UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title
Prefrontal and Subcortical Dynamics of Value-Based Decision Making

Permalink
https://escholarship.org/uc/item/98v9p7z0

Author
Balewski, Zuzanna Zofia

Publication Date
2022

Peer reviewed|Thesis/dissertation

Prefrontal and Subcortical Dynamics of Value-Based Decision Making

by

Zuzanna Balewski

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Neuroscience

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Joni Wallis, Chair
Assistant Professor Anne Collins
Professor David Foster
Professor Frederic Theunissen

Summer 2022

Prefrontal and Subcortical Dynamics of Value-Based Decision Making

Abstract

Prefrontal and Subcortical Dynamics of Value-Based Decision Making

by

Zuzanna Balewski

Doctor of Philosophy in Neuroscience

University of California, Berkeley

Professor Joni Wallis, Chair

The orbitofrontal cortex (OFC) is a region critical to value-based decision making. Previous work from our lab showed that the neural population in OFC flip-flops between representing the available options while animals are deliberating between them. While these value dynamics were predictive of subjects' choices, they did not build toward the chosen option preceding the choice. We hypothesised that a downstream region with connections to both OFC and motor regions, for example the anterior cingulate cortex (ACC) or head of the caudate nucleus (CdN), may be integrating over these OFC value dynamics, leading to the choice. We recorded large neural populations simultaneously from OFC and each of these regions while two monkeys used a bidrectional lever to select between pairs of pictures, each associated with a unique probabilistic reward. In the ACC, we indeed observed a motor preparation signal that was influenced by OFC value dynamics on individual trials. However, in the CdN we observed a rapid choice direction signal, locked to the start of the trial instead of the motor response, that was independent of OFC value dynamics. Overall, our results are consistent with parallel, rather than serial, processing underlying value-based decision making.

To Cleo

Contents

List of Figures

List of Tables

Acknowledgments

This work would not have come together without the support and guidance of many mentors and friends.

I would first like to thank Joni Wallis for giving me much needed structure and direction over the last few years. Her unfaltering optimism and belief in science pulled me through multiple world-ending crises.

I would also like to thank Eric Knudsen for teaching me everything I know about primate electrophysiology, and generously volunteering his time and energy every time I asked for help.

I feel very fortunate to have learned from and worked with many incredible scientists, both at Berkeley and beyond. Thank you to Erin Rich, whose work inspired all of my projects, and the rest of the lab: Feng-Kuei Chiang, Thom Elston, Celia Ford, Eric Hu, Lauren Meckler, Nathan Munet. Thank you to Ev Fedorenko, who first welcomed me into the world of research. And thank you to many people who offered invaluable advice, particularly Tanner Dixon, Will Liberti, Pratiksha Thaker, and Ellen Zippi.

I am especially grateful for Kevin Yu, who patiently discussed every miniscule update on my work, made whimsical monkey illustrations for my presentations, and added a lot of joy to my life. I would also like to thank my family, Ewa, Jan, and Weronika Balewski, for blindly admiring my graphs, feeding me, and always supporting me for the last 3 (!) decades.

Finally, I would like to thank my thesis committee, Anne Collis, David Foster, and Frederic Theunissen, for always providing thoughtful feedback; the OLAC veterinary and husbandry teams for excellent animal care; and the HWNI faculty and administrative staff, especially Candace Groskreutz and Leleña Avila, for fostering a supportive graduate community.

Chapter 1

Introduction

I am often faced with decisions that involve comparing disparate options. For example, I may need to choose between eating breakfast and catching the bus to work; in the long run, I may consider buying a car or just a louder alarm to help me wake up earlier. What are the neural computations underlying these complex decisions?

The orbitofrontal cortex (OFC) is a region critical to this behavior. Patients with OFC lesions are impaired in decision-making despite otherwise unaffected cognitive function (Eslinger & Damasio, 1985; Szczepanski & Knight, 2014). Lesion and microstimulation studies from rats and monkeys indicate that OFC is critical for learning the expected outcomes associated with reward-predictive cues (Ballesta et al., 2020; Burke et al., 2008; Knudsen & Wallis, 2020; Rudebeck & Murray, 2011), especially when new information needs to be computed or integrated (Murray & Rudebeck, 2018; Stalnaker et al., 2015).

Historically, when only a handful of cells could be recorded at a time, the relationship between behavior and neuronal responses was usually examined through trial-averaged firing rates. Studies have shown that OFC neurons encode the value of the chosen options, typically independent of sensory information like stimulus identity or motor movement required to indicate the choice (Kennerley & Wallis, 2009; Padoa-Schioppa, 2011; Padoa-Schioppa & Assad, 2006).

However, by only examining average neuronal responses aligned to task events (e.g.: option presentation, motor movement), we may have been missing important dynamics in the evolution of individual decisions. The cognitive processes underlying deliberative decisions may vary from trial to trial, but are difficult to study directly because they are unobservable from animals' behavior alone (Wallis, 2018).

To tackle this problem, Rich and Wallis (2016) leveraged larger, simultaneously recorded OFC populations to first find neural patterns corresponding to different value representations, then infer the single-trial dynamics through this value representation space while monkeys deliberated between two available options. OFC value representations flip-flopped between the available options in a unique pattern on each trial; the strength and duration of these representations during the deliberative period was predictive of subjects' choices and response times. This serial representation of available options may reflect some deliberative

comparison process.

Importantly, Rich and Wallis (2016) found that subjects' choices could not be predicted from OFC population responses immediately before or during choice execution. In other words, the internal value representation did not build toward the chosen option preceding the motor movement. They hypothesized that a downstream region may integrate over these OFC value comparisons to drive the ultimate choice or motor output (Rich & Wallis, 2016; Wallis, 2018), analogous to the noisy sensory information process observed in the lateral interparietal area (Gold & Shadlen, 2001; Mazurek et al., 2003; Roitman & Shadlen, 2002) in the visual system.

In this thesis, we set out to find such an integrator region. We looked for evidence that a motor preparation signal was influenced by the simultaneous OFC value representation. If such a relationship exists, we expected to find it in a region that is (1) anatomically connected to both OFC and motor regions, and (2) critical to choice selection. We explored two such candidate regions: the head of the caudate nucleus (CdN) and the anterior cingulate cortex (ACC). Both regions are anatomically well-positioned for this role, with strong bidirectional connections to OFC (*CdN*: Clarke et al., 2014; Ferry et al., 2000; Haber et al., 1995; *ACC*: Barbas and Pandya, 1989; Carmichael and Price, 1996) and premotor or motor areas (*CdN*: Inase et al., 1996; Takada et al., 2001, *ACC*: Dum and Strick, 1993; Morecraft et al., 2012; Van Hoesen et al., 1993. Neurons in both regions encode value and direction information (*CdN*: Grahn et al., 2008; Hori et al., 2020; Kawagoe et al., 1998; Lauwereyns et al., 2002; *ACC*: Cai and Padoa-Schioppa, 2012; Hunt et al., 2018; Kennerley and Wallis, 2009; Luk and Wallis, 2013). Finally, microstimulation of CdN can bias subjects toward a specific reward-predictive cue (Nakamura & Hikosaka, 2006; Santacruz et al., 2017) and damage to the ACC produces decision-making impairments (Camille et al., 2011; Kennerley & Walton, 2011; Kennerley et al., 2006).

We simultaneously recorded large populations of neurons from OFC and either CdN or ACC while two monkeys viewed pairs of pictures and indicated their preferred option with a lever. We examined single-trial population dynamics in all three regions, adapting the approach from Rich and Wallis (2016). We found a motor preparation signal in ACC that was indeed influenced by OFC value dynamics, consistent with the hypothesized downstream integrator model. In contrast, the CdN choice direction signal was locked to fast, value-guided saccades early in the trial, not the motor choice response. Additionally, this fast direction signal was independent of OFC value dynamics, suggesting these regions reflect parallel value processes in the brain.

Outline

Chapter 2, *A model-based approach for targeted neurophysiology in the behaving non-human primate*, was previously published as Knudsen et al., 2019 and included here with minor modifications. To obtain reliable single-trial signals required for the subsequent experiments, we needed to collect large (∼100 cells) neural recordings simultaneously from multiple brain regions. We developed a modular recording platform that allowed us to pre-

cisely target brain regions of interest with multiple multi-channel probes (maximally 256 channels; Plexon: V- and K-probes) and could be easily 3D printed and assembled in-house.

Chapter 3, *Fast and slow contributions to decision making in corticostriatal circuits*, was previously published as Balewski, 2022 and included here with minor modifications. We observed a rapid, phasic signal in the head of the caudate nucleus (CdN) that predicted the choice response and closely aligned with the subjects' initial orienting eye movements. We replicated the OFC value dynamics reported by Rich and Wallis (2016), but found no relationship with the rapid CdN direction response, suggesting these signals reflect independent cognitive processes.

Chapter 4, *Value dynamics in orbitofrontal cortex drive the choice response in anterior cingulate cortex during decision-making*. We observed a gradual increase in representational strength of the choice direction in ACC preceding the motor movement. This motor preparation signal was biased in favor of the simultaneous OFC value representation, consistent with integration over OFC value dynamics.

Chapter 2

A model-based approach for targeted neurophysiology in the behaving non-human primate

Knudsen, Eric B., Balewski, Zuzanna Z., & Wallis, Joni D.

Acute neurophysiology in the behaving primate typically relies on traditional manufacturing approaches for the instrumentation necessary for recording. For example, our previous approach consisted of distributing single microelectrodes in a fixed plane situated over a circular patch of frontal cortex using conventionally-milled recording grids. With the advent of robust, multisite linear probes, and the introduction of commercially-available, high-resolution rapid prototyping systems, we have been able to improve upon traditional approaches. Here, we report our methodology for producing flexible, magnetic resonance informed recording platforms that allow us to precisely target brain structures of interest, including those that would be unreachable using previous methods. We have increased our single-session recording yields by an order of magnitude and recorded neural activity from widely-distributed regions using only a single recording chamber. This approach speeds data collection, reduces the damage done to neural tissue over the course of a single experiment, and reduces the number of surgical procedures experienced by the animal.

2.1   Introduction

Deciphering how behavior arises from the combined activity of millions of neurons relies on our ability to measure neural activity. Early studies used serial recordings of single neurons to correlate firing rates and behavioral variables (Fuster, 1973; Niki & Watanabe, 1979). While the behavior of individual neurons is still a rich source of insight into behavior (Hosokawa et al., 2013; Kennerley et al., 2011), it is increasingly clear that the collective

dynamics emerging from neural populations can relate to behavior in ways not observable at the level of individual neurons (Lara & Wallis, 2014; Mante et al., 2013; Rich & Wallis, 2016). Although chronically implanted arrays provide a means to collect multineuronal datasets (Ganguly et al., 2011), acute neurophysiology remains a critical tool for studying neural activity in relation to behavior in the head-fixed primate model. Acute recordings allow a broad sampling of the neural environment from day to day, removing sampling bias from analyses. Further, they also enable the study of circuit level activity over the scale of tens of millimeters, and the testing of a many such hypotheses across several experiments within a single subject.

As acute probe technology matures (Jun et al., 2017), parallel efforts are necessary to optimize how these technologies are used (Baden et al., 2015; Chen et al., 2017; Patel et al., 2014). Here we provide details of our own efforts to develop a system that allows us to perform recordings from up to eight multisite (e.g. 16-32) commercially-available recording probes. This system is fully customizable such that the number, location, and even individual probe trajectories can be fully tailored to the recording target(s) of interest with precision and repeatability. First, we describe the imaging pipeline we use to develop our anatomical models. Next, we present the basic shuttle and tower designs that can be used in conjunction with traditional planar grids. Third, we demonstrate how we build upon our initial designs to both increase probe density and target distant (30 - 35 mm apart) brain regions within a single recording chamber. Finally, we report neuronal yields from our ongoing experiments.

2.2 Methods

Segmenting MRIs for anatomical models

For each subject, we obtained anatomical images with resolution of 1 x 0.84 x 0.84 mm using a 3T Siemens TIM/Trio MRI scanner with a 2-channel receive-only head coil (Fig. 2.1a). We manually traced regions of interest in the frontal lobe such as orbitofrontal cortex (OFC), dorsolateral prefrontal cortex (dlPFC), and the striatum (STR) in Slicer3D (Fedorov et al., 2012) using the Paxinos primate atlas (Paxinos et al., 2000) as a reference. We used this neuroanatomical model in conjunction with a cranial model (Gray Matter Research, Bozeman, MT) to position custom-fit polyether-ether-ketone (PEEK) (Mulliken et al., 2015) unilateral acute recording chambers. Chambers were milled using traditional approaches and surgically implanted over the left hemisphere in two subjects (C and V) and the right hemisphere of one subject (T).

Initial design components

We previously performed acute recordings using 8 to 16 individually drivable single tungsten microelectrodes lowered bilaterally across two recording chambers (Kennerley et al., 2011; Rich & Wallis, 2016). We have since transitioned to multisite linear probes (V/K-
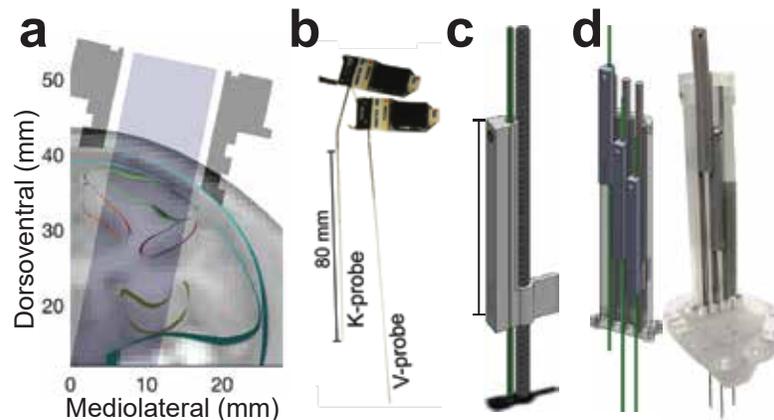
Figure 2.1: **Probe microdrive components.** (a) Cross-section of anatomical model with MRI slice (subject C, coronal section 24 mm anterior to the interaural line) with cortical regions outlined. Chamber profile in dark grey; accessible area in light grey. (b) Multisite linear probes (left: K-probe, right: V-probe) with headstages and groundwires attached. K-probes have a 20° bend 8 mm above electrode shank. (c) Rendering of single shuttle and drive screw. Green shows track for electrodes. (d) Render (left) and photo (right) of 3 drives assembled into a linear tower and placed on a planar grid (clear). Hypodermic needles serve as guide tubes.

probes, Plexon Inc., Dallas, Tx; Fig. 2.1b) configured with 16 or 32 recording sites spaced 100 μm apart. There are two primary challenges to using these probes for high channel count neurophysiology (i.e. 3+ within one brain area). First, whereas metal microelectrodes are solid and can bend to some degree, allowing drives to be angled outwards to increase density (Patel et al., 2014), multisite probes are hollow, making it critical to insert them orthogonal to the plane of entry to prevent probe damage. Second, standard microelectrodes use a single pin connection to route neural signals to an electrode interface board where a headstage connects, allowing for relatively close (1.5 - 2 mm) spacing of electrodes. The multisite probes interface directly with the headstage, increasing the footprint of a single probe to roughly 11 by 3 mm.

To work within these constraints, we set out to design a modular system that could (1) actuate a probe with a similar precision to our old system (i.e. 1 mm = 3 turns), (2) provide a stable platform for recordings from multiple probes deep within the brain for up to two hours, (3) be modular such that blocks could be reconfigured to cover the breadth of brain available within the craniotomy, and (4) be 3D-printable for rapid and flexible production. To achieve (1), we designed (SolidWorks, Dassault Systèmes, Vélizy-Villacoublay, France) a new shuttle that adds a long vertical arm capable of supporting the length of the thicker hypodermic tubing, which supports the probe shank, at the connector-adjacent end, and secured with a set screw (Fig. 2.1c). The shuttle interfaces with an 0-80 drive screw that achieves ~330 μm/turn resolution. This drive assembly is slotted into a tower (Fig. 2.1d) that keeps each drive stabilized via a retaining cap. In the tower shown, we record from 3 probes spaced at 3 mm between probes, at depths of up to 16 - 20 mm (with V probes,
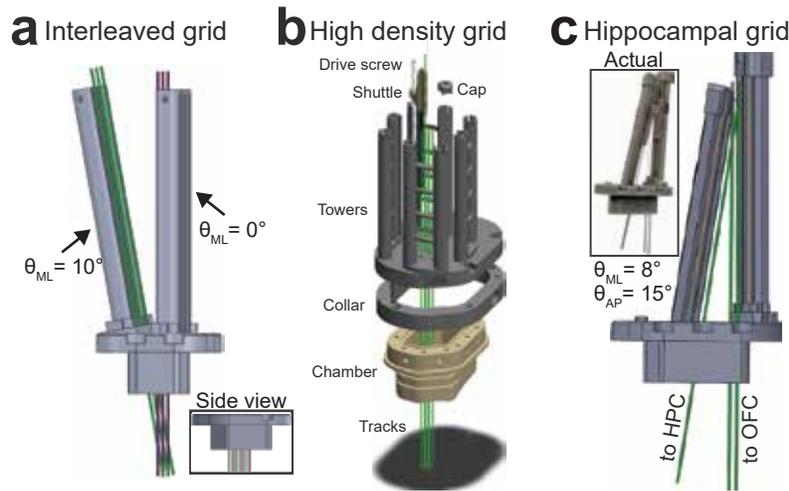
Figure 2.2: **Grid arrangements.** **(a)** Linear dense interleaved grid in coronal profile; inset shows side view. **(b)** High density grid. One shuttle, screw, and cap shown for clarify. Green tracks show electrode trajectories. Grid footprint slots into mating collar that is mounted onto chamber. **(c)** Hippocampal grid render; inset shows photo of assembled system.

> 40 mm with K-probes) from the cortical surface (Fig. 2.1d, green tracks). The tower assembly is then positioned on a planar grid according to experimental needs (Fig 1d, right, 750 μm resolution shown). Each designed component was printed in-house using the Form 2 stereolithographic (SLA) printer (Form Labs, Cambridge, MA).

Increasing probe density

The size of the neural population we can record simultaneously is constrained by the density with which we can arrange the electrodes. For example, the closest possible spacing of 6 electrodes using our initial design is either a linear arrangement within a single tower spanning 15 mm, or split across two towers, covering a 6 x 5.25 mm area. These configurations are inadequate for lowering a large number of probes into a single target region.

One strategy to increase density is to decrease the inline spacing between electrodes. We achieve this by angling opposing towers such that their trajectories cross at the target region as shown in Fig. 2.2a. Here, we offset one tower 1.5 mm in front of the other and tilt its base by 10° medially, resulting in approximately 1.5 mm spacing between 6 electrodes lowered to OFC, effectively halving the recording area.

A second strategy we are currently piloting is to increase the density per square area of tissue (Fig. 2.2b). This design will enable us to lower 8 multisite probes within a roughly 2.5 x 2.5 mm area, a comparable footprint to conventional silicon arrays (e.g. Utah array). In order to lower 8 probes in such close proximity, we require a new shuttle design in which the electrode track is offset 12 mm from the microdrive screw. This necessitates extra stabilization, achieved using a nonconductive carbon fiber rod. Due to the extra space
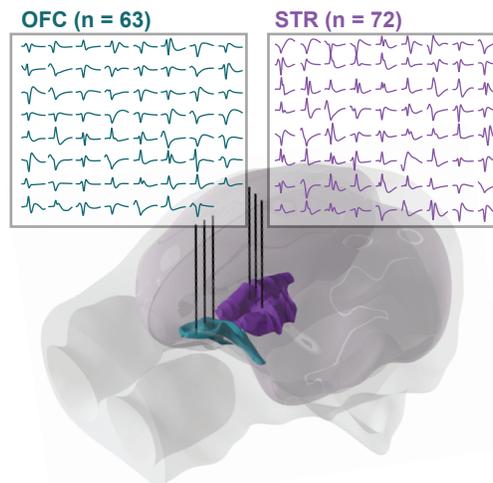
Figure 2.3: **Example neurons.** Average waveforms for well-isolated single units from six probes lowered to OFC (cyan) and striatum (purple; subject V).

required by the elongated shuttles, we developed an adaptor ("collar") that mates with the recording chamber, accommodating the larger footprint required by the 8 electrode towers.

Targeting distant brain regions

The custom-designed shuttles provide much greater flexibility with respect to the trajectory with which the probe enters the brain. This is useful for targeting regions around the craniotomy that would be otherwise unreachable with a direct approach. In addition, tailoring individual probe trajectories ensures we can enter cortex perpendicular to the cortical layers, improving our ability to study cortical layer-specific computations (Bastos et al., 2018). Finally, we can target brain regions separated by relatively large distances (30 mm), as shown in Fig. 2.2c. In one subject we performed simultaneous recordings from OFC in the frontal lobe and the hippocampus in the temporal lobe from a single chamber, a first in the behaving primate. Other studies have recorded from hippocampus and other prefrontal structures (e.g. lateral PFC), but did so through multiple recording chambers (Brincat & Miller, 2015) approach is straightforward: start and end locations of the desired track are specified via the MRI anatomical model, from which the necessary AP and ML angles are back calculated and translated into a blank recording grid template. The footprint for the tower is laid down around the tracks and either printed as a single piece (Fig.2.2a) or modularly (Fig. 2.2c).

2.3 Results

Fig. 2.3 shows an example yield of neurons simultaneously recorded from OFC and the striatum. Here we used six probes in the standard probe/tower configuration (Fig. 1). On average, we recorded approximately 1 - 1.5 well isolated single units per recording site during the first several insertions (mean $\pm$ s.e.m. units per 16-channel probe; subject V: 16 ± 1.6, subject C: 15 ± 1.6, subject T: 15 ± 1.8). Although we are still piloting our 8-probe within-region system, using the other systems highlighted in this report, we have yielded 100+ unit counts within the frontal lobe in multiple sessions; we anticipate that lowering 8 32-channel probes into cortex (256 channels) will yield upwards of 200 neurons within a single region based on current data.

2.4 Discussion

We outlined a general approach to recording increasingly large ensembles of neurons in the behaving non-human primate. Starting with anatomical scans, we model brain regions of interest in order to localize a form fitting chamber. Once implanted, we have the option to use generic planar grids combined with modular multi-probe shuttle and tower groups. To increase density of recording, we can interleave probes along a single axis, or cluster several probes in close proximity by increasing the effective footprint of our recording chamber. To record from distant brain regions simultaneously, we use our models to generate a number of different trajectories that can be easily printed using commercially-available 3D printing techniques.

