# Value signals in orbitofrontal cortex predict economic decisions on a trial-to-trial basis

**Vincent B. McGinty[1*] & Shira M. Lupkin[1,2]**
1. Center for Molecular and Behavioral Neuroscience, Rutgers University - Newark
2. Behavioral and Neural Sciences Graduate Program, Rutgers University - Newark
*Correspondence: vince.mcginty@rutgers.edu, 197 University Ave, Newark, NJ 07102

**Acknowledgements**: W.T. Newsome for funding, material support, and thoughtful discussions; A. Rangel for thoughtful discussions; J. Brown, E. Carson, S. Fong, A. McCormick, M. Ortiz, J. Powell, J. Sanders, D. Siegel for technical assistance; and R. Ebitz, D. Headley, G. Karpov, D. Kimmel, M. Rosenberg-Lee, K. Peterson, and D. Sharma for helpful comments on the manuscript.

**Funding**: This work was supported by the Howard Hughes Medical Institute (W.T. Newsome), National Institutes of Health Grant K01-DA-036659-01 (V.B.M.), the Busch Biomedical Foundation (V.B.M.), and a Whitehall Foundation fellowship (V.B.M.).

**Author contributions**: VBM designed the research; VBM and SML performed the research. VBM and SML analyzed the data. VBM and SML wrote the paper.

**Data and code availability**: After peer review is complete and the final version of this study is accepted for publication, the data and code that support the findings of this study will be available from the corresponding author upon reasonable request.

## ABSTRACT

Neuroeconomics seeks to explain how neural activity contributes to decision behavior. For value-based decisions, the primate orbitofrontal cortex (OFC) is thought to have a key role; however, the mechanism by which single OFC cells contribute to choices is still unclear. Here, we show for the first time a trial-to-trial relationship between choices and population-level value representations in OFC, defined by the weighted sum of activity from many individual value-coding neurons.

## INTRODUCTION

One of the key ways of identifying decision mechanisms is to show a relationship between neural variability and behavioral variability. The canonical example is in the primate extrastriate area MT, where trial-to-trial variability in the firing of motion-sensitive neurons correlates with trial-to-trial variability in motion judgements, even over trials in which the same noisy motion stimulus is repeated many times[1]. From a mechanistic standpoint, this fine-grained relationship between neural activity and behavior – known as 'choice probability' – is extremely informative. In concert with other findings[2–4] it evinces the likely causal role of MT neurons in motion perception, and, more broadly, provides the foundation for a robust computational framework for understanding how motion evidence is represented by neurons and transformed into a decision[5–7].

For economic choices, defined as those based upon value or preference, the OFC is thought to have a similar causal role: encoding evidence in the form of value representations that are ultimately transformed into a decision[8–12]. Substantial converging evidence supports this hypothesis. First, perturbation of OFC activity disrupts value-based choices[13–18]. Second, a large fraction of OFC neurons encodes economic value[8–10,19,20], a function for which this region appears to be specialized within the frontal lobe[21]. Of particular significance are the findings that the sensitivity of value coding neurons matches the behavioral sensitivity of the subjects measured in the same sessions[10], and that network-level fluctuations in OFC value representations correlate with decision reaction times in single trials.[11] However, while these studies suggest some link between value representation and decision behavior, one key finding has been conspicuously absent: value-based choice probability. Even in large and rigorously analyzed samples of OFC cells, a significant trial-to-trial correlation between value coding and value-based choices has yet to be found[22–24]. This is especially puzzling; if OFC value signals mediate value-based choices, there must be some measurable relationship between them.

Moreover, it is a major barrier to progress in the field, because until this relationship is identified, it will be difficult to fully validate the causal role of OFC, or to establish the kind of robust computational framework for value-based decisions like that available for perceptual judgements.

What can explain the absence of value-based choice probability in OFC? One possibility is that choices are not mediated by value signals, but by some other form of decision-related activity found in OFC, such as the encoding of spatial variables[25]. Another possibility, suggested by the elegant work of Conen and Padoa-Schioppa[24], is that the influence of any single OFC neuron is so small that single-cell choice probabilities cannot be detected above chance levels.

