Fast and slow contributions to decision making in corticostriatal circuits

Zuzanna Z. Balewski[1], **Eric B. Knudsen**[1], **Joni D. Wallis**[1,2,*]

[1]Helen Wills Neuroscience Institute, University of California at Berkeley, Berkeley, California, 94720, USA

[2]Department of Psychology, University of California at Berkeley, Berkeley, California, 94720, USA

Summary

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.

eTOC blurb

Valuable stimuli attract our attention, but how this interacts with subsequent decisions remains unclear. Balewski et al. found a rapid, phasic signal in the caudate that guided eye movements and biased subsequent choices. This signal was independent of value representations in the orbitofrontal cortex that are involved in value-guided decision-making.

Introduction

To navigate complex decisions, we rely on a combination of fast, autonomous intuitions and slower, deliberative judgements (Evans and 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 and 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 and 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 and Assad, 2006), and the region is critical to value-based decisions (Ballesta et al.; Knudsen and Wallis, 2020), especially when new information needs to be computed or integrated (Murray and 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 and 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.

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. 1a). Both subjects learned 16 pictures, which were associated with one of four juice amounts and four reward probabilities (Fig. 1b). As expected, subjects preferred larger juice amounts and more certain rewards (Fig. 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. 1d). In contrast, the first saccades following picture onset were fast (subject C: free trials median = 162 ms, forced trials median 206 = ms; subject G: 189 ms, 203 ms; Fig. 1d), yet overwhelmingly predicted the chosen picture even on free trials (subject C: 84% of free trials; subject 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. 1, Fig. 1e). We estimated subjective picture value as expected value (Eq. 2), which better explained choices (subject C: 76% explained variance; subject G: 86%) compared with other reported value functions (Eq. 3) (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. 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 (subject C: $\beta = -0.39$, CPD = 11%, $p < 1 \times 10^{-5}$; subject 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 (subject C: $\beta = -0.04$, CPD = 0.2%, $p < 0.0001$; subject G: $\beta = -0.03$, CPD = 0.1%, $p = 0.01$; Fig. 1g). The onset of the first saccade was also strongly dependent on the maximum value, with faster saccades to more valuable pictures (subject C: $\beta = -0.6$, CPD = 27%, $p < 1 \times 10^{-5}$; subject G: $\beta = -0.54$, CPD = 22%, $p = 1 \times < 10^{-5}$). The effect of value difference on saccade timing however was weak and inconsistent (subject C: $\beta = 0.06$, CPD = 0.4%, $p < 1 \times 10^{-5}$; subject G: $\beta = 0.01$, CPD = 0%, $p = 0.58$; Fig. 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. 2a). 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. 4, see Methods). Overall, the firing rates of 69% of 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. 2b); we found no anatomical organization among these neurons (Fig. S1). Most neurons were significantly predicted by maximum value, choice direction, or both (Fig. 2c–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. 3a). Repeating this procedure with trials synced to the lever movement, we observed significantly lower overall decoding accuracy leading to choice (Fig. 3b). 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. 3c). 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. 2e). Since most CdN neurons also encoded value, we examined how the choice direction representation varied with maximum picture value (Fig. 3d). We quantified this relationship by modelling the logit-transformed decoding strength as a linear function of maximum value (Fig. 3e) 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 log10(response times) with logit-transformed decoding strength and maximum value as predictors (Fig. 3f). 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. 1d, 3c). 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 (subject C: median = 160 ms, G: 170 ms) was very similar to first saccade movement times (C: median = 162 ms, G: 189 ms; Fig. 4a). Similar to first saccades, the onset of the first direction state occurred significantly earlier for trials with pictures of larger maximum value (Fig. 4b; linear regression, C: $\beta = -0.07$, $p < 1 \times 10^{-12}$; 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. 4c).

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. 4c). Of course, we cannot rule out 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. 4d), and for most trials (C: 62%, G: 71%) the decoded state did not change (Fig. 4e).

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.

5a). 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. 5b). Compared to CdN neurons, fewer OFC neurons were modulated by choice direction and trial type (Fig. 5c–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 and 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 and 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. 6a). 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. 6b). 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. 6c). 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. S2a). 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. 6d). 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. 6e).

In sum, value decoding in OFC correlated with choice behavior on average but vacillated 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. 7a). 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. 7b). 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. 7c). 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. S2b). 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. 7d).

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. 7e). 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.

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 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 and 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 and 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 and O'Doherty, 2010). 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 (Steinmetz et al., 2018; Steinmetz et al., 2019).

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 and 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 and 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 and Rich, 2011). Value signals are nearly ubiquitous throughout the brain (Vickery et al., 2011), yet they often subserve very different functions (Wallis and 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.

STAR Methods

RESOURCE AVAILABILITY

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

MATERIALS AVAILABILITY—This study did not generate any new unique reagents.

DATA AND CODE AVAILABILITY—Any data, code, and additional information supporting the current work are available from the lead contact on request.

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.

METHOD DETAILS

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° × 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; subject 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 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. 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 and 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: delta (2–4 Hz), theta (4–8Hz), alpha (8–12 Hz), beta (12–30 Hz), gamma (30–60 Hz), and high gamma (70–200 Hz). Analytic amplitudes were obtained from Hilbert transforms in the pass bands.

