Neuronal representation of the decisional reference point in monkeys

Duc Nguyen[1], Erin L. Rich[3], Joni D. Wallis[4], Kenway Louie[1,2] & Paul W. Glimcher[2,1]

[1] Center for Neural Science, New York University.

[2] School of Medicine, New York University.

[3] Icahn School of Medicine, Mount Sinai.

[4] Department of Neuroscience, University of California at Berkeley.

Corresponding at dmn9844@nyu.edu

ABSTRACT:

The reward reference point serves as a hidden benchmark for evaluating options in decision-making. Despite extensive behavioral evidence for reference-dependence, no neural representation of the reference point has been discovered. We analyzed single neuron activity from macaque monkeys performing a decision-making task designed to orthogonalize objective reward values from the reference point. Regression analyses of neuronal activity across six frontal brain regions identified a robust neural representation of the reference point in the ventral bank of anterior cingulate cortex (vbACC). Activity in the dorsal bank of anterior cingulate cortex and the dorsolateral prefrontal cortex, in contrast, encoded the reference-dependent subjective values of the rewards offered or obtained on each trial. The temporal dynamics of these signals and connections between these regions suggest a dedicated neural circuit implementing reference-dependent reward encoding, with the vbACC serving as the reference point signal source modulating activity in other frontal value-encoding areas.

Keywords: reward, expectation, reference point, reference-dependent, frontal cortex, decision making

INTRODUCTION

All values are inherently subjective. Consider twins who each receive a $10,000 salary raise and differ only in their expectations. The first one, expecting only a $5,000 raise, is overjoyed. The second, expecting a $20,000 raise, is disappointed by the same amount and quits her job. In modern mathematical and behavioral studies of decision-making, this expectation is one instantiation of the *reference point* - a hidden benchmark against which individuals evaluate the subjective values of real-world outcomes.

The introduction of the reference point into choice theories marked a crucial advance. Before the 1970s, decision-making theories centered around expected utility, which employed a subjective notion of value (utility) that was not context or expectation dependent [1-3]. However, evidence from the 1960's and 1970's suggested that the subjective values of rewards were indeed dependent on context, or the expectations that derive from it [4, 5]. Inspired from these findings, in 1979, Kahneman and Tversky published their influential descriptive theory of decision-making (Prospect Theory), the central novelty of which was a reference-point-dependent subjective value, successfully explaining behaviors that existing neoclassical theories could not [6-8]. Specifically, the theory proposed that all objective rewards are converted to subjective values by comparison to this fluctuating hidden benchmark: at low reference point, a small reward is considered desirable; but at high reference point, the same reward becomes less satisfying or even undesirable. Subsequent studies have shown that nearly all aspects of decision-making show a dependence on the reference point [8-12], a finding compatible with perceptual psychology results [13-15]. Consequently, nearly all modern choice theories now incorporate a reference point of some kind, albeit with varying specifications [16-22].

Inspired by these insights, neurobiological studies have recently shown evidence of various kinds of reference-dependence in the neural representation of rewards, gains, and losses [23, 24]. Reference-dependent adaptations of neural signals to reward history [25, 26] and adaptive representations of value have been observed in various brain regions, including the dopaminergic system [27-29], orbitofrontal cortex [30, 31], and anterior cingulate cortex [32, 33]. Collectively, these findings provide clear evidence that reference points, or expectations, strongly influence the neural circuits associated with reward value coding.

Despite extensive behavioral and neural evidence for reference dependence, the neuronal origin of the reference point remains largely unidentified. There are hints of a reference-point-related signal in the cingulate cortex which have been observed during foraging tasks in monkeys [34], rodents [35], and humans [36, 37], but the direct observation of a neural signal encoding the reference point itself has remained elusive. In this study, we aimed to (1) identify where in the brain the reference point signal is explicitly and homogenously represented in neural activity, and (2) identify areas where this reference point signal modulates value-related representations. To achieve this, we reanalyzed data from previously published neuronal recordings by Chien, Wallis, and

Rich (2023) [38]. In their study, macaque monkeys performed a task in which they could either earn rewards or avoid losses. The unique design of this task allowed us to specifically separate reward expectations on each trial (the reference point by definition) from the values of rewards on each trial.

Our analysis revealed three key findings. First, neurons on the ventral bank of the anterior cingulate cortex (vbACC), below the cingulate sulcus, appear to encode the reference point homogenously and stably throughout the decision-making process. Second, neurons in the dorsal bank of the anterior cingulate cortex (dbACC) encoded a cue valence (potential reward) signal modulated by the reference point. Third, neurons in the dorsolateral prefrontal cortex (dlPFC) encoded an outcome (earned reward) signal, also modulated by the reference point. These effects followed an expected pattern: as the reference point increased, the neural representation of the subjective value signals in the dbACC and the dlPFC decreased – as would be expected from behavioral data. Interestingly, the progression of these signals from reference point to the subjective evaluation of earned rewards mirrored an anatomical transition from ventral to dorsal regions. These findings suggest a neurobiological substrate for the influence of the reference point on decision-making.

RESULTS

To investigate neural signals related to the reference point, we reanalyzed data from previously published studies [38, 39]. Earlier analyses demonstrated that multiple frontal brain regions in monkeys encode various decision variables during decision-making tasks. However, the data had not been specifically examined with respect to the expectation of future reward, which following Kahneman and Tversky we take to be the formal definition of the reference point.

Two subjects (both male *Macaca mulatta*, age 6 and 10) were trained to perform an instructed response task. Each recording session consisted of several hundred consecutive trials. Briefly after the start of each trial, a fixation point appeared, requiring the monkeys to fixate at the center of the screen for 650ms. Following fixation, a visual cue appeared, indicating whether the monkey stood to gain or lose a reward in this trial and prompted the subjects to move a joystick either left or right (as indicated by the visual cue) in order to achieve the best possible outcome on that trial (either gaining reward or not losing reward, depending on the stimulus). Finally, depending on cue valence and the monkey's response, feedback was provided showing whether the reward amount had increased, decreased, or remained unchanged. Feedback was represented by the length of the expected-reward bar that remained visible at the bottom of the task screen throughout the trial. Its length indicated the cumulative amount of reward subjects earned at that time point and served as a reference point against which individual trial outcomes could be evaluated (Fig. 1A).

A total of 4 trial-type cues were used in this task, differentiated by correct response direction (left/ right) and valence (positive/ negative) in a 2×2 design. Positive

cues (one left, one right) increased the reward if the response was correct and gave no reward if incorrect. Negative cues (one left, one right) gave no reward if correct and reduced the reward if incorrect (Fig. 1B).

Subjects completed six sequential trials in each block before receiving their total accumulated juice reward, an amount proportional to the length of the horizontal expected-reward bar displayed at the end of the sixth trial (Fig. 1C). This expected-reward bar, representing the subjects' current endowment, served as a reference point against which individual trial outcomes could be evaluated. Based on the distribution of reference points, we either performed a parametric regression on length or grouped trials across all sessions into 3 groups for visualization: low, medium and high reference levels (Fig. 1D). We recorded single neuron activities from six frontal cortex. Of these, neurons who had sufficient activities for statistical analyses and brain origin confidently identified were used for further analysis (Fig. 1E and Methods).

Behavioral performance is reference-dependent

Behavior in this task has been previously described [38, 39]. Overall, both subjects performed well, with subject M making fewer errors. Response times varied across trials (Fig. 2A), which guided our selection of the time windows for our neural analyses (see Methods).

Consistent with previous studies [38, 39], we found that the monkeys' behavioral performance exhibited reference dependence. We fitted a multivariable regression model for reaction time and a multivariable logistic regression model for percent correct. Our analysis showed that all factors we examined significantly influenced both reaction time and probability of a correct response (Table 1). Particularly, the length of the expected-reward bar (the reference point signal) and cue valence were the two most potent modulators of reaction time and the probability of a correct response. Higher reference levels were linked to better performance, including higher probability of a correct response and lower reaction times (Fig. 2B-C). Importantly, these results capture the key characteristics of reference-dependent behaviors in animals [40, 41].

The ventral bank of the anterior cingulate cortex encodes the reference point signal

To examine the effects of the reference point on neural activity, we performed a series of regression analyses to model trial-by-trial firing rates (see Fig. 3A and Methods). We analyzed neural activity within a 500ms time window (101-600ms after an event) across three key epochs: trial start, cue onset, and feedback. Each successive epoch contained additional information and reflected the dynamic evolution of the decision-making process.

To examine the effect of the reference point in isolation, we first focused on the trial start epoch, where the expected-reward bar length was the only available information (Fig. 3B). During this epoch, subjects had no knowledge of the impending

trial type or the outcome of their decision, ensuring that the signal was not confounded by other task-related factors.

A univariate regression analysis revealed that the reference point (expected-reward bar length) was consistently encoded in the ventral bank of the anterior cingulate cortex (vbACC). This conclusion was supported by two analyses of the neural population (n = 156). First, the average regression coefficient of the reference point across all neurons significantly deviated from zero (t-test, $p = 0.0057$), indicating a consistent impact of the reference point on this area (Fig. 3C). Second, we examined the percentage of the population with statistically significant coefficients for the reference point. Among these, we found a significantly higher proportion of neurons with negative versus positive encoding of the reference point (χ^2, $p < 0.001$), further supporting the homogenous impact of reference point on this brain region (Fig. 3D). Our findings were robust across four different measures of firing rate (see Fig. S1A and Methods). The vbACC was the only area in this dataset to homogenously encode the reference point, as shown by the biased distribution of coefficient values (Fig. S1B) and the proportion of neurons with significant coefficients (Fig. S1C).

For each block, the monkeys' goal was to increase the length of the expected-reward bar across trials. This raised the possibility of collinearity between the expected-reward bar length and trial number within a block. To address this, we conducted two additional tests. First, we calculated the variance-inflation-factor (VIF) and found the VIF to be 1.0575. This magnitude of collinearity is typically considered to be low [42], reflecting the fact that expected-reward bar length could be increased or decreased across trials. Second, we reanalyzed the effect of reference point on the neural activities after regressing out the effect of the trial-number (see Methods) and confirmed the significant and consistent impact of the reference point on vbACC activity (Fig. S2).

Overall, our analyses indicate that the neural population in vbACC homogenously encodes a representation of the reference point throughout the trial. This can be seen in population average firing rates, as illustrated by population dynamics separated by different reference levels (Fig. 3E). These firing rate plots show that the impact of the reference point remains significant until the end of the trial, as shown by a repeated-measures ANOVA (Fig. S3A-B). Univariate regression analyses confirm that this homogeneous encoding persists across most epochs, except during the phasic transient elicited by cue onset (Fig. S3C). Together, these findings suggest that the vbACC is a region that encodes the reference point.

The dorsal bank of the anterior cingulate cortex encodes a reference-dependent cue valence signal

Next, we focused on the cue onset epoch, during which subjects were presented with new information, consisting of a visual cue that indicates a potential future gain or loss. Importantly, during this epoch the expected-reward bar remained visible (Fig. 4A).

Therefore, in this epoch we conducted a multivariable regression analysis to assess how the reference point influenced the perception of cue valence (see Methods).

We found that the dorsal bank of the anterior cingulate cortex (dbACC) exhibited homogeneous encoding of the reference-dependent cue valence signal (as measured by the interaction term between reference-point (expected-reward bar length) and cue-valence; see Methods). This conclusion was supported by two analyses of the neural population (n = 154). First, the average regression coefficient for the reference-dependent cue valence signal significantly deviated from zero (t-test, $p = 0.0011$), indicating a consistent impact of the reference point on cue valence-related firing rates (Fig. 4B). Second, within the subpopulation with significant coefficients, we found an uneven distribution of neurons encoding this signal, such that neuron in dbACC tended to have negative coefficients (χ^2, $p = 0.0269$), further supporting the homogeneity of this representation in the population (Fig. 4C). Finally, these findings were robust regardless of the firing rate measure used in the analyses (Fig. S4).

Notably, we did not observe homogeneous encoding of cue valence itself in the dbACC. Instead, neurons encoded cue valence with both positive and negative coefficients, and we only detected a bias in the reference-modulated effect. Although the response instruction signal (cue-direction) was also introduced in this epoch, no brain region from which we have recordings showed uniform encoding or a reference-modulated effect in response to it. This is an interesting effect, as the response direction was instructed by the same cue that indicated potential outcome valence, but only the latter, the information about potential value, was represented in relation to the current expected-reward bar length. Overall, the reference-dependent cue valence signal was the most significant and robust effect observed in this epoch compared to other effects (Fig. S5).