We reasoned that if the latter hypothesis were true, it may be possible to observe significant choice probabilities using population-level value signals – i.e. value representations derived from an ensemble of many simultaneously recorded OFC neurons, which should generally be more informative than signals from individual cells[11,26]. To test this hypothesis, we recorded from populations of of OFC neurons in monkeys making value-based choices, and correlated their decisions with value representations calculated at the population level. Consistent with prior studies using multi-neuronal methods in OFC, we identify population-level representations of economic value[11,21,23,27]; however, unlike prior work, we also identify a robust relationship between these value representations and the decisions made in individual trials.

## RESULTS

Two macaque monkeys performed a novel economic choice task. In every trial, two choice targets were shown on the left and right sides of the task display, each associated with the delivery of a fixed juice reward, between 1-5 drops (Fig. 1a,b). There were 12 physically unique targets used in a given session, but only 5 levels of juice reward; the term ‘target identity’

89 therefore refers to a target's unique physical appearance, and 'target value' is used to refer to its
90 reward association. An important feature of this task is that after an initial fixation period,
91 saccadic eye movements were unrestricted – and in fact were required for optimal performance.
92 The reason is that in each trial the two choice targets were obscured from view until the
93 monkeys fixated upon them directly, an effect achieved through the use of visual crowding (Fig.
94 1c)[28,29] and a dynamic display that was contingent upon gaze (see Methods, and also Hunt et
95 al.[21] for a similar approach). The monkeys therefore viewed the targets sequentially, by fixating
96 them one at a time. Monkeys could view the targets in any order and for any amount of time,
97 and eye tracking was used to determine when each target was viewed. The monkeys were free
98 to choose at any time, and did so by first lifting a central response lever (extinguishing the
99 targets), and then pressing the left or right lever within a 500ms deadline (Fig. 1a, bottom). The
100 task therefore combines two features that were advantageous for neural analysis: First, decision
101 times were recorded in every trial, permitting pre-decision neural activity to be distinguished
102 from post-decision activity. Second, the target viewing times indicated when the target values
103 became known to the monkey in each trial, and were therefore used as temporal reference
104 points for analysis of neural value coding.
105
106 The monkeys viewed both targets at least once in the vast majority of trials (99.3% for Monkey
107 C, 93.2% for Monkey K), and their choices were consistent with both target values being
108 considered in nearly every trial. They almost always chose the larger reward when the targets
109 differed in value by 3 or 4 drops of juice, but chose left or right at near-chance levels when the
110 targets were of equal value (Fig. 1d). In addition, the decision reaction times (RTs) and the
111 mean number of on-target fixations per trial both increased as a function of the difference in
112 target values (Fig. 1e,f). The graded nature of the choice outcomes, RTs, and fixations are
113 consistent with a deliberative process dependent upon the target values. Moreover, the
114 extremely low rate of suboptimal choices in trials with large value differences suggests that the

more variable choices in trials with small or no value differences can be attributable to variability in the choice process itself, rather than to lapses, exploration, or perceptual failure[30].