There are several advantages to our approach over those reported by others (Jun et al., 2017) and options developed by commercial suppliers of multisite linear probes (Plexon, NeuroNexus, etc.). First, that we know of, our approach is the only one that enables high density, high channel count (>100 channel) electrophysiology in deep structures of the primate brain. Orbitofrontal cortex sits at the base of the frontal lobe superior to the eye orbits, approximately 15 mm below the cortical surface, as does much of the striatum; hippocampus in primates is located about twice as deep at 30 mm. Thus, as investigators begin to hypothesize about the network properties of large populations of neurons in these deep regions, it is critical that they be able to collect the data. Unlike surface recordings, these regions cannot be accessed through high channel count chronic arrays.

The next advantage to our approach is the decreased insertion to yield ratio. To collect data from 100 neurons with our approach requires 5-6 probe insertions, while yielding the same with conventional microelectrodes would optimistically require 30-50 insertions (assuming 2-3 units per electrode). Not only do we benefit from collecting the data simultaneously, allowing for a more diverse set of analyses (Rich & Wallis, 2016), the damage done to the neural tissue is drastically reduced. This increases the amount of data that can be collected from a single subject thereby reducing the number of subjects needed. Finally, the flexibility our approach affords allows us to implant and maintain only a single unilateral recording

chamber. Our previous work (Hosokawa et al., 2013; Kennerley et al., 2011; Lara & Wallis, 2014; Mante et al., 2013; Rich & Wallis, 2016) relied on bilateral recording chambers to increase yields (16-20 neurons maximum); by switching to a single implant, we reduce the maintenance requirements of the sterile environment inside the chamber(s), and reduce the likelihood that either chamber gets infected, necessitating a removal of both.

Rapid prototyping technology has enabled a quiet revolution in neuroscience (Baden et al., 2015; Chen et al., 2017; Freedman et al., 2016; Headley et al., 2015; Kloosterman et al., 2009; Patel et al., 2014). Combined with medical imaging and computer aided design, much of the guesswork that previously dictated the precision of neurophysiological experimentation has been eliminated. While recording technology continues to advance (**10**), it is only natural that new pipelines are developed. Our flexible approach has allowed us to reliably improve single neuron yields compared to our previous work. This ability allows us to test new hypotheses about both population-level dynamics of prefrontal cortex (Lara & Wallis, 2014; Rich & Wallis, 2016) and how PFC interacts with far reaching regions as part of the larger network, all while reducing damage to neural tissue caused by repeated microelectrode insertions.

We presented here one methodological approach to using multiple commercially-available linear multisite recording probes for acute primate neurophysiology. Our approach combines MR imaging, computer aided design, and rapid prototyping techniques to enable precisely-targeted high channel count distributed and concentrated neural recording.

Resource availability

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Eric Knudsen (eric.knudsen@berkeley.edu).

2.5 Acknowledgments

Chapter 3

Fast and slow contributions to decision making in corticostriatal circuits

Balewski, Zuzanna Z., Knudsen, Eric B. & Wallis, Joni D.

We make complex decisions using both fast judgements and slower, more deliberative reasoning. For example, during value-based decision-making, animals make rapid value-guided orienting eye movements after stimulus presentation that bias the upcoming decision. The neural mechanisms underlying these processes remain unclear. To address this, we recorded from the caudate nucleus and orbitofrontal cortex while animals made value-guided decisions. Using population-level decoding, we found a rapid, phasic signal in caudate that predicted the choice response and closely aligned with animals' initial orienting eye movements. In contrast, the dynamics in orbitofrontal cortex were more consistent with a deliberative system serially representing the value of each available option. The phasic caudate value signal and the deliberative orbitofrontal value signal were largely independent from each other, consistent with value-guided orienting and value-guided decision-making being independent processes.

3.1 Introduction

To navigate complex decisions, we rely on a combination of fast, autonomous intuitions and slower, deliberative judgements (Evans & Stanovich, 2013; Kahneman, 2011). Recent studies have begun to highlight a complex interplay between fast and slow processes of decision making (McClure et al., 2004). When subjects are presented with a choice, initial eye movements appear to reflect rapid value processing; both humans (Shimojo et al., 2003) and monkeys (Cavanagh et al., 2019) quickly orient their gaze to high value items within

150 ms. This quick orientation then biases the subsequent slower deliberation by increasing the likelihood that fixated items will be chosen (Cavanagh et al., 2019; Shimojo et al., 2003). During learning, this orienting fixation lags changes in choice behavior (Cavanagh et al., 2019), suggesting that these behaviors reflect two distinct processes that are at least partially dissociable.

The neural systems involved in implementing these fast and slow mechanisms remain unclear. The basal ganglia are involved in forming and executing habitual behaviors (Atallah et al., 2014; Balleine & Dickinson, 1998; Graybiel, 2008). The caudate nucleus (CdN) in particular is implicated in orienting value-guided fixations (Hikosaka et al., 2006; Lauwereyns et al., 2002; Watanabe & Hikosaka, 2005) and microstimulation of CdN can bias choice behavior towards a given reward-predictive cue (Santacruz et al., 2017). In contrast, the prefrontal cortex, and particularly the orbitofrontal cortex (OFC), is more strongly associated with deliberative value-based decision-making (Padoa-Schioppa, 2011; Wallis, 2007). OFC neurons primarily encode the value of chosen options (Kennerley et al., 2009; Padoa-Schioppa & Assad, 2006), and the region is critical to value-based decisions (Ballesta et al., 2020; Knudsen & Wallis, 2020), especially when new information needs to be computed or integrated (Murray & Rudebeck, 2018; Stalnaker et al., 2015). Furthermore, when animals consider two options, the value representation in the OFC vacillates between the two idiosyncratically on each trial, predicting both the speed and optimality of the final choice (Rich & Wallis, 2016), consistent with deliberative decision processes occurring in OFC. Given the strong, bidirectional connectivity between the head of CdN and OFC (Clarke et al., 2014; Ferry et al., 2000; Haber et al., 1995), this circuit may mediate the interaction between fast and deliberative valuation processes.

In the current study, we had two goals. First, we wanted to determine whether the CdN population dynamics were consistent with orienting eye movements during a value-based decision-making task. Second, we wanted to contrast these dynamics with flip-flopping value representations in OFC to investigate the interaction between rapid and deliberative processes in decision-making. To this end, we simultaneously recorded from the head of the CdN and OFC while two monkeys freely viewed pairs of pictures and used a lever to select their preferred associated juice outcome. CdN activity predicted choice direction rapidly after trial onset, independent of the relatively slow response times, and correlated with the direction of the first saccade, consistent with a role in mediating the effects of gaze on choice behavior. However, the flipping-flopping dynamics of OFC value representations were not influenced by eye movements nor the CdN direction signal, but were predictive of choice time, suggesting that the two systems reflect parallel valuation processes that both contribute to choice.

3.2   Results

We taught two monkeys (subjects C and G) to use a bidirectional lever to select either one (forced choice trials) or between two (free choice trials) available pictures for the correspond-

ing juice outcome (Fig. 3.1a). Both subjects learned 16 pictures, which were associated with one of four juice amounts and four reward probabilities (Fig. 3.1b). As expected, subjects preferred larger juice amounts and more certain rewards (Fig. 3.1c). Their lever response times were relatively slow (subject C: free trials, median = 666 ms, forced trials, median = 700 ms; subject G: 454 ms, 457 ms) and broadly distributed (Fig. 3.1d). In contrast, the first saccades following picture onset were fast (C: free trials median = 162 ms, forced trials median 206 = ms; G: 189 ms, 203 ms; Fig. 3.1d), yet overwhelmingly predicted the chosen picture even on free trials (C: 84% of free trials; G: 91%).

We modeled each subjects' choice behavior on free trials by fitting a soft-max decision function, which depended on the value difference between the two pictures and included a bias term for the first saccade direction (Eq. 3.1, Fig. 3.1). We estimated subjective picture value as expected value (Eq. 3.2), which better explained choices (C: 76% explained variance; G: 86%) compared with other reported value functions (Eq. 3.6, Fig. 3.2) (Hosokawa et al., 2013). Subjects selected the higher expected value option on 90% (subject C) and 92% (subject G) of free trials. Both subjects were more likely to select the option they looked at first, as indicated by the x-axis shift for the two curves depending on the direction of the subjects' first saccade. To simplify subsequent neuronal analyses, we binned the 16 pictures into groups of four from lowest to highest value; picture value (1-4) refers to these groupings (Fig. 3.1f).

To examine the relationship between picture value, choice responses, and initial saccade times, we considered the effect of the more valuable picture (maximum value) and the overall difficulty, or value difference, of the choice. Because each metric is separately correlated with response time, we performed a linear regression on response times with both as predictors to separate their unique contributions to behavior. Subjects responded significantly faster as the maximum value increased (C: $\beta = -0.39$, CPD = 11%, $p < 1 \times 10^{-5}$; G: $\beta = -0.40$, CPD = 12%, $p < 1 \times 10^{-5}$). There was also a small but significant effect of value difference, with larger differences predicting faster responses (C: $\beta = -0.04$, CPD = 0.2%, $p < 0.0001$; G: $\beta = -0.03$, CPD = 0.1%, $p = 0.01$; Fig. 3.1g). The onset of the first saccade was also strongly dependent on the maximum value, with faster saccades to more valuable pictures (C: $\beta = -0.6$, CPD = 27%, $p < 1 \times 10^{-5}$; G: $\beta = -0.54$, CPD = 22%, $p < 1 \times 10^{-5}$). The effect of value difference on saccade timing however was weak and inconsistent (C: $\beta = 0.06$, CPD = 0.4%, $p < 1 \times 10^{-5}$; G: $\beta = 0.01$, CPD = 0%, $p = 0.58$; Fig. 3.1h). Since maximum value, but not value difference, strongly correlated with both subjects' behavior, we focused only on this metric of picture value in the neural data.

Caudate neurons encode choice direction rapidly and phasically

We recorded single neurons and the local field potential (LFP) from the caudate nucleus (CdN) using up to 4 acute multisite probes per session (Fig. 3.3a). To investigate the relationship between choice behavior and neuronal firing rates, we performed linear regressions on the firing rates in sliding time windows with maximum value, choice direction, and trial type (forced or free) as predictors (Eq. 3.7, see Methods). Overall, the firing rates of 69% of
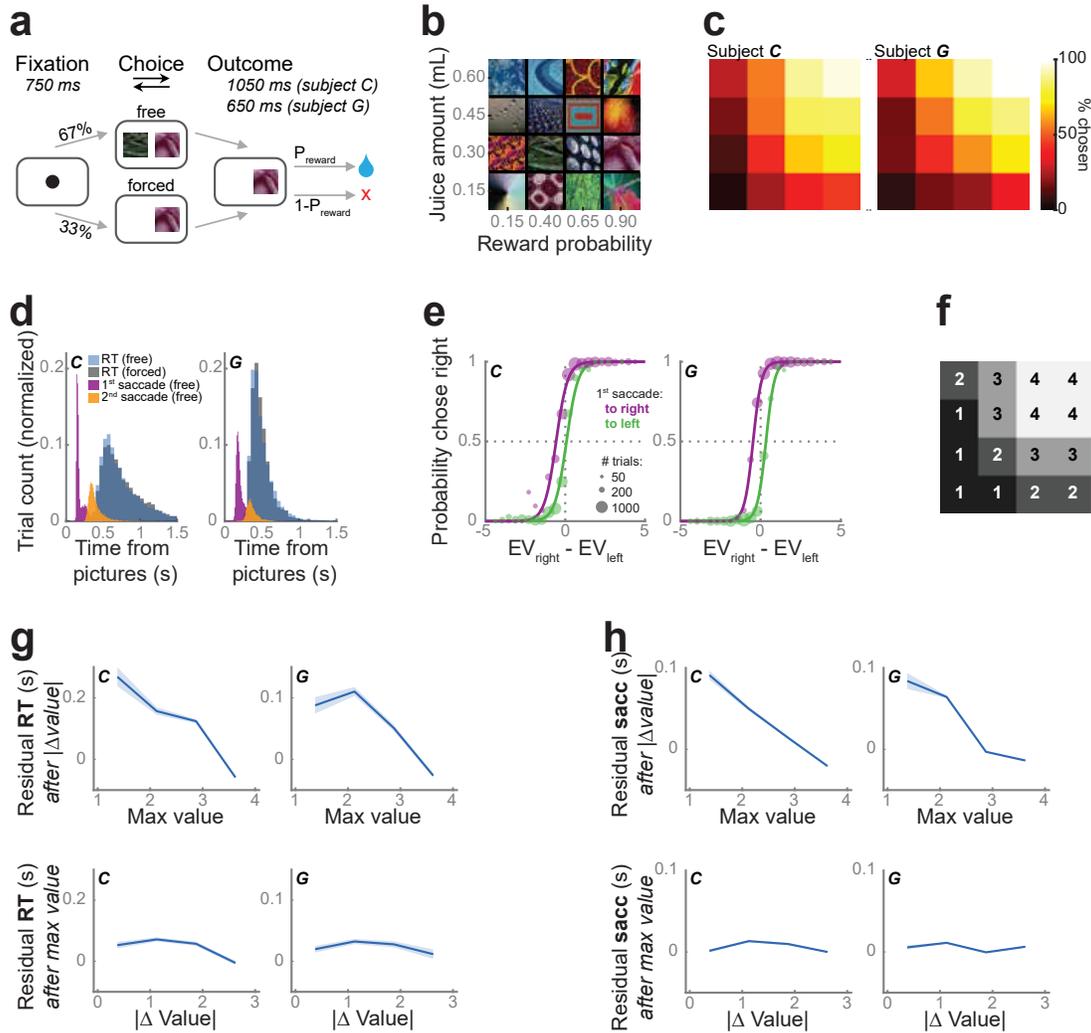
Figure 3.1: **Task and behavior.** **(a)** Subjects fixated a cue to initiate each trial. They were presented with one (forced choice trials) or two (free) pictures. After indicating their choice with a lever, they received the corresponding juice amount with the probability P_{reward}. **(b)** Subjects learned 16 pictures associated with different juice amounts and reward probabilities. **(c)** Subjects were more likely to choose pictures associated with larger and more probable rewards. Using the same arrangement as (b), the color indicates the percent of free trials on which each picture was chosen. **(d)** Lever response time distributions for free (blue) and forced (gray) trials. First (purple) and second (orange) saccade distributions from free trials. **(e)** The likelihood of choosing the right picture as a function of the difference in expected value (EV) of the two pictures, separated by the first saccade direction. Circle size represents the number of trials. Lines show the model fit. Both subjects were more likely to select an option when they first looked at it. **(f)** We split the pictures into equal-sized groups from lowest expected value (value bin = 1) to highest (value bin = 4). Groupings were the same for both subjects. Same arrangement as (b). **(g)** The unique effects of maximum value and value difference on reaction times. To visualize the unique contributions of each of these predictors, $log_{10}(response\,times)$ were modeled as a linear function of one parameter, and the residuals (mean ± s.e.m.) were plotted as a function of the other parameter. **(h)** The effect of maximum value and value difference on the timing of the first saccade. Conventions as in (g).
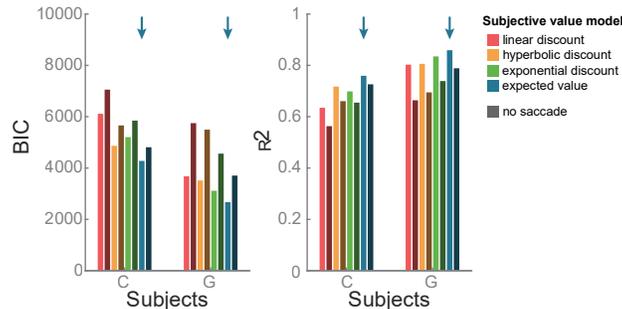
Figure 3.2: **Subjective value model comparison.** Models fitting subjects' choice behavior with different value functions (Eq. 3.2-3.5), and including (bright colors) or excluding (dark colors) the first saccade bias term, w_2. The best model, quantified with lowest BIC (left) and highest explained variance (right), was expected value including the saccade bias term (blue arrows).

344 neurons and 68% of 1259 neurons from subjects C and G, respectively, were predicted by at least one term during the first 500 ms following picture onset (Fig. 3.3b); we found no anatomical organization among these neurons (Fig. 3.4). Most neurons were significantly predicted by maximum value, choice direction, or both (Fig. 3.3c-e). In fact, we observed more neurons that encoded both value and direction than would be expected by their independent prevalence (χ^2 test; C: $\chi^2 = 25$, d.f. $= 1$, $p < 1 \times 10^{-7}$; G: $\chi^2 = 185$, d.f. $= 1$, $p < 1 \times 10^{-10}$).

To better understand the temporal evolution of information in CdN neurons, we looked at the dynamics of CdN population-level activity on individual trials. For each session, we trained linear discriminant analysis (LDA) decoders to predict the choice direction (left or right) from neuronal firing rates and LFP magnitude in sliding windows. We used a subset of free trials to validate decoder performance with a leave-one-out (LOO) procedure (see Methods). Accuracy rose sharply immediately after picture onset, peaking at 220 ms ($74 \pm 1\%$ across sessions) and 275 ms ($82 \pm 1\%$) after picture onset for subjects C and G, respectively (Fig. 3.5a). Repeating this procedure with trials synced to the lever movement, we observed significantly lower overall decoding accuracy leading to choice (Fig. 3.5b). This confirmed that the encoding of choice direction in CdN was driven by the onset of the pictures and did not reflect response preparation.

We next applied decoder weights to held-out free trials, iterating the procedure multiple times to maximize the number of trials (see Methods). We used the posterior probability for the choice direction as a proxy for representation strength (Fig. 3.5c). We observed a phasic rise in choice direction representation locked to the picture onset. Even when subjects took longer to respond, choice direction encoding was rapid. These dynamics, though readily observable from the neural population on individual trials, were not apparent in most individual neurons (Fig. 3.3e). Since most CdN neurons also encoded value, we examined how the choice direction representation varied with maximum picture value (Fig. 3.5d). We quantified this relationship by modelling the logit-transformed decoding strength as a linear
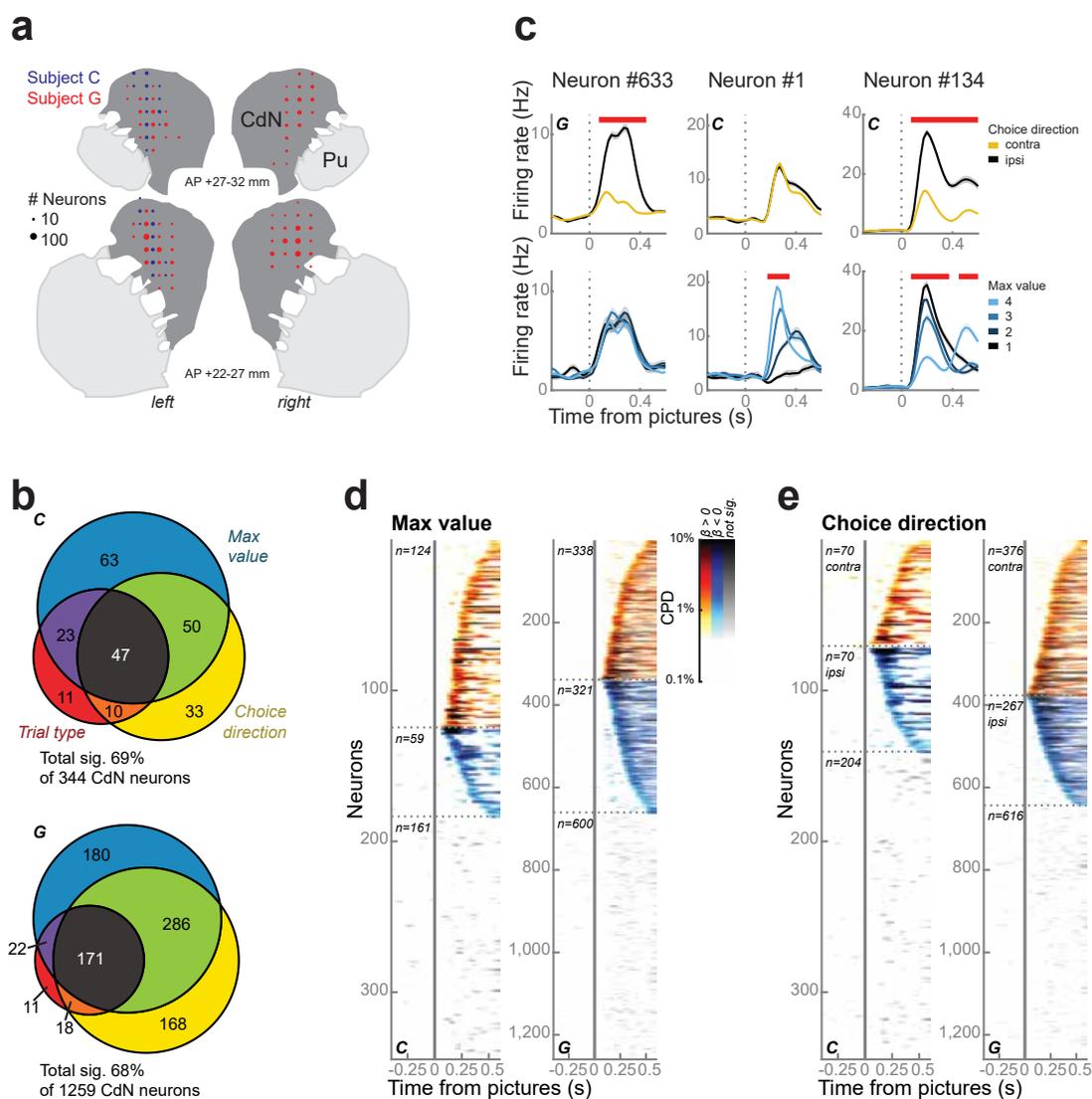
Figure 3.3: **CdN neurons.** **(a)** Reconstruction of CdN recording sites on coronal slices. Circles represent the number of neurons. **(b)** CdN neuron firing rates were modelled as a linear combination of maximum value, choice direction, and trial type. Most neurons were significantly predicted by at least one factor during the first 500 ms following picture onset. **(c)** Example neurons illustrating prevalent response types, including significant encoding of choice direction (left), maximum value (middle), or both (right). Panels within each column show firing rates (mean ± s.e.m.) averaged over choice direction (top) or maximum value (bottom row). Red bars at the top of each panel indicate periods of significance as determined from the regression. **(d)** Single neuron encoding of maximum value across the entire population. Each horizontal line indicates the percentage of variance in each neuron's firing rate explained by maximum value, as determined by the coefficient of partial determination (CPD). Neurons were grouped according to whether they showed a positive (hot colors) or negative (cold colors) relationship between firing rate and value, and ordered by the first significant time bin. Data from non-significant neurons are shown using grayscale. **(e)** Single neuron encoding of choice direction across the entire population. Conventions as in (d).
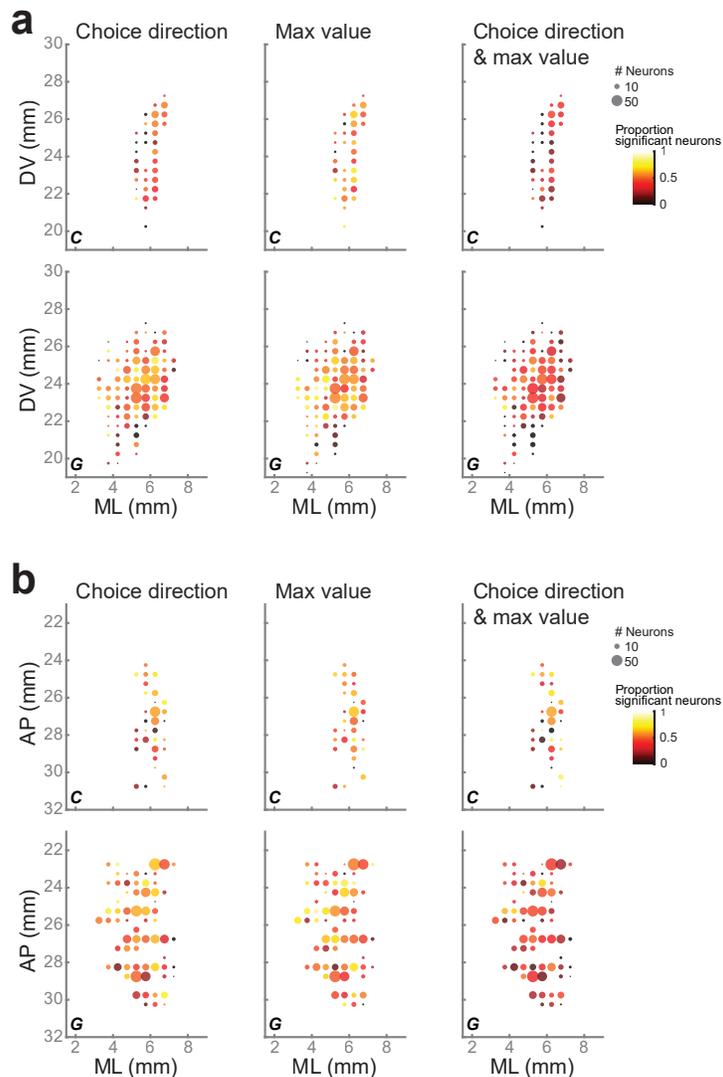
Figure 3.4: **Anatomical distribution of CdN neuron encoding.** We investigated the anatomical distribution of CdN neurons that were significantly predicted by choice direction, maximum value, or both during the first 500 ms after picture onset (Fig. 3.3b). For each group, we built a logistic regression for the proportion of significant neurons, predicted by the three anatomical axes (ML, AP, DV). We did not find significant organization along any gradient (subjects C and G: $p > 0.1$ for all models). **(a)** For visualization, we projected the recording sites in the ML-DV plane; we pooled across hemispheres for subject G. The circle size reflects the number of neurons. In each panel column, the circle color indicates the proportion of neurons significantly predicted by that task variable. **(b)** We projected the recording sites in the ML-AP plane. Same conventions as (a).