The impact of the reference point on a predictive reward signal as the decision problem unfolded is revealed by the neural dynamics of dbACC activity. In this area, the population activity reflected different reference levels in the cue onset epoch but not during the preceding epochs (Fig. 4D). When further separating these responses into different cue valence trial types, we observed a difference in dbACC population activities responding to positive vs. negative cues at low reference levels. However, consistent with the significant interaction in our regression analyses, this valence difference diminished as reference level increased, as visualized by the reference-dependent cue valence signal impact on the dbACC population (Fig. 4E).

Dorsolateral prefrontal cortex encodes a reference-dependent outcome signal

Lastly, we investigated reference-related signals during the feedback epoch. During this epoch, subjects were informed whether their cumulative earned reward changed or remained the same based on their action (Fig. 5A). We used a multivariable regression analysis to evaluate the impact of the reference point on population neural activities during this interval.

After controlling for all available information (a total of 15 regressors were used in the model for this epoch, see Methods), we identified reference-dependent feedback encoded in the dorsolateral prefrontal cortex (dlPFC). We found that the reference point consistently influenced encoding of the outcome (earned reward) in dlPFC activity, measured by the interaction term between reference-point and outcome. This conclusion was supported by two analyses of the neural populations (n = 159). First, the average regression coefficient for the reference-dependent outcome value significantly deviated from zero (t-test, p = 0.0182), indicating a consistent impact of the reference point on outcome encoding (Fig. 5B). Second, when we examined the subpopulation with significant coefficients, there were more positively encoding neurons than negatively encoding (χ^2, p = 0. 0474), further supporting the homogeneity of the representation in this population (Fig. 5C). These findings were robust regardless of the firing rate metric used in the analyses (Fig. S6).

We also examined the effects of other task variables on neural activity during the feedback epoch. The reference-dependent outcome signal was one of the most significant and robust effects observed (Fig. S7). No other task variables modeled by our regression showed a reference point influence in this epoch.

In our analysis of average firing rates, we observed that the dlPFC population reflected reference levels only during the feedback epoch (Fig. 5D). When analyzing responses to specific feedback types, we found that the dlPFC strongly responded to negative outcome (loss of earned reward), particularly during the feedback epoch, but similar to reference-dependent cue valence in dbACC, this response diminished as the reference level increased (Fig. 5E-F).

Overall, as the decision-making trial progressed, we observed a dynamic evolution of the reference point signal across different monkey frontal brain regions. Initially, the vbACC appeared to track the reference point. Next, the dbACC responded to cue valence modulated by this reference point. Finally, the dlPFC responded to the outcome, also modulated by the reference point (Fig. 6A).

DISCUSSION

Summary

Extensive behavioral evidence has demonstrated that human and animal decision-making is profoundly influenced by individual expectations. These expectations, functioning as reference points in the valuation process, also influence neural representations as seen in electrophysiological studies in humans [43] and animals [24, 44], and in functional imaging studies in humans [23, 45]. However, a neural representation of the reference point itself has not yet been identified, nor has functional imaging localized it to specific brain areas. In this study, we identified a neuronal population in the vbACC that appears to specifically encode current endowments that set the expectation of future rewards, an instantiation of the behavioral economic reference point. Moreover, we found that during this task the

reference point modulates two other decision-related reward signals: one in the dbACC, encoding cue valence (potential gain or loss), and another in the dlPFC, encoding the outcome (amount of reward gained or lost). As the reference point value increased, the reward-related neural activity in these two areas decreased, a pattern consistent with the well-known behavioral phenomenon of diminishing marginal value relative to expectation [46]. Together, these results demonstrate the existence of a neural representation of the reference point during this task and highlight its broad influence on the neural circuitry involved in decision-making.

Homogenous encoding

Based on behavioral results and reference-dependent behavioral economic theories, the reference point is expected to have a consistent negative relationship to subjective value encoding [16, 21]. We hypothesized that, given this unform effect on downstream value computations, a neural circuit representing reference point information would show a consistent, homogenous population-level modulation.

In our study, we indeed observed systematic biases in representation: the vbACC is negatively correlated with the reference point, and the dbACC, and dlPFC consistently encoded reward signals suppressed by the reference point. This biased modulation contrasts with scenarios where a neural population exhibits both positive and negative modulation by a decision variable, but no overall modulation on average. Given the predominance of mixed representations in frontal cortical brain areas [47, 48], the relative homogeneity of reference point and reference-dependent signals suggest those may play a particularly important role in downstream computations.

A proposed neural circuity for reference-dependence

Our findings are compatible with the existence of a dedicated neural circuit for encoding the reference point and its influence. It had been shown that the vbACC projects primarily to the dbACC and dlPFC [49], and both the vbACC and dbACC (*largely* Brodmann area 24) project to the dlPFC [50, 51]. Given these specific projections, along with our findings of reference-dependent signals in the dbACC and dlPFC, it is plausible that these areas could form a representational circuit for subjective reward value. Based on these data, we hypothesize that the vbACC serves as a source of the reference point signal, modulating value encoding in other brain regions, dbACC and dlPFC for example (Fig. 6B).

A key question is how the reference point signal originates in the vbACC. The vbACC receives inputs primarily from amygdala [52, 53], parahippocampal and entorhinal cortices [54], frontal eye field/ premotor area [51, 55], and anterior insula [56]. These connections are known to carry emotion-, reward-, and sensory-related signals, which together could contribute to forming expectations or reference points. Future research will be necessary to investigate how these signals integrate to establish the reference point within the vbACC.

Previous neural evidence of reference point and/or expectation in the ACC

In behavioral ecology, the marginal value theorem (MVT) posits that foraging animals should leave a patch when the marginal rate of resource intake equals the average rate in the environment [57]. This requires animals to track their expectation of environmental availability, a quantity equivalent to the reference point [58]. Studies have shown that animals and humans possess brain subpopulations encoding this global reward rate, akin to the reference point in MVT theory. Specifically, in foraging tasks the ACC encodes a decision variable that signals when to leave the patch in monkeys [34], rodents [35], and humans [36, 37], aligning with our findings on the neuronal origin of the reference point. Consistent with our findings, Hayden and colleagues found that the monkey ACC (including data from regions we identify separately as dbACC and vbACC) encodes a signal inversely correlated with this foraging reference point. Thus, vbACC activity may represent a general reference point signal across different types of tasks.

The ACC has long been recognized as a crucial component in decision-making, particularly in relation to expectation. It has been shown that the ACC is involved in expectation-related functions, such as encoding expected outcomes [59-64] and being modulated by reward expectations and history [32, 65-67]. These findings provide additional support for the neural representation of the reference point within the ACC. The variability in findings regarding reference-point (expectation) itself versus reference-dependent encoding of other information may stem from a lack of differentiation between the dorsal and ventral banks of the ACC (above and below the sulcus), which have distinct anatomical structures and functions [68].

Differences between Ventral and Dorsal bank of the ACC

Few studies have distinguished between the dorsal and ventral banks of the ACC [33, 38, 69]. Interestingly, Amemori and Graybiel showed that both the vbACC and dbACC encode subjective values during decision-making, with the vbACC encoding these values more uniformly and exhibiting a stronger negative bias than the dbACC, consistent with our observations. Importantly, they demonstrated that microstimulation of the vbACC, and not dbACC, “increased negative decision-making”, as shown by reduced tolerance toward negative stimuli during an approach-avoidance task. In the context of our study, one might hypothesize that stimulating the vbACC is analogous to lowering the reference point, leading to change in subjective value and behaviors. Excitingly, stimulating the vbACC indeed changed the choice behavior. However, due to the complexity of their task, which required weighing both positive and negative information for each option, and the absence of a derived utility function, it remains unclear whether this behavioral change was specifically due to a shift in the reference point as we would hypothesize. Nonetheless, the distinct effects of vbACC versus dbACC stimulation raise specific questions for future research about the unique role of the vbACC in decision-making, specifically encoding the reference point.

Potential different neural mechanisms of reference-dependence

Neural adaptation broadly describes changes in neural circuits that reflect the statistical properties of the environment. In sensory systems, studies have shown that these circuits adapt to statistical characteristics such as the mean, range, and variance of stimuli [70-72]. A similar mechanism has been proposed for value encoding systems [73, 74]. There is substantial evidence supporting this idea in the dopaminergic system, where dopamine neurons encode updates to learned expectations, demonstrating a history-as-reference dependence [27, 28, 75, 76]. Recent work has also demonstrated range-adaptation in other value encoding networks such as the OFC in primates [30, 31]; the related area, vmPFC, in human [45, 76]; and the ACC in both human and primates [33, 45].

Interestingly, while both the OFC and ACC undergo range adaptation, our results showed that only the ACC, and not the OFC, reflected a reference-point signal in our task. This suggests two possible neural mechanisms may underlie reference-dependent signals: distributed (intracortical) versus centralized (intercortical) representations of the reference point. In the distributed model, individual brain regions adjust their baselines or tuning curves according to the distribution of values, aligning with adaptation. In the centralized model, certain regions specialize in tracking the reference point and transmit these signals to other value-encoding areas. Consequently, these regions integrate both the reference point and option values during decision-making. It is possible that our study observed the latter mechanism of the reference-dependence, while the former have been demonstrated through adaptation in other areas, such as the dopaminergic systems and the OFC. Future research is needed to further differentiate these two types of context modulation and identify which brain regions are influenced by them.

To summarize, the decisional reference point is a central concept in choice theories, yet its neural instantiation has remained unclear. Our study provides the first direct evidence of this signal in the neural populations of the vbACC. The observed homogenous encoding in this region, combined with its connectivity to other brain areas, offers critical insights into a potential neural circuit that underlies reference dependence. These findings lay the groundwork for future animal research aimed at establishing the causal role of the reference point through stimulation and pharmacological inhibition. Additionally, the homogenous encoding observed in the vbACC makes it possible to apply our findings to human brain studies using fMRI, providing a useful starting point for further research.

ACKNOWLEDGEMENTS

We thank all colleagues in the Glimcher lab for their support and helpful comments. We especially thank Drs. Erin L. Rich and Joni D. Wallis for sharing the data and for constructive comments on the project as well as the manuscript. This work was supported by the NIMH T32 Training program to D.N., National Institutes of Health - Office of Extramural Research Grants R01DA19028 and P01NS040813 to J.D.W., R01MH121480 and a Hilda and Preston Davis Foundation grant to E.L.R., and NIH U19 NS107616 to K.L..

AUTHOR CONTRIBUTIONS

Conceptualization - ELR, JDW, and PWG

Data curation - ELR, and JDW

Formal analysis - DN, KL, and PWG

Funding acquisition - JDW, and PWG

Investigation - ELR, and JDW

Methodology - ELR, and JDW

Project administration - ELR, JDW, KL, and PWG

Resources - ELR, and JDW

Software - DN, and ELR

Supervision - ELR, KL, and PWG

Validation - DN, ELR, JDW, KL, and PWG

Visualization - DN

Writing – original draft - DN, KL, and PWG

Writing – review & editing – DN, ELR, JDW, KL, and PWG

DECLARATION OF INTERESTS

The authors declare no competing interests.

METHODS

Subjects and Behavioral Task

The subjects were two male rhesus macaques (Macaca mulatta), named M and C, aged 6 and 10 years old, and weighing 11 and 14.5 kg at the time of recording, respectively. The monkeys were trained to perform an instructed response task, where they moved a bidirectional joystick left or right in response to visual cues displayed at the center of a computer screen. The task involved four different cues, differing by correct response direction (left/right) and valence (positive/negative) in a 2×2 design. Two cues (one left and one right) could lead to an increased earned reward if the response was correct, and no change if incorrect (positive valence). The other two cues (one left and one right) resulted in a reduced earned reward if the response was incorrect, and no change if correct (negative valence) (Fig. 1B). Each cue represented a unique combination of valence and response instruction. Changes in earned rewards were reflected by adjustments in the length of the expected-reward bar displayed on the screen. The magnitude of the increases and decreases was equal and fixed across trials. To ensure that sufficient trials resulted in negative feedback (equivalent to error response), 15% of positive cue trials ended in no change (the worse outcome) even if the monkey made a correct response, and 15% of negative cue trials ended in decreased earned reward.