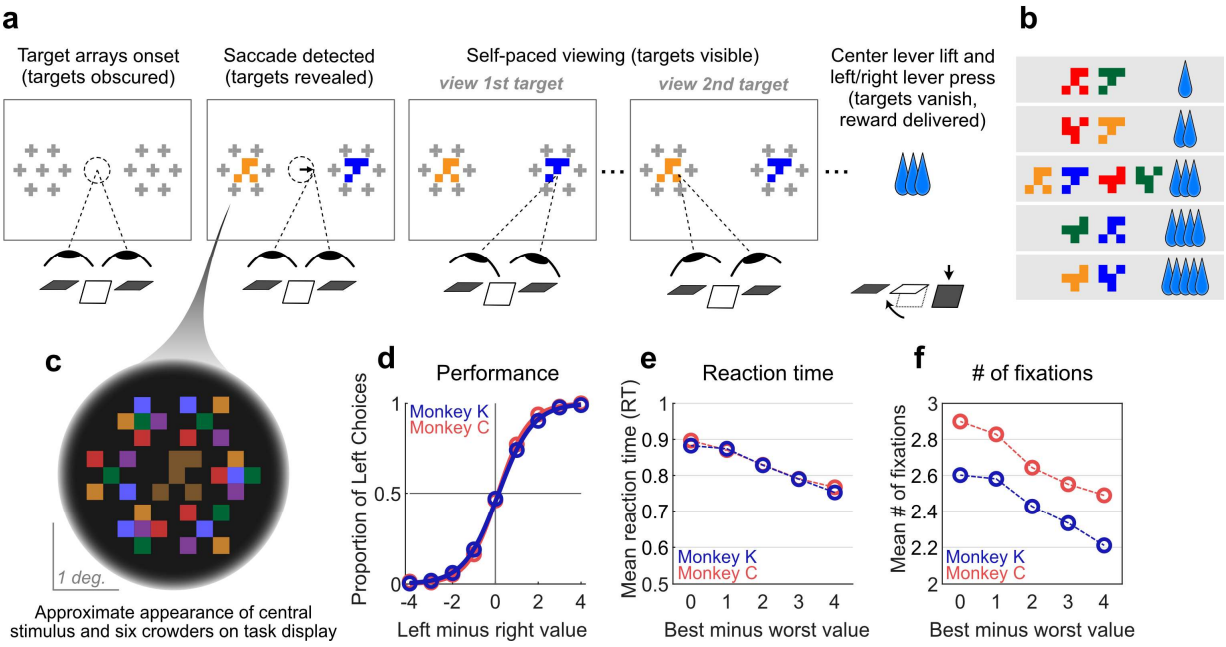



**Figure 1: Decision-making task and performance in two monkeys**. **(a)** Abbreviated task sequence, not shown to scale; see Figure S1 for full task sequence. The yellow and blue glyphs are choice targets, and the six gray "+" shapes indicate the location of crowders designed to obscure the targets until they are viewed (fixated) directly. For clarity, in this panel the crowders are illustrated in gray at reduced scale; on the actual task display, crowders were multicolored and the same size as the targets, as in panel (c). The interval between target arrays onset and center lever lift defines the decision reaction time (RT). **(b)** Example set of 12 choice targets, organized into 5 groups corresponding to the 5 levels of associated juice reward. (**c**) Close-up view of a single target array, consisting of a central yellow choice target surrounded by six non-task-relevant crowders. Two such arrays appear on the display, each centered 7.5° from the display center. **(d-f)** Choices, reaction time, and fixations over 10,972 trials in 16 sessions for Monkey C and 9,617 trials in 16 sessions for Monkey K. **(d)** Fraction of left choices as a function of the left minus right target value. Smooth lines give logistic fit (Eqn. 1). Logistic regression estimates for effect of left minus right value: Monkey K, 1.25 (SE 0.02, n = 9,617 trials); Monkey C, 1.42 (SE 0.03, n = 10,972 trials). **(e)** Mean reaction times decreased as a function of trial difficulty, defined as the absolute difference in target values. SEMs are too small to plot, and are smaller than 0.004s for all data points. A linear fit of RT as a function of difficulty gives the following average slopes: Monkey K -0.035 (SE 0.001); Monkey C, -0.035 (SE 0.001). **(f)** Mean number of fixations performed per trial, as a function of trial difficulty. SEMs are too small to plot, and are smaller than 0.03 for all data points. A linear fit gives the following average slopes: Monkey K, -0.106 (SE 0.007); Monkey C, -0.120 (SE 0.006).

120 OFC neural activity was recorded using multi-channel linear arrays (Fig. 2a-c, Fig. S2), and cells
121 were isolated using semi-automated spike-sorting procedures (see Methods). On average, 45
122 cells were isolated per session (16 sessions and 848 cells for Monkey K, 16 sessions and 602
123 cells for Monkey C).
124
125 Because the target values in each trial were revealed serially (by sequential viewing, see Fig.
126 1a), we characterized neural activity with variables defined by the order in which the targets
127 were viewed, consistent with prior studies using serial designs[21,31,32]. The target values were
128 characterized by the value of the first target viewed by the monkey in each trial (variable name:
129 '1st value'), and the value of the second target viewed in each trial (variable name: '2nd value'),
130 in units of juice drops (Fig. 1b). Choice outcome was characterized by the variable '1st chosen',
131 which was defined as *true* when the first-viewed target was chosen, and *false* otherwise. So that
132 all variables were defined for every trial, neural analyses used only trials in which monkeys
133 viewed both targets (see above).
134
135 We quantified the fraction OFC cells encoding these variables at each moment in time
136 throughout the trial. Consistent with prior work[21], the encoding of the variable '1st value'
137 increased following the viewing of the first target (Fig. 2d); and encoding of '2nd value'
138 increased following the viewing of the second target (Fig. 2e). Notably, after viewing the second
139 target, a large fraction of cells continued to encode '1st value' (black line in Fig. 1e), consistent
140 with prior findings that the OFC maintains a representation of recently viewed items[21,31,32]. A
141 small fraction of cells encoded '1st chosen arising primarily after the monkey viewed the second
142 target (Fig. 2e, gray line).