QUANTIFICATION AND STATISTICAL ANALYSIS

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 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,

$$P_R = \frac{1}{1 + exp(w_1 * (V_R - V_L) + w_2 * S + w_3)}$$

Eq. 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

$$expected\ value = juice\ probability * reward\ amount$$

Eq. 2:

which did not add any free parameters to the model, and discount functions from Hosokawa et al (2013), which all added one free parameter. We estimated all free parameters by maximizing the log likelihood of the full model:

$$-\sum N_R * \ln\left(P_{R\ |\ \varDelta V}\right) - \sum N_L * \ln\left(1 - P_{R\ |\ \varDelta V}\right)$$

Eq. 3:

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, AIC, 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*:

$$FR = \beta_0 + \beta_1 * trial\ type + \beta_2 * choice\ direction + \beta_3 * max\ value + \beta_4 * trial\ number$$

Eq. 4:

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 (subject C: n = 581, G: n = 383). We calculated 99.9% confidence intervals over the means of 10,000 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 ± 1.7%) and 5 sessions for CdN (46 ± 1.7%). However, only 1 of these sessions had acceptable decoding in both areas. For subject G, all 5 sessions with probes in OFC met these criteria for both regions (OFC: 58 ± 2.5%, CdN: 53 ± 1.3%), likely due to larger neuron yields. 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. All comparisons were two-tailed.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Alexander GE, and DeLong MR (1985). Microstimulation of the primate neostriatum. II. Somatotopic organization of striatal microexcitable zones and their relation to neuronal response properties. J. Neurophysiol. 53, 1417–1430. 10.1152/jn.1985.53.6.1417. [PubMed: 4009227]

Anderson BA (2016). The attention habit: how reward learning shapes attentional selection. Ann. N. Y. Acad. Sci. 1369, 24–39. 10.1111/nyas.12957. [PubMed: 26595376]

Anderson BA (2019). Neurobiology of value-driven attention. Curr Opin Psychol 29, 27–33. 10.1016/j.copsyc.2018.11.004. [PubMed: 30472540]

Anderson BA, Laurent PA, and Yantis S (2014). Value-driven attentional priority signals in human basal ganglia and visual cortex. Brain Res. 1587, 88–96. 10.1016/j.brainres.2014.08.062. [PubMed: 25171805]

Aosaki T, Kimura M, and Graybiel AM (1995). Temporal and spatial characteristics of tonically active neurons of the primate's striatum. J. Neurophysiol. 73, 1234–1252. 10.1152/jn.1995.73.3.1234. [PubMed: 7608768]

Atallah HE, McCool AD, Howe MW, and Graybiel AM (2014). Neurons in the ventral striatum exhibit cell-type-specific representations of outcome during learning. Neuron 82, 1145–1156. 10.1016/j.neuron.2014.04.021. [PubMed: 24908491]

Balleine BW, and Dickinson A (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. Neuropharmacology 37, 407–419. [PubMed: 9704982]

Balleine BW, and O'Doherty JP (2010). Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. Neuropsychopharmacology 35, 48–69. 10.1038/npp.2009.131. [PubMed: 19776734]

Ballesta S, Shi W, Conen KE, and Padoa-Schioppa C (2020). Values encoded in orbitofrontal cortex are causally related to economic choices. Nature. 10.1038/s41586-020-2880-x.

Cavanagh SE, Malalasekera WMN, Miranda B, Hunt LT, and Kennerley SW (2019). Visual fixation patterns during economic choice reflect covert valuation processes that emerge with learning. Proc Natl Acad Sci U S A 116, 22795–22801. 10.1073/pnas.1906662116. [PubMed: 31636178]

Clarke HF, Cardinal RN, Rygula R, Hong YT, Fryer TD, Sawiak SJ, Ferrari V, Cockcroft G, Aigbirhio FI, Robbins TW, and Roberts AC (2014). Orbitofrontal dopamine depletion upregulates caudate dopamine and alters behavior via changes in reinforcement sensitivity. J. Neurosci. 34, 7663–7676. 10.1523/JNEUROSCI.0718-14.2014. [PubMed: 24872570]

Creed M, Pascoli VJ, and Luscher C (2015). Addiction therapy. Refining deep brain stimulation to emulate optogenetic treatment of synaptic pathology. Science 347, 659–664. 10.1126/science.1260776. [PubMed: 25657248]

Della Libera C, and Chelazzi L (2006). Visual selective attention and the effects of monetary rewards. Psychol. Sci. 17, 222–227. [PubMed: 16507062]

Evans JS, and Stanovich KE (2013). Dual-Process Theories of Higher Cognition: Advancing the Debate. Perspect Psychol Sci 8, 223–241. 10.1177/1745691612460685. [PubMed: 26172965]

Fernando AB, and Robbins TW (2011). Animal models of neuropsychiatric disorders. Annu Rev Clin Psychol 7, 39–61. 10.1146/annurev-clinpsy-032210-104454. [PubMed: 21219191]

Ferry AT, Ongur D, An X, and Price JL (2000). Prefrontal cortical projections to the striatum in macaque monkeys: evidence for an organization related to prefrontal networks. J. Comp. Neurol. 425, 447–470. 10.1002/1096-9861(20000925)425:3<447::aid-cne9>3.0.co;2-v. [PubMed: 10972944]

Ghazizadeh A, Griggs W, and Hikosaka O (2016). Ecological Origins of Object Salience: Reward, Uncertainty, Aversiveness, and Novelty. Front Neurosci 10, 378. 10.3389/fnins.2016.00378. [PubMed: 27594825]

Graybiel AM (2008). Habits, rituals, and the evaluative brain. Annu. Rev. Neurosci. 31, 359–387. 10.1146/annurev.neuro.29.051605.112851. [PubMed: 18558860]

Haber SN, Kunishio K, Mizobuchi M, and Lynd-Balta E (1995). The orbital and medial prefrontal circuit through the primate basal ganglia. J. Neurosci. 15, 4851–4867. [PubMed: 7623116]

Hikosaka O, Nakamura K, and Nakahara H (2006). Basal ganglia orient eyes to reward. J. Neurophysiol. 95, 567–584. [PubMed: 16424448]

Hori Y, Minamimoto T, and Kimura M (2009). Neuronal encoding of reward value and direction of actions in the primate putamen. J. Neurophysiol. 102, 3530–3543. 10.1152/jn.00104.2009. [PubMed: 19812294]

Hosokawa T, Kennerley SW, Sloan J, and Wallis JD (2013). Single-neuron mechanisms underlying cost-benefit analysis in frontal cortex. J. Neurosci. 33, 17385–17397. 10.1523/JNEUROSCI.2221-13.2013. [PubMed: 24174671]

Hwang J, Mitz AR, and Murray EA (2019). NIMH MonkeyLogic: Behavioral control and data acquisition in MATLAB. J. Neurosci. Methods 323, 13–21. 10.1016/j.jneumeth.2019.05.002. [PubMed: 31071345]

Inase M, Sakai ST, and 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, 283–296. 10.1002/(SICI)1096-9861(19960916)373:2<283::AID-CNE10>3.0.CO;2-M. [PubMed: 8889928]

Kahneman D (2011). Thinking, Fast and Slow (Farrar, Straus and Giroux).