The subjects completed six trials within each block before accumulated reward was delivered (Fig. 1C). This number of trials was calibrated to maintain the subjects' motivation throughout the block while maximizing variation in the expected-reward bar length. Rewards were fruit juice, and the amount directly proportional to the length of the expected-reward bar at the end of the trial block. A trial was considered complete when the subject made a joystick response to the visual stimulus, regardless of the outcome. Incomplete trials, where the monkey failed to make a joystick response or failed to successfully hold initial fixation, did not count toward block completion. After juice delivery, the expected-reward bar length was reset to a value of 2 at the start of the next block.

Behavioral contingencies were implemented using MonkeyLogic software [77], and eye movements were tracked with an infrared system (iSCAN). Each subject was surgically implanted with two recording chambers and a titanium head post to stabilize the head position during recording. The precise implantation sites were determined based on stereotaxic coordinates of target areas identified on 1.5T MRI scans of each subject's brain. All procedures were carried out following the guidelines of the National Institutes of Health Guide for the Care and Use of Laboratory Animals and the recommendations of the University of California at Berkeley Animal Care and Use Committee.

Each recording session consisted of hundreds of consecutive trials (average 646.90 ± 12.43). At the beginning of each trial, only the expected-reward bar that

445 indicated the accumulated reward was visible, followed by a fixation point at the center of the screen. Subjects were required to hold fixation within a 1.3° radius of the fixation point for 650 ms to initiate the trial. After fixation, one of four visual cues, which were images of natural scenes approximately 2°×3° in size, was presented. Subjects responded by moving a joystick fixed to the front of their chair either left or right. To
450 discourage random responses, any response made within 150 ms of stimulus onset triggered a 5-second timeout. After the response, feedback was provided through changes in the length of the expected-reward bar (Fig. 1A).

In this setup, the expected-reward bar remained visible throughout each trial and its length carried over from one trial to the next within a block, so that the expected-
455 reward bar consistently indicated the total amount of reward the subjects had accumulated and served as the reference point in classic economic decision-making. This design ensured that subjects were continuously aware of their current earned reward. Additionally, by fixing the amounts of reward increases and decreases, we controlled the magnitude of gains and losses, enabling us to examine the effect of
460 reference levels without confounding subjective perceptions of value.

Neural recording and general processing

Neural recordings were acquired using a Plexon MAP system, with 4 to 20 tungsten microelectrodes (FHC Instruments), independently positioned in the targeted brain regions during each session (see Fig. 1E). The targeted brain regions were
465 broadly labeled as follows: vbACC (areas 24a & 24b), dbACC (areas 24b & 24c), mOFC (area 14), OFC (area 13), dlPFC (areas 9, 46, & 9/46), and vlPFC (area 47/12). Precisely, our dbACC (vbACC) target was the dorsal (ventral) bank of the cingulate sulcus within the rostral (pre and perigenual) ACC. Areal boundaries were defined by sulci visible on MRIs, as shown in Fig 1E. During recording, electrodes were placed in
470 cortical layers by acoustic mapping of gray-white matter boundaries of each electrode path. After recording, final placement was determined on MR images for each subject using stereotaxic coordinates and depth of each electrode. If a neuron could not be confidently assigned to one of the anatomical locations, it was excluded from further analysis.

475 All well-isolated units were recorded, with no pre-screening. The recorded waveforms were digitized, and single units were identified using Offline Sorter (Plexon). After spike sorting, we excluded units with an average firing rate below 1 Hz across the entire session, as these neurons did not show sufficient activity for reliable statistical analysis.

480 **Behavioral analyses**

Behavioral performance, including percent correct and reaction time, were analyzed using a logistic multivariable regression and a general multivariable regression model, respectively. Both models included the same six regressors and an intercept term. The regressors were: (1) *refPoint*: reference point (length of the expected-reward

bar at the beginning of each trial); (2) cue valence as an indicator variable (1 for positive, 0 for negative); (3) cue direction as an indicator variable (1 for left, 0 for right response); (4) trial number within a block; (5) trial number within a session; and (6) subject ID as an indicator variable (1 for subject C, 0 for subject M):

For reaction time:

$$\begin{aligned}\text{ReactionTime} = {} & \beta_0 + \beta_1\text{refPoint} + \beta_2\text{cueValence} + \beta_3\text{cueDirection} \\ & + \beta_4\text{trialWithinBlock} + \beta_5\text{trialWithinSession} + \beta_6\text{subjectID} + \varepsilon\end{aligned}$$

For percent correct:

$$\text{PC} = \frac{1}{1 + e^{-t}}$$

where $t = \beta_0 + \beta_1\text{refPoint} + \beta_2\text{cueValence} + \beta_3\text{cueDirection} + \beta_4\text{trialWithinBlock} + \beta_5\text{trialWithinSession} + \beta_6\text{subjectID} + \varepsilon$

Single-unit regression

Neural analyses in three main epochs

For each neuron, we calculated its firing rate during three event epochs: trial start, cue onset, and feedback. For each epoch, spikes were aligned to the event onset for all trials, we then averaged spike counts in 150 ms time bins stepped at 25ms, with bin centers ranging from 100ms preceding to 700 ms following event onset. We analyzed neural activities within a 500ms time window beginning 100ms after event onset (101-600ms post-event; Fig. 3A). This window was chosen a priori based on the reaction time distribution in the cue onset epoch (Fig. 2A), which was selected to approximately capture the time preceding median responses of both subjects. The same analysis window was used for all epochs. Only completed trials were included in the analysis.

Within this window, we averaged spike count for mean firing rate (mean FR), a standard calculation in previous studies [38, 39]. We used this activity to fit multivariable regression models for each neuron for each epoch. Models used are as follows:

- Trial start epoch:

$$\text{FR} = \beta_0 + \beta_1\text{refPoint} + \varepsilon$$

where ε is the error term, *refPoint* corresponds to length of the expected-reward bar at the beginning of each trial. We reported population analysis for β_1 in the main text.

- Cue onset epoch:

$$\begin{aligned}\text{FR} = {} & \beta_0 + \beta_1\text{refPoint} + \beta_2\text{cueValence} + \beta_3\text{cueDirection} + \beta_4\text{refPoint} \times \text{cueValence} \\ & + \beta_5\text{refPoint} \times \text{cueDirection} + \beta_6\text{cueValence} \times \text{cueDirection} \\ & + \beta_7\text{refPoint} \times \text{cueValence} \times \text{cueDirection} + \varepsilon\end{aligned}$$

where ε is the error term, *refPoint* corresponds to length of the expected-reward bar at the beginning of each trial; *cueValence* and *cueDirection* corresponds to valence (positive/ negative) and correct response (left/ right) of the cue of each trial. We reported population analysis for β_4 in the main text.

- Feedback epoch:

$$\begin{aligned}\mathrm{FR} = {} & \beta_0 + \beta_1\mathrm{refPoint} + \beta_2\mathrm{cueValence} + \beta_3\mathrm{cueDirection} + \beta_4\mathrm{outcome} \\ & + \beta_5\mathrm{refPoint} \times \mathrm{cueValence} + \beta_6\mathrm{refPoint} \times \mathrm{cueDirection} + \beta_7\mathrm{refPoint} \times \mathrm{outcome} \\ & + \beta_8\mathrm{cueValence} \times \mathrm{cueDirection} + \beta_9\mathrm{cueValence} \times \mathrm{outcome} + \beta_{10}\mathrm{cueDirection} \times \mathrm{outcome} \\ & + \beta_{11}\mathrm{refPoint} \times \mathrm{cueValence} \times \mathrm{cueDirection} \\ & + \beta_{12}\mathrm{refPoint} \times \mathrm{cueValence} \times \mathrm{outcome} \\ & + \beta_{13}\mathrm{refPoint} \times \mathrm{cueDirection} \times \mathrm{outcome} \\ & + \beta_{14}\mathrm{cueValence} \times \mathrm{cueDirection} \times \mathrm{coutome} \\ & + \beta_{15}\mathrm{refPoint} \times \mathrm{cueValence} \times \mathrm{cueDirection} \times \mathrm{outcome} + \varepsilon\end{aligned}$$

where ε is the error term, *refPoint* corresponds to length of the expected-reward bar at the beginning of each trial; *cueValence* and *cueDirection* corresponds to valence (positive/ negative) and correct response (left/ right) of the cue of each trial; and *outcome* refers to earned-reward/ change in expected-reward-bar (+1/0/-1). We reported population analysis for β_7 in the main text.

The model coefficient is influenced by the firing rate profile of each neuron, which can vary with firing rate range and response timing across populations. In addition to mean FR, we also calculated other types of firing rates, including: (1) standardized mean FR, where the mean FR of each neuron was normalized based on the variance of these values across all trials within a session before fitting the model; (2) peak FR, which is the peak firing rate in a 150 ms window within a 600ms time window beginning right after event onset (1-600ms post-event); and (3) standardized peak FR, where peak FR was normalized based on the variance of these values across all trials within a session before fitting the model (see Fig. S1A). Peak FR was included to account for neurons with different response delays, where the mean FR in the analysis window could be higher or lower due to earlier or later responses. Standardizing these FR values prevents impact of outliers when analyzing population-level statistics.

To assess a brain area's specific response to a model regressor, we examined the βs (coefficients) and the corresponding *p-value*s (Fig. 3A) of all neurons belonging to that area.

Population analysis

Neurons were grouped based on their recorded brain location, encompassing six different areas across sessions (Fig. 1E).

First, to assess an area's homogeneous, or population average, response to a specific task variable, we evaluated the mean ± SEM of all β values for neurons in each area, using a two-tailed one-sample t-test against zero. The p-value of this t-test was included to demonstrate that, as a population, neurons responded uniformly to certain variables (see Fig. 3C).

Second, we calculated the percentage of the population (PoP) with significant coefficients for each area by applying a significance level of 0.01 to the β value of each neuron's regressor. Neurons meeting this criterion were divided into two groups based on the sign of their β (positive/negative). The PoP for each area was calculated as:

- Positive PoP: $\frac{(\text{number of neurons with } \textbf{positive} \text{ and significant } \beta)}{(\text{total number of neurons})}$
- Negative PoP: $\frac{(\text{number of neurons with } \textbf{negative} \text{ and significant } \beta)}{(\text{total number of neurons})}$

A χ^2 test was used to determine whether there was a significant bias in the percentage of positive versus negative populations.

Population Neural Dynamics

For each neuron in each event epoch, spikes were aligned to the event onset for all trials, we then averaged spike counts in 150 ms time bins stepped at 25ms, with bin centers ranging from 1000ms preceding cue presentation to 1000 ms following event onset. We then grouped trials into 3 reference-level categories (Fig. 1D) and calculated the mean FR in each one.

For all neurons in each brain area, we continued averaging FR across all neurons in each reference-level category, resulting in mean ± SEM for each time bin.

Further analyses for vbACC

Sequential regression analysis in trial start epoch for vbACC

For each neuron in the vbACC, and for each firing rate measurement type, we performed a sequential regression analysis. First, we applied a univariate regression to account for the influence of trial counting until reward reception:

$$\text{FR} = \beta_0 + \beta_1 \text{trialth} + \varepsilon$$

where ε is the error term, *trialth* corresponds to trial-number in the current block (1-6). We calculated the residuals of the above model, defined as the difference between real and predicted firing rate, to fit the second univariate regression:

$$\text{residuals} = \beta_0 + \beta_1 \text{refPoint} + \varepsilon$$

where *residuals* is defined above, *refPoint* corresponds to length of the expected-reward bar at the beginning of each trial. **Figure S2** showed the result of population analysis from the second regression model.

For testing the consistency of reference point in vbACC through the trial

First, using the population neural dynamics described above, we calculated the average FR within a 500ms time window beginning 100ms after event onset (for fixation, it is the fixation start) for different reference levels. We then performed repeated-measurement ANOVA for each epoch to test the difference of neural activity due to reference levels (Fig. S3B).

Second, for each neuron in the vbACC, we calculated its firing rate during five event epochs: trial start, fixation, cue onset, response, and feedback. For each epoch, analyses were done as described in “Single-unit regression” section but using only univariate regression with *refPoint* as the sole regressor. Figure S3C showed the result of population analysis for these univariate regressions.

General statistics

P-values less than or equal to 0.05 were considered significant (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, n.s. $p > 0.05$). All figures show mean ± SEM, except otherwise noted.