Figure 3.5: **Direction decoding from CdN population.** (a) Population decoding of choice direction (left or right) on free trials from CdN. Gray lines represent accuracy from the leave-one-out validation procedure for each session; the mean across sessions is shown in black. The mean cho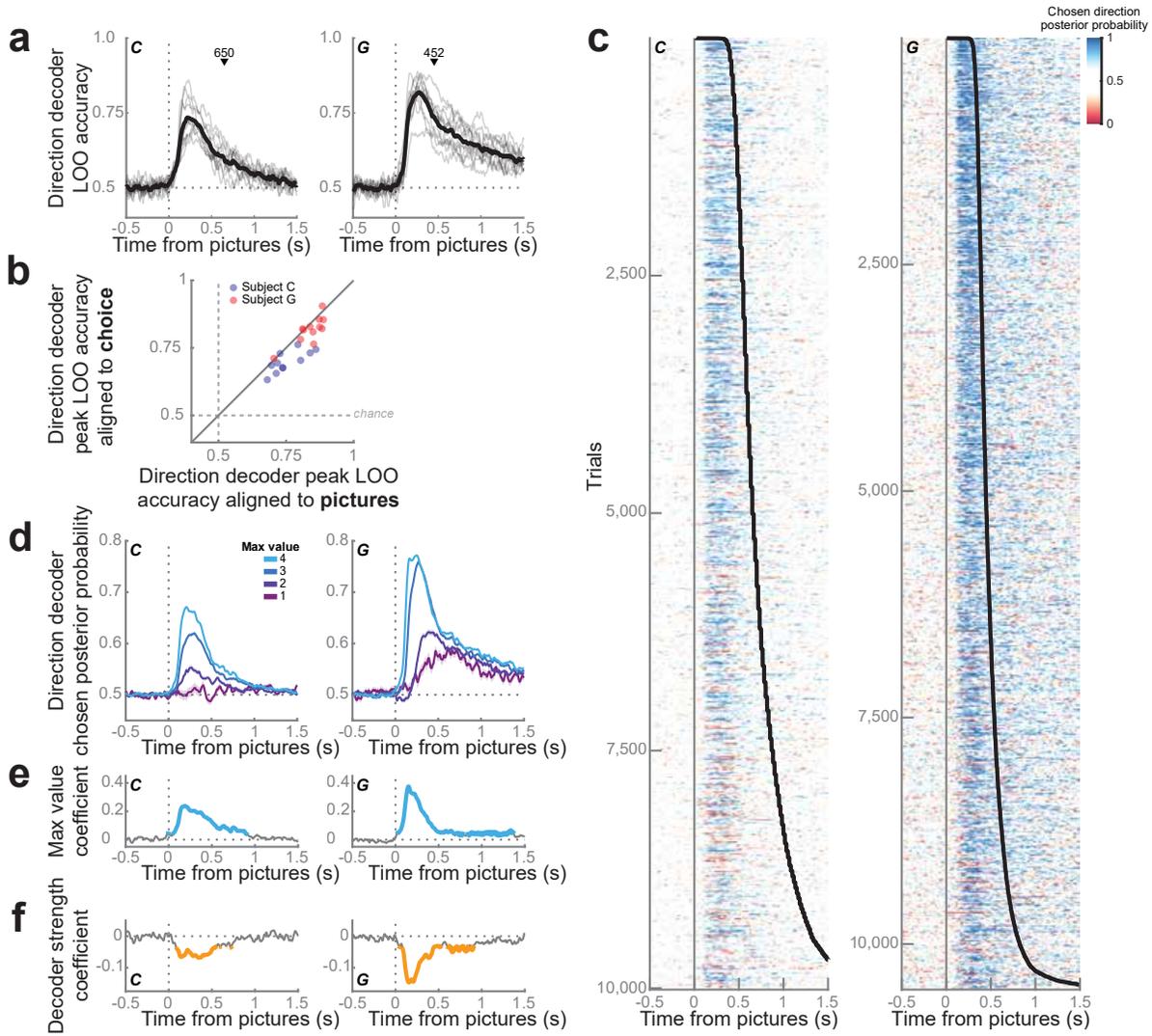ice time (black triangle) is shown for reference. (b) Peak LOO accuracy of the direction decoder for each session, trained on trials aligned to the onset of the pictures (x-axis) compared to choice (y-axis). Most sessions for both subjects fall below the diagonal equality line, indicating better decoding performance for trials aligned to picture onset (C: $68 \pm 1\%$, paired $t(10) = 4.7$, $p < 0.001$; G: $80 \pm 1\%$, paired $t(11) = 2.5$, $p = 0.02$). (c) We used the LDA weights to compute the posterior probability for the choice direction on held-out free trials (1 = chosen direction, 0 = unchosen direction). Each row corresponds to a single trial, ordered by increasing response time (black line). (d) Choice direction representation strength increased with maximum picture value. Lines represent mean $\pm$ s.e.m. choice direction posterior probability across trials grouped by maximum value. (e) We performed a linear regression to verify the observation in (d) by modelling choice direction strength as a function of maximum value. Blue lines indicate significant periods, $p < 0.001$. (f) Regression of response times as a function of choice direction strength after partialing out the effect of maximum value. Orange lines indicate significant periods, $p < 0.001$.

function of maximum value (Fig. 3.5e) and calculating the resulting beta coefficients and coefficient of partial determination (CPD), which is the percentage of the overall variance in decoding strength that could be explained by value. The strength of decoding of choice direction was strongly predicted by value, peaking at around 190 ms ($\beta = 0.24$, CPD = 5.7%, $p < 1 \times 10^{-15}$) and 150 ms ($\beta = 0.38$, CPD = 14%, $p < 1 \times 10^{-15}$) for subjects C and G, respectively.

We next asked if choice direction decoding strength predicted any additional variance in response time beyond what was already explained by the maximum picture value. We performed a linear regression on log_{10}(response times) with logit-transformed decoding strength and maximum value as predictors (Fig. 3.5f). Indeed, decoding strength significantly predicted response times even once the effect of maximum value was accounted for (C: $\beta = -0.07$, CPD = 0.55%, $p < 1 \times 10^{-13}$; G: $\beta = -0.15$, CPD = 2.3%, $p < 1 \times 10^{-15}$).

Phasic caudate responses predict first saccades

While the strength of direction decoding was correlated with maximum picture value and response times, the latency appeared to follow picture onset, coinciding with the first saccade (Fig. 3.1d, 3.5c). To quantify the onset of direction decoding on each trial, we defined direction states to be periods of at least four consecutive time bins (spanning 35 ms) when the posterior probability exceeded 99% of a null distribution constructed from values during the fixation period (C: 99% threshold = 0.74; G: 0.84).

The distribution of first state onset times (C: median = 160 ms, G: 170 ms) was very similar to first saccade movement times (C: median = 162 ms, G: 189 ms; Fig. 3.6a). Similar to first saccades, the onset of the first direction state occurred significantly earlier for trials with pictures of larger maximum value (Fig. 3.6b; linear regression, C: $\beta = -0.07$, p ¡ 1 x 10-12; G: $\beta = -0.18$, $p < 1 \times 10^{-15}$. The first saccade location was highly predictive of the decoded direction (C: same on 68% of trials; G: 80%) and greater than would be expected by chance (χ^2 test; C: $\chi^2 = 318$, d.f. = 1, $p < 1 \times 10^{-15}$; G: $\chi^2 = 304$, d.f. = 1, $p < 1 \times 10^{-15}$). To examine the relationship between the CdN response and the saccade direction, independent of the effect of value, we examined trials where both choice options had the same value and compared trials where the first saccade was to the ultimately chosen option versus those where the first saccade was to the unchosen option. When the first saccade was to the chosen option, CdN direction decoding strength was significantly greater and faster than when the first saccade was to the unchosen option, and the subjects also responded more quickly (Fig. 3.6c).

These results suggest that the phasic CdN response reflects the first saccade direction and that it effects the overall time for the animals to make their choice. However, it is also not a simple oculomotor movement signal. Subject C tended to saccade first to the chosen picture and then to the unchosen picture before moving the lever. The decoded direction did not abruptly change from chosen to unchosen around the time of the second saccade, suggesting the information decoded from CdN is related to the processes guiding the initial orienting eye movement, but not subsequent movements (Fig. 3.6c). Of course, we cannot rule out
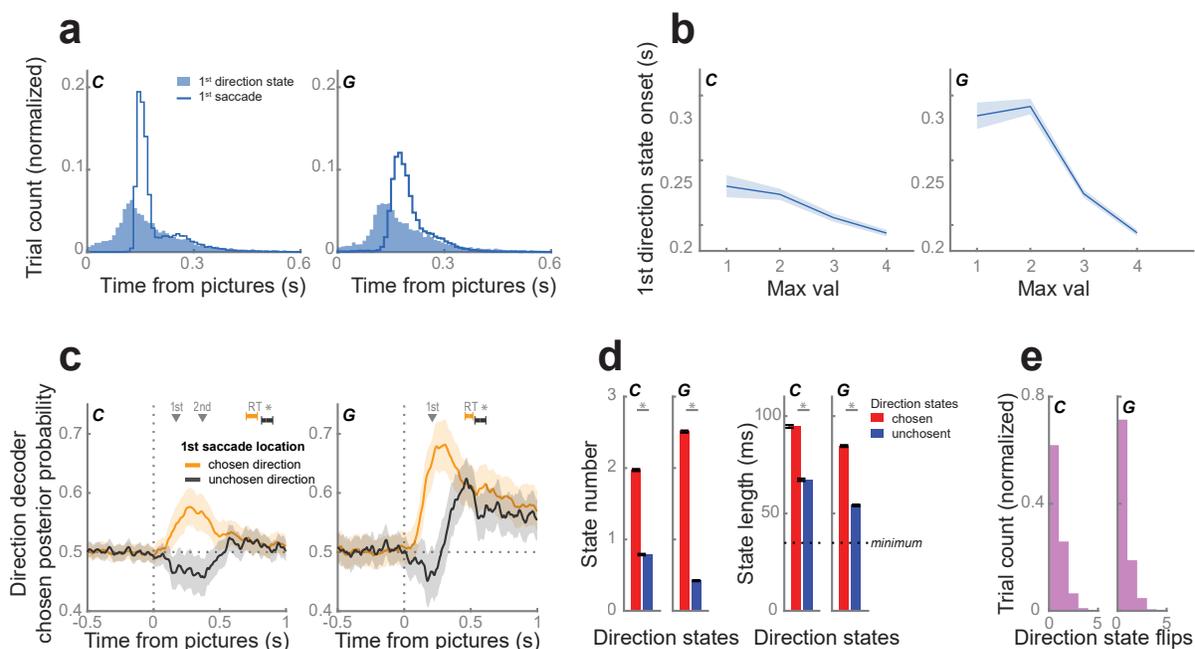
Figure 3.6: **CdN population tracks first saccade.** **(a)** We defined direction states as sustained periods of decoding of choice direction. The distribution of first state onsets (shaded blue) coincided with the distribution of first saccade movements (blue line). **(b)** First direction state onsets occurred significantly earlier on trials with higher maximum picture values. **(c)** We compared the average choice direction posterior probability for trials where the first saccade location was to the choice direction (matched trials; orange) and the unchosen direction (mismatched trials; black), restricting our analysis to trials where the pictures were of equal value. Dark lines represent the means calculated from 10,000 bootstrapped samples. Subjects responded more quickly (C: 116 ms, G: 83 ms) on matched trials. Shading and error bars indicate 99.9% CIs. Gray triangles indicate the mean time of the 1st (both subjects) and 2nd saccade (subject C only) across both trial groups. **(d)** During the 800 ms immediately following picture onset, chosen states (red) were more prevalent (right) and longer duration (left) than unchosen (blue) direction states. Asterisks indicate $p < 1 \times 10^{-15}$ determined from paired t-tests. **(e)** Distribution of number of flips between chosen and unchosen direction states per trial.

more high-level motoric control: for example, the phasic signal might reflect the preparation of a sequence of eye movements. In addition, we note that we could not perform the same analysis in subject G, since he only typically made a single saccade before making his choice. However, in both subjects it was clear that the decoded choice direction typically did not change later in the trial. Chosen direction states dominated unchosen direction states in quantity and duration (Fig. 3.6d), and for most trials (C: 62%, G: 71%) the decoded state did not change (Fig. 3.6e).

In sum, direction decoding from CdN reflected similar timing and content as the initial saccade. This suggests the decoder is classifying information correlated with the same fast valuation system reflected in overt eye movements (Cavanagh et al., 2019; Kim et al., 2020).

Orbitofrontal activity reflects a more deliberative decision-making process

To compare OFC and CdN contributions to our subjects' decision-making, we recorded single neurons and LFP from OFC using up to 4 acute multisite probes per session (Fig. 3.7a). We repeated the single neuron analyses described above to investigate the relationship between behavior and neuronal firing rates. The firing rates of 57% of 235 neurons and 54% of 498 neurons from subjects C and G, respectively, were predicted by at least one term during the first 500 ms following picture onset (Fig. 3.7b). Compared to CdN neurons, fewer OFC neurons were modulated by choice direction and trial type (Fig. 3.7c-e).

Because we found little evidence of direction information in individual neurons, we looked at the value dynamics of the OFC population and compared them to population dynamics in CdN. We trained an LDA decoder on forced trials to predict picture value (1 - 4) from neuronal firing rates and LFP magnitudes (Rich & Wallis, 2016). We applied the decoder weights to free trials in sliding windows to obtain the posterior probability for each value across time. Task events, including fixation and picture onsets, systematically biased the baseline rates of each value level at these moments in the trial. To remove these artifacts, we used the percent change in posterior probability from these baselines as a proxy for representation strength. We defined value states as periods of at least four consecutive time bins (spanning 35 ms) when the decoding strength exceeded twice the baseline rate.

Unlike the direction states decoded from CdN, value states in OFC flip-flopped between multiple picture values idiosyncratically over the course of each trial, replicating our previous study (Rich & Wallis, 2016). Decoding strength for the chosen and unchosen values peaked at 190 ms and 305 ms for subject C, and 310 ms and 250 ms for subject G, respectively. In the first 800 ms from picture onset, we observed a mean of 5.2 ± 0.03 and 8.5 ± 0.04 value states for subjects C and G, respectively. Chosen value states were significantly more prevalent and of longer duration than unchosen states, which in turn were more prevalent and longer than unavailable value states (Fig. 3.8a). For trials where at least one chosen or unchosen value state was identified (C: 98.3%; G: 100%), the median number of flips between them was 1 and 3 for subjects C and G, respectively (Fig. 3.8b). Decoding strength was predictive of response time: responses were faster on trials with stronger chosen value decoding and slower on trials with stronger unchosen value decoding (Fig. 3.8c). However, similar to Rich and Wallis (2016), there was no evidence that the decoder needed to be in a particular state at the time of the choice response. Indeed, the difference between chosen and unchosen value strength around the time of choice was weak and inconsistent between subjects (Fig. 3.10a). Value decoding in OFC was also not linked to gaze location. The strength of decoding of either the chosen or unchosen value did not depend on the location of the first saccade (Fig. 3.8d). Furthermore, there was no relationship between the phasic CdN choice direction signal and the OFC value signal. We grouped trials with chosen or unchosen initial value states in OFC and found no difference in the strength of the CdN direction decoding (Fig. 3.8e).

In sum, value decoding in OFC correlated with choice behavior on average but vacillated
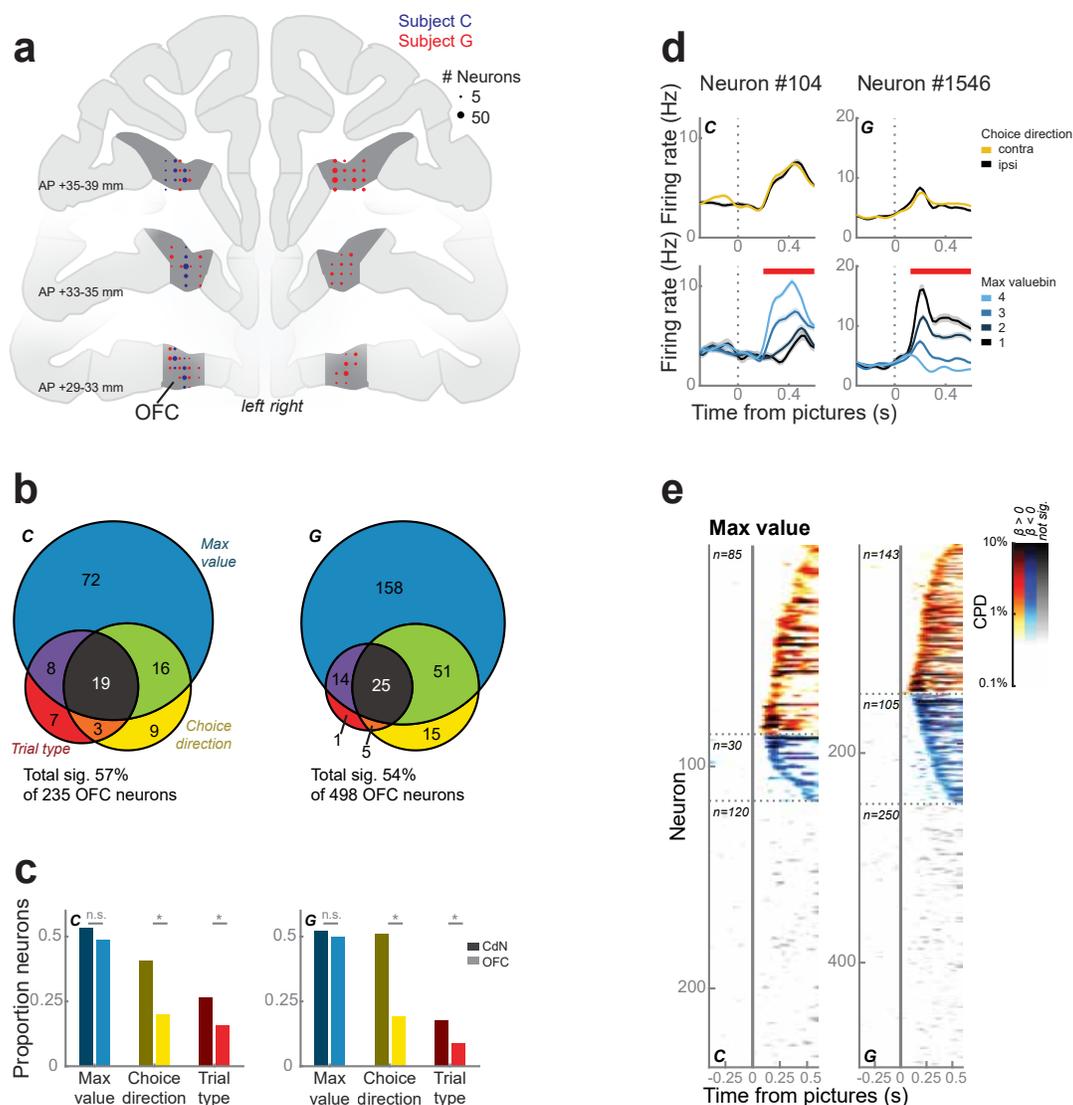
Figure 3.7: **OFC neurons.** (a) Reconstruction of OFC recording sites on coronal slices. Circles represent the number of neurons. (b) OFC neuron firing rates were modelled as a linear combination of maximum value, choice direction, and trial type. About half of the neurons were significantly predicted by at least one factor during the first 500 ms after picture onset, of which the majority was dominated by maximum value. (c) Proportion of neurons in CdN (dark colors) and OFC (light colors) significantly predicted by each factor (from Figs. (b) and 3.3b). Asterisks indicate $p < 0.01$ as determined by a χ^2 test. There was no difference between the areas in the prevalence of value-encoding neurons, but neurons encoding choice direction and trial type were more common in CdN. (d) Two example OFC neurons that were modulated by maximum value, but not choice direction. Conventions as in Fig. 3.3c. (e) Single neuron encoding of maximum value across the entire OFC population. Conventions as in Fig. 3.3d.