Kawagoe R, Takikawa Y, and Hikosaka O (1998). Expectation of reward modulates cognitive signals in the basal ganglia. Nat. Neurosci. 1, 411–416. [PubMed: 10196532]

Kelly RM, and Strick PL (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. 10.1016/s0079-6123(03)43042-2. [PubMed: 14653187]

Kennerley SW, Dahmubed AF, Lara AH, and Wallis JD (2009). Neurons in the frontal lobe encode the value of multiple decision variables. J Cogn Neurosci 21, 1162–1178. 10.1162/jocn.2009.21100. [PubMed: 18752411]

Kim HF, Griggs WS, and Hikosaka O (2020). Long-Term Value Memory in the Primate Posterior Thalamus for Fast Automatic Action. Curr. Biol. 30, 2901–2911 e2903. 10.1016/j.cub.2020.05.047. [PubMed: 32531286]

Knudsen EB, Balewski ZZ, and Wallis JD (2019). A model-based approach for targeted neurophysiology in the behaving non-human primate. Int IEEE EMBS Conf Neural Eng 2019, 195–198. 10.1109/NER.2019.8716968. [PubMed: 31367267]

Knudsen EB, and Wallis JD (2020). Closed-loop theta stimulation in the orbitofrontal cortex prevents reward-based learning. Neuron 106, 537–547 e534. 10.1016/j.neuron.2020.02.003. [PubMed: 32160515]

Lauwereyns J, Watanabe K, Coe B, and Hikosaka O (2002). A neural correlate of response bias in monkey caudate nucleus. Nature 418, 413–417. [PubMed: 12140557]

Le Pelley ME, Pearson D, Griffiths O, and Beesley T (2015). When goals conflict with values: counterproductive attentional and oculomotor capture by reward-related stimuli. J Exp Psychol Gen 144, 158–171. 10.1037/xge0000037. [PubMed: 25420117]

McClure SM, Laibson DI, Loewenstein G, and Cohen JD (2004). Separate neural systems value immediate and delayed monetary rewards. Science 306, 503–507. [PubMed: 15486304]

McHaffie JG, Stanford TR, Stein BE, Coizet V, and Redgrave P (2005). Subcortical loops through the basal ganglia. Trends Neurosci. 28, 401–407. 10.1016/j.tins.2005.06.006. [PubMed: 15982753]

Monosov IE (2020). How Outcome Uncertainty Mediates Attention, Learning, and Decision-Making. Trends Neurosci. 43, 795–809. 10.1016/j.tins.2020.06.009. [PubMed: 32736849]

Muranishi M, Inokawa H, Yamada H, Ueda Y, Matsumoto N, Nakagawa M, and Kimura M (2011). Inactivation of the putamen selectively impairs reward history-based action selection. Exp. Brain Res. 209, 235–246. 10.1007/s00221-011-2545-y. [PubMed: 21298425]

Murray EA, and Rudebeck PH (2018). Specializations for reward-guided decision-making in the primate ventral prefrontal cortex. Nat. Rev. Neurosci. 19, 404–417. 10.1038/s41583-018-0013-4. [PubMed: 29795133]

Nakamura K, and Hikosaka O (2006). Facilitation of saccadic eye movements by postsaccadic electrical stimulation in the primate caudate. J. Neurosci. 26, 12885–12895. 10.1523/JNEUROSCI.3688-06.2006. [PubMed: 17167079]

Padoa-Schioppa C (2011). Neurobiology of economic choice: a good-based model. Annu. Rev. Neurosci. 34, 333–359. 10.1146/annurev-neuro-061010-113648. [PubMed: 21456961]

Padoa-Schioppa C, and Assad JA (2006). Neurons in the orbitofrontal cortex encode economic value. Nature 441, 223–226. [PubMed: 16633341]

Paxinos G, Huang X, and Toga AW (2000). The rhesus monkey brain in stereotaxic coordinates (Academic press).

Pearson D, Donkin C, Tran SC, Most SB, and Le Pelley ME (2015). Cognitive control and counterproductive oculomotor capture by reward-related stimuli. Visual Cognition 23, 41–66. 10.1080/13506285.2014.994252.

Rich EL, and Wallis JD (2016). Decoding subjective decisions from orbitofrontal cortex. Nat. Neurosci. 19, 973–980. 10.1038/nn.4320. [PubMed: 27273768]

Santacruz SR, Rich EL, Wallis JD, and Carmena JM (2017). Caudate microstimulation increases value of specific choices. Curr. Biol. 27, 3375–3383 e3373. 10.1016/j.cub.2017.09.051. [PubMed: 29107551]

Scangos KW, Khambhati AN, Daly PM, Makhoul GS, Sugrue LP, Zamanian H, Liu TX, Rao VR, Sellers KK, Dawes HE, et al. (2021). Closed-loop neuromodulation in an individual with treatment-resistant depression. Nat. Med. 10.1038/s41591-021-01480-w.