REFERENCES

1. Samuelson, P.A., *A Note on Measurement of Utility.* Review of Economic Studies, 1936. **4**: p. 155-161.
2. Von Neumann, J. and O. Morgenstern, *Theory of games and economic behavior*. 1944, Princeton: Princeton University Press. xviii, 625 pages.
3. Houthakker, H.S., *Revealed Preference and the Utility Function*. Economica-New Series, 1950. **17**(66): p. 159-174.
4. Allais, M., *Behavior of the Rational Man before Risk - Criticism of American School Postulates and Axioms*. Econometrica, 1953. **21**(4): p. 503-546.
5. Ellsberg, D., *Risk, Ambiguity, and the Savage Axioms.* Quarterly Journal of Economics, 1961. **75**(4): p. 643-669.
6. Kahneman, D. and A. Tversky, *Prospect Theory - Analysis of Decision under Risk*. Econometrica, 1979. **47**(2): p. 263-291.
7. Tversky, A. and D. Kahneman, *The Framing of Decisions and the Psychology of Choice*. Science, 1981. **211**(4481): p. 453-458.
8. Tversky, A. and D. Kahneman, *Loss Aversion in Riskless Choice - a Reference-Dependent Model*. Quarterly Journal of Economics, 1991. **106**(4): p. 1039-1061.
9. Markowitz, H., *The Utility of Wealth*. Journal of Political Economy, 1952. **60**(2): p. 151-158.
10. Rabin, M., *Risk aversion and expected-utility theory: A calibration theorem*. Econometrica, 2000. **68**(5): p. 1281-1292.
11. Simonson, I. and A. Tversky, *Choice in Context - Tradeoff Contrast and Extremeness Aversion*. Journal of Marketing Research, 1992. **29**(3): p. 281-295.
12. Wakker, P.P., *Prospect theory : for risk and ambiguity*. 2010, Cambridge: Cambridge University Press. 1 online resource (xiii, 503 pages).
13. Helson, H., *Adaptation-Level as Frame of Reference for Prediction of Psychophysical Data*. American Journal of Psychology, 1947. **60**(1): p. 1-29.
14. Parducci, A., *Category Judgment - a Range-Frequency Model*. Psychological Review, 1965. **72**(6): p. 407-418.
15. Rutledge, R.B., et al., *A computational and neural model of momentary subjective well-being*. Proceedings of the National Academy of Sciences of the United States of America, 2014. **111**(33): p. 12252-12257.
16. Tversky, A. and D. Kahneman, *Advances in Prospect-Theory - Cumulative Representation of Uncertainty*. Journal of Risk and Uncertainty, 1992. **5**(4): p. 297-323.
17. Parducci, A., *Happiness, pleasure, and judgment : the contextual theory and its applications*. 1995, Mahwah, N.J.: L. Erlbaum Associates. ix, 225 pages.
18. Delquié, P. and A. Cillo, *Expectations, disappointment, and rank-dependent probability weighting*. Theory and Decision, 2006. **60**(2-3): p. 193-206.
19. Koszegi, B. and M. Rabin, *A model of reference-dependent preferences*. Quarterly Journal of Economics, 2006. **121**(4): p. 1133-1165.
20. Koszegi, B. and M. Rabin, *Reference-dependent risk attitudes*. American Economic Review, 2007. **97**(4): p. 1047-1073.
21. Glimcher, P.W. and A.A. Tymula, *Expected subjective value theory (ESVT): A representation of decision under risk and certainty*. Journal of Economic Behavior & Organization, 2023. **207**: p. 110-128.
22. Chan, H.K. and T. Toyoizumi, *A multi-stage anticipated surprise model with dynamic expectation for economic decision-making*. Scientific Reports, 2024. **14**(1).

23. Tom, S.M., et al., *The neural basis of loss aversion in decision-making under risk*. Science, 2007. **315**(5811): p. 515-518.
24. Yang, Y.P., X.J. Li, and V. Stuphorn, *Primate anterior insular cortex represents economic decision variables proposed by prospect theory*. Nature Communications, 2022. **13**(1).
25. Ludvig, E.A., C.R. Madan, and M.L. Spetch, *Extreme Outcomes Sway Risky Decisions from Experience*. Journal of Behavioral Decision Making, 2014. **27**(2): p. 146-156.
26. Rigoli, F., et al., *Learning Contextual Reward Expectations for Value Adaptation*. Journal of Cognitive Neuroscience, 2018. **30**(1): p. 50-69.
27. Bayer, H.M. and P.W. Glimcher, *Midbrain dopamine neurons encode a quantitative reward prediction error signal*. Neuron, 2005. **47**(1): p. 129-141.
28. Tobler, P.N., C.D. Fiorillo, and W. Schultz, *Adaptive coding of reward value by dopamine neurons*. Science, 2005. **307**(5715): p. 1642-1645.
29. Glimcher, P.W., *Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis (vol 108, pg 15647, 2011)*. Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(42): p. 17568-17568.
30. Padoa-Schioppa, C., *Range-Adapting Representation of Economic Value in the Orbitofrontal Cortex*. Journal of Neuroscience, 2009. **29**(44): p. 14004-14014.
31. Kobayashi, S., O.P. de Carvalho, and W. Schultz, *Adaptation of Reward Sensitivity in Orbitofrontal Neurons*. Journal of Neuroscience, 2010. **30**(2): p. 534-544.
32. Seo, H. and D. Lee, *Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game*. Journal of Neuroscience, 2007. **27**(31): p. 8366-8377.
33. Cai, X.Y. and C. Padoa-Schioppa, *Neuronal Encoding of Subjective Value in Dorsal and Ventral Anterior Cingulate Cortex*. Journal of Neuroscience, 2012. **32**(11): p. 3791-3808.
34. Hayden, B.Y., J.M. Pearson, and M.L. Platt, *Neuronal basis of sequential foraging decisions in a patchy environment*. Nature Neuroscience, 2011. **14**(7): p. 933-U165.
35. Kane, G.A., et al., *Rat Anterior Cingulate Cortex Continuously Signals Decision Variables in a Patch Foraging Task*. J Neurosci, 2022. **42**(29): p. 5730-5744.
36. Kolling, N., et al., *Neural Mechanisms of Foraging*. Science, 2012. **336**(6077): p. 95-98.
37. Zacharopoulos, G., G. Maio, and D.E.J. Linden, *Dissecting the neurocomputational bases of patch-switching*. Cerebral Cortex, 2023. **33**(12): p. 7930-7940.
38. Chien, J.M., J.D. Wallis, and E.L. Rich, *Abstraction of Reward Context Facilitates Relative Reward Coding in Neural Populations of the Macaque Anterior Cingulate Cortex*. Journal of Neuroscience, 2023. **43**(33): p. 5944-5962.
39. Rich, E.L. and J.D. Wallis, *Medial-lateral Organization of the Orbitofrontal Cortex*. Journal of Cognitive Neuroscience, 2014. **26**(7): p. 1347-1362.
40. Flaherty, C.F., *Incentive Contrast - a Review of Behavioral-Changes Following Shifts in Reward*. Animal Learning & Behavior, 1982. **10**(4): p. 409-440.
41. Niv, Y., et al., *Tonic dopamine: opportunity costs and the control of response vigor*. Psychopharmacology, 2007. **191**(3): p. 507-520.
42. Kutner, M.H., C. Nachtsheim, and J. Neter, *Applied linear regression models*. 4th ed. 2004, Boston ; New York: McGraw-Hill/Irwin. 701 p.
43. Shih, W.Y., et al., *Electrophysiological population dynamics reveal context dependencies during decision making in human frontal cortex*. Nat Commun, 2023. **14**(1): p. 7821.
44. Constantinople, C.M., A.T. Piet, and C.D. Brody, *An Analysis of Decision under Risk in Rats*. Current Biology, 2019. **29**(12): p. 2066-+.
45. Cox, K.M. and J.W. Kable, *BOLD subjective value signals exhibit robust range adaptation*. J Neurosci, 2014. **34**(49): p. 16533-43.

46. Bernoulli, D., *Specimen theoriae novae de mensura sortis*. 1967, Farnborough: Gregg. 5 60 1 23-36 p.
47. Rigotti, M., et al., *The importance of mixed selectivity in complex cognitive tasks*. Nature, 2013. **497**(7451): p. 585-590.
48. Fusi, S., E.K. Miller, and M. Rigotti, *Why neurons mix: high dimensionality for higher cognition*. Current Opinion in Neurobiology, 2016. **37**: p. 66-74.
49. Pandya, D.N., G.W. Vanhoesen, and M.M. Mesulam, *Efferent Connections of the Cingulate Gyrus in the Rhesus-Monkey*. Experimental Brain Research, 1981. **42**(3-4): p. 319-330.
50. Morecraft RJ, T.J., *Cingulorontal interaction and the cingulate motor areas*, in *Cingulate neurobiology and disease*, V. BA, Editor. 2009, Oxford University Press: Oxford. p. 113-144.
51. Pandya, D.N. and L.A. Vignolo, *Intra- and interhemispheric projections of the precentral, premotor and arcuate areas in the rhesus monkey*. Brain Res, 1971. **26**(2): p. 217-33.
52. Vogt, B.A. and D.N. Pandya, *Cingulate Cortex of the Rhesus-Monkey .2. Cortical Afferents*. Journal of Comparative Neurology, 1987. **262**(2): p. 271-289.
53. Carmichael, S.T. and J.L. Price, *Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys*. Journal of Comparative Neurology, 1995. **363**(4): p. 642-664.
54. Rosene, D.L. and G.W. Van Hoesen, *Hippocampal efferents reach widespread areas of cerebral cortex and amygdala in the rhesus monkey*. Science, 1977. **198**(4314): p. 315-7.
55. Tang, W., et al., *A connectional hub in the rostral anterior cingulate cortex links areas of emotion and cognitive control*. Elife, 2019. **8**.
56. Mesulam, M.M. and E.J. Mufson, *Insula of the old world monkey. III: Efferent cortical output and comments on function*. J Comp Neurol, 1982. **212**(1): p. 38-52.
57. Charnov, E.L., *Optimal foraging, the marginal value theorem*. Theor Popul Biol, 1976. **9**(2): p. 129-36.
58. Hayden, B.Y., *Economic choice: the foraging perspective*. Current Opinion in Behavioral Sciences, 2018. **24**: p. 1-6.
59. Kennerley, S.W., et al., *Optimal decision making and the anterior cingulate cortex*. Nature Neuroscience, 2006. **9**(7): p. 940-947.
60. Sallet, J., et al., *Expectations, gains, and losses in the anterior cingulate cortex*. Cognitive Affective & Behavioral Neuroscience, 2007. **7**(4): p. 327-336.
61. Kennerley, S.W., et al., *Neurons in the Frontal Lobe Encode the Value of Multiple Decision Variables*. Journal of Cognitive Neuroscience, 2009. **21**(6): p. 1162-1178.
62. Kolling, N., et al., *Multiple signals in anterior cingulate cortex*. Current Opinion in Neurobiology, 2016. **37**: p. 36-43.
63. Hayden, B.Y., J.M. Pearson, and M.L. Platt, *Fictive reward signals in the anterior cingulate cortex*. Science, 2009. **324**(5929): p. 948-50.
64. Shidara, M. and B.J. Richmond, *Anterior cingulate: Single neuronal signals related to degree of reward expectancy*. Science, 2002. **296**(5573): p. 1709-1711.
65. Amiez, C., J.P. Joseph, and E. Procyk, *Anterior cingulate error-related activity is modulated by predicted reward*. European Journal of Neuroscience, 2005. **21**(12): p. 3447-3452.
66. Kolling, N., et al., *Value, search, persistence and model updating in anterior cingulate cortex*. Nature Neuroscience, 2016. **19**(10): p. 1280-1285.
67. Akam, T., et al., *The Anterior Cingulate Cortex Predicts Future States to Mediate Model-Based Action Selection*. Neuron, 2021. **109**(1).
68. Vogt, B.A., *Midcingulate cortex: Structure, connections, homologies, functions and diseases*. Journal of Chemical Neuroanatomy, 2016. **74**: p. 28-46.