Figure 3.8: **Value dynamics in OFC. (a)** We trained an LDA decoder to classify picture value (1 - 4) on forced trials from OFC and applied the decoder weights to free trials in sliding windows. Value states are sustained periods of confident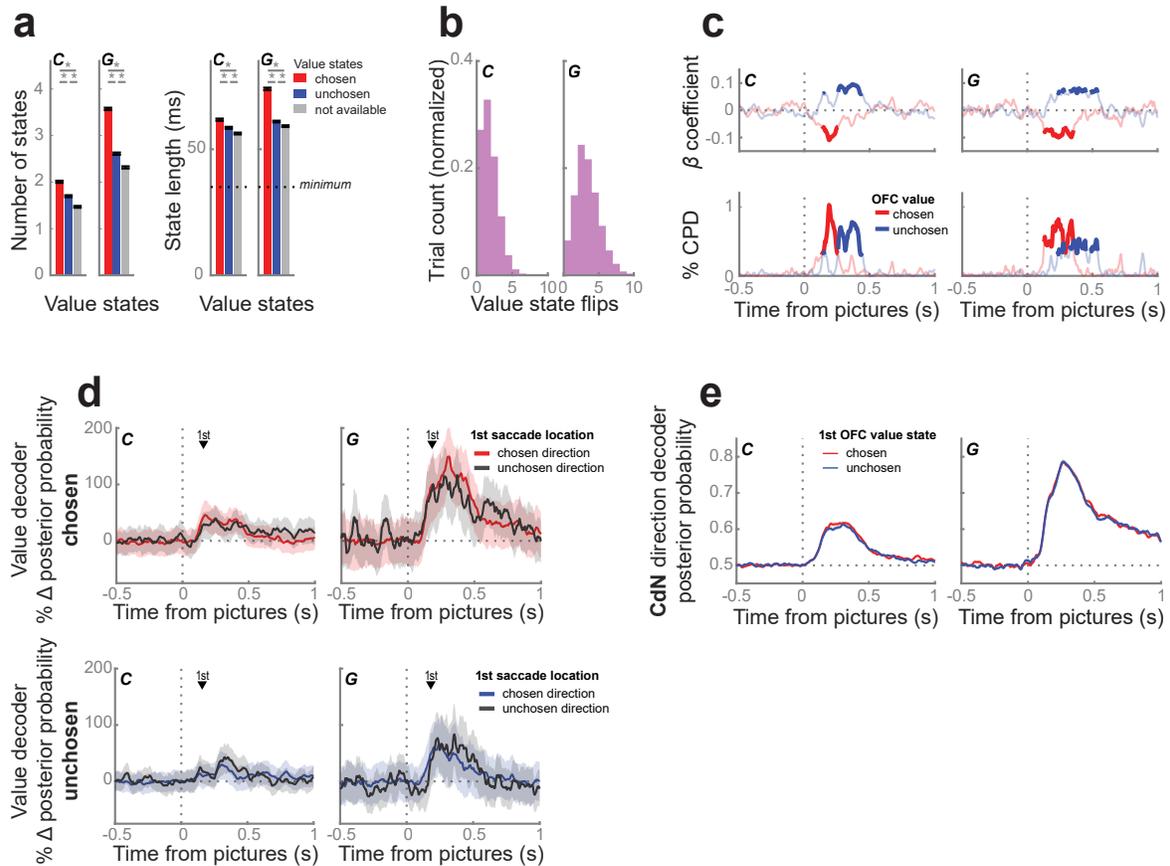 decoding of value. The prevalence (left) and duration (right) of value states favored chosen (red) over unchosen (blue) over unavailable (gray) value states. Asterisks indicate $p < 0.001$ as determined from a 1-way ANOVA with post-hoc t-tests. **(b)** Distribution of number of flips between chosen and unchosen value states per trial. **(c)** Response times were faster when the chosen value (red) was decoded more strongly, and slower when the unchosen value (blue) was decoded more strongly. We built a linear regression for response time with chosen and unchosen decoding strength as predictors. Top: regression coefficients; bottom: % CPD. Bold lines indicate $p < 0.001$. **(d)** We compared the average decoding strength of the chosen value (top) or the unchosen value (bottom) for trials where the first saccade location was to the chosen direction (shaded) and the unchosen direction (black line), using the same procedure as in Figure 3.6c. There was no difference in decoding strength of either value given the first saccade direction. Shading indicates bootstrapped 99.9% confidence intervals. The mean 1st saccade times (black triangles) are shown for reference. **(e)** We observed no difference in the strength (mean $\pm$ s.e.m.) of the CdN direction decoder between trials with an initial chosen (red) or unchosen (blue) value state in OFC.

between the representation of the chosen and unchosen state on single trials, consistent with our previous study. The OFC value signal was not linked to overt eye movements nor CdN direction decoding, suggesting it reflects a more deliberative valuation process that is separate from the rapid CdN orienting response.

Value decoding in caudate resembles dynamics in OFC

Since many CdN neurons also encoded value, we repeated the value decoding analyses from OFC in the CdN. We found that CdN value dynamics were qualitatively similar to those observed in OFC. Decoding strength for the chosen and unchosen values peaked at 170 ms and 175 ms for subject C, and 265 ms and 260 ms for subject G, respectively. In the first 800 ms from picture onset, we observed 4.5 ± 0.03 and 6.5 ± 0.05 value states for subjects C and G, respectively. Chosen states were again more frequent and longer than unchosen states and unavailable values (Fig. 3.9a). For trials where at least one chosen or unchosen value state was identified (C: 96.1%; G: 92%), the median number of flips between them was 1 and 2 for subjects C and G, respectively (Fig. 3.9b). As in OFC, strong encoding of the chosen state produced significantly faster response times, while strong encoding of the unchosen state produced significantly slower response times (Fig. 3.9c). However, on the slowest 25% of trials, there was no difference between the chosen and unchosen value strength following picture onset or around choice (Fig. 3.10b). Unlike direction decoding, value decoding in CdN was not linked to gaze location. We again compared trials where the first saccade and choice direction matched those when the directions were mismatched. We found no difference in chosen or unchosen value decoding strength across these trial types (Fig. 3.9d).

Given that we saw vacillation of value states in both OFC and CdN, we next asked whether these dynamics correlated between the two structures. We restricted our analysis to those sessions with successful value decoding simultaneously in both regions (C: 1 session; G: 5 sessions, see Methods). We correlated separately the chosen and unchosen value strengths decoded from OFC and CdN in each trial over time (Fig. 3.9e). We tested these correlations against null distributions of correlation values from permuted trials. In both subjects, within the first 500 ms after picture onset, both chosen and unchosen value dynamics were significantly correlated, although the effect was rather weak and inconsistent. We also performed a cross-correlation on these time series to examine whether there might be a lag between the two regions, but there was no evidence of a consistent lead or lag between the two regions.

3.3 Discussion

We observed rapid and phasic encoding of the choice response in the head of the CdN that was synced to the presentation of the choice options. While difficult to infer from trial-averaged neural firing rates, these dynamics were readily observed from the neural population
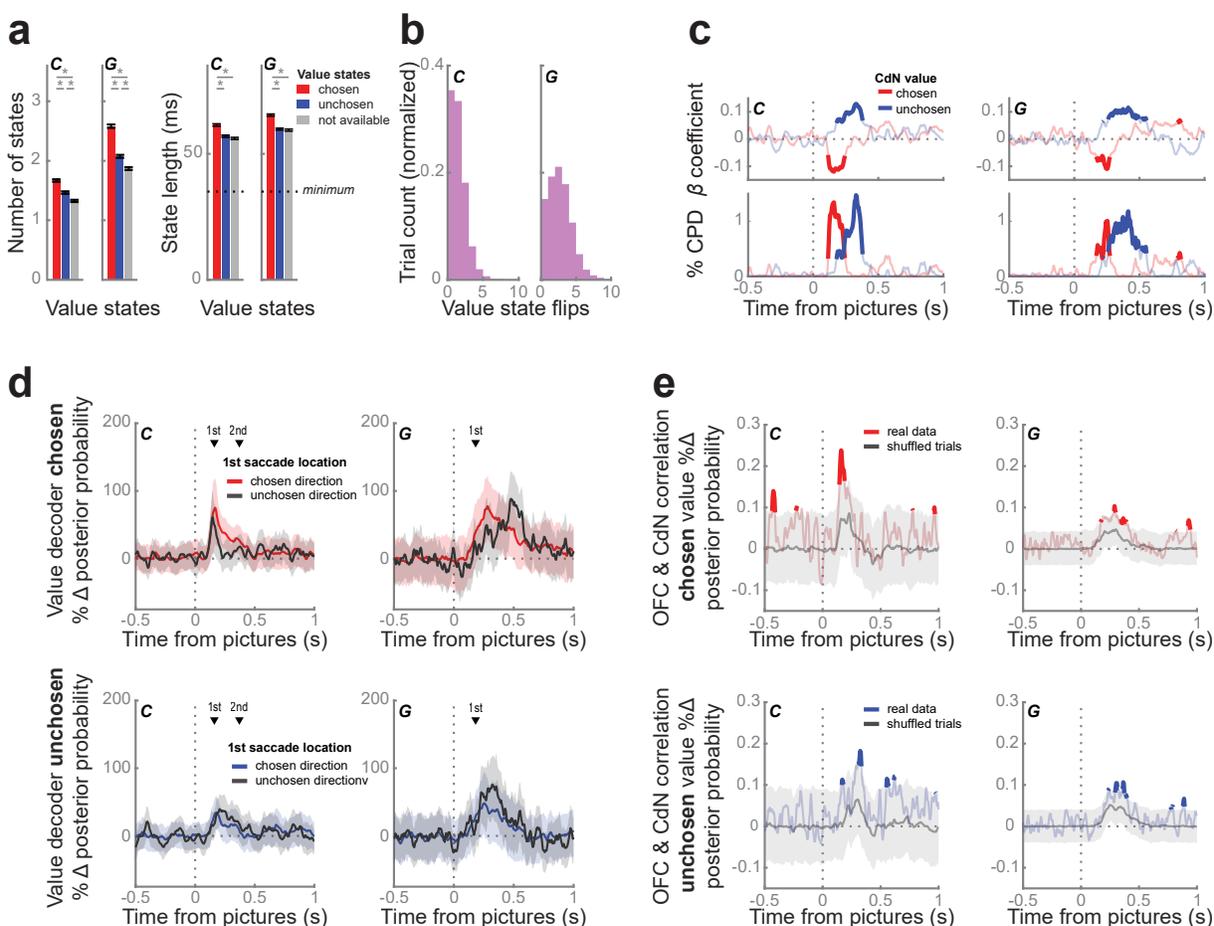
Figure 3.9: **Value dynamics in CdN. (a)** We trained an LDA decoder to classify picture value from CdN using the same procedure as in Figure 6a. The prevalence (left) and duration (right) of value states favored chosen (red) over unchosen (blue) and unavailable (gray) value states. Asterisks indicate p ¡ 0.001 as determined from a 1-way ANOVA with post-hoc tests. **(b)** Distribution of number of flips between chosen and unchosen value states per trial. **(c)** Response times were faster when the chosen value (red) was decoded more strongly, and slower when the unchosen value (blue) was decoded more strongly, using the same regression and visualization as Figure 6c. **(d)** We compared the average value decoding strength of the chosen (top) or the unchosen value (bottom) for trials where the first saccade location was to the chosen (shaded) or the unchosen direction (black line), using the same procedure as Figures 4c and 6d. Decoding strength of either value did not depend on the direction of the first saccade. Shading indicates bootstrapped 99.9% confidence intervals. The mean 1st saccade times (black triangles) are shown for reference. **(e)** We correlated the chosen (top, red) and unchosen (bottom, blue) value strength decoded from OFC and CdN on individual trials at each time point. We built a null distribution of correlation values from 10,000 permutations of trials; the gray line represents the median across these permutations and the shaded area indicates the 1% and 99% bounds. There was a weak, but significant, correlation in both animals during the first 500 ms of the choice period indicating that OFC and CdN tended to encode the same value at the same time.
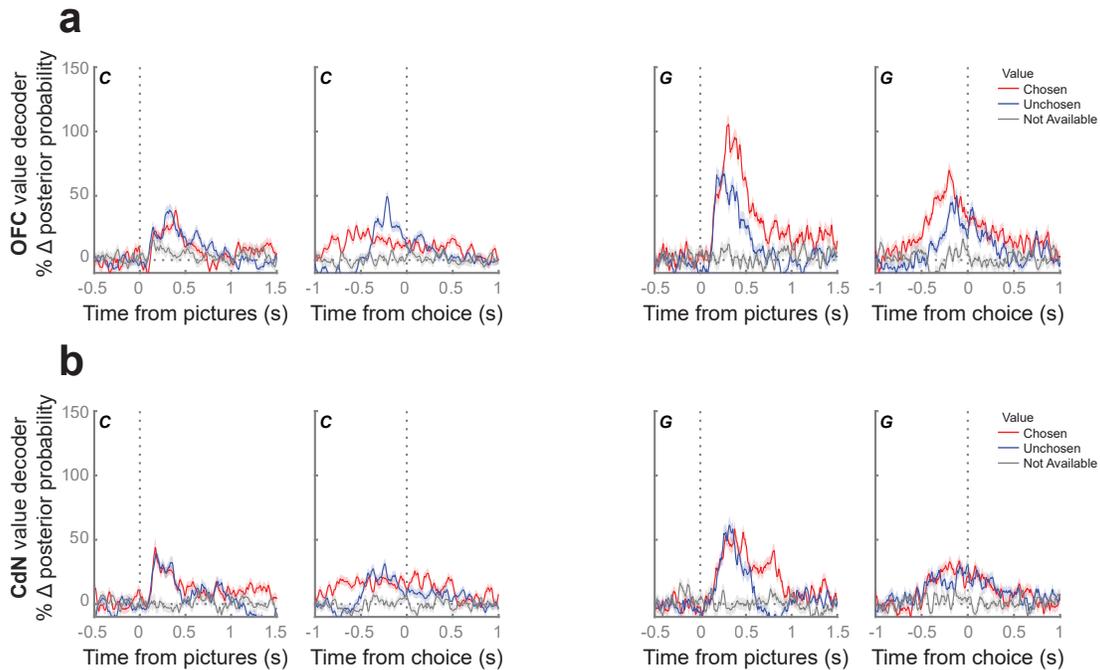
Figure 3.10: **Value decoding after picture onset and around choice.** We aimed to compare the relative strength of decoding chosen and unchosen values following picture onset and around the time of choice for both regions. We restricted this analysis to the slowest 25% of correct trials to maximally separate any phasic response following picture presentation to neural responses related to the preparation of the choice response. **(a)** Average decoding strength from OFC. The difference between chosen and unchosen value strength around the time of choice was inconsistent between subjects. **(b)** Average decoding strength from CdN. There was no difference between chosen and unchosen value strength around the time of choice.

on individual trials. Despite this early signal that predicted the choice, both animals took much longer to make their final decision. The direction signals in the CdN were strikingly similar to the subjects' initial saccade in that they occurred within 200 ms of the presentation of the choice options and were directed to the chosen item on the majority of trials. Our results are consistent with behavioral findings which suggest that there are two independent systems responsible for value-guided attentional capture and value-guided choice (Cavanagh et al., 2019). Our results suggest that CdN, but not OFC, is an important component of the value-guided attentional capture system.

Striatal encoding of salient environmental stimuli

In recent years, there has been increasing appreciation of the role that value plays in guiding attention (Anderson, 2016, 2019; Della Libera & Chelazzi, 2006). The basal ganglia appear to play a particularly important role in guiding saccades to rewarded targets (Hikosaka et al., 2006). For example, in a task where monkeys fixated peripheral cues that

predicted different sizes of reward, CdN neuronal activity was spatially selective and typically stronger when the cue predicted a larger reward (Kawagoe et al., 1998; Lauwereyns et al., 2002). In humans, neuroimaging results have shown that CdN is more active when humans must avoid being distracted by reward-associated cues (Anderson et al., 2014) suggesting that it may also play an important role in value-guided attentional capture. A possible mechanism for this involves projections to CdN from the superior colliculus via the thalamus (McHaffie et al., 2005). Neurons in the suprageniculate nucleus, which relays information from the superior colliculus to CdN, discriminate between pictures that predict high and low values of reward (Kim et al., 2020). This would also help explain why the CdN response is so rapid, since it conceivably only involves three synapses from the retina. Outputs from the CdN to the superior colliculus via the substantia nigra can then generate saccades to the target (McHaffie et al., 2005). The entire system can function to generate saccades to orient gaze towards ecologically relevant visual stimuli (Ghazizadeh et al., 2016).

Although CdN responses have been extensively studied in the context of peripheral presentation of reward-predictive cues, it was unclear how these findings would translate to more complex decision-making tasks. Our results show that the fast, phasic CdN response is also present during choice and may make an important contribution to value-based decision-making where it could mediate the biasing effect of saccades on choice. However, CdN also showed the kind of value vacillation that we previously observed in OFC. Thus, CdN appears to multiplex two signals in the dynamics of its firing that could make distinct contributions to choice. The first is the rapid, phasic directional signal that could mediate the effects of value-guided attentional capture, and the second is the vacillation in the value representation that corresponds to deliberative valuation. While our evidence for the multiplexing of fast and slow decision signals in CdN is correlational, this framework may help to explain previously reported effects of causal manipulations of this region. In a visually guided saccade task, CdN microstimulation following the saccade caused subsequent saccades in that direction to occur more rapidly, an effect that mimicked the effects of natural rewards on saccades (Nakamura & Hikosaka, 2006). In addition, we have shown that CdN microstimulation during a value-guided decision-making task can selectively increase the value of a specific stimulus independent of the choice response (Santacruz et al., 2017). An intriguing possibility is that the timing of microstimulation in the CdN relative to the presentation of the choice options may be able to differentially effect value-guided attention from value-guided decision-making.

It will also be important to determine how these signals are initially learned. CdN has frequently been associated with goal-directed, model-based reward learning, while the more lateral putamen has been associated with habitual, model-free responses (Balleine & O'Doherty, 2009). However, several features of value-guided attentional capture, suggest that it is more consistent with a habitual, as opposed to goal-directed response. For example, it occurs even when subjects are unaware of the underlying reward schedules (Pearson et al., 2015) and when it conflicts with the goals of a task (Le Pelley et al., 2015), neither of which is consistent with a deliberative process. With respect to the influence of value-guided attentional capture specifically on decision-making, here too the behavioral evidence seems

to favor a model-free, habitual response. For example, valuable options on which animals are overtrained show an attentional capture effect which biases subsequent choices. In contrast, while novel choices also demonstrate attentional capture, they do not bias choice (Cavanagh et al., 2019). The fact that value-based attentional capture is only seen in the overtrained situation is consistent with a model-free mechanism, and the fact that the behavioral effects on choice are different from those of novel choice options, suggests that it is also at least partly independent of information sampling (Monosov, 2020; Traner et al., 2021). To understand these mechanisms and how they map onto striatal circuitry it will be important to compare neural responses between CdN and the putamen, as well as use behavioral probes that can test whether behaviors are goal-directed or habitual. For example, one could devalue one of the rewards and examine how quickly the attentional capture effect changes relative to the animal's choice behavior.

Relationship to frontal deliberative decision-making

In contrast to the highly stereotyped early phasic direction response in CdN, OFC population value dynamics are idiosyncratic across trials. We replicated the main findings from Rich and Wallis (2016): the value representation flip-flopped between the available pictures, more valuable states were represented more frequently and for longer duration, and the strength of these representations predicted subjects' response times. Importantly, these dynamics were not influenced by overt eye movements. We saw very similar dynamics in the value encoding in CdN, along with a correlation with the states on OFC. The interareal correlation was rather weak and inconsistent. The weakness of the correlation may be real, but it could also reflect a relative lack of power in our analyses. Decoding simultaneously from two deep structures is technically challenging since it requires recording many neurons. These are exactly the kinds of experiments that will be made possible by new recording probes that have many hundreds of contacts (N. Steinmetz, 2019; N. A. Steinmetz et al., 2018).

It also remains unclear how the value vacillation is translated into choice. We previously speculated that a motor area downstream of OFC could integrate the value vacillation to determine the correct response, given that more valuable states are represented more frequently and for longer duration (Wallis, 2018). We identified three likely candidates based on their connections with OFC and the presence of both value and motor signals: anterior cingulate cortex, dorsolateral prefrontal cortex, and the striatum. We can now eliminate CdN as a possibility since the encoding of the choice response was clearly not a gradual accumulation of evidence from the value vacillation. However, it is conceivable that we might have seen integration of value signals or stronger correlations with OFC vacillation if we had recorded from the putamen rather than CdN. The putamen is a key striatal region involved in arm movement control (Inase et al., 1996; Kelly & Strick, 2004; Takada et al., 1998) and our animals were using arm movements to indicate their choice response. Furthermore, OFC projections into the striatum are not limited to CdN but extend into ventromedial putamen (Ferry et al., 2000). Neurons in the putamen encode action and value information (Hori

et al., 2009), and inactivation of this region disrupts action selection that is dependent on reward history (Muranishi et al., 2011).

Dysfunction of corticostriatal circuitry is a component of many neuropsychiatric disorders (Fernando & Robbins, 2011). The corticostriatal network is increasingly the target of implantable devices that can meaningfully interact with the circuit to treat these disorders (Creed et al., 2015; Scangos et al., 2021; Shanechi, 2019). Our results demonstrate both the challenges and potential of these approaches (Wallis, 2011). Value signals are nearly ubiquitous throughout the brain (Vickery et al., 2011), yet they often subserve very different functions (Wallis & Kennerley, 2010). Our results show that CdN multiplexes two distinct value signals with different dynamics. Understanding those dynamics potentially allows devices to target different value-related processes.