Shanechi MM (2019). Brain-machine interfaces from motor to mood. Nat. Neurosci. 22, 1554–1564. 10.1038/s41593-019-0488-y. [PubMed: 31551595]

Shimojo S, Simion C, Shimojo E, and Scheier C (2003). Gaze bias both reflects and influences preference. Nat. Neurosci. 6, 1317–1322. 10.1038/nn1150. [PubMed: 14608360]

Stalnaker TA, Cooch NK, and Schoenbaum G (2015). What the orbitofrontal cortex does not do. Nat. Neurosci. 18, 620–627. 10.1038/nn.3982. [PubMed: 25919962]

Steinmetz NA, Koch C, Harris KD, and Carandini M (2018). Challenges and opportunities for large-scale electrophysiology with Neuropixels probes. Curr. Opin. Neurobiol. 50, 92–100. 10.1016/j.conb.2018.01.009. [PubMed: 29444488]

Steinmetz NA, Zatka-Haas P, Carandini M, and Harris KD (2019). Distributed coding of choice, action and engagement across the mouse brain. Nature 576, 266–273. 10.1038/s41586-019-1787-x. [PubMed: 31776518]

Takada M, Tokuno H, Nambu A, and 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, 114–128. 10.1007/s002210050384. [PubMed: 9628410]

Traner MR, Bromberg-Martin ES, and Monosov IE (2021). How the value of the environment controls persistence in visual search. PLoS Comput Biol 17, e1009662. 10.1371/journal.pcbi.1009662. [PubMed: 34905548]

Vickery TJ, Chun MM, and Lee D (2011). Ubiquity and specificity of reinforcement signals throughout the human brain. Neuron 72, 166–177. 10.1016/j.neuron.2011.08.011. [PubMed: 21982377]

Wallis JD (2007). Orbitofrontal cortex and its contribution to decision-making. Annu. Rev. Neurosci. 30, 31–56. [PubMed: 17417936]

Wallis JD (2018). Decoding Cognitive Processes from Neural Ensembles. Trends Cogn Sci. 10.1016/j.tics.2018.09.002.

Wallis JD, and Kennerley SW (2010). Heterogeneous reward signals in prefrontal cortex. Curr. Opin. Neurobiol. 20, 191–198. 10.1016/j.conb.2010.02.009. [PubMed: 20303739]

Wallis JD, and Rich EL (2011). Challenges of interpreting frontal neurons during value-based decision-making. Frontiers in neuroscience 5, 124. 10.3389/fnins.2011.00124. [PubMed: 22125508]

Watanabe K, and Hikosaka O (2005). Immediate changes in anticipatory activity of caudate neurons associated with reversal of position-reward contingency. J. Neurophysiol. 94, 1879–1887. 10.1152/jn.00012.2005. [PubMed: 15872072]

Highlights

- Rapid, phasic signals in caudate predict choice behavior during value-guided choices

- The phasic caudate response aligned with value-guided eye movements

- Fast caudate signals were independent of vacillating value representations in OFC
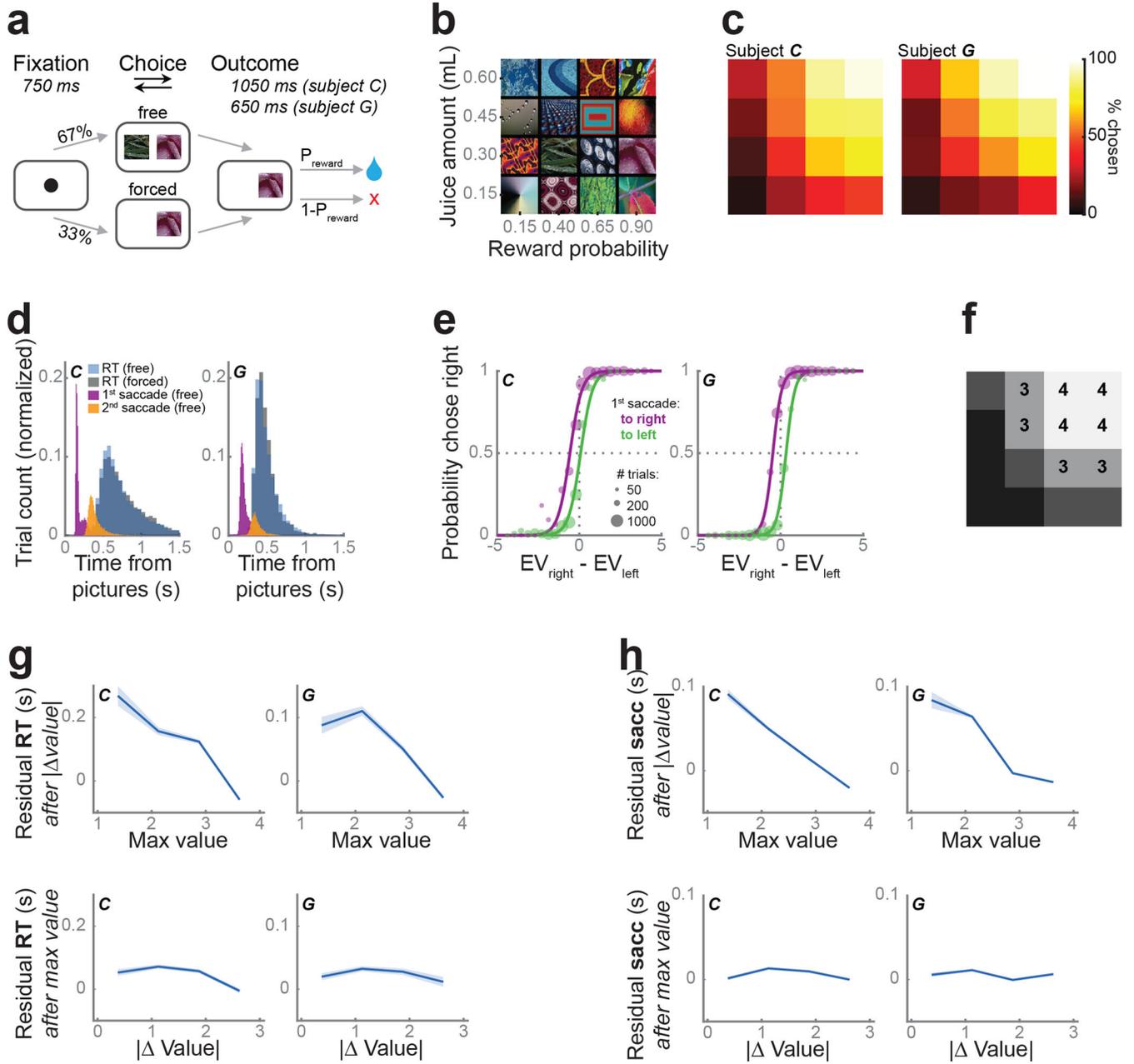
Figure 1.