69. Amiez, C., et al., *Modulation of feedback related activity in the rostral anterior cingulate cortex during trial and error exploration*. Neuroimage, 2012. **63**(3): p. 1078-1090.
70. Brenner, N., W. Bialek, and R. de Ruyter van Steveninck, *Adaptive rescaling maximizes information transmission*. Neuron, 2000. **26**(3): p. 695-702.
71. Dean, I., N.S. Harper, and D. McAlpine, *Neural population coding of sound level adapts to stimulus statistics*. Nat Neurosci, 2005. **8**(12): p. 1684-9.
72. Nagel, K.I. and A.J. Doupe, *Temporal processing and adaptation in the songbird auditory forebrain*. Neuron, 2006. **51**(6): p. 845-59.
73. Louie, K. and P.W. Glimcher, *Efficient coding and the neural representation of value*. Year in Cognitive Neuroscience, 2012. **1251**: p. 13-32.
74. Rangel, A. and J.A. Clithero, *Value normalization in decision making: theory and evidence*. Curr Opin Neurobiol, 2012. **22**(6): p. 970-81.
75. Park, S.Q., et al., *Adaptive coding of reward prediction errors is gated by striatal coupling*. Proc Natl Acad Sci U S A, 2012. **109**(11): p. 4285-9.
76. Rigoli, F., et al., *The influence of contextual reward statistics on risk preference*. Neuroimage, 2016. **128**: p. 74-84.
77. Asaad, W.F. and E.N. Eskandar, *A flexible software tool for temporally-precise behavioral control in Matlab*. Journal of Neuroscience Methods, 2008. **174**(2): p. 245-258.

Table 1: Details of regression models on behavioral performance.

A. Percent correct

Regressor	Coefficient (SE)	p-value
intercept	2.0624 (0.0809)	3.15×10^{-143}
Expected-reward bar length / Reference point	0.1531 (0.0180)	1.96×10^{-17}
Valence (positive-cue indicator)	1.2505 (0.0385)	4.22×10^{-231}
Direction (left-response indicator)	0.2956 (0.0366)	6.13×10^{-16}
Trial number within a block	0.0467 (0.0116)	6.13×10^{-5}
Trial number within session	-5.79×10^{-4} (0.0001)	3.57×10^{-10}
Subject (subject C indicator)	-1.7430 (0.0434)	0

B. Reaction time

Regressor	Coefficient (SE)	p-value
intercept	684.64 (20.66)	2.11×10^{-237}
Expected-reward bar length / Reference point	-71.76 (4.76)	3.94×10^{-51}
Valence (positive-cue indicator)	-157.23 (9.69)	3.56×10^{-59}
Direction (left-response indicator)	57.66 (9.62)	2.07×10^{-9}
Trial number within a block	-15.68 (3.21)	1.06×10^{-6}
Trial number within session	0.35 (0.02)	9.10×10^{-44}
Subject (subject C indicator)	304.62 (9.65)	9.99×10^{-216}

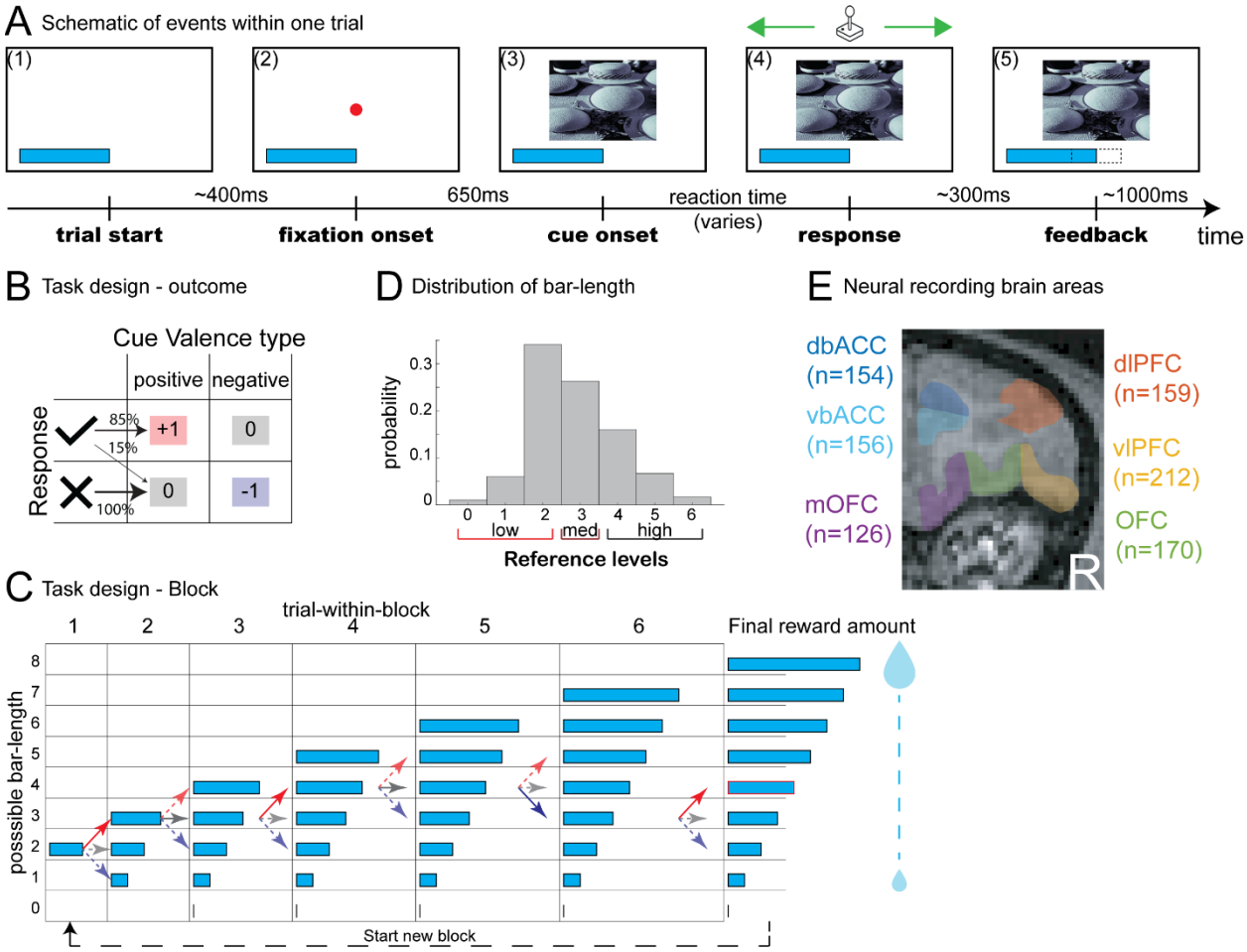


Figure 1: Instructed response task with accumulating reward.

A. Schematic diagram of instructed response task. (1) The trial started after a random inter-trial-interval. The expected-reward bar remained visible throughout a block of 6 trials and only disappeared during reward delivery at the end of the block. Its length indicated the cumulative amount of reward subjects earned at that time point and served as a reference point against which individual trial outcomes could be evaluated. (2) Fixation point onset instructing subjects to fixate on it to start the task. (3) One of four familiar pictures were randomly chosen to appear on the screen. (4) Subjects responded by moving a joystick to left or right. (5) After a ~300ms delay, subjects receive feedback via change (or no change) of the expected-reward bar length, indicating whether they gained/ lost reward or nothing changed.

B. Task design: Outcome depends on subjects' responses and cue valence. A total of 4 visual cues were used as stimulus. For two cues of positive valence, a correct response (check) probabilistically increased the reward amount (top-left) and an error (X) resulted in no change (bottom-left). For the two of negative valence cues, a correct response resulted in no change (top-right) and errors probabilistically resulted in decreased reward amount (bottom-right). The bar length could not

decrease below 0, so when the current bar length was 0, errors always produced no change.

C. Task design: Subjects completed blocks of 6 trials before reward realization. Blocks always started at the expected-reward bar length of 2. Subjects proceeded through 6 trials before receiving the amount of reward accumulated by the end of the 6th trial. The volume of juice reward at the end of the block was proportional to the length of the bar. Arrows indicate potential reward amount earned after finishing each trial, smooth arrows showed an example of one block.

D. Distribution of lengths of expected-reward bar at the beginning of each trial. We clustered trials into low, medium and high reference level trials based on the initial bar-length of that one.

E. Schematic diagram of recorded regions of the brain. Shown on the coronal section of the right hemisphere. In subject M, dbACC, vbACC, mOFC, and OFC neurons were recorded in the left hemisphere, and dlPFC, vlPFC, and OFC neurons were recorded on the right. In subject C, dbACC, vbACC, mOFC, and OFC neurons were recorded in the right hemisphere and dlPFC, vlPFC, and OFC neurons were recorded on the left.

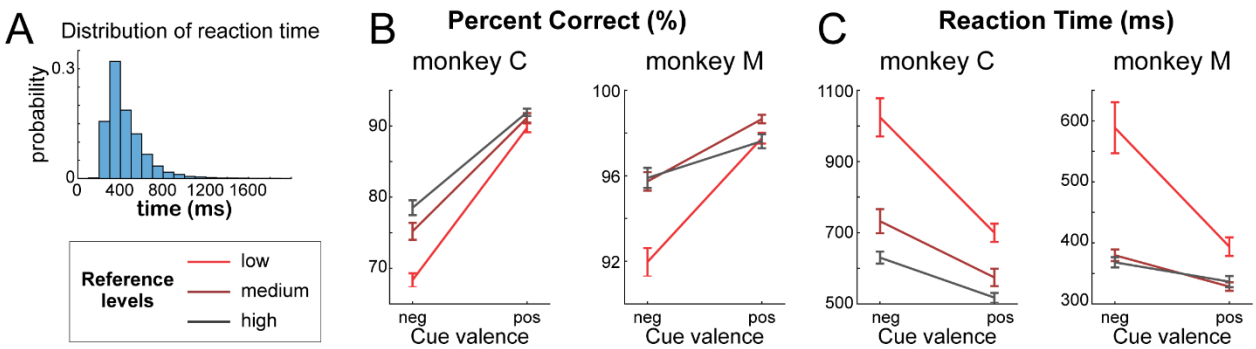


Figure 2: Task performance is reference-dependent. See table 1 for complete model and statistics.

A. Distribution of reaction time (time to respond after cue onset).
B. Percent correct of subjects separated by trials of different reference levels and cue valence.
C. Reaction time of subjects separated by trials of different reference levels and cue valence.

Error bars represent mean ± SEM.

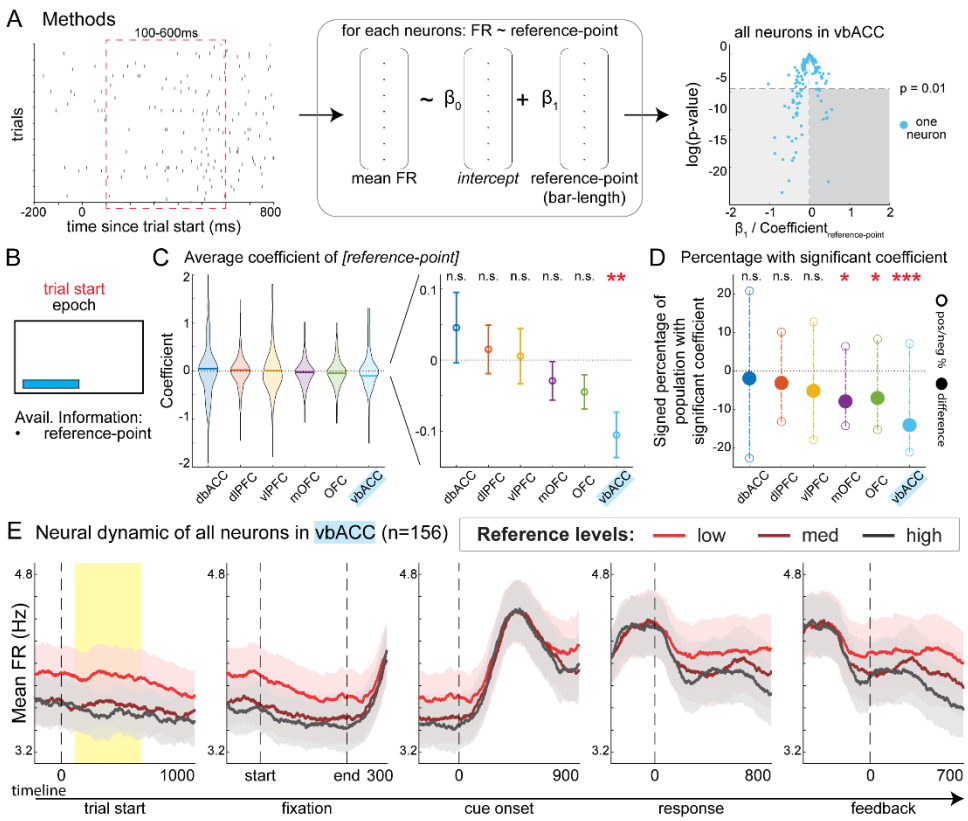


Figure 3: Reference point signal in ventral bank of anterior cingulate cortex (vbACC) during trial start epoch.