3.4 Methods

Experimental model and subject details

All procedures were carried out as specified in the National Research Council guidelines and approved by the Animal Care and Use Committee and the University of California, Berkeley. Two male rhesus macaques (subjects C and G, respectively) aged 6 and 4 years, and weighing 10 and 7 kg at the time of recording were used in the current study. Subjects sat head-fixed in a primate chair (Crist Instrument, Hagerstown, MD) and manipulated a bidirectional lever located on the front of the chair; eye movements were tracked with an infrared eye-tracking system (SR Research, Ottawa, Ontario, CN). Stimulus presentation and behavioral conditions were controlled using the MonkeyLogic toolbox (Hwang et al., 2019). Subjects had unilateral (subject C) or bilateral (subject G) recording chambers implanted, centered over the frontal lobe.

Task Design

Subjects performed a task in which they were required to choose between pairs of pictures or single pictures. Each trial started with a 0.5° fixation cue in the center of the screen, which subjects fixated continuously for 750 ms to initiate picture presentation. On free choice trials (67% of trials), two pictures 2.5° x 2.5° appeared 6° on either side of the fixation cue; on forced choice trials (33%), one picture appeared on either side. Subjects used a bidirectional lever to indicate the left or right choice. Both subjects learned to associate 16 pictures with one of four amounts of juice (subject C: 0.15, 0.3, 0.45, 0.6 mL; subject G: 0.1, 0.2, 0.3, 0.4 mL) and one of four reward probabilities (C: 0.15, 0.4, 0.65, 0.9; G: 0.1, 0.37, 0.63, 0.9). Amounts and probabilities were titrated for each subject so that one dimension did not dominate choice behavior. The selected picture remained on the screen while the corresponding juice amount was delivered probabilistically. Trials were separated by a 1000 ms intertrial interval. Gaze position was sampled at 500 Hz. Subject C completed 10,358

free trials and 5,397 forced trials over 11 sessions; subject G completed 10,793 and 5,873 trials over 12 sessions.

Neurophysiological recordings

Subjects were fitted with head positioners and imaged in a 3T MRI scanner. From the MR images, we constructed 3D models of each subjects' skull and target brain areas (Paxinos et al., 2000). Subjects were implanted with custom radiolucent recording chambers fabricated from polyether ether ketone (PEEK). During each recording session, up to 8 multisite linear probes (16- or 32- channel V probes with 75, 100, or 200 µm contact spacing, Plexon, Dallas, TX) were lowered into CdN (head of the caudate nucleus, AP +22 to +32 mm) and OFC (areas 11 and 13, AP +29 to +39 mm) in each chamber. Electrode trajectories were defined in custom software, and the appropriate microdrives were 3D printed (Form 2 and 3, Formlabs, Cambridge, MA) (Knudsen et al., 2019). Lowering depths were derived from the MR images and verified from neurophysiological signals via grey/white matter transitions. Neural signals were digitized using a Plexon OmniPlex system, with continuous spike-filtered signals (200 Hz - 6 kHz) acquired at 40 kHz and local field-filtered signals acquired at 1 kHz.

We recorded neuronal activity over the course of 11 sessions for subject C and 12 sessions for subject G (Table 3.1). Units were manually isolated using 1400 ms waveforms, thresholded at 4 standard deviations above noise (Offline Sorter, Plexon). We restricted our analysis to neurons with a mean firing rate across the session > 1 Hz firing rates. In addition, to ensure adequate isolation of neurons, we excluded neurons where > 0.2% of spikes were separated by < 1100 ms. Sorting quality was subjectively ranked on a 1 - 5 scale to separate well-isolated single neurons from possible multi-units. None of the results reported in the manuscript depended on the isolation quality of our neurons, and so we included all neurons in our analyses. CdN neurons were additionally labeled as putative phasic or tonic neurons from waveform shape, firing rate, and ISI distribution shape (Alexander & DeLong, 1985; Aosaki et al., 1995). We did not find any significant differences in the functional properties of these two classes of neurons and so we included all CdN neurons for the analyses reported in the manuscript. Only LFP channels with at least one unit were also included. In total, we recorded 344 and 1259 neurons and 247 and 899 LFP channels from CdN in subjects C and G, respectively; we recorded 235 and 498 neurons and 189 and 383 LFP channels from OFC.

Raw LFPs were notch filtered at 60, 120, and 180 Hz and bandpassed using finite impulse response filters in six frequency bands: δ (2-4 Hz), θ (4-8Hz), α (8-12 Hz), β (12-30 Hz), γ (30-60 Hz), and high γ (70-200 Hz). Analytic amplitudes were obtained from Hilbert transforms in the pass bands.

Behavioral analysis

We defined saccades as eye movements whose velocity exceeded 6 standard deviations from the mean velocity during fixation. The saccade time was defined by the peak velocity

Subject	Session	CdN neurons	OFC neurons
C	rec06	38 *	19
	rec07	39 *	24 †
	rec17	33 ‡	32 †
	rec20	41 *	17 †
	rec47	20 *	33 †
	rec49	40 ‡	20
	rec50	22 ‡	28
	rec51	31 ‡	13
	rec52	36 ‡	22
	rec53	36 *	7
	rec54	18 *	20 †
	total:	**344**	**235**
G	rec14	123 *	–
	rec15	182 ‡	111 †
	rec16	67 *	–
	rec17	146 *	–
	rec18	103 ‡	112 †
	rec19	114 ‡	96 †
	rec20	89 ‡	102 †
	rec21	100 ‡	77 †
	rec22	82 *	–
	rec23	93 *	–
	rec24	97 *	–
	rec25	63 *	–
	total:	**1259**	**498**

Table 3.1: **Neural data.** Number of recording sessions and neurons recorded per session in the two subjects. ‡ Sessions with successful value and direction decoding; † successful value decoding; * successful direction decoding.

within each movement, and direction was identified by the subsequent eye position.

We fit choice behavior for each subject using a soft-max decision rule,

$$P_R = \frac{1}{1 + e^{w_1(V_R - V_L) + w_2 S + w_3}} \tag{3.1}$$

which attempts to predict the probability that the subject will select the picture on the right, P_R, with three free parameters: the inverse temperature, w_1, which determines the stochasticity of the choice behavior as a function of the difference in value, V, between the right and left picture, a saccade bias term, w_2, which accounts for influence of the first gaze location, S, after picture onset, and a side bias term, w_3, which accounts for any preference for one choice direction.

We considered multiple subjective value functions, including expected value

$$expected\,value = juice\,amount \times reward\,probability \tag{3.2}$$

which did not add any free parameters to the model, and three discount functions from Hosokawa et al (2013)

$$value_{linear} = juice\,amount - w_4\,(1 - reward\,probability) \tag{3.3}$$

$$value_{hyperbolic} = \frac{juice\,amount}{1 + w_4(1 - reward\,probability)} \tag{3.4}$$

$$value_{exponential} = juice\,amount\,e^{w_4(1 - reward\,probability)} \tag{3.5}$$

which all added one free parameter, w_4.

We estimated all free parameters by maximizing the log likelihood of the full model:

$$-\sum N_R \ln(P_{R|\Delta V}) - \sum N_L \ln(1 - P_{R|\Delta V}) \tag{3.6}$$

where N is the number of trials where the animal selected either the right or left option. The expected value model was the best fit for both subjects' behavior, when models were compared with BIC and overall explained variance.

Single neuron regression analysis

For each neuron, in overlapping 100 ms windows shifted by 25 ms, we performed the following linear regression on firing rate, FR:

$$FR = \beta_0 + \beta_1 * trial\,type + \beta_2 * choice\,direction + \beta_3 * max\,value + \beta_4 * trial\,number \tag{3.7}$$

with binary variables for *trial type* (+1 free, -1 forced) and *choice direction* (+1 contralateral, -1 ipsilateral), continuous variables for the maximum value of the presented pictures (*max value*: 1 - 4), and *trial number* as a noise parameter to absorb potential variance due to neuronal drift over the recording session. Significance was defined as $p < 0.01$ for 100 ms (four consecutive time bins). To compare the proportion of significant neurons between CdN and OFC, we performed a χ^2 test for each factor (trial type, choice direction, and max value) independently.

Direction decoding with single trial resolution

Choice direction was decoded from CdN only. For each session, we trained linear discrimination analysis (LDA) decoders to predict the choice direction (left or right) from neural activity during free trials. We synced the trials to picture onset or choice response, and the decoders were trained in sliding windows of 20 ms stepped by 5 ms. To reduce the dimensionality of the input features in each window, we performed PCA across trials separately for normalized neuron firing rates and normalized magnitudes from the δ LFP band (2 - 4 Hz); higher frequency LFP bands did not meaningfully improve decoding accuracy. We then restricted inputs to the top PCs that explained 95% of the variance within each group.

We randomly split all free trials into separate training and held-out sets. Trials in the training set were constrained such that there was an equal number of each offer pair where the chosen value level was greater than the unchosen value (i.e., left 3 vs. right 4, left 4 vs. right 3, 2 vs. 4, 4 vs. 2, etc.). The held-out set contained all remaining trials. We used a k-fold leave-one-out validation procedure on the trials in the training set to assess the accuracy of the decoders. We then used the entire training set to compute the posterior probabilities for left and right direction on each trial in the held-out set. To maximally use all trials in a session, we repeated the procedure with 25 random splits of training and held-out sets. For each trial, we averaged the posterior probabilities across all instances when it was in the held-out set. To ease interpretation of reported results, we relabeled the decoder classes (left and right) as chosen direction and unchosen direction for each trial.

We identified sustained periods of confident decoding as states. We built a null distribution of posterior probability values from the fixation period (-500 to 0 ms before picture onset) to define a threshold that exceeded 99% of this distribution. On each trial, the posterior probability for a given direction needed to surpass this threshold for at least four consecutive time bins (spanning 35 ms, accounting for bin overlap) to be considered a state.

We bootstrapped confidence intervals to compare decoding strength between trials where the first saccade was to the chosen (matched) or unchosen (mismatched) direction, restricted to trials where the choice was between pictures of equal value. Note that we had about eight times more matched trials compared to mismatched trials, and mismatched trials were more likely to occur for lower picture values. To account for these differences, we ensured that the random samples of matched trials had the same value distribution and total number as the mismatched trials (C: n = 581, G: n = 383). We calculated 99.9% confidence intervals over the means of 10,000 bootstrapped samples drawn from each trial group.

Value decoding with single trial resolution

Picture value was decoded from OFC and CdN, separately for each session. It is difficult to train a decoder for picture value from free trials, where two pictures are available, without making assumptions about how and when each picture is being represented (e.g., only the picture on the right; only the best picture). Instead, following the approach in Rich and Wallis (2016), for each session we trained an LDA decoder to predict the value (1 - 4) from

neural activity on forced trials, where a single picture was available. We then used these decoder weights to compute the posterior probability of each value on every free trial in sliding windows of 20 ms stepped by 5 ms. Trials were synced to picture onset or lever response. The neural input and dimensionality reduction was the same as for direction decoding.

We trained one decoder from forced trials on neural activity averaged between 100 - 500 ms, for each session. We randomly sampled these trials to uniformly represent all pictures on either side of the screen and used a k-fold LOO validation procedure. We repeated this procedure with 50 random samples and averaged classification accuracy across all instances. We applied these decoder weights to each free trial in sliding windows of 20 ms stepped by 5 ms. We repeated this procedure with 50 random samples of forced trials and averaged the posterior probabilities on free trials across all instances.

Only sessions with good decoding accuracy on forced choice trials were included in the analyses, defined as $> 40\%$ overall accuracy on forced trials (chance $= 25\%$). For subject C, 5 sessions met these criteria for OFC (mean accuracy $44 \pm 1.7\%$) and 5 sessions for CdN ($46 \pm 1.7\%$). However, only 1 of these sessions had acceptable decoding in both areas. To minimize the number of sessions with poor value decoding, broadly predicted by the number of value-encoding sessions, in both areas in the second subject, we implanted subject G bilaterally to increase the number of probes per session. Indeed, for subject G, all 5 sessions with probes in OFC met these criteria for both regions (OFC: $58 \pm 2.5\%$, CdN: $53 \pm 1.3\%$). In the remaining sessions, no probes were lowered to OFC and CdN value decoding was not analyzed to match trial counts from subject C.

We observed that chance rates of decoding each value deviated from the expected 25%, likely due to global differences in population firing rates during different phases of the task. To correct for this, we defined a unique baseline rate for each value level by averaging the posterior probability corresponding to that level across trials when that picture value was not available. We reported posterior probabilities as percent change from these baselines. We refer to this metric as decoding strength.

To ease interpretation of reported results, we relabeled the decoder classes (1 - 4) as chosen value, unchosen value, and unavailable for each trial. Only free trials where the chosen value was greater than the unchosen value were analyzed (i.e., correct choice). For all visualizations and analyses, one of the remaining unavailable levels was randomly selected for each trial.

We identified sustained periods of confident decoding as states. On each trial, the decoding strength for a given value level needed to exceed 200% (i.e., double the baseline rate) for at least four consecutive time bins (spanning 35 ms) to be considered a state.

Using sessions with good value decoding from both OFC and CdN, we asked if decoded value dynamics were synchronized in the two regions (C: 1 session, G: 5 sessions). We separately correlated the chosen value and unchosen value decoder strengths from OFC and CdN at each time point. We built a null distribution of correlation values at each time point from 10,000 permutations of the trials; trials were shuffled within value comparison (e.g., 1 vs. 2, 1 vs. 3, etc.)

Statistics

All statistical tests are described in the main text or the corresponding figure legends. Error bars and shading indicate standard error of the mean (s.e.m.) unless otherwise specified. All terms in regression models were normalized and had maximum variance inflation factors of 1.7. The coefficient of partial determination (CPD) was used to quantify the percentage of overall variance uniquely explained by each term. All comparisons were two-tailed.

Resource availability

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Joni Wallis (wallis@berkeley.edu).

3.5   Acknowledgments

ZZB and JDW designed the experiments and wrote the manuscript. EBK edited the manuscript. ZZB and EBK collected the data. ZZB analyzed the data. JDW supervised the project. The authors declare no competing interests. One or more of the authors of this paper self-identifies as a member of the LGBTQ+ community. While citing references scientifically relevant for this work, we also actively worked to promote gender balance in our reference list.

Chapter 4

Value dynamics in orbitofrontal cortex drive the choice response in anterior cingulate cortex during decision-making

Balewski, Zuzanna Z., Knudsen, Eric B. & Wallis, Joni D.

4.1 Introduction

During decision-making, neurons in orbitofrontal cortex (OFC) sequentially represent the value of each option in turn (Balewski et al., 2022; Rich & Wallis, 2016). Although these dynamics can predict both the optimality and the speed of the decision, there is no evidence that OFC represents a specific option at the time of the choice response, raising the question as to how the OFC value dynamics are translated into a decision. One clue is that more valuable options tend to be represented more frequently and for longer duration (Balewski et al., 2022; Rich & Wallis, 2016). Consequently, a downstream region that integrated the OFC value dynamics would be able to select the more valuable option (Wallis, 2018), analogous to the noisy sensory information accumulation process observed in the lateral interparietal area (Gold & Shadlen, 2001; Mazurek et al., 2003; Roitman & Shadlen, 2002). Anterior cingulate cortex (ACC) is a potential locus for such an integrator, since it connects to both OFC (Barbas & Pandya, 1989; Carmichael & Price, 1996) and motor areas (Dum & Strick, 1993; Van Hoesen et al., 1993), contains neurons that encode both value and actions (Cai & Padoa-Schioppa, 2012; Hunt et al., 2018; Kennerley & Wallis, 2009; Luk & Wallis, 2013), and produces decision-making impairments when damaged (Camille et al., 2011; Kennerley & Walton, 2011; Kennerley et al., 2006). To test this hypothesis, we carried out high-channel count recordings from OFC and ACC while monkeys performed a value-based decision-making task.

4.2 Results

We taught two monkeys (subjects C and G) to use a bidirectional lever to select either one (forced choice trials) or between two (free choice trials) available pictures for the corresponding juice outcome (Fig. 4.1a-b). Both subjects preferred larger juice amounts and more certain rewards associated with the 16 pictures (Fig. 4.1c). We modeled each subjects' choice behavior on free trials with a soft-max decision function that estimated subjective value as expected value (Eq. 4.1-4.2; C: 78% explained variance; G: 84% explained variance). Subjects selected the higher value option on 92% (subject C) and 91% (subject G) of free trials. To simplify subsequent neuronal analyses, we binned the 16 pictures into groups of four from lowest to highest value (Fig. 4.1d).

Subjects' lever response times were relatively slow (C: free trials, median = 649 ms, forced trials median = 666 ms; G: 439 ms, 448 ms) and broadly distributed (Fig. 4.1e). To examine the relationship between picture value and response times, we modeled response times as a linear function of the best picture (maximum value) and the choice difficulty (value difference). Subjects responded significantly faster when maximum value increased (C: $\beta = -0.37$, CPD = 10.5%, $p < 1 \times 10^{-15}$; G: $\beta = -0.34$, CPD = 9.1%, $p < 1 \times 10^{-15}$; Fig. 4.1f). There was also a small but significant decrease in response times with larger value difference (C: $\beta = -0.06$, CPD = 0.3%, $p < 1 \times 10^{-8}$; G: $\beta = -0.05$, CPD = 0.2%, $p < 1 \times 10^{-3}$, Fig. 4.1f). Since maximum value, not value difference, was predominantly correlated with both subjects' behavior, we focused on this metric of picture value in the neural data.

ACC neurons encode choice direction preceding movement

We recorded single neurons and LFP from anterior cingulate cortex (ACC) using up to four acute multisite probes per session (Fig. 4.4a). To investigate the relationship between subject behavior and neuronal firing rates leading to the decision, we performed linear regressions on firing rates in 100 ms overlapping time windows with maximum value, choice direction (ipsilateral or contralateral to the recording location), and trial type (forced or free) as predictors (Eq. 4.3; see Methods). Overall, 70% of 453 neurons and 54% of 527 neurons from subjects C and G, respectively, significantly encoded at least one term during the last 500 ms preceding choice (Fig 4.2a). Encoding of value was most common, followed by encoding of choice direction.

To better understand the temporal evolution of information in ACC neurons leading to the decision, we analyzed ACC population activity on individual trials (Balewski et al., 2022). For each session, we trained linear discriminant analysis (LDA) decoders to predict the choice direction (left or right) from neuronal firing rates and LFP magnitudes in overlapping 20 ms windows. We used a subset of free trials to assess decoder performance with a k-fold validation procedure (C: $68 \pm 2\%$ accuracy across sessions; G: $69 \pm 2\%$; chance: 50%). We applied the trained decoder weights to held-out free trials (see Methods) and used the posterior probability for the chosen direction (DIR_{ACC}) as a proxy for representation
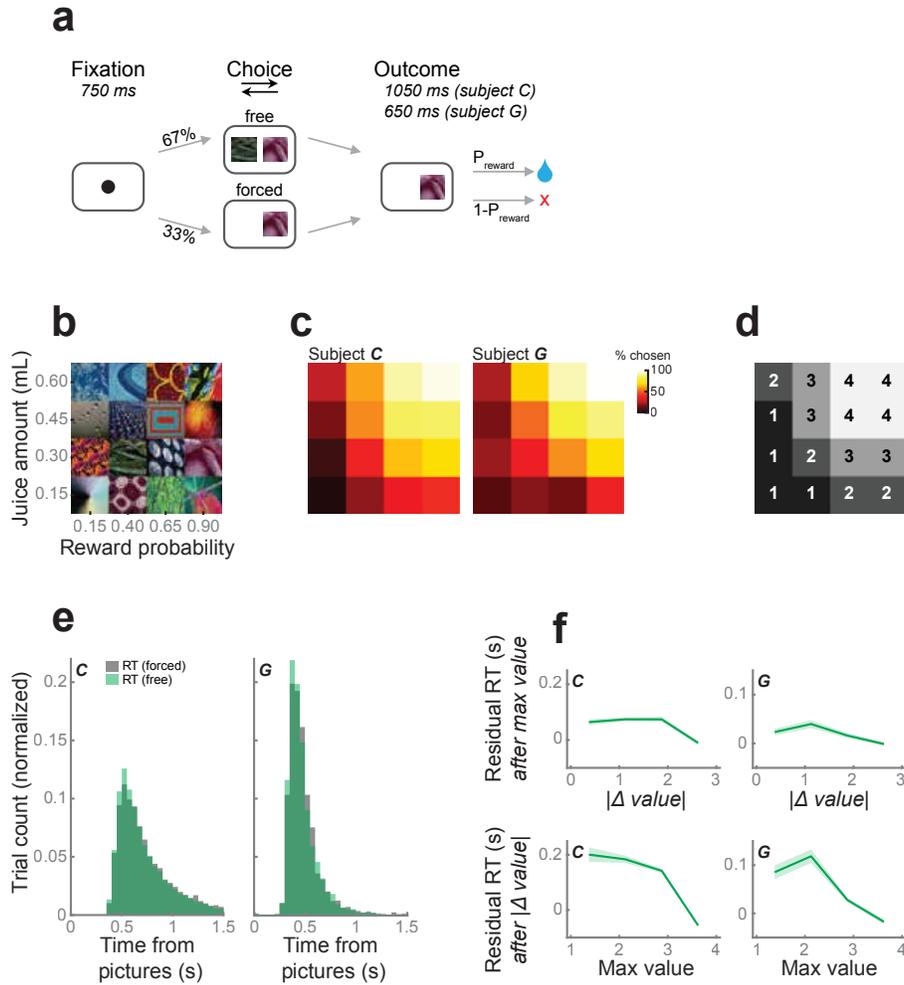
Figure 4.1: **Task and behavior.** **(a)** Subjects fixated a cue to initiate each trial. They were presented with one (forced choice) or two (free choice) pictures. After indicating their choice with a lever, they received the corresponding juice amount with the probability P_{reward}. **(b)** Subjects learned 16 pictures associated with different juice amounts and reward probabilities. **(c)** were more likely to choose pictures associated with larger and more probable rewards. Using the same arrangement as (b), the color indicates the percent of free trials on which each picture was chosen when available. **(d)** Expected value was used to approximate subjective value. We split the pictures into equal-sized groups from lowest expected value (value bin = 1) to highest (value bin = 4). Groupings were the same for both subjects; same arrangement as (b). **(e)** Choice response times on free (green) and forced (gray) trials. **(f)** The effects of maximum value and value difference on reaction times. To visualize the unique contributions of each of these predictors, response times were modeled as a linear function of one parameter and the residuals (mean $\pm$ s.e.m.) were plotted as a function of the other parameter.
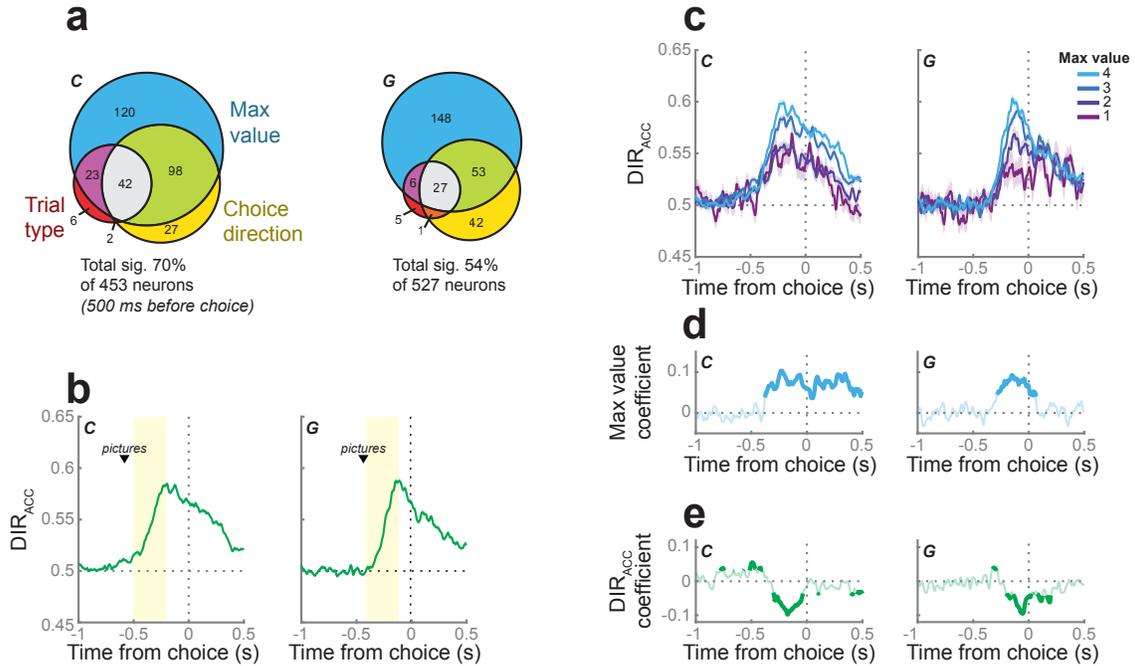
Figure 4.2: **ACC direction decoding. (a)** ACC firing rates were modeled as a linear combination of maximum value (blue), choice direction (yellow), and trial type (red) in overlapping 100 ms windows stepped by 25 ms. The majority of neurons were significantly predicted by at least one factor during the 500 ms preceding choice. Significance was defined as $p < 0.01$ for at least 100 ms. **(b)** Average posterior probability of decoding the chosen direction (DIR_{ACC}) on free trials from ACC neuronal firing rates and LFP magnitudes. DIR_{ACC} gradually increased over 300 ms (yellow), peaking before the lever movement. Black triangles indicate the median picture onset. **(c)** DIR_{ACC} increased with maximum picture value. Lines represent mean $\pm$ s.e.m. during free trials grouped by maximum value. **(d)** We performed a linear regression to quantify the relationship between DIR_{ACC} and maximum value. Bold lines indicate significance at $p < 0.01$. Decoding of choice direction was significantly stronger for larger values. **(e)** Regression of response times as a function of DIR_{ACC}, after partialling out the effect of maximum value. Bold lines indicate significance $p < 0.01$. Stronger decoding of choice direction predicted faster response times.