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).
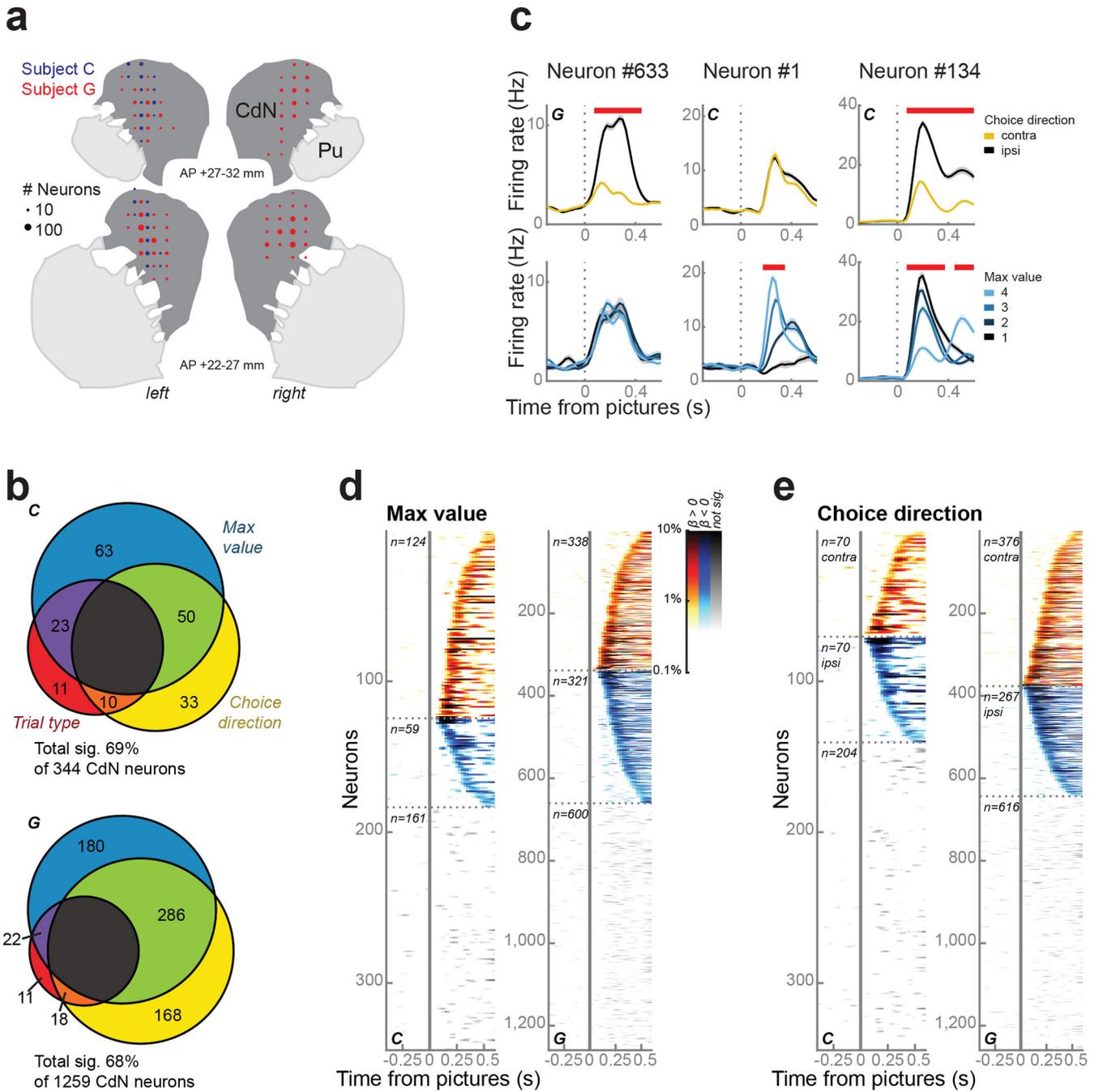
Figure 2.
CdN neurons.