A. Schematic of analysis plan. (Left) Time window for analysis. (Middle) For each neuron, we averaged spike count in this window to fit regression models. (Right) From this, we obtained a coefficient value (β) and its p-value for each neuron for each regressor and use this for population analysis.
B. Schematic of trial start epoch. The only information available to subjects is the current amount of cached reward, which acts as a reference point. Therefore, we fitted a regression model with bar-length.
C. Distribution (left) and mean ± SEM (right) for all coefficient values of reference point (*bar-length*) of neurons grouped by brain area. P-value of one-sample t-test for vbACC: 0.0057.
D. The percentage of each population that had significant coefficient values for reference point (*bar-length*) per brain area, divided into positive and negative coefficients. Filled circle indicates the difference between the overall percentage of neurons with positive and negative coefficients. P-value of χ^2 test for positive versus negative: mOFC: 0.0384; OFC: 0.0434; vbACC: 3.45×10^{-4}.
E. Average firing rate of all neurons in vbACC across all events of a trial. Line plots indicate mean and shaded areas indicate SEM.

*p ≤ 0.05, **p ≤ 0.01, ***p ≤0.001, n.s. p > 0.05. Error bars represent mean ± SEM.

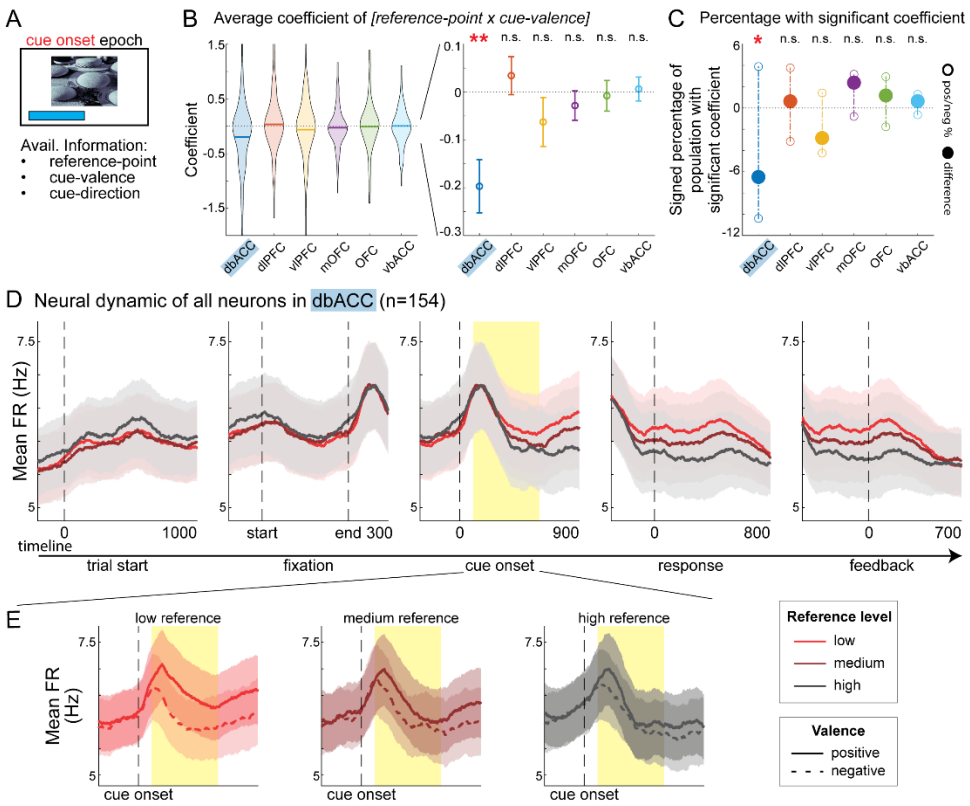


Figure 4: Reference-dependent cue valence signal in dorsal bank of anterior cingulate cortex (dbACC) during cue-onset epoch.

A. Schematic of cue-onset epoch. Available information to subjects is current amount of reward that acts as a reference point, and the 2 properties of cue: cue valence and cue direction. For this epoch, we fitted a multivariable regression model that consists of the above 3 parameters and their interaction terms.
B. Distribution (left) and mean ± SEM (right) for all coefficient values of reference-dependent cue valence signal (interaction term [*reference-point* × *cue-valence*]) of neurons grouped by brain area. P-value of one-sample t-test for dbACC: 0.0011.
C. The percentage of each population that had significant coefficient value for reference-dependent cue valence signal (interaction term [*reference-point* × *cue-valence*]) per brain area. P-value of χ^2 test for dbACC: 0.0269.
D. Average firing rate of all neurons in dbACC across all events of a trial, separated by reference levels. Line plots indicate mean and shaded areas indicate SEM. Same in panel E.
E. Average neural dynamics within cue-onset epoch of all neurons in dbACC, separated by reference levels and cue valence type.

*p ≤ 0.05, **p ≤ 0.01, ***p ≤0.001, n.s. p > 0.05. Error bars represent mean ± SEM.

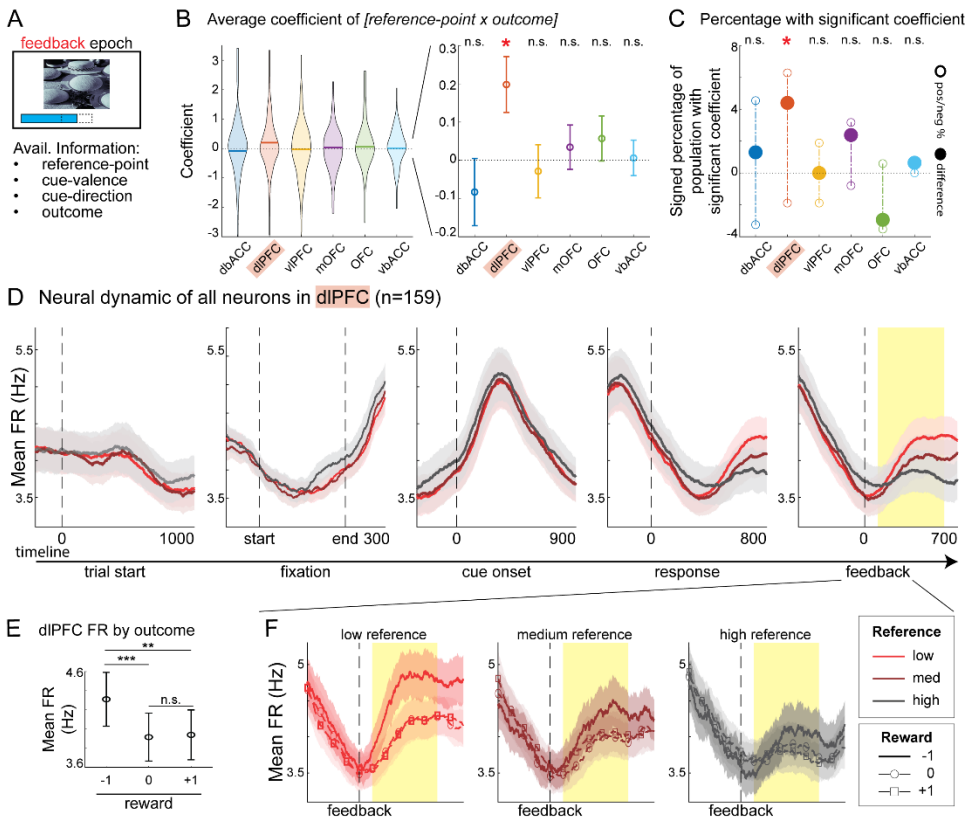


Figure 5: Reference-dependent outcome signal in dorsolateral prefrontal cortex (dlPFC) during feedback epoch.

A. Schematic of feedback epoch. Available information to subjects includes the current amount of reward that acts as a reference point, the 2 properties of the cue (cue valence and cue direction), and the outcome (change in earned reward). For this epoch, we fitted a multivariable regression model that consists of the above 4 parameters and their interaction terms.
B. Distribution (left) and mean ± SEM (right) for all coefficient values of reference-dependent outcome signal (interaction term [*reference-point* × *outcome*]) of neurons grouped by brain area. P-value of one-sample t-test for dlPFC: 0.0182.
C. The percentage of each population that had significant coefficient value for reference-dependent outcome signal (interaction term [*reference-point* × *outcome*]) per brain area. P-value of χ^2 test for dlPFC: 0.0474.
D. Average firing rate of all neurons in dlPFC across all events of a trial, separated by reference levels. Lines indicate mean and shaded areas indicate SEM. Same in panel F.
E. The average firing rate of all neurons in dlPFC, showing it responded specifically to lost reward. P-values for (-1) vs. (0): 2.60×10^{-5}; (-1) vs. (+1): 0.0011.
F. Average neural dynamics within the feedback epoch of all neurons in dlPFC, separated by reference levels and earned reward amount.

*p ≤ 0.05, **p ≤ 0.01, ***p ≤0.001, n.s. p > 0.05. Error bars represent mean ± SEM.

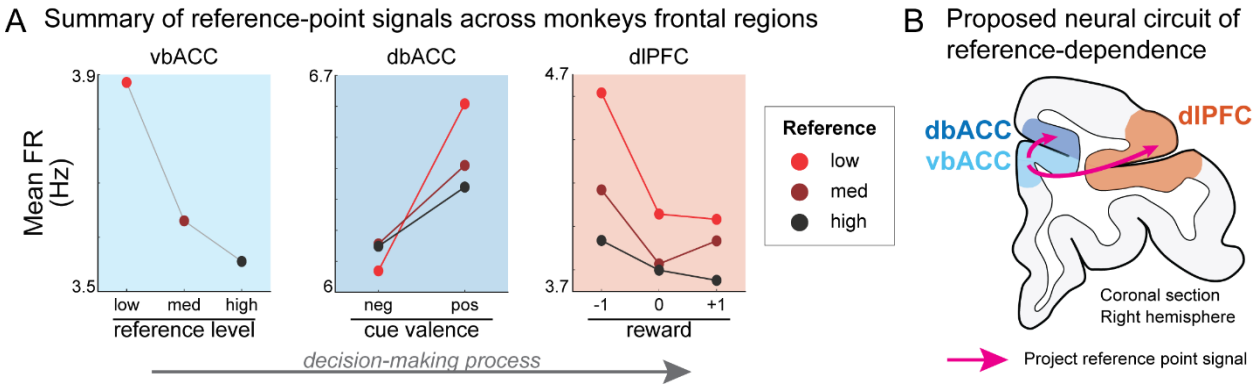


Figure 6: Summary of reference point signal in monkeys' frontal regions.

A. Summary of homogeneous reference point signal in monkeys' frontal areas. As decision-making trial progressed: (1) vbACC appeared to track the reference point, (2) dbACC encoded reference-dependent cue valence signal, and (3) dlPFC encoded reference-dependent outcome signal.
B. A proposed neural circuit of reference-dependence, in which the vbACC serves as the source of reference point signal, modulating decision variables signal in dbACC and dlPFC.

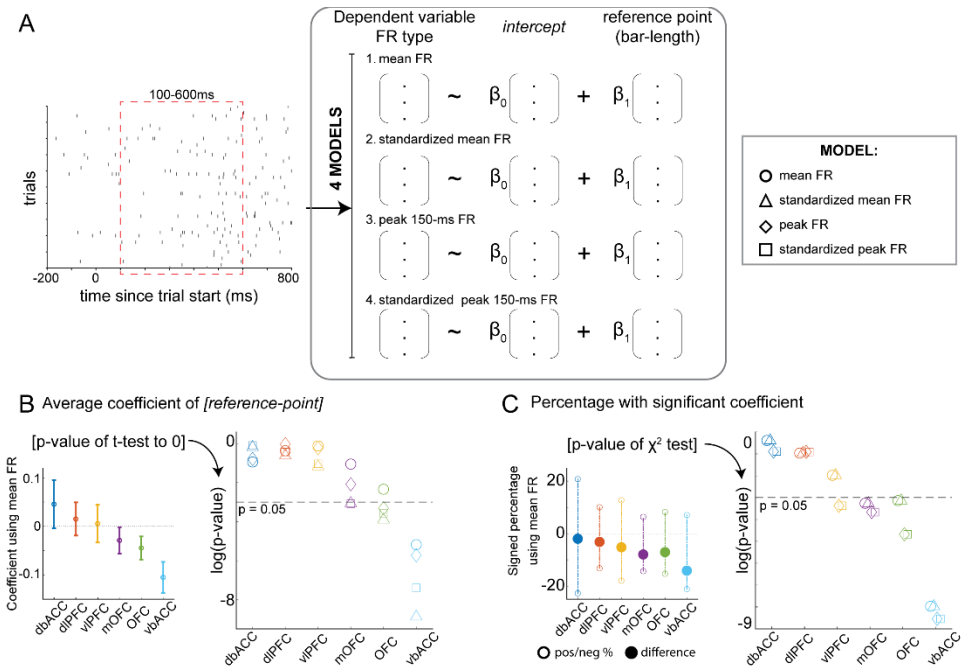


Figure S1: Reference point signal in vbACC during trial start epoch is robust regardless of FR calculation.