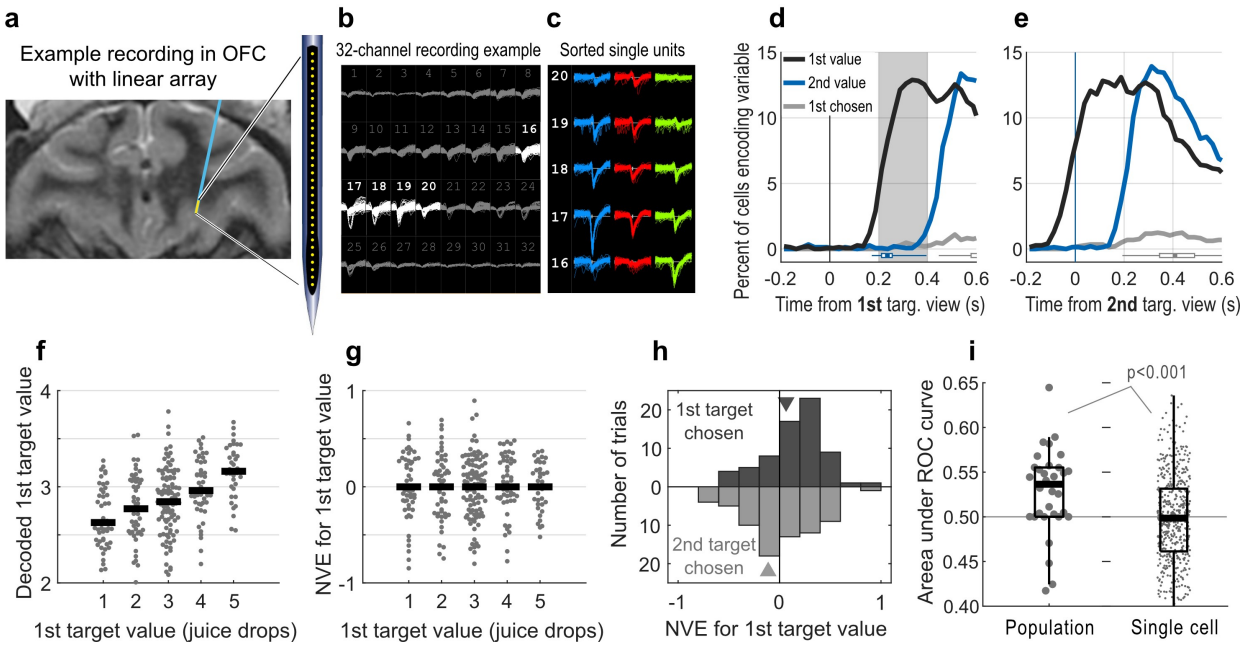



**Figure 2: OFC recording and choice probability analysis (a)** Example placement of electrode array in OFC. See also Figure S2. **(b)** Traces from an example recording. Numerals give array channels. Highlighted channels are shown in panel (c). **(c)** Blue, red, and green traces show three sorted units. **(d-e)** Fraction of cells with significant encoding of three task variables in a linear model ($p < 0.001$ uncorrected, see Eqns. 1 and 2). Data are combined across monkeys. In (d) and (e), the data are aligned to the viewing of the first and second target in each trial, respectively. The blue and gray box plots above the x-axis indicate the onsets of second target viewing and of the decision RT, respectively. The shading gives the epoch used for the analyses in panels (f) through (i). **(f)** Estimates of the value of the first-viewed target, decoded from neural population activity recorded during the shaded epoch in panel (d). Data are from a single representative session. Each dot shows the weighted sum of spiking observed across all recorded cells in a held-out test trial (Eqn. 4, n = 284 test trials total). Trials are grouped according to the reward associated with the first-viewed target; horizontal spread of dots within each group are for visualization purposes only. Horizontal black lines indicate group means. **(g)** The neurally-decoded value estimates in panel (f) were transformed into *normalized variable estimates* (NVEs, see Methods), which quantify the degree to which the decoded values are more or less than expected given the nominal value of the first-viewed target. In other words, NVEs quantify the vertical spread within each group in panel (f). Conventions are the same as in panel (f). **(h)** Variability in the NVEs predicted variability in choices. From the data in panel (g), trials in which the targets differed by 0 or 1 drop of juice were selected (n = 141 out of 284 trials), and were then divided according to the decision outcome (first/second target chosen in the upper/lower histograms, respectively). Arrows give the distribution medians. In this example, the ability of an ideal observer to discriminate between the upper and lower distributions corresponds to an area under the ROC curve of 0.644. **(i)** Choice probabilities aggregated across sessions and monkeys. On average, choice probabilities derived from population-based NVEs (left, 32 sessions) were significantly larger than choice probabilities derived from single cells (right, 515 cells with non-zero weights in Eqn. 3). Gray line indicates chance-level performance. P-value indicates outcome of two-sample t-test.