(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) Three example neurons illustrating prevalent response types, including significant encoding of choice direction (left column), maximum value (middle), or both (right). Panels within each column show firing rates (mean ± s.e.m.) averaged over choice direction (top row) 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 data from a single neuron. Color intensity indicates the percentage of variance in the neuron's firing rate explained by maximum value as determined by the coefficient of partial determination. 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).
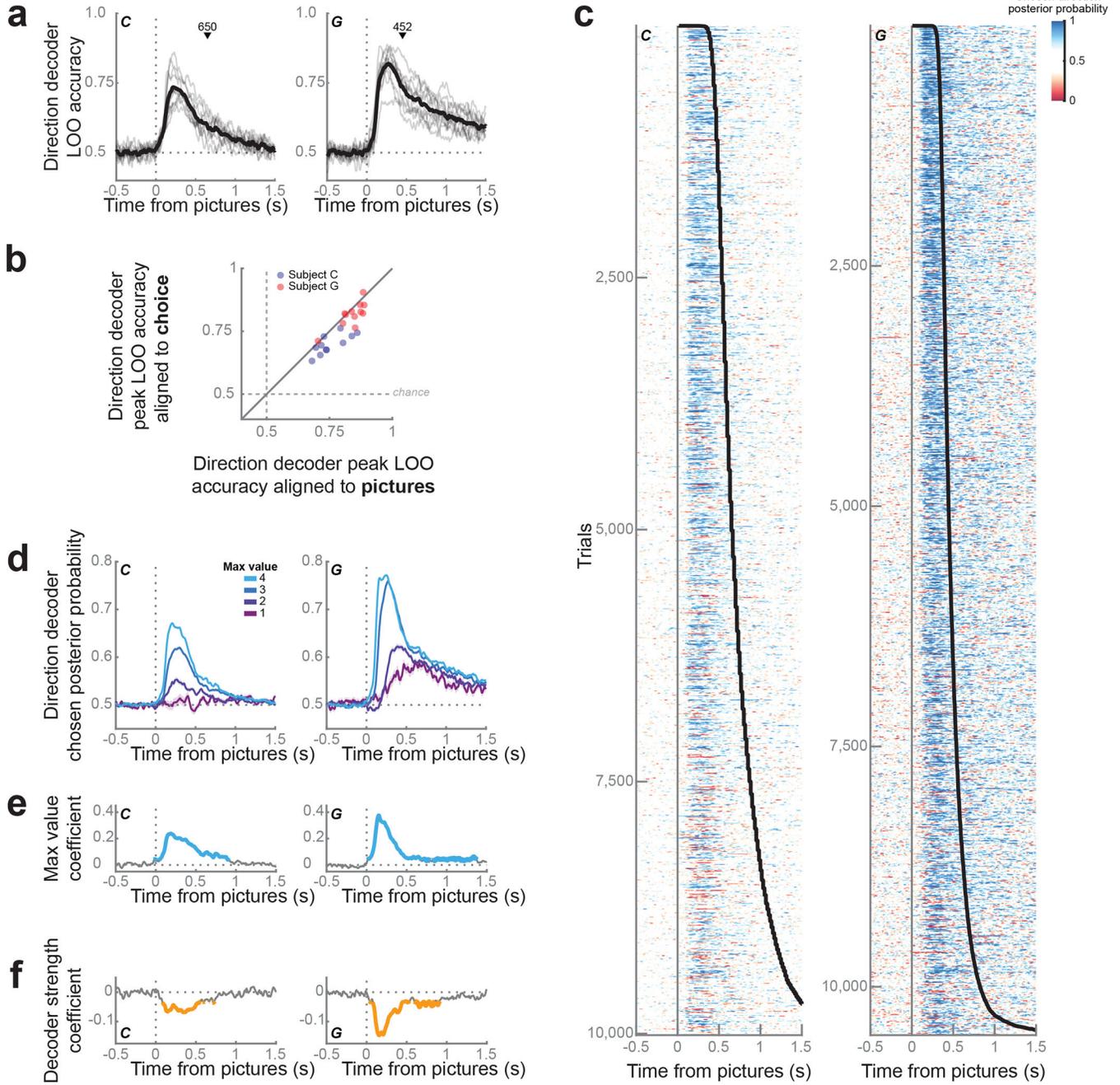
Figure 3.
Direction decoding from CdN population.

(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 choice 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 ± 1%, paired $t(10) = 4.7$, $p < 0.001$; G: 80 ± 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 ± 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, $p < 0.001$.
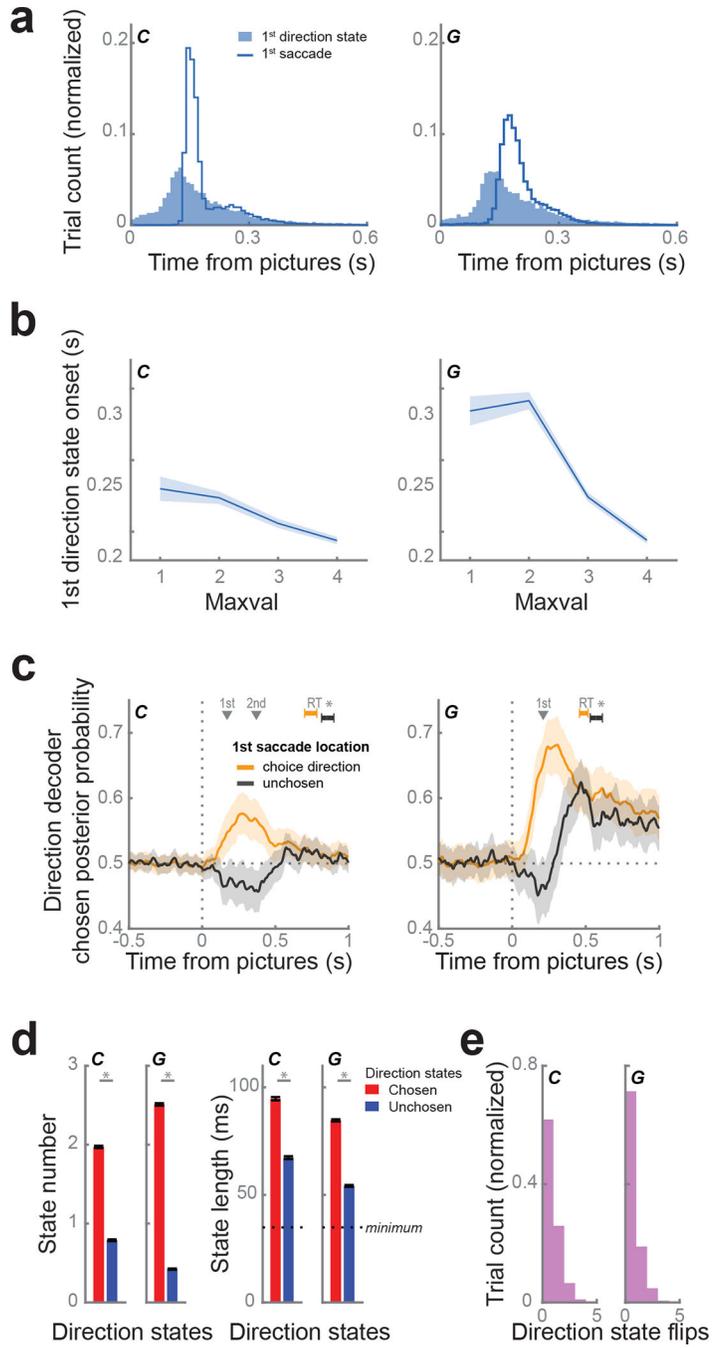
Figure 4.
CdN population tracks first saccade.