A. Analysis plan. (Left) Time window for analysis. (Middle) For each neuron, we calculated firing-rate (FR) in 4 different ways: (1) average spike count as shown in figure 3 (mean FR); (2) standardized mean FR across trials (standardized mean FR); (3) calculated peak FR of 150ms window (peak FR); (4) standardized peak FR across trials (standardized peak FR). Each of these types of FR were used to fit regression models.
B. (Left) Mean ± SEM for coefficient values of reference point, adapted from figure 3C. For each model, we tested the average coefficient values of all neurons of each brain area against 0 using a two-tail t-test. (Right) P-values of the t-test of each model for each brain area.
C. (Left) The percentage of population that had significant coefficient value for reference point signal per brain area, adapted from figure 3D. For each model, we tested whether there is a significant difference between positive and negative percentage using a χ^2 test. (Right) P-values of χ^2 test for each model for each brain area.

B-right-panel and C-right-panel are supplements to figures 3C and D.

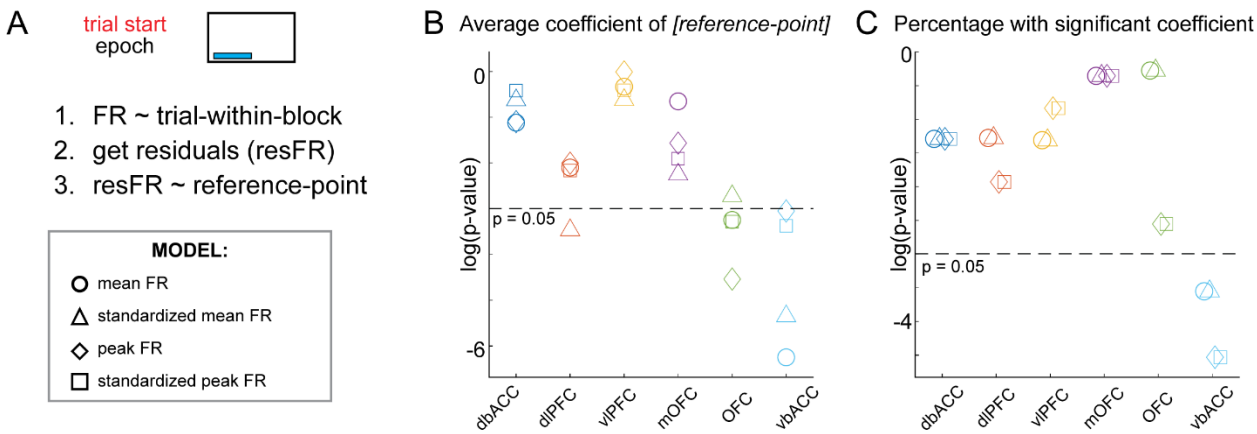


Figure S2: Reference point signal in vbACC after removing effect of trial-number within block.

A. Schematic of analysis. We regressed out the effect of trial-number-within-block (trial-within-block) before calculating the residuals of that model and used it to fit regression model with reference-point.
B. P-values of the t-test of each model for each brain area for coefficient values of reference point. After removing the effect of trial-within-block.
C. P-value of χ^2 test of each model for each brain area for percentage of population that had significant coefficient value for reference point. After removing the effect of trial-within-block.

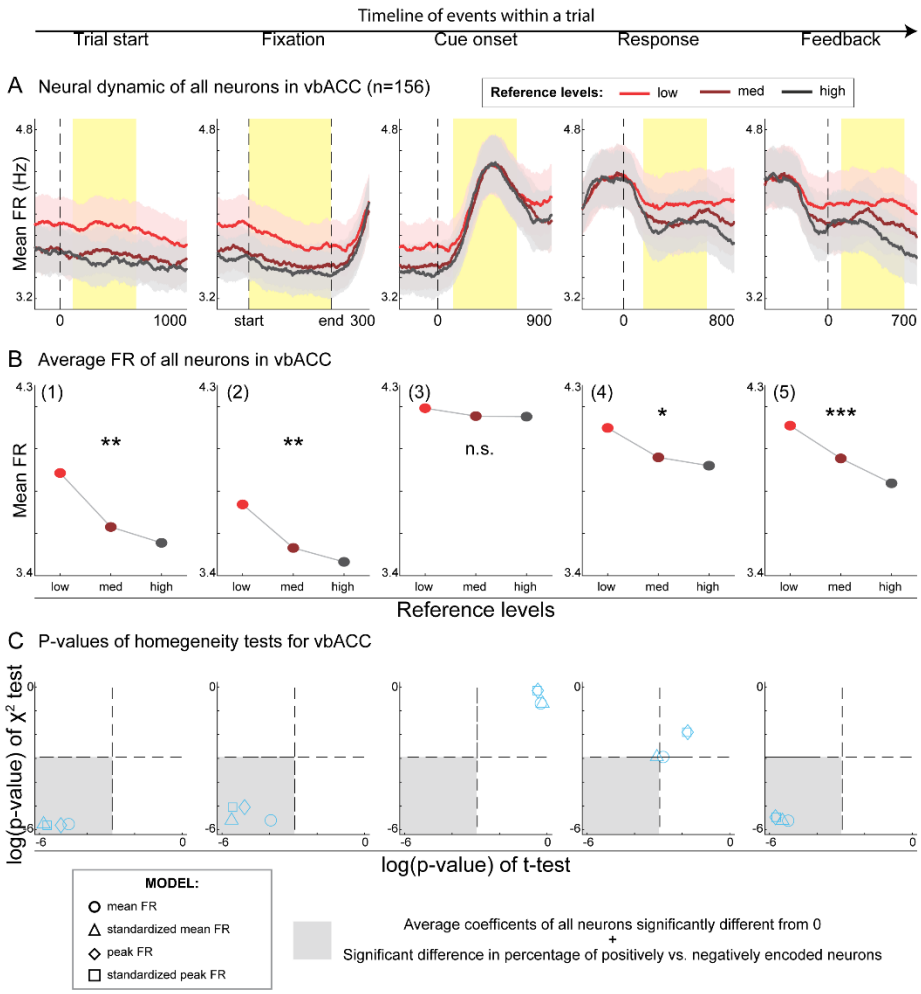


Figure S3: Lasting homogenous encoding of reference point signal in vbACC

A. (from Fig. 3E) Average firing rate of all neurons in vbACC across all events of a trial. Line indicates mean and shaded area indicates SEM. Yellow shaded area indicates the time window used for analysis.
B. Average firing rate within analysis windows across all events separated based on reference levels. P-values of repeated-measured ANOVA test for each epoch: (1) 0.0027; (2) 0.0091; (3) 0.8637; (4) 0.0210; (5) 8.38×10^{-5}.
C. P-values of the t-test and p-value of χ^2 test for reference point encoding of each model across all events. Shaded areas indicated significant homogeneous encoding of reference point. For visualization purpose, here we displayed (p-value + 0.003) to keep the y-axis bounded between [-5.8, 0.003] and roughly equal range for above and below p-value of 0.05.

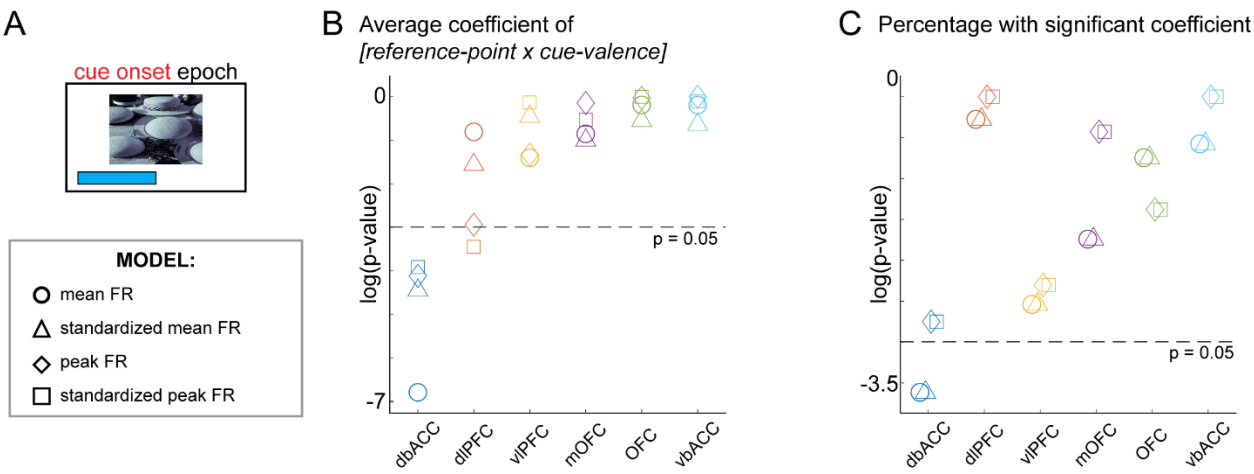


Figure S4: Reference-dependent cue valence signal in dbACC during cue-onset epoch is robust regardless of FR calculation.

A. Schematic of cue-onset epoch.
B. P-values of the t-test of each model for each brain area for coefficient values of reference-dependent cue valence signal (interaction term [*reference-point* × *cue-valence*]).
C. P-value of χ^2 test of each model for each brain area for percentage of population that had significant coefficient value for reference-dependent predicted reward signal (interaction term [*reference-point* × *cue-valence*]).

B and C are supplements to figures 4B and C.

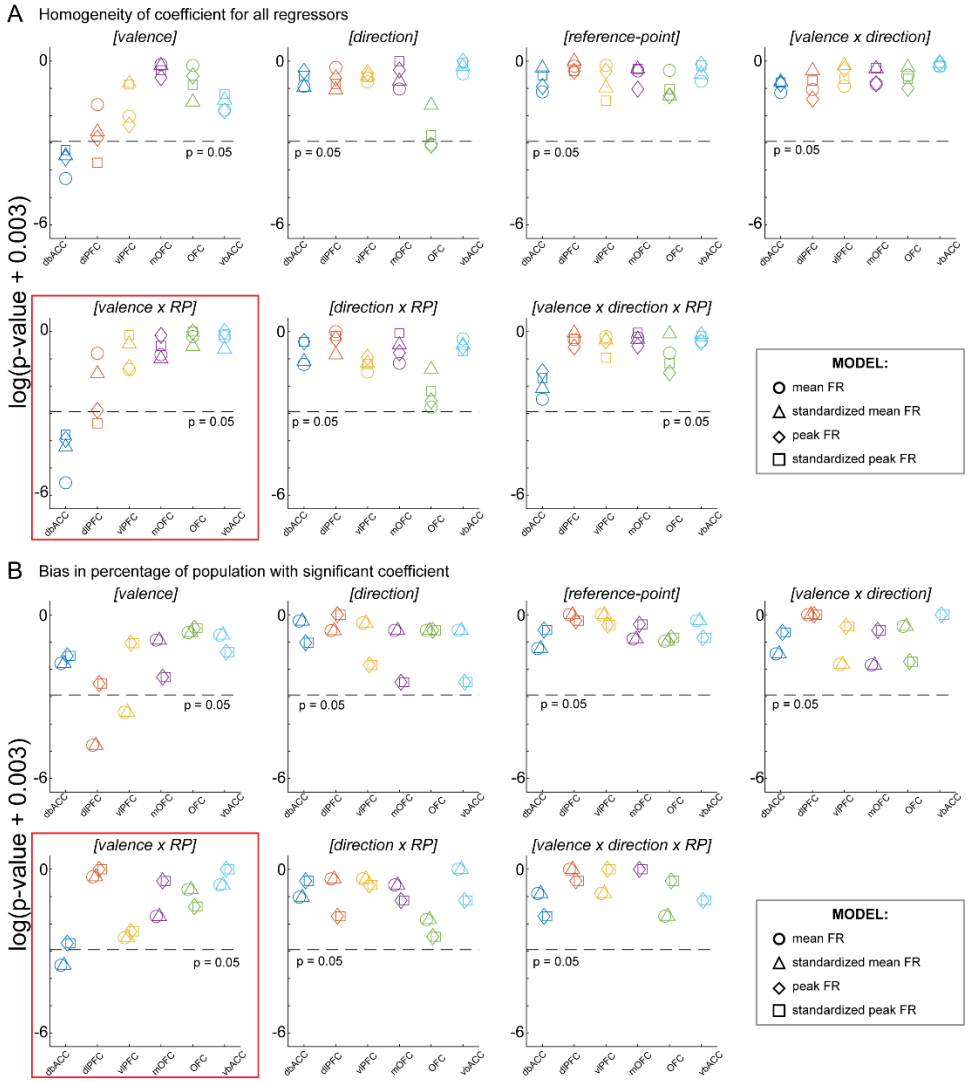


Figure S5: Results of all regressors during cue-onset epoch.