strength (Fig. 4.2b). We observed a gradual rise in DIR_{ACC} over approximately 300 ms, peaking before the lever movement (C: 195 ms; G: 100 ms). DIR_{ACC} preceding the lever movement was significantly stronger for higher picture values (Fig. 4.2c, d).

We examined whether DIR_{ACC} predicted any variance in response times, beyond what was already explained by maximum picture value (Fig. 4.1e). We performed a linear regression on response times with DIR_{ACC} and maximum avlue as predictors. Indeed, DIR_{ACC} significantly predicted response times preceding the lever movement even after the effect of maximum value was accounted for (Fig. 4.2e). Importantly, this relationship was unique to the end of the trial, suggesting this direction signal is related to motor preparation rather than spatial value representation. Although we observed similar encoding of task variables (Fig. 4.4a) and even better accuracy in decoding choice direction (Fig. 4.4b) early in the
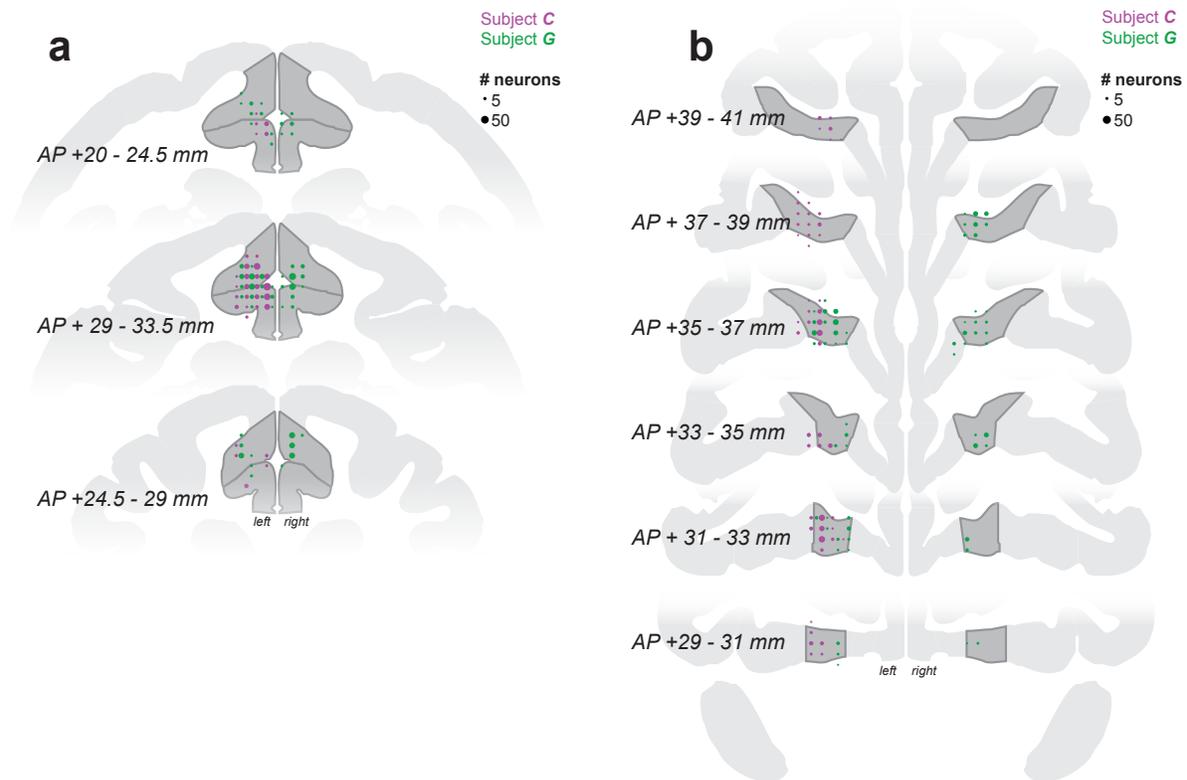
Figure 4.3: **Recording locations. (a)** Reconstruction of ACC recording sites on coronal slices. Circle sizes represent the number of neurons. CS = cingulate sulcus. **(b)** Reconstruction of OFC recording sites on coronal slices. Circle sizes represent the number of neurons.

trial, the DIR_{ACC} during the 500 ms following picture onset was a consistently weaker predictor of response times (Fig. 4.4c).

OFC value states drive ACC direction signal

We next investigated whether DIR_{ACC} ramping prior to the lever movement was affected by the dynamics of the value signal decoded simultaneously from OFC. Value representations in OFC vacillate between the available options when animals deliberate (Balewski et al., 2022; Rich & Wallis, 2016). We hypothesized that, if the ACC population direction response reflects an integration of these OFC value fluctuations, DIR_{ACC} would be affected by the value of the picture currently being represented by OFC (Fig. 4.5a).

First, we decoded value states, or sustained periods of confident decoding, on individual free trials (see Methods) from neuronal firing rates (C: 454 neurons; G: 429) and LFP from up to four multisite probes in OFC per session (Fig. 4.3b). Consistent with our previous
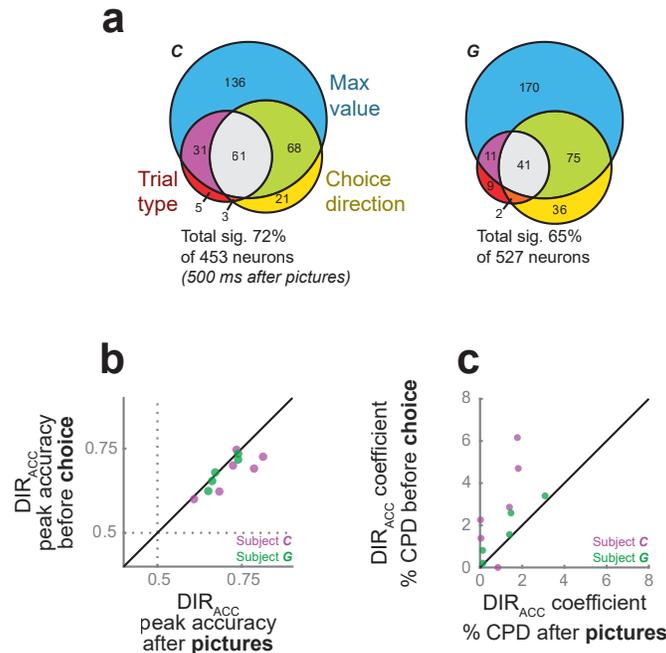
Figure 4.4: **ACC encoding and decoding details.** **(a)** The response profile of ACC neurons early in the trial (500 ms following picture onset) was similar to late in the trial (Fig. 4.11d). Same conventions as Fig. 4.1d. Neurons that encoded maximum value (C: 53.2%; G: 38.7%) or choice direction (C: 25.4%; G: 18.8%) in both trial epochs were more prevalent than expected by chance (Chi^2 test; C: value $Chi^2 = 130.8$, df $= 1$, $p < 1 \times 10^{-10}$, direction $Chi^2 = 141.6$, df $= 1$, $p < 1 \times 10^{-10}$; G: value $Chi^2 = 162.6$, df $= 1$, $p < 1 \times 10^{-10}$, direction $Chi^2 = 203.9$, df $= 1$, $p < 1 \times 10^{-10}$). **(b)** Peak DIR_{ACC} accuracy was better following picture onset than preceding choice (Wilcoxon signed-rank test for difference; subjects C and G combined due to small number of sessions: $p = 0.02$). Each circle represents one session. Dashed lines indicate chance decoding. **(c)** DIR_{ACC} following picture onset was a weaker predictor of response times than DIR_{ACC} preceding choice, measured with %CPD from the regression model in Fig. 4.2e (Wilcoxon signed-rank test for difference; subjects C and G combined: $p = 0.01$). Each circle represents one session.

findings (Balewski et al., 2022; Rich & Wallis, 2016), we observed that OFC value states corresponding to the available pictures (chosen and unchosen) flip-flopped idiosyncratically on each trial (Fig. 4.6a-c), and the strength of value decoding was predictive of response times (Fig. 4.6d).

We compared DIR_{ACC} at the onset of chosen and unchosen OFC value states (Fig. 4.5b) in sessions with simultaneous ACC and OFC recordings (C: n = 5 sessions; G: n = 5). We restricted the analysis to value states in the 300 ms period of rising DIR_{ACC} prior to lever movement (Fig. 4.2b). We found that DIR_{ACC} was stronger in the 100 ms following a chosen OFC value state compared to an unchosen value state (two-sample t-tests; C: $t(4273) = 3.1$, $p = 0.0018$; G: $t(5366) = 2.8$, $p = 0.0054$). We also asked whether chosen OFC value states led to faster increases in DIR_{ACC} compared to unchosen states. In other words, we tested the interaction between time (100 ms before and after value state onset) and value state
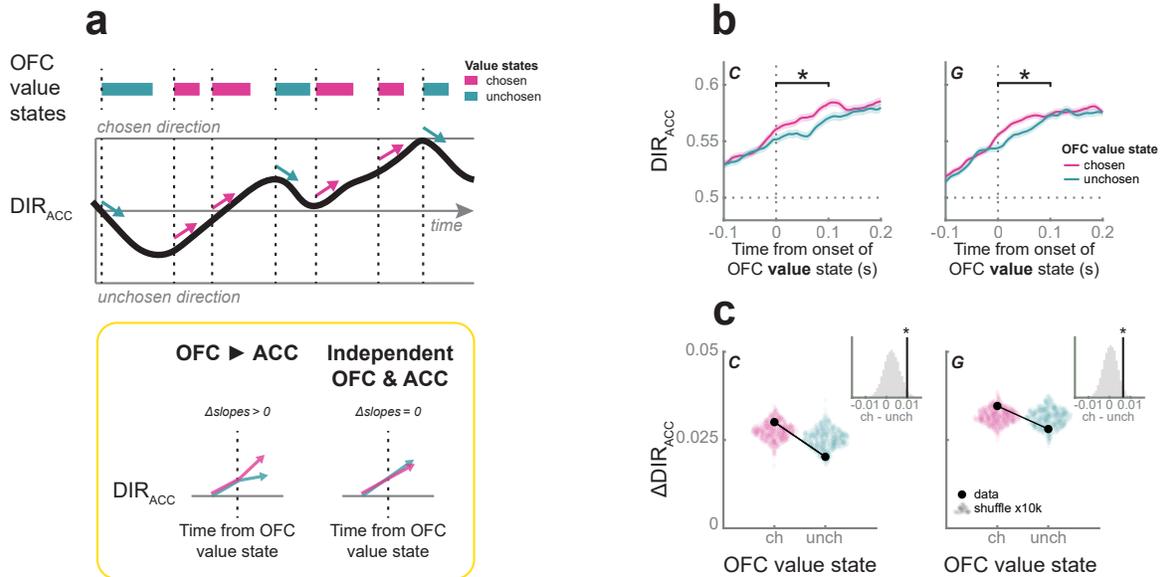
Figure 4.5: **OFC value states influence ACC direction signal.** **(a)** Model of how OFC value state integration might affect the ramping ACC choice direction signal (top). We predict that DIR_{ACC} will increase faster following chosen (high value) than unchosen (low value) OFC value states. The null hypothesis is that OFC value and DIR_{ACC} are independent signals (yellow box). **(b)** DIR_{ACC} synced to the onset of chosen (magenta) or unchosen (cyan) OFC value states during the period of rising DIR_{ACC} (Fig. 4.2c, yellow). DIR_{ACC} was stronger in the 100 ms following chosen OFC value states; $p < 0.01$, as determined from two-sample t-tests. **(c)** Mean ΔDIR_{ACC} (black) across trials at chosen or unchosen OFC state onsets. ΔDIR_{ACC} is the change in average DIR_{ACC} 100 ms before and after the value state onset. Swarms represent the null distribution of mean ΔDIR_{ACC} from 10,000 shuffled trials. The difference in ΔDIR_{ACC} following chosen compared to unchosen OFC value states was significantly greater than expected from the null distributions (inset); asterisks indicate significance at $p < 0.05$.

type (chosen and unchosen). Indeed, ΔDIR_{ACC} was significantly larger following a chosen than an unchosen OFC value state, compared to a null distribution from shuffled trials (C: $p = 0.03$; G: $p = 0.04$; Fig. 4.5c).

ACC value dynamics

Since a large number of ACC neurons encode value, we also characterized the single-trial value dynamics in ACC. Using the same approach as in OFC (see Methods), we successfully decoded value from ACC on forced trials (C: $56 \pm 4\%$ accuracy across sessions; G: $52 \pm 4\%$; chance: 25%) and observed flip-flopping between chosen and unchosen value states on free trials (Fig. 4.7a-c). The strength of value decoding in ACC was also predictive of response time (Fig. 4.7d). However, the value dynamics decoded simultaneously from OFC and ACC were not significantly correlated with each other (Fig. 4.8a).

Finally, we investigated the relationship between value and direction signals decoded from
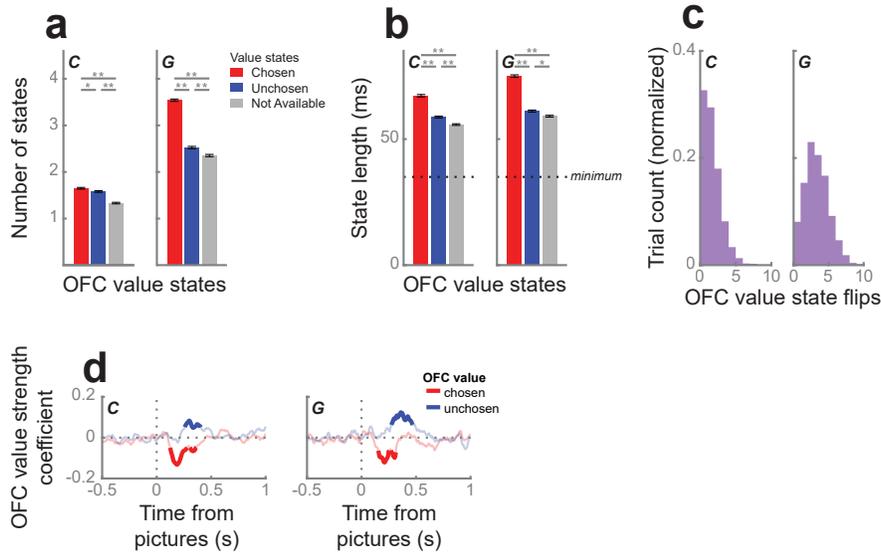
Figure 4.6: **OFC value decoding. (a)** We trained a linear discriminant analysis (LDA) decoder to predict the value (1 - 4) on forced trials from OFC neuronal firing rates and LFP magnitudes, and applied the decoder weights to compute the posterior probability for each value on individual free trials in overlapping windows of 20 ms stepped by 5 ms. States are sustained periods ($\geq$ 35 ms) of confidence decoding of a specific value. We observed more chosen (red) than unchosen (blue) than unavailable (gray) value states, replicating our previous results Rich, 2016. Single asterisk indicates significance at $p < 0.01$; double asterisks indicate significance $p < 0.0001$, as determined from a 1-way ANOVA with post-hoc t-tests. **(b)** Chosen value states were of significantly longer duration than unchosen value states. Conventions are the same as (a). **(c)** Distribution of number of flips between chosen and unchosen value states per trial. The median number of flips was 1.2 ± 0.02 and 2.8 ± 0.03 for subjects C and G, respectively. **(d)** We built a linear regression for response time with chosen and unchosen decoding strength as predictors. Response times were significantly faster when the chosen value (red) was decoded more strongly, and significantly slower when the unchosen value (blue) was decoded more strongly. Bold lines indicate significance at $p < 0.001$.

the same ACC populations on each session. We anticipated that this relationship would be even stronger within than across regions. Surprisingly, we did not see a strong relationship between the timing of ACC value states and DIR_{ACC}. Although DIR_{ACC} was stronger in the 100 ms following chosen compared to unchosen value states (two-sample t-tests; C: $t(5837) = 4.5$, $p < 1 \times 10^{-5}$; G: $t(5218) = 4.0$, $p < 1 \times 10^{-4}$; Fig. 4.8b), it was also stronger in the preceding 100 ms in subject G, suggesting that it was not causal to DIR_{ACC} ramping. The interaction between time (before and after state onset) and value state (chosen and unchosen) was not significant in both subjects (C: $p = 0.03$; G: $p = 0.397$; Fig. 4.8c).

4.3   Discussion

In summary, our results show that during value-based decision-making motor preparation signals in ACC are affected by which value the OFC is currently representing. Because
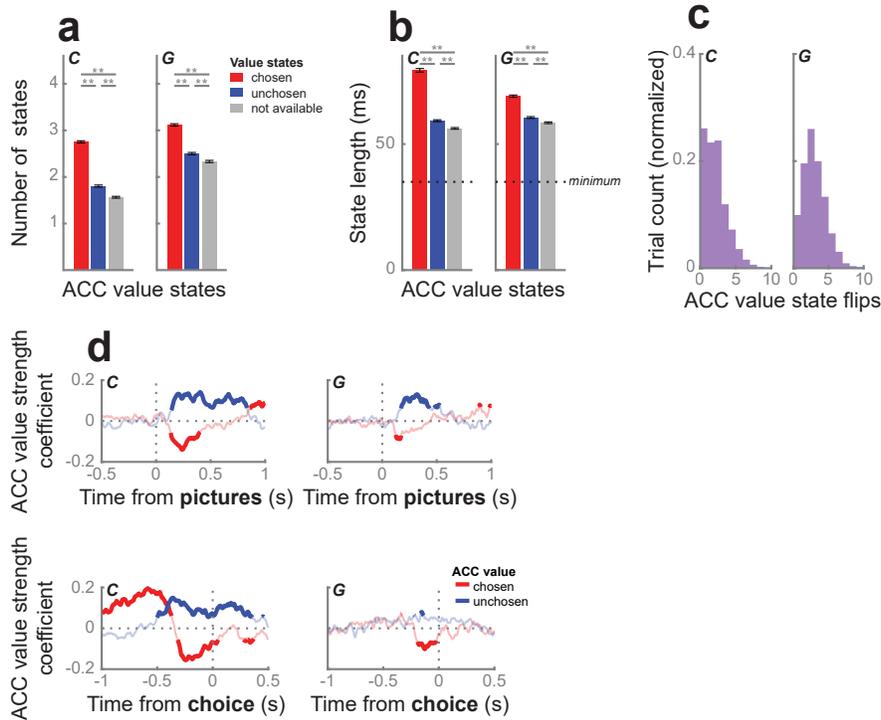
Figure 4.7: **ACC value decoding. (a)** We observed significantly more chosen (red) than unchosen (blue) than unavailable (gray) value states. Conventions are the same as Fig. 4.6a. **(b)** Chosen value states were of significantly longer duration than unchosen value states. Conventions are the same as Fig. 4.6b. **(c)** Distribution of number of flips between chosen and unchosen value states per trial. The mean number of flips was 1.7 ± 0.02 and 2.5 ± 0.03 for subjects C and G, respectively. **(d)** Response times were faster when the chosen value (red) was decoded more strongly and slower when the unchosen value (blue) was decoded more strongly, both early (top) and late (bottom) in the trial. Regression model and conventions same as Fig. 4.6d.