(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 (subject 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.
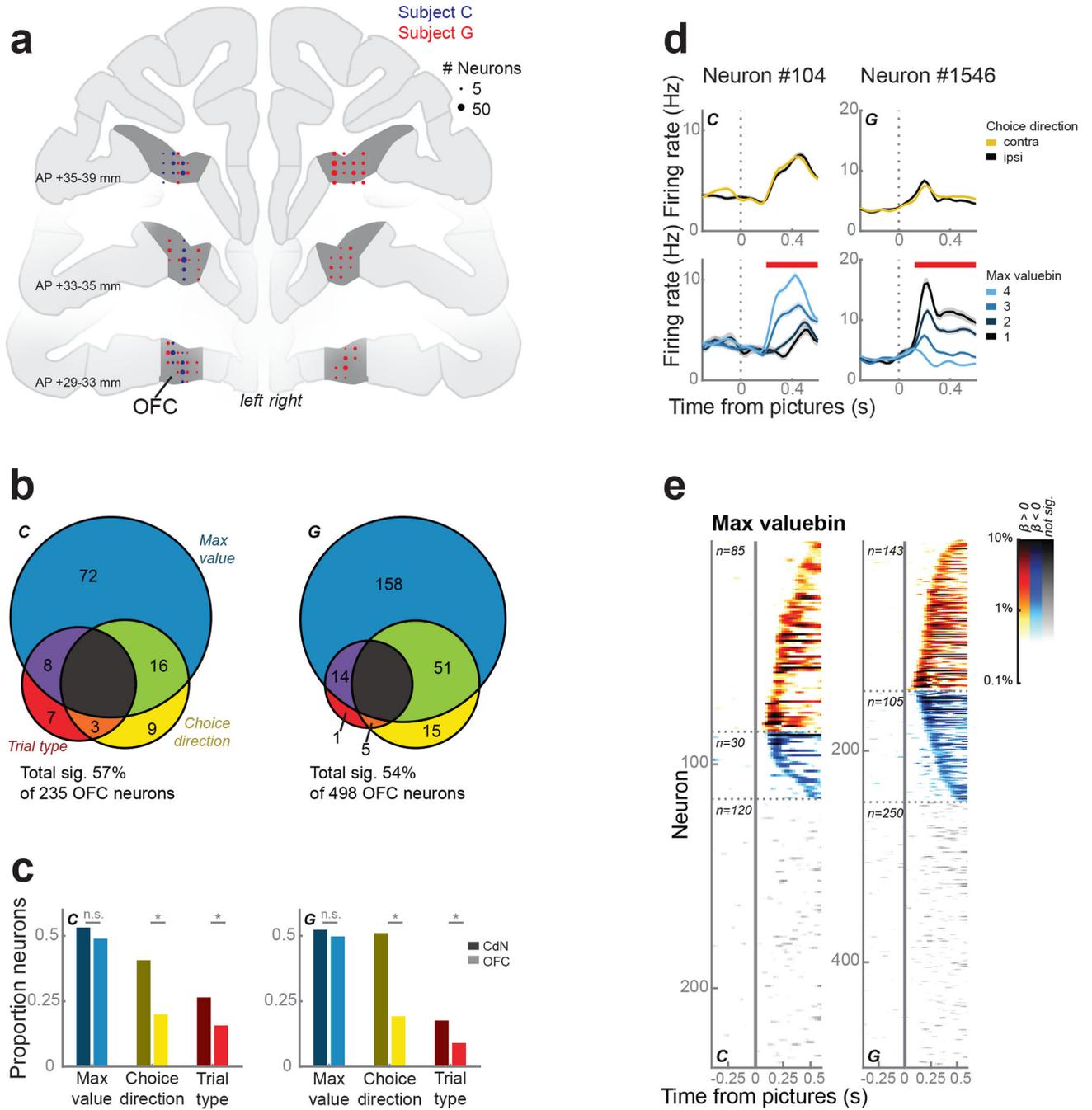
Figure 5.

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. 2b and 5b). 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. 2c.

(e) Single neuron encoding of maximum value across the entire OFC population. Conventions as in Fig. 2d.
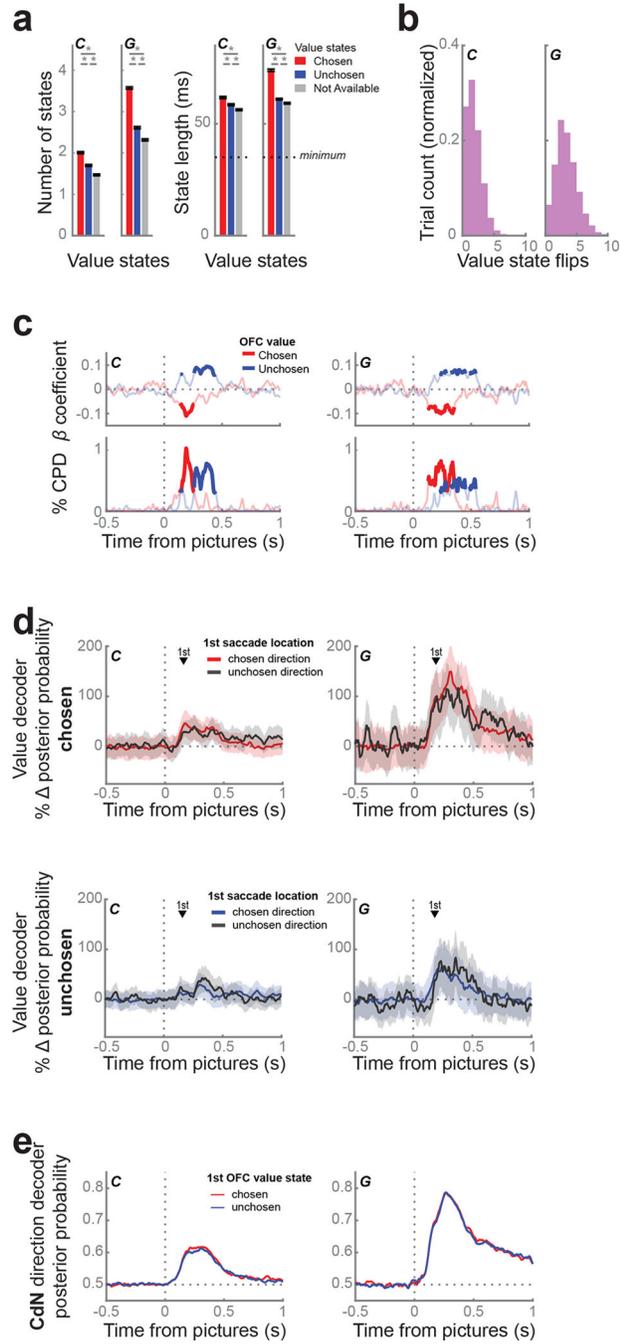
Figure 6.