A. P-values of one-sample t-test of each model for each brain area for coefficient values of all regressors used. Title refers to the regressor in the models (RP = reference-point). For visualization purpose, here we displayed (p-value + 0.003) to keep the y-axis bounded between [-5.8, 0.003] and roughly equal range for above and below p-value of 0.05. Same with B.
B. P-value of χ^2 test of each model for each brain area for percentage of population that had significant coefficient value for all regressors used. Title refers to the regressor in the models (RP = reference-point).

Red square indicates the main analysis shown in main figures.

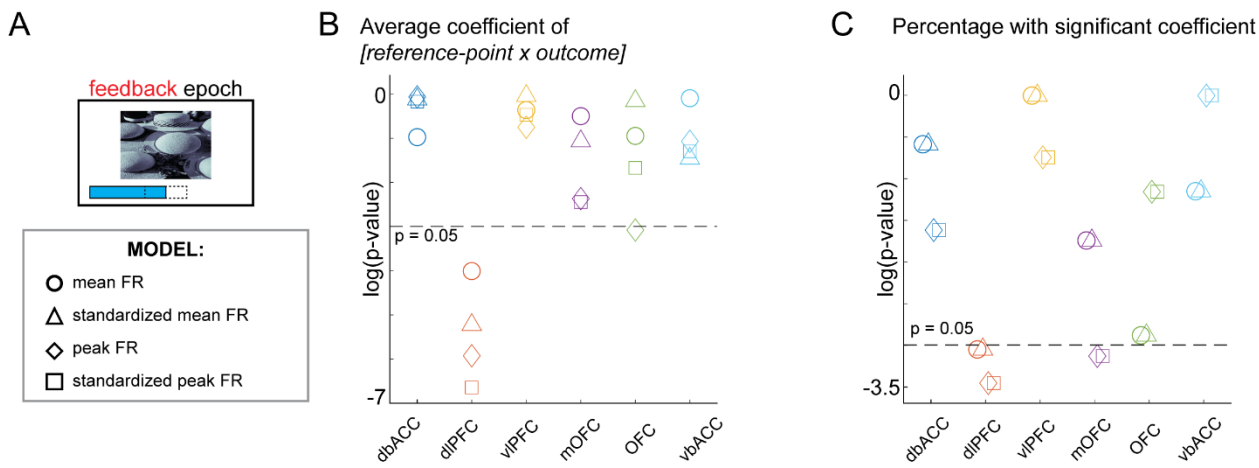


Figure S6: Reference-dependent outcome signal in dlPFC during feedback epoch is robust regardless of FR calculation.

A. Schematic of feedback epoch.
B. P-values of the t-test of each model for each brain area for coefficient values of reference-dependent outcome signal (interaction term [*reference-point* × *outcome*]).
C. P-value of χ^2 test of each model for each brain area for percentage of population that had significant coefficient value for reference-dependent outcome signal (interaction term [*reference-point* × *outcome*]).

B and C are supplements to figures 5B and C.

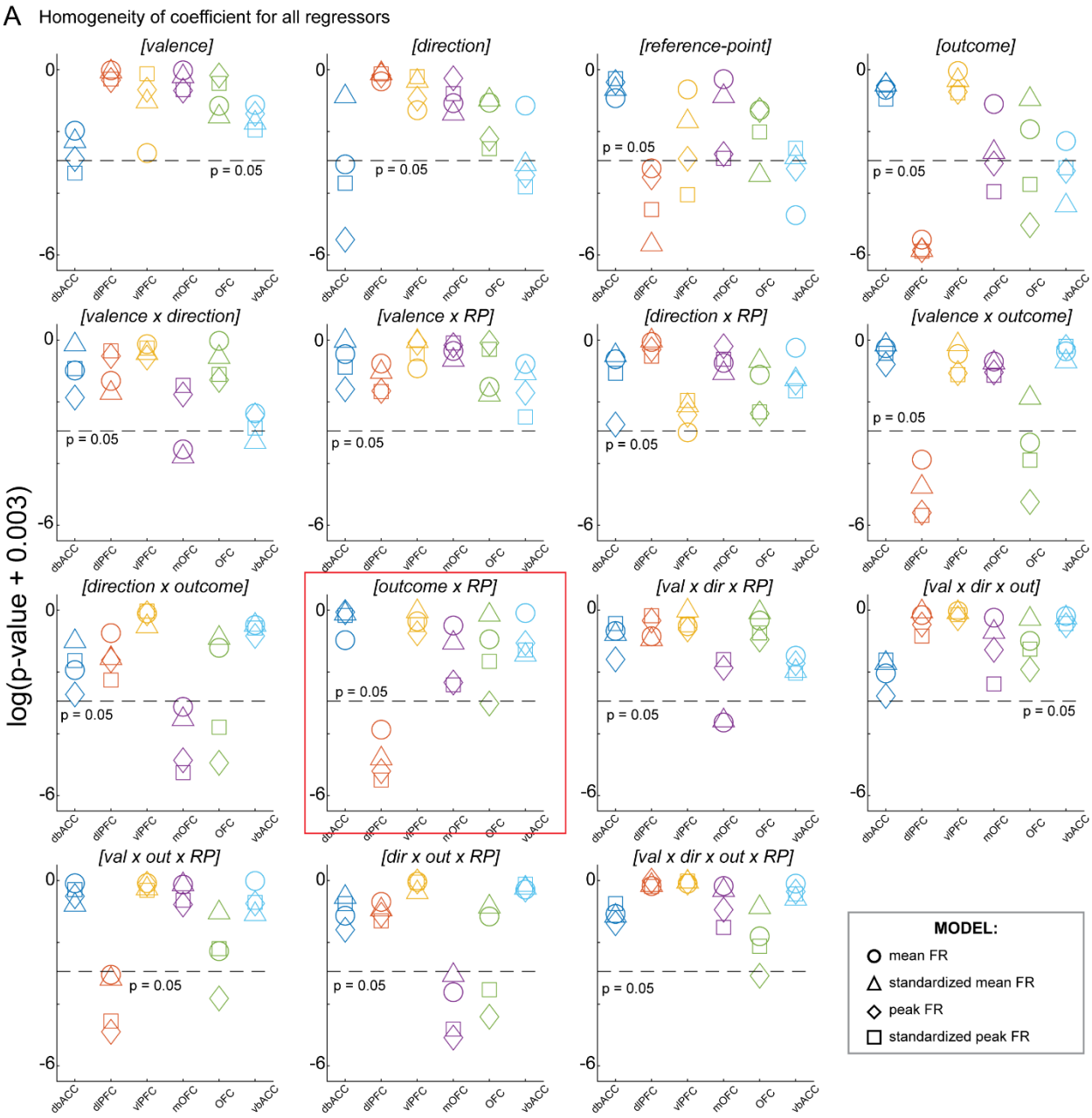
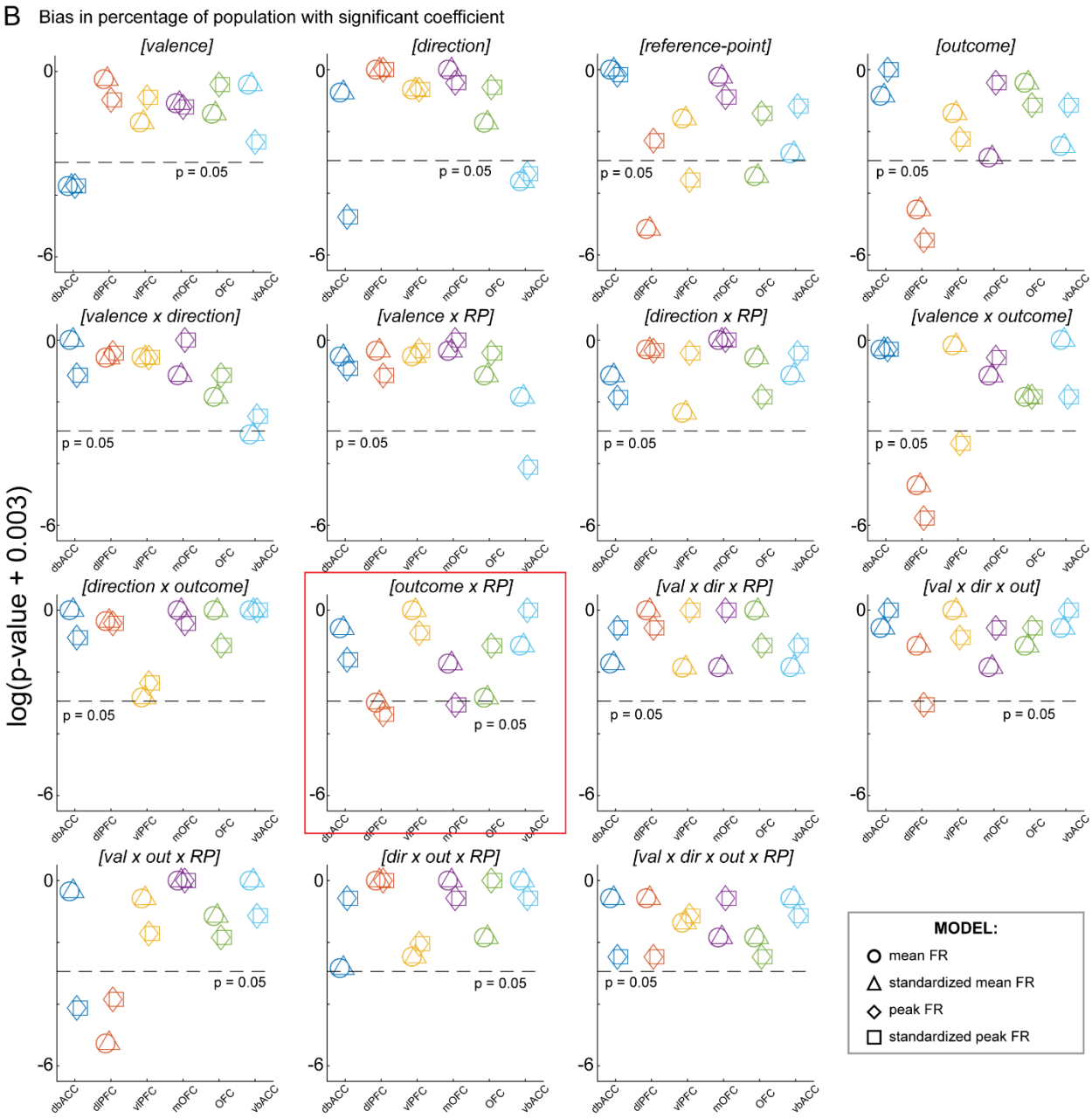


Figure S7: Results of all regressors during feedback epoch.

A. P-values of one-sample t-test of each model for each brain area for coefficient values of any regressors used. Title refers to the regressor in the models (RP = reference-point; val = cue-valence; dir = cue-direction; out = outcome). For visualization purpose, here we displayed (p-value + 0.003) to keep the y-axis bounded between [-5.8, 0.003] and roughly equal range for above and below p-value of 0.05. Same with B.



B. P-value of χ^2 test of each model for each brain area for percentage of population that had significant coefficient value for all regressors used. Title refers to the regressor in the models. (RP = reference-point; val = cue-valence; dir = cue-direction; out = outcome).

Red square indicates the main analysis shown in main figures.