OFC represents higher values more frequently and for longer duration than lower values, this interaction between the two areas might lead ACC motor preparation to be biased towards choosing the more valuable option. Our results are also consistent with previous findings in ACC. For example, motor preparation signals can also be affected by the arrival of new information about the value of a choice option (Hunt et al., 2018). Furthermore, ACC may not be the only area involved in this accumulation of evidence: similar signals have been observed in dorsolateral prefrontal cortex (Lin et al., 2020). More generally, our results are consistent with the notion that decision-making involves value comparison and motor preparation processes that occur in parallel in the brain (Cisek, 2012; Cisek & Kalaska, 2010), rather than involving serial computations (Cai & Padoa-Schioppa, 2014; Padoa-Schioppa, 2011). Indeed, our results show that the interaction between value signals and motor preparation occurs on a moment-to-moment basis.

One outstanding issue is why ACC does not behave exactly like an accumulator of evi-
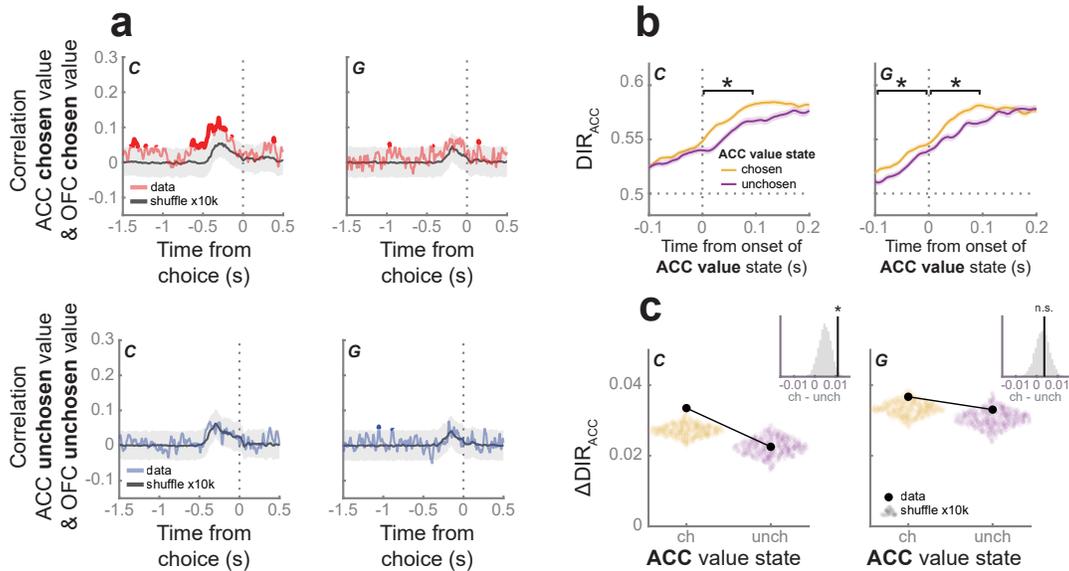
Figure 4.8: **ACC value correlations. (a)** ACC and OFC value strength (top: chosen, bottom: unchosen) was not significantly correlated. There was a weak correlation for chosen value states in subject C, but this was not replicated in subject G, and there was no correlation for unchosen value states. Gray line and shading indicate the median and 99% confidence interval from the null distribution of correlations from 10,000 shuffled trials. Bold colored lines indicate significance $p < 0.01$. **(b)** DIR_{ACC} synced to the onset of chosen (orange) or unchosen (purple) ACC value states during the period of rising DIR_{ACC}. DIR_{ACC} was stronger in the 100 ms following chosen ACC value states, as well as the preceding 100 ms for subject G. Same conventions as Fig. 4.5b. **(c)** Mean ΔDIR_{ACC} (black) across trials at chosen or unchosen ACC state onsets, and the difference in ΔDIR_{ACC} following chosen compared to unchosen ACC value states (inset). Same statistical tests and conventions as Fig. 4.5c.

dence. Specifically, when the chosen option is represented in OFC, we might expect motor preparation signals in ACC to reverse direction rather than simply decelerate. There are several possible explanations. First, since chosen values states are more frequent, misclassification of value states is more likely to lead to some chosen states been incorrectly labeled as unchosen states, leading to a bias to observe an increase in DIR_{ACC} as opposed to a decrease. Second, ACC may be integrating value signals from multiple brain regions which could limit the influence of OFC dynamics. For example, we have recently demonstrated the existence of two distinct value signals in the striatum, related to attentional capture and deliberative decision-making (Balewski et al., 2022). Finally, there may be distinct subpopulations in ACC related to the preparation of different movements, as occurs in other areas engaged in motor preparation (Cisek & Kalaska, 2005; Klaes et al., 2011). Our current design, in which there were only two different motor responses, was not ideal for detecting different subpopulations. For example, we cannot differentiate whether the population is encoding 'left' vs. 'right' or whether it might be encoding 'left' vs. 'not left'.

In conclusion, our results are consistent with ACC accumulating value evidence from

OFC in order to determine the optimal choice response. Our results also demonstrate the advantages of using high-channel count recordings combined with population-level analysis to decode the dynamics of cognitive processes (Wallis, 2018). We used the encoding of value in OFC to understand the dynamics of decision-making, which then enabled us to better understand neural signals in ACC.

4.4 Methods

Experimental model and subject details

All procedures were carried out as specified in the National Research Council guidelines and approved by the Animal Care and Use Committee at the University of California, Berkeley. Two male rhesus macaques (subjects C and G, respectively) aged 6 and 4 years, and weighing 10 and 7 kg at the time of recording were used in the current study. Subjects sat head-fixed in a primate chair (Crist Instrument, Hagerstown, MD) and manipulated a bidirectional lever located on the front of the chair. Eye movements were tracked with an infrared system (SR Research, Ottawa, Ontario, CN). Stimulus presentation and behavioral conditions were controlled using the MonkeyLogic toolbox (Hwang et al., 2019). Subjects had unilateral (subject C) or bilateral (subject G) recording chambers implanted, centered over the frontal lobe.

Task Design

Subjects performed a task in which they were required to choose between pairs of pictures or single pictures. Subjects fixated continuously for 750 ms on a central 0.5° cue to initiate the presentation of one (forced choice trials, 33%) or two (free choice trials, 67%) 2.5° x 2.5° pictures, presented 6° to either side of the fixation cue. We sampled gaze position at 500 Hz. Subjects used a bidirectional lever to indicate a left or right choice, and the selected picture remained on the screen while the corresponding juice amount was delivered probabilistically. The juice amounts (subject C: 0.15, 0.3, 0.45, 0.6 mL; subject G: 0.1, 0.2, 0.3, 0.4 mL) and reward probabilities (C: 0.15, 0.4, 0.65, 0.9; subject G: 0.1, 0.37, 0.63, 0.9) uniquely associated with each picture were titrated for each subject so that the subjects considered both dimensions during their choices. Subject C completed 12,033 free and 6,307 forced trials over 13 sessions; subject G completed 4,696 free and 2,577 forced trials over 5 sessions.

Neurophysiological recordings

Subjects were fitted with head positioners and imaged in a 3T MRI scanner. From the MR images, we constructed 3D models of each subjects' skull and target brain areas (Paxinos et al., 2000). Subjects were implanted with custom radiolucent recording chambers fabricated from polyether ether ketone (PEEK). During each recording session, up to eight

multisite linear probes (16- or 32- channel V probes with 75, 100, or 200 µm contact spacing, Plexon, Dallas, TX) were lowered into ACC (AP +21 to +37 mm) and OFC (areas 11 and 13, AP +30 to +40 mm). Electrode trajectories were defined in custom software, and the appropriate microdrives were 3D printed (Form 2 and 3, Formlabs, Cambridge, MA) (Knudsen et al., 2019). Lowering depths were derived from the MR images and verified from neurophysiological signals via gray/white matter transitions. Neural signals were digitized using a Plexon OmniPlex system, with continuous spike-filtered signals (200 Hz - 6 kHz) acquired at 40 kHz and local field-filtered signals acquired at 1 kHz.

We recorded neuronal activity over the course of 13 sessions for subject C and 5 sessions for subject G (Table 4.1). Units were manually sorted using 1400 ms waveforms, thresholded at 4 standard deviations above noise (Offline Sorter, Plexon). We restricted our analysis to neurons with a mean firing rate of > 1 Hz across the session. To ensure adequate isolation of neurons, we excluded neurons where $> 0.2\%$ of spikes were separated by < 1100 ms. Sorting quality was subjectively ranked on a 1 - 5 scale to separate well-isolated single neurons from possible multi-units. None of the results reported in the manuscript depended on the isolation quality of our neurons, so we included all neurons in our analysis. It was impossible to distinguish dorsal and ventral ACC in subject C due to his anatomy but none of the reported results indicated a dorso-ventral gradient, so we combined all ACC neurons from both sides of the sulcus. In total, we recorded 453 and 517 neurons from ACC in subjects C and G, respectively, and 454 and 429 neurons from OFC.

Only LFP channels with at least one unit were analyzed. Raw LFPs were notch filtered at 60, 120, and 180 Hz and bandpassed using finite impulse response filters in six frequency bands: δ (2-4 Hz), θ (4-8Hz), α (8-12 Hz), β (12-30 Hz), γ (30-60 Hz), and high γ (70-200 Hz). Analytic amplitudes and phase angles were obtained from Hilbert transforms in the pass bands. In total, we recorded 302 and 402 LFP channels from ACC in subjects C and G, respectively, and 341 and 360 LFP channels from OFC.

Behavioral analysis

We estimated the subjective value for each picture as the expected value, V:

$$expected\ value = juice\ amount \times reward\ probability \qquad (4.1)$$

We then fit choice behavior using a soft-max decision rule:

$$P_R = \frac{1}{1 + e^{w_1(V_R - V_L) + w_2 S + w_3}} \qquad (4.2)$$

This modeled the probability that the subject will select the picture on right, P_R, with three free parameters: the inverse temperature, w_1, which determines the stochasticity as a function of the value difference between the right and left picture; a saccade bias term, w_2, which accounts for the influence of the first gaze location, S, after picture onset (Cavanagh et al., 2019); and a side bias term, w_3, which accounts for any preference for one choice direction.

Subject	Session	ACC neurons	OFC neurons
C	rec09	–	29 [†]
	rec11	23	–
	rec16	22	31
	rec33	–	33 [†]
	rec36	–	33 [†]
	rec41	–	38
	rec42	–	52
	rec60	107 [‡]	42 [†]
	rec61	92 [‡]	46 [†]
	rec62	60 [‡]	17 [†]
	rec63	41 [‡]	44
	rec65	47 [‡]	34 [†]
	rec66	61 [‡]	55 [†]
	total:	**453**	**454**
G	rec28	99 [‡]	107 [†]
	rec29	128 [‡]	70 [†]
	rec31	112 [‡]	95 [†]
	rec32	83 [‡]	85 [†]
	rec33	95 [‡]	72 [†]
	total:	**517**	**429**

Table 4.1: **Neural data.** Number of recording sessions and neurons recorded per session in the two subjects. [‡] Sessions with successful value and direction decoding; [†] successful value decoding.

We estimated all free parameters by maximizing the log likelihood of the full model. We defined saccades as eye movements whose velocity exceeded 6 standard deviations from the mean velocity during fixation. The saccade time was defined by the peak velocity within each movement, and direction was identified by the subsequent eye position.

Single neuron regression analysis

For each neuron, we examined the relationship between firing rate, FR, and task variables in a linear regression:

$$FR = \beta_0 + \beta_1 * trial\,type + \beta_2 * choice\,direction + \beta_3 * max\,value + \beta_4 * trial\,number \quad (4.3)$$

with binary variables for $trial\,type$ (+1 free, -1 forced) and $choice\,direction$ (+1 contralateral to the neuron, -1 ipsilateral), continuous variables for the $maximum\,value$ of the presented pictures (max value: 1 - 4), and $trial\,number$, a nuisance parameter to absorb potential variance due to neuronal drift over the recording session. We repeated this regression in overlapping 100 ms windows shifted by 25 ms, and defined significance as $p < 0.01$ for

at least 100 ms (four consecutive time bins). We ensured that this criterion produced an acceptable false discovery rate, by confirming that $< 1.1\%$ of neurons reached significance during the fixation epoch. We performed this analysis with trials synced to picture onset and choice.

Direction decoding with single trial resolution

For each session, we trained linear discrimination analysis (LDA) decoders to predict the choice direction (left or right) from ACC neural activity during free trials in overlapping windows of 20 ms stepped by 5 ms (Balewski et al., 2022). We synced the trials to either picture onset or choice response. To reduce the dimensionality of the input features, we performed PCA across trials separately for normalized neuron firing rates and normalized amplitudes from the δ LFP band (2 - 4 Hz); higher frequency LFP bands did not meaningfully improve decoding accuracy. We restricted the decoder inputs to the top PCs that explained 95% of the variance within each group.

We randomly split all free trials into separate training and held-out sets. The training set was constrained such that there were an equal number of each offer pair where the chosen value level was greater than the unchosen value (i.e. left 3 vs chosen right 4, chosen left 4 vs right 3, 2 vs chosen 4, chosen 4 vs 2, etc.). The held-out set contained all remaining trials where the chosen value was equal to or greater than the unchosen value (i.e. correct trials). We used a k-fold validation procedure on the training set to assess decoding accuracy. We then used the entire training set to obtain decoder weights to compute the posterior probability for left vs. right direction on each trial in the held-out set.

To maximally use all trials in a session, we repeated the procedure with 25 random splits of training and held-out sets. For each trial, we averaged the computed posterior probabilities across all instances when it was in a held-out set. To ease interpretation of reported results, we relabeled the decoder class (left and right) as chosen direction and unchosen direction as appropriate for each trial; chosen direction decoding strength is referred to as DIR_{ACC} in the main text. Only sessions with good decoding accuracy on the training set of free trials were included in the analyses, defined as $> 60\%$ overall accuracy (chance = 50%).

Value decoding with single trial resolution

For each session, we trained an LDA decoder to predict the value (1 - 4) from either OFC or ACC neural activity during forced trials (Rich & Wallis, 2016). We averaged neuron firing rates and all LFP band magnitudes following picture onset (100 - 400 ms) or preceding choice (-300 to 0 ms). We performed the same dimensionality reduction as detailed for direction decoding. For the training set, we randomly sampled forced trials to uniformly represent all pictures on either side of the screen and used a k-fold validation procedure to assess decoding accuracy. We then used all of the training set to obtain decoder weights to compute the posterior probability for each value on all free trials in overlapping windows of 20 ms stepped by 5 ms. To maximally use all trials in a session, we repeated the procedure with 50 random

samples of forced training set trials and averaged the posterior probabilities for free trials. Only sessions with good decoding accuracy on forced trials were included in the analyses, defined as $> 40\%$ overall accuracy (C: $54 \pm 4\%$; G: $52 \pm 4\%$; chance $= 25\%$).

We observed that chance rates of decoding value on free trials deviated from the expected 25%, likely due to global differences in population firing rates during different phases of the task. To correct for this, we defined a unique baseline rate for each value level by averaging the posterior probability corresponding to that level across trials when the picture value was not available. We reported posterior probabilities as percent change from these baselines; we refer to this metric as decoding strength.

To ease interpretation of reported results, we relabeled the decoder classes (1 - 4) as chosen value, unchosen value, and unavailable as appropriate for each trial. Only free trials where the chosen value was greater than the unchosen value (i.e. correct trials) were analyzed. For all visualizations and analyses, one of the two unavailable value levels was randomly selected for each trial.

We identified sustained periods of confident decoding as states. On each trial, the decoding strength for a given value level needed to surpass 200% (i.e., double the baseline rate) for at least four consecutive time bins (spanning 35 ms) to be considered a state.

Change in DIR_{ACC} following OFC value states

We defined the change in direction strength decoded from ACC (ΔDIR_{ACC}) as the difference between the average strength (DIR_{ACC}) 100 ms before and after the onset of an OFC value state. We averaged ΔDIR_{ACC} within trials with more than one chosen or unchosen OFC value state; we restricted the analysis to value states in the period of rising DIR_{ACC} leading to choice (Fig. 4.2c), i.e. 300 ms before the peak for each subject.

We compared the difference between the average DIR_{ACC} following OFC chosen value states (C: 57% of free trials; G:83%) and OFC unchosen value states (C: 52% of free trials; G: 68%). To assess statistical significance we built a null distribution over this metric for 10,000 trial shuffles, where trial labels were randomly permuted within each value comparison (1 vs. 2, 1 vs. 3, 1 vs. 4, etc.) per session.

Statistics

All statistical tests are described in the main text or the corresponding figure legends. Error bars and shading indicate standard error of the mean (s.e.m.) unless otherwise specified. Probabilities and response times were transformed with *logit* and log_{10} functions, respectively. All terms in regression models were normalized, and had maximum variance inflation factors of 1.7. All comparisons were two-sided.

Resource availability

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Joni Wallis (wallis@berkeley.edu).

4.5 Acknowledgments

Chapter 5

Conclusion

We began this thesis with the goal of understanding how orbitofrontal cortex (OFC) value dynamics (Rich and Wallis, 2018; replication in chapter 3) are translated into decisions. We hypothesized that a downstream region, such as the head of the caudate nucleus (CdN; Chapter 3) or the anterior cingulate cortex (ACC; Chapter 4), may be integrating over these serial value representations of the available option. We recorded large neural populations (Chapter 2) simultaneously from OFC and either candidate region while two monkeys used a bidirectional lever to select between pairs of pictures, each associated with a unique probabilistic reward. We successfully predicted subjects' choices and response times from the neural activity in both ACC and CdN, but at vastly different times in the trial, suggesting different computational roles for this regions. We observed that the motor preparation signal was indeed influenced by the OFC value signal. In contrast in the CdN, we observed a rapid choice responses, which was locked to the picture presentation instead of motor response and independent of OFC value. This chapter will summarize the findings from our experiments and address remaining questions and future directions.

CdN supports fast, value-guided attention

We observed rapid encoding of the choice response in the CdN, locked to the picture presentation instead of motor response. This signal was strongly connected with the subjects' initial saccades, which also occurred within 200 ms of the start of the trial and were directed to the chosen item on the majority of trials. While CdN responses have been studied extensively int he context of individual reward-predictive cues (Anderson et al., 2014; Hikosaka et al., 2006; Kawagoe et al., 1998; Lauwereyns et al., 2002), our results show that this fast direction signal is also present during choice. Overall, this rapid CdN direction signal is consistent with a value-guided orienting response.

In contrast to the highly stereotyped early phasic direction response in CdN, OFC population dynamics are idiosyncratic across trials, likely reflecting slower, more deliberative reasoning. We replicated the main findings from Rich and Wallis, 2016: value representation flip-flopped between the available pictures, predicting both the optimality and speed

of the decision. OFC value signals were independent of the CdN direction response, and were also not influenced by overt eye movements. These findings are consistent with value-guided orienting and value-guided decision-making being parallel processes processes (see also: Cavanagh et al., 2019; Evans and Stanovich, 2013; Kahneman, 2011; McClure et al., 2004.

OFC value integration in ACC

The strength of the choice direction signal in ACC increased gradually leading up to the motor response. We observed that this motor preparation signal was biased toward the picture corresponding to the simultaneous OFC value representation on each trial, consistent with our integration model. Our results are consistent with previous findings motor preparation signals in ACC can be affected by the arrival of new information about the value of a choice option (Hunt et al., 2018).

One useful approach for further investigating this relationship may be a causal manipulation of OFC value dynamics in real time, for example with closed-loop microstimulation (Jadhav et al., 2012; Knudsen & Wallis, 2020; Zhou et al., 2019). We predict that selectively enhancing or disrupting OFC value states would bias both the ACC direction signal and motor response times toward or against the corresponding picture, respectively; we predict initial saccades and the CdN direction signal would be unchanged.

This relationship with OFC value may not be unique to ACC. Future experiments will need to explore other brain regions that also fit the candidate profile we established (Chapter 1). For example, the dorsolateral prefrontal cortex also has strong reciprocal connections with both OFC (Barbas & Pandya, 1989) and motor areas (Young & Shapiro, 2011), and neurons here transition from encoding value to motor responses over the evolution of the trial (Cai & Padoa-Schioppa, 2014; Hunt et al., 2015). The anterior striatum also receives strong projections from OFC (Ferry et al., 2000; Haber et al., 1995); while we ruled out the head of the caudate nucleus as an OFC value integrator (Chapter 3), the putamen remains a strong candidate. This region is critical to arm movement control (Inase et al., 1996; Kelly & Strick, 2004; Takada et al., 1998) and neurons here also encode both value and motor information (Hori et al., 2009).

High-channel recordings

Over the course of these experiments, we advanced existing recording methods to dramatically expand the number of simultaneously recorded neurons, peaking over 100 neurons each from two regions. We leveraged these ensembles to investigate neural dynamics that were not observable from trial-averaged single neuron responses: individual ACC and CdN neuronal response profiles were very similar, but the single-trial dynamics revealed important temporal differences in the encoding of direction response information. As more sophisticated probes (N. Steinmetz, 2019; N. A. Steinmetz et al., 2018) and imagine methods (Bollimunta et al., 2021; Kondo et al., 2018; Oguchi et al., 2021) continue to increase the density of

recorded neural populations, we will need to advance more sophisticated analysis methods to understand circuit- and systems-level population dynamics.

Parallel value processing

Value-based decision making is often presented as a serial process of choice evaluation, comparison, and motor action (Cai & Padoa-Schioppa, 2014; Padoa-Schioppa, 2011). In contrast, our results are more consistent with all of these elements happening in parallel in the span of the deliberative period. Value and direction signals from all three brain regions (OFC, ACC, and CdN) were not necessarily correlated with each other, but were all predictive of subjects' choices and response times.

While we made progress in understanding how OFC value signals are translated into actions, we opened a series of new questions about which regions beyond ACC may serve as OFC value integrators; whether there are other value circuits, like the parallel attention signal evident in CdN and initial saccades, that influence the choice or motor response; and how these parallel signals all ultimately drive the behavior.

Bibliography

Alexander, G. E., & DeLong, M. R. (1985). Microstimulation of the primate neostriatum. II. somatotopic organization of striatal microexcitable zones and their relation to neuronal response properties. *J. Neurophysiol.*, *53*(6), 1417–1430.

Anderson, B. A. (2016). The attention habit: How reward learning shapes attentional selection. *Ann. N. Y. Acad. Sci.*, *1369*(1), 24–39.

Anderson, B. A. (2019). Neurobiology of value-driven attention. *Curr Opin Psychol*, *29*, 27–33.

Anderson, B. A., Laurent, P. A., & Yantis, S. (2014). Value-driven attentional priority signals in human basal ganglia and visual cortex. *Brain Res.*, *1587*, 88–96.

Aosaki, T., Kimura, M., & Graybiel, A. M. (1995). Temporal and spatial characteristics of tonically active neurons of the primate's striatum. *J. Neurophysiol.*, *73*(3), 1234–1252.