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 4c. 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 ± s.e.m.) of the CdN direction decoder between trials with an initial chosen (red) or unchosen (blue) value state in OFC.
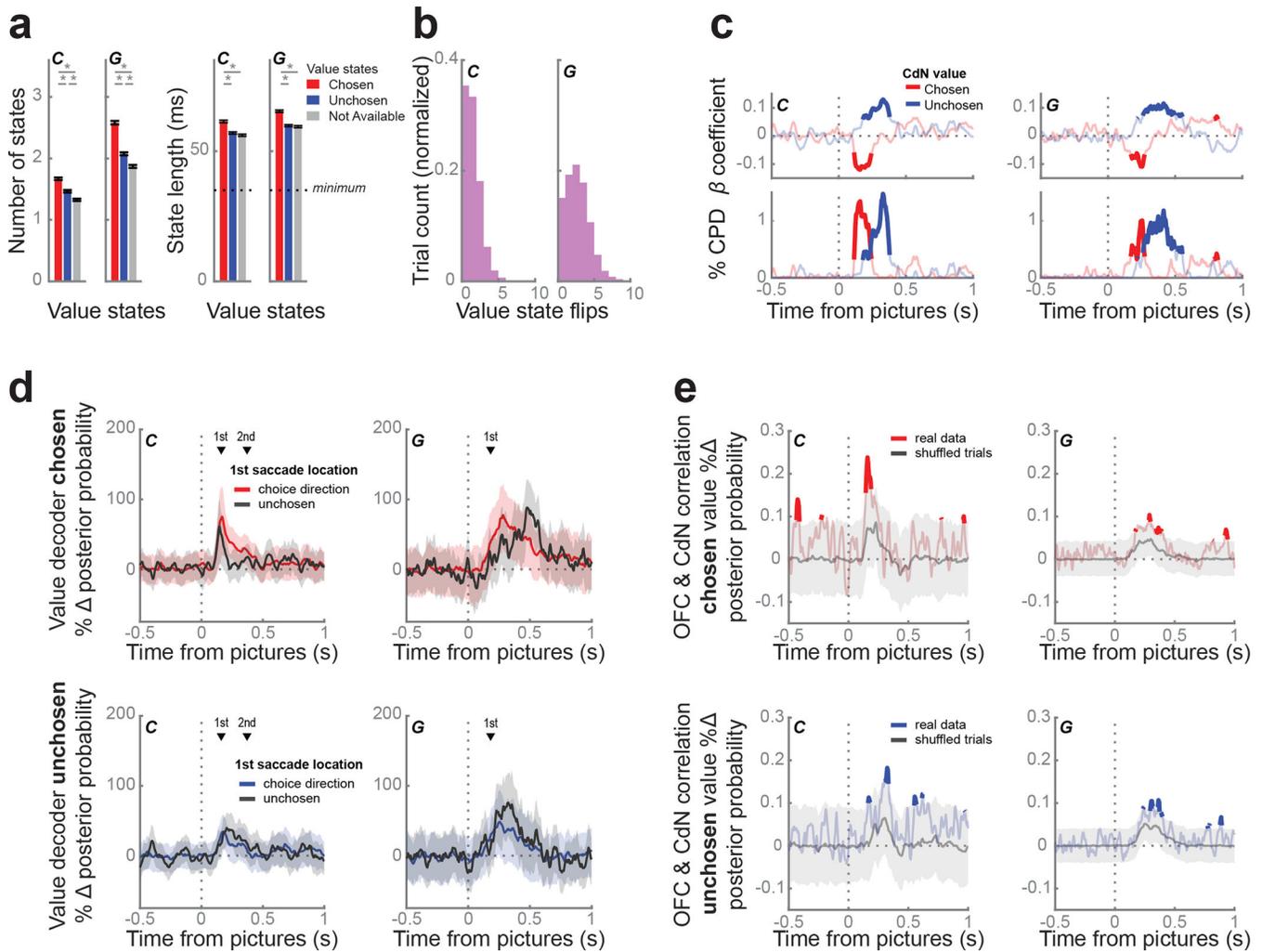
Figure 7.

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.

Key Resources table:

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental models: Organisms/strains		
Non-human primate (*macaca mulatta*)	UC Davis	N/A
PEEK inserts	Graymatter Research	https://www.graymatter-research.com
Form2 and Form3 3D printers	Formlabs	https://formlabs.com
Plexon V-Probes	Plexon	https://plexon.com
Software and algorithms		
MATLAB	The Mathworks	https://www.mathworks.com/products/matlab.html
Omniplex	Plexon	https://plexon.com
Offline Sorter	Plexon	https://plexon.com
Monkey Logic Toolbox for MATLAB	NIH	https://www.brown.edu/Research/monkeylogic
3D Slicer	BWH	https://www.slicer.org
Eyelink 1000	Eyelink	https://www.sr-research.com/eyelink-1000-plus/
Fast smoothing function for MATLAB	Tom O'Haver	https://www.mathworks.com/matlabcentral/fileexchange/19998-fast-smoothing-function
Statistics and Machine Learning Toolbox for MATLAB	The Mathworks	https://www.mathworks.com/products/statistics.html
Original code	This paper	https://doi.org/10.5281/zenodo.6408604