).

Skaggs, W. E., McNaughton, B. L., Wilson, M. A., & Barnes, C. A. (1996). Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences. Hippocampus, 6(2), 149-172.

Sleezer, B. J., Castagno, M. D., & Hayden, B. Y. (2016). Rule encoding in orbitofrontal cortex and striatum guides selection. Journal of Neuroscience, 36(44), 11223–11237.

Smallwood, J., Tipper, C., Brown, K., Baird, B., Engen, H., Michaels, J. R., ... & Schooler, J. W. (2013). Escaping the here and now: evidence for a role of the default mode network in perceptually decoupled thought. *Neuroimage*, *69*, 120-125.

Smith, D. V., Hayden, B. Y., Truong, T. K., Song, A. W., Platt, M. L., & Huettel, S. A. (2010). Distinct value signals in anterior and posterior ventromedial prefrontal cortex. Journal of Neuroscience, 30(7), 2490-2495.

Stephens, D. W. (1989). Variance and the value of information. *The American Naturalist*, *134*(1), 128-140.

Stephens, D. W., & Krebs, J. R. (1986). *Foraging theory*. Princeton university press.

Strait, C. E., Blanchard, T. C., & Hayden, B. Y. (2014). Reward value comparison via mutual inhibition in ventromedial prefrontal cortex. *Neuron*, *82*(6), 1357-1366.

Strait, C. E., Sleezer, B. J., & Hayden, B. Y. (2015). Signatures of value comparison in ventral striatum neurons. PLoS biology, 13(6), e1002173.

Strait, C. E., Sleezer, B. J., Blanchard, T. C., Azab, H., Castagno, M. D., & Hayden, B. Y. (2016). Neuronal selectivity for spatial positions of offers and choices in five reward regions. Journal of neurophysiology, 115(3), 1098-1111.

Sugrue, L. P., Corrado, G. S., & Newsome, W. T. (2004). Matching behavior and the representation of value in the parietal cortex. *science*, *304*(5678), 1782-1787.

Tang, C., Herikstad, R., Parthasarathy, A., Libedinsky, C., & Yen, S. C. (2020). Minimally dependent activity subspaces for working memory and motor preparation in the lateral prefrontal cortex. *eLife*, *9*, e58154.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *science*, *211*(4481), 453-458.

Vassena, E., Krebs, R. M., Silvetti, M., Fias, W., & Verguts, T. (2014). Dissociating contributions of ACC and vmPFC in reward prediction, outcome, and choice. *Neuropsychologia*, *59*, 112-123.

Vlaev, I., Chater, N., Stewart, N., and Brown, G. D. A. (2011). Does the brain calculate value? Trends in Cognitive Science 15: 546–554.

Wallis, J. D., & Kennerley, S. W. (2010). Heterogeneous reward signals in prefrontal cortex. Current opinion in neurobiology, 20(2), 191-198.

Wang, M. Z., & Hayden, B. Y. (2017). Reactivation of associative structure specific outcome responses during prospective evaluation in reward-based choices. Nature communications, 8(1), 1-13.

Watson, K. K., & Platt, M. L. (2012). Social signals in primate orbitofrontal cortex. *Current Biology*, *22*(23), 2268-2273.

Weissman, D. H., Roberts, K. C., Visscher, K. M., & Woldorff, M. G. (2006). The neural bases of momentary lapses in attention. *Nature neuroscience*, *9*(7), 971-978.

Winecoff, A., Clithero, J. A., Carter, R. M., Bergman, S. R., Wang, L., & Huettel, S. A. (2013). Ventromedial prefrontal cortex encodes emotional value. *Journal of Neuroscience*, *33*(27), 11032-11039.

Wise, S. P., Murray, E. A., & Gerfen, C. R. (1996). The frontal cortex-basal ganglia system in primates. Critical Reviews™ in Neurobiology, 10(3-4).

Wunderlich, K., Rangel, A., & O'Doherty, J. P. (2009). Neural computations underlying action-based decision making in the human brain. Proceedings of the National Academy of Sciences, 106(40), 17199-17204.

Yoo, S. B. M., & Hayden, B. Y. (2018). Economic choice as an untangling of options into actions. Neuron, 99(3), 434-447.

Yoo, S. B. M., & Hayden, B. Y. (2020). The transition from evaluation to selection involves neural subspace reorganization in core reward regions. *Neuron*, *105*(4), 712-724.

Yoo, S. B. M., Sleezer, B. J., & Hayden, B. Y. (2018). Robust encoding of spatial information in orbitofrontal cortex and striatum. Journal of cognitive neuroscience, 30(6), 898-913.

Zou, Q., Ross, T. J., Gu, H., Geng, X., Zuo, X. N., Hong, L. E., ... & Yang, Y. (2013). Intrinsic resting‐state activity predicts working memory brain activation and behavioral performance. *Human brain mapping*, *34*(12), 3204-3215.