Atallah, H. E., McCool, A. D., Howe, M. W., & Graybiel, A. M. (2014). Neurons in the ventral striatum exhibit cell-type-specific representations of outcome during learning. *Neuron*, *82*(5), 1145–1156.

Baden, T., Chagas, A. M., Gage, G. J., Marzullo, T. C., Prieto-Godino, L. L., & Euler, T. (2015). Open labware: 3-D printing your own lab equipment. *PLoS Biol.*, *13*(3), e1002086.

Balewski, Z. Z., Knudsen, E. B., & Wallis, J. D. (2022). Fast and slow contributions to decision-making in corticostriatal circuits. *Neuron*, *0*(0).

Balleine, B. W., & Dickinson, A. (1998). Goal-directed instrumental action: Contingency and incentive learning and their cortical substrates. *Neuropharmacology*, *37*(4-5), 407–419.

Balleine, B. W., & O'Doherty, J. P. (2009). Human and rodent homologies in action control: Corticostriatal determinants of Goal-Directed and habitual action. *Neuropsychopharmacology*, *35*(1), 48–69.

Ballesta, S., Shi, W., Conen, K. E., & Padoa-Schioppa, C. (2020). Values encoded in orbitofrontal cortex are causally related to economic choices. *Nature*, *588*(7838), 450–453.

Barbas, H., & Pandya, D. N. (1989). Architecture and intrinsic connections of the prefrontal cortex in the rhesus monkey. *J. Comp. Neurol.*, *286*(3), 353–375.

Bastos, A. M., Loonis, R., Kornblith, S., Lundqvist, M., & Miller, E. K. (2018). Laminar recordings in frontal cortex suggest distinct layers for maintenance and control of working memory. *Proc. Natl. Acad. Sci. U. S. A.*, *115*(5), 1117–1122.

Bollimunta, A., Santacruz, S. R., Eaton, R. W., Xu, P. S., Morrison, J. H., Moxon, K. A., Carmena, J. M., & Nassi, J. J. (2021). Head-mounted microendoscopic calcium imaging in dorsal premotor cortex of behaving rhesus macaque. *Cell Rep.*, *35*(11), 109239.

Brincat, S. L., & Miller, E. K. (2015). Frequency-specific hippocampal-prefrontal interactions during associative learning. *Nat. Neurosci.*, *18*(4), 576–581.

Burke, K. A., Franz, T. M., Miller, D. N., & Schoenbaum, G. (2008). The role of the orbitofrontal cortex in the pursuit of happiness and more specific rewards. *Nature*, *454*(7202), 340–344.

Cai, X., & Padoa-Schioppa, C. (2012). Neuronal encoding of subjective value in dorsal and ventral anterior cingulate cortex. *J. Neurosci.*, *32*(11), 3791–3808.

Cai, X., & Padoa-Schioppa, C. (2014). Contributions of orbitofrontal and lateral prefrontal cortices to economic choice and the good-to-action transformation. *Neuron*, *81*(5), 1140–1151.

Camille, N., Tsuchida, A., & Fellows, L. K. (2011). Double dissociation of stimulus-value and action-value learning in humans with orbitofrontal or anterior cingulate cortex damage. *J. Neurosci.*, *31*(42), 15048–15052.

Carmichael, S. T., & Price, J. L. (1996). Connectional networks within the orbital and medial prefrontal cortex of macaque monkeys. *J. Comp. Neurol.*, *371*(2), 179–207.

Cavanagh, S. E., Malalasekera, W. M. N., Miranda, B., Hunt, L. T., & Kennerley, S. W. (2019). Visual fixation patterns during economic choice reflect covert valuation processes that emerge with learning. *Proc. Natl. Acad. Sci. U. S. A.*, *116*(45), 22795–22801.

Chen, X., Possel, J. K., Wacongne, C., van Ham, A. F., Klink, P. C., & Roelfsema, P. R. (2017). 3D printing and modelling of customized implants and surgical guides for non-human primates. *J. Neurosci. Methods*, *286*, 38–55.

Cisek, P. (2012). Making decisions through a distributed consensus. *Curr. Opin. Neurobiol.*, *22*(6), 927–936.

Cisek, P., & Kalaska, J. F. (2005). Neural correlates of reaching decisions in dorsal premotor cortex: Specification of multiple direction choices and final selection of action. *Neuron*, *45*(5), 801–814.

Cisek, P., & Kalaska, J. F. (2010). Neural mechanisms for interacting with a world full of action choices. *Annu. Rev. Neurosci.*, *33*, 269–298.

Clarke, H. F., Cardinal, R. N., Rygula, R., Hong, Y. T., Fryer, T. D., Sawiak, S. J., Ferrari, V., Cockcroft, G., Aigbirhio, F. I., Robbins, T. W., & Roberts, A. C. (2014). Orbitofrontal dopamine depletion upregulates caudate dopamine and alters behavior via changes in reinforcement sensitivity. *J. Neurosci.*, *34*(22), 7663–7676.

Creed, M., Pascoli, V. J., & Lüscher, C. (2015). Refining deep brain stimulation to emulate optogenetic treatment of synaptic pathology. *Science*, *347*(6222), 659–664.

Della Libera, C., & Chelazzi, L. (2006). Visual selective attention and the effects of monetary rewards. *Psychol. Sci.*, *17*(3), 222–227.

Dum, R. P., & Strick, P. L. (1993). Cingulate motor areas. In B. A. Vogt & M. Gabriel (Eds.), *Neurobiology of cingulate cortex and limbic thalamus: A comprehensive handbook* (pp. 415–441). Birkhäuser Boston.

Eslinger, P. J., & Damasio, A. R. (1985). Severe disturbance of higher cognition after bilateral frontal lobe ablation: Patient EVR. *Neurology*, *35*(12), 1731–1741.

Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-Process theories of higher cognition: Advancing the debate. *Perspect. Psychol. Sci.*, *8*(3), 223–241.

Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.-C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F., Sonka, M., Buatti, J., Aylward, S., Miller, J. V., Pieper, S., & Kikinis, R. (2012). 3D slicer as an image computing platform for the quantitative imaging network. *Magn. Reson. Imaging*, *30*(9), 1323–1341.

Fernando, A. B. P., & Robbins, T. W. (2011). Animal models of neuropsychiatric disorders. *Annu. Rev. Clin. Psychol.*, *7*, 39–61.

Ferry, A. T., Ongür, D., An, X., & Price, J. L. (2000). Prefrontal cortical projections to the striatum in macaque monkeys: Evidence for an organization related to prefrontal networks. *J. Comp. Neurol.*, *425*(3), 447–470.

Freedman, D. S., Schroeder, J. B., Telian, G. I., Zhang, Z., Sunil, S., & Ritt, J. T. (2016). OptoZIF drive: A 3D printed implant and assembly tool package for neural recording and optical stimulation in freely moving mice. *J. Neural Eng.*, *13*(6), 066013.

Fuster, J. M. (1973). Unit activity in prefrontal cortex during delayed-response performance: Neuronal correlates of transient memory. *J. Neurophysiol.*, *36*(1), 61–78.

Ganguly, K., Dimitrov, D. F., Wallis, J. D., & Carmena, J. M. (2011). Reversible large-scale modification of cortical networks during neuroprosthetic control. *Nat. Neurosci.*, *14*(5), 662–667.

Ghazizadeh, A., Griggs, W., & Hikosaka, O. (2016). Ecological origins of object salience: Reward, uncertainty, aversiveness, and novelty. *Front. Neurosci.*, *10*, 378.

Gold, J. I., & Shadlen, M. N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends Cogn. Sci.*, *5*(1), 10–16.

Grahn, J. A., Parkinson, J. A., & Owen, A. M. (2008). The cognitive functions of the caudate nucleus. *Prog. Neurobiol.*, *86*(3), 141–155.

Graybiel, A. M. (2008). Habits, rituals, and the evaluative brain. *Annu. Rev. Neurosci.*, *31*(1), 359–387.

Haber, S. N., Kunishio, K., Mizobuchi, M., & Lynd-Balta, E. (1995). The orbital and medial prefrontal circuit through the primate basal ganglia. *J. Neurosci.*, *15*(7 Pt 1), 4851–4867.

Headley, D. B., DeLucca, M. V., Haufler, D., & Paré, D. (2015). Incorporating 3d-printing technology in the design of head-caps and electrode drives for recording neurons in multiple brain regions. *J. Neurophysiol.*, *113*(7), 2721–2732.

Hikosaka, O., Nakamura, K., & Nakahara, H. (2006). Basal ganglia orient eyes to reward. *J. Neurophysiol.*, *95*(2), 567–584.

Hori, Y., Mimura, K., Nagai, Y., Fujimoto, A., Oyama, K., Kikuchi, E., Inoue, K.-I., Takada, M., Suhara, T., Richmond, B. J., & Minamimoto, T. (2020). *Single caudate neurons encode temporally discounted value for formulating motivation for action.*

Hori, Y., Minamimoto, T., & Kimura, M. (2009). Neuronal encoding of reward value and direction of actions in the primate putamen. *J. Neurophysiol.*, *102*(6), 3530–3543.

Hosokawa, T., Kennerley, S. W., Sloan, J., & Wallis, J. D. (2013). Single-neuron mechanisms underlying cost-benefit analysis in frontal cortex. *J. Neurosci.*, *33*(44), 17385–17397.

Hunt, L. T., Behrens, T. E. J., Hosokawa, T., Wallis, J. D., & Kennerley, S. W. (2015). Capturing the temporal evolution of choice across prefrontal cortex. *Elife, 4.*

Hunt, L. T., Malalasekera, W. M. N., de Berker, A. O., Miranda, B., Farmer, S. F., Behrens, T. E. J., & Kennerley, S. W. (2018). Triple dissociation of attention and decision computations across prefrontal cortex. *Nat. Neurosci.*, *21*(10), 1471–1481.

Hwang, J., Mitz, A. R., & Murray, E. A. (2019). NIMH MonkeyLogic: Behavioral control and data acquisition in MATLAB. *J. Neurosci. Methods*, *323*, 13–21.

Inase, M., Sakai, S. T., & Tanji, J. (1996). Overlapping corticostriatal projections from the supplementary motor area and the primary motor cortex in the macaque monkey: An anterograde double labeling study. *J. Comp. Neurol.*, *373*(2), 283–296.

Jadhav, S. P., Kemere, C., German, P. W., & Frank, L. M. (2012). Awake hippocampal Sharp-Wave ripples support spatial memory. *Science*, *336*(6087), 1454–1458.

Jun, J. J., Steinmetz, N. A., Siegle, J. H., Denman, D. J., Bauza, M., Barbarits, B., Lee, A. K., Anastassiou, C. A., Andrei, A., Aydın, Ç., Barbic, M., Blanche, T. J., Bonin, V., Couto, J., Dutta, B., Gratiy, S. L., Gutnisky, D. A., Häusser, M., Karsh, B., . . . Harris, T. D. (2017). Fully integrated silicon probes for high-density recording of neural activity. *Nature*, *551*(7679), 232–236.

Kahneman, D. (2011). *Thinking, fast and slow.* Macmillan.

Kawagoe, R., Takikawa, Y., & Hikosaka, O. (1998). Expectation of reward modulates cognitive signals in the basal ganglia. *Nat. Neurosci.*, *1*(5), 411–416.

Kelly, R. M., & Strick, P. L. (2004). Macro-architecture of basal ganglia loops with the cerebral cortex: Use of rabies virus to reveal multisynaptic circuits. *Prog. Brain Res.*, *143*, 449–459.

Kennerley, S. W., Behrens, T. E. J., & Wallis, J. D. (2011). Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nat. Neurosci.*, *14*(12), 1581–1589.

Kennerley, S. W., Dahmubed, A. F., Lara, A. H., & Wallis, J. D. (2009). Neurons in the frontal lobe encode the value of multiple decision variables.

Kennerley, S. W., & Wallis, J. D. (2009). Evaluating choices by single neurons in the frontal lobe: Outcome value encoded across multiple decision variables. *Eur. J. Neurosci.*, *29*(10), 2061–2073.

Kennerley, S. W., & Walton, M. E. (2011). Decision making and reward in frontal cortex: Complementary evidence from neurophysiological and neuropsychological studies. *Behav. Neurosci.*, *125*(3), 297–317.

Kennerley, S. W., Walton, M. E., Behrens, T. E. J., Buckley, M. J., & Rushworth, M. F. S. (2006). Optimal decision making and the anterior cingulate cortex. *Nat. Neurosci.*, *9*(7), 940–947.

Kim, H. F., Griggs, W. S., & Hikosaka, O. (2020). Long-Term value memory in the primate posterior thalamus for fast automatic action. *Curr. Biol.*, *30*(15), 2901–2911.e3.

Klaes, C., Westendorff, S., Chakrabarti, S., & Gail, A. (2011). Choosing goals, not rules: Deciding among rule-based action plans. *Neuron, 70*(3), 536–548.

Kloosterman, F., Davidson, T. J., Gomperts, S. N., Layton, S. P., Hale, G., Nguyen, D. P., & Wilson, M. A. (2009). Micro-drive array for chronic in vivo recording: Drive fabrication. *J. Vis. Exp.*, (26).

Knudsen, E. B., Balewski, Z. Z., & Wallis, J. D. (2019). A model-based approach for targeted neurophysiology in the behaving non-human primate. *Int. IEEE EMBS Conf. Neural Eng.*, *2019*, 195–198.

Knudsen, E. B., & Wallis, J. D. (2020). Closed-Loop theta stimulation in the orbitofrontal cortex prevents Reward-Based learning. *Neuron, 106*(3), 537–547.e4.

Kondo, T., Saito, R., Otaka, M., Yoshino-Saito, K., Yamanaka, A., Yamamori, T., Watakabe, A., Mizukami, H., Schnitzer, M. J., Tanaka, K. F., Ushiba, J., & Okano, H. (2018). Calcium transient dynamics of neural ensembles in the primary motor cortex of naturally behaving monkeys. *Cell Rep.*, *24*(8), 2191–2195.e4.

Lara, A. H., & Wallis, J. D. (2014). Executive control processes underlying multi-item working memory. *Nat. Neurosci.*, *17*(6), 876–883.

Lauwereyns, J., Watanabe, K., Coe, B., & Hikosaka, O. (2002). A neural correlate of response bias in monkey caudate nucleus. *Nature, 418*(6896), 413–417.

Le Pelley, M. E., Pearson, D., Griffiths, O., & Beesley, T. (2015). When goals conflict with values: Counterproductive attentional and oculomotor capture by reward-related stimuli. *J. Exp. Psychol. Gen.*, *144*(1), 158–171.

Lin, Z., Nie, C., Zhang, Y., Chen, Y., & Yang, T. (2020). Evidence accumulation for value computation in the prefrontal cortex during decision making. *Proc. Natl. Acad. Sci. U. S. A.*

Luk, C.-H., & Wallis, J. D. (2013). Choice coding in frontal cortex during stimulus-guided or action-guided decision-making. *J. Neurosci.*, *33*(5), 1864–1871.

Mante, V., Sussillo, D., Shenoy, K. V., & Newsome, W. T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature, 503*(7474), 78–84.

Mazurek, M. E., Roitman, J. D., Ditterich, J., & Shadlen, M. N. (2003). A role for neural integrators in perceptual decision making. *Cereb. Cortex, 13*(11), 1257–1269.

McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science, 306*(5695), 503–507.

McHaffie, J. G., Stanford, T. R., Stein, B. E., Coizet, V., & Redgrave, P. (2005). Subcortical loops through the basal ganglia. *Trends Neurosci.*, *28*(8), 401–407.

Monosov, I. E. (2020). How outcome uncertainty mediates attention, learning, and Decision-Making. *Trends Neurosci.*, *43*(10), 795–809.

Morecraft, R. J., Stilwell-Morecraft, K. S., Cipolloni, P. B., Ge, J., McNeal, D. W., & Pandya, D. N. (2012). Cytoarchitecture and cortical connections of the anterior cingulate and adjacent somatomotor fields in the rhesus monkey. *Brain Res. Bull.*, *87*(4-5), 457–497.

Mulliken, G. H., Bichot, N. P., Ghadooshahy, A., Sharma, J., Kornblith, S., Philcock, M., & Desimone, R. (2015). Custom-fit radiolucent cranial implants for neurophysiological recording and stimulation. *J. Neurosci. Methods*, *241*, 146–154.

Muranishi, M., Inokawa, H., Yamada, H., Ueda, Y., Matsumoto, N., Nakagawa, M., & Kimura, M. (2011). Inactivation of the putamen selectively impairs reward history-based action selection. *Exp. Brain Res.*, *209*(2), 235–246.

Murray, E. A., & Rudebeck, P. H. (2018). Specializations for reward-guided decision-making in the primate ventral prefrontal cortex. *Nat. Rev. Neurosci.*, *19*(7), 404–417.

Nakamura, K., & Hikosaka, O. (2006). Facilitation of saccadic eye movements by postsaccadic electrical stimulation in the primate caudate. *J. Neurosci.*, *26*(50), 12885–12895.

Niki, H., & Watanabe, M. (1979). Prefrontal and cingulate unit activity during timing behavior in the monkey. *Brain Res.*, *171*(2), 213–224.

Oguchi, M., Jiasen, J., Yoshioka, T. W., Tanaka, Y. R., Inoue, K., Takada, M., Kikusui, T., Nomoto, K., & Sakagami, M. (2021). Microendoscopic calcium imaging of the primary visual cortex of behaving macaques. *Sci. Rep.*, *11*(1), 17021.

Padoa-Schioppa, C. (2011). Neurobiology of economic choice: A good-based model. *Annu. Rev. Neurosci.*, *34*, 333–359.

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

Patel, S. R., Ghose, K., & Eskandar, E. N. (2014). An open source 3-d printed modular micro-drive system for acute neurophysiology. *PLoS One*, *9*(4), e94262.

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

Pearson, D., Donkin, C., Tran, S. C., Most, S. B., & Le Pelley, M. E. (2015). Cognitive control and counterproductive oculomotor capture by reward-related stimuli. *Vis. cogn.*, *23*(1), 41–66.

Rich, E. L., & Wallis, J. D. (2016). Decoding subjective decisions from orbitofrontal cortex. *Nat. Neurosci.*, *19*(7), 973–980.

Roitman, J. D., & Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci.*, *22*(21), 9475–9489.

Rudebeck, P. H., & Murray, E. A. (2011). Dissociable effects of subtotal lesions within the macaque orbital prefrontal cortex on Reward-Guided behavior. *J. Neurosci.*, *31*(29), 10569–10578.

Santacruz, S. R., Rich, E. L., Wallis, J. D., & Carmena, J. M. (2017). Caudate microstimulation increases value of specific choices. *Curr. Biol.*, *27*(21), 3375–3383.e3.

Scangos, K. W., Khambhati, A. N., Daly, P. M., Makhoul, G. S., Sugrue, L. P., Zamanian, H., Liu, T. X., Rao, V. R., Sellers, K. K., Dawes, H. E., Starr, P. A., Krystal, A. D., &

Chang, E. F. (2021). Closed-loop neuromodulation in an individual with treatment-resistant depression. *Nat. Med.*, 1–5.

Shanechi, M. M. (2019). Brain–machine interfaces from motor to mood. *Nat. Neurosci.*, *22*(10), 1554–1564.

Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nat. Neurosci.*, *6*(12), 1317–1322.

Stalnaker, T. A., Cooch, N. K., & Schoenbaum, G. (2015). What the orbitofrontal cortex does not do. *Nat. Neurosci.*, *18*(5), 620–627.

Steinmetz, N. (2019). Large-scale electrophysiology with neuropixels: Scientific advances and future directions. *IBRO Rep.*, *6*, S43.

Steinmetz, N. A., Koch, C., Harris, K. D., & Carandini, M. (2018). Challenges and opportunities for large-scale electrophysiology with neuropixels probes. *Curr. Opin. Neurobiol.*, *50*, 92–100.

Szczepanski, S. M., & Knight, R. T. (2014). Insights into human behavior from lesions to the prefrontal cortex. *Neuron*, *83*(5), 1002–1018.

Takada, M., Tokuno, H., Hamada, I., Inase, M., Ito, Y., Imanishi, M., Hasegawa, N., Akazawa, T., Hatanaka, N., & Nambu, A. (2001). Organization of inputs from cingulate motor areas to basal ganglia in macaque monkey. *Eur. J. Neurosci.*, *14*(10), 1633–1650.

Takada, M., Tokuno, H., Nambu, A., & Inase, M. (1998). Corticostriatal projections from the somatic motor areas of the frontal cortex in the macaque monkey: Segregation versus overlap of input zones from the primary motor cortex, the supplementary motor area, and the premotor cortex. *Exp. Brain Res.*, *120*(1), 114–128.

Traner, M. R., Bromberg-Martin, E. S., & Monosov, I. E. (2021). How the value of the environment controls persistence in visual search. *PLoS Comput. Biol.*, *17*(12), e1009662.

Van Hoesen, G. W., Morecraft, R. J., & Vogt, B. A. (1993). Connections of the monkey cingulate cortex. In B. A. Vogt & M. Gabriel (Eds.), *Neurobiology of cingulate cortex and limbic thalamus: A comprehensive handbook* (pp. 249–284). Birkhäuser Boston.

Vickery, T. J., Chun, M. M., & Lee, D. (2011). Ubiquity and specificity of reinforcement signals throughout the human brain. *Neuron*, *72*(1), 166–177.

Wallis, J. D. (2007). Neuronal mechanisms in prefrontal cortex underlying adaptive choice behavior. *Ann. N. Y. Acad. Sci.*, *1121*, 447–460.

Wallis, J. D. (2011). Cross-species studies of orbitofrontal cortex and value-based decision-making. *Nat. Neurosci.*, *15*(1), 13–19.

Wallis, J. D., & Kennerley, S. W. (2010). Heterogeneous reward signals in prefrontal cortex. *Curr. Opin. Neurobiol.*, *20*(2), 191–198.

Wallis, J. D. (2018). Decoding cognitive processes from neural ensembles. *Trends Cogn. Sci.*, *22*(12), 1091–1102.

Watanabe, K., & Hikosaka, O. (2005). Immediate changes in anticipatory activity of caudate neurons associated with reversal of position-reward contingency. *J. Neurophysiol.*, *94*(3), 1879–1887.

Young, J. J., & Shapiro, M. L. (2011). Dynamic coding of Goal-Directed paths by orbital prefrontal cortex. *J. Neurosci.*, *31*(16), 5989–6000.

Zhou, A., Santacruz, S. R., Johnson, B. C., Alexandrov, G., Moin, A., Burghardt, F. L., Rabaey, J. M., Carmena, J. M., & Muller, R. (2019). A wireless and artefact-free 128-channel neuromodulation device for closed-loop stimulation and recording in non-human primates. *Nat Biomed Eng*, *3*(1), 15–26.