If neural representations of the target values contribute to choice, they should exhibit choice probability: a trial-to-trial relationship between variability in neural activity and variability in the choices. Whereas prior studies computed choice probabilities for single neurons[22,24], here choice probability was measured with respect to population-level value signals, by leveraging the simultaneous nature of the neural recordings (mean 45 cells per session).

We first considered neural activity within the shaded epoch in Fig. 2d, in which '1st value' is the only variable represented. To find the population-level representation of '1st value' in this epoch, we computed a weighted sum of activity of all of the simultaneously recorded cells in a given session. The weights for each cell were obtained from the fitted coefficients of a linear model, in which '1st value' is explained as a linear function of the activity of all the cells recorded in a session (Eqn. 3). In other words, the weights quantify each cell's unique contribution to explaining variance in '1st value'. Critically, Eqn. 3 uses LASSO regression, meaning that cells making only a negligible contribution to explained variance are assigned weights of exactly zero[33]. For example, in the shaded epoch in 2d, non-zero weights were assigned to 16.1 cells per session on average, or 515 cells total across all 32 sessions in both monkeys. To find the weighted sums, the firing of each cell was multiplied by its respective weight, and added together (Eqn. 4), over a set of held-out test trials that were not used to fit the weights.

The weighted sums provide out-of-sample estimates of the value of the first target, decoded from the neural activity on each trial. Naturally, as the value of the first target increases, the neurally-decoded estimates of target value also increase, evident in the means of these estimates when they are grouped by '1st value', as in the single-session example in Fig. 2f. Note, however, that over repeated presentations of targets with the same value, the neurally-decoded estimates have substantial across-trial variability, evident in the wide vertical spread *within* each of the groups in Fig. 2f.

170

171 Trial-to-trial variability in decoded value is the key measure that we use to assess the
172 relationship between neural variability and decision variability (choice probability). To quantify
173 neural variability, the decoded values from Eqn. 4 were transformed into *normalized variable*
174 *estimates* (NVE), which measures the degree to which the decoded values exceed or fall short
175 of their expected magnitude given the identity of the target (see Methods). For example, Fig. 2f
176 shows the neurally-decoded values for a single session, and Fig. 2g shows these data
177 transformed into their corresponding NVEs. Thus, in Fig. 2f the data reflect mean differences
178 across value levels as well as trial-to-trial variability within each value level; but in Fig. 2g the
179 data reflect only the variability across trials – i.e. the residual variability left after taking target
180 values into account.

181

182 The NVEs were then were used to compute choice probability, defined as the accuracy of an
183 ideal observer using neural activity to classify binary decision outcomes. Here, decisions were
184 classified in terms of whether the first or second target was chosen, and accuracy was
185 quantified by the area under the receiver operating characteristic (ROC) curve, for which chance
186 performance is 0.5 (Fig. 2h). Importantly, only difficult trials were used for classification – those
187 in which the targets differed in value by 0 or 1 – because these were the only trials in which
188 choices were sufficiently variable (Fig. 1d).

189

190 As shown in Fig. 2h-i, variability in the neurally-decoded value signals predicted variability the
191 outcome of the monkey's choices. Specifically, higher NVEs for '1st value' predicted a greater
192 likelihood of choosing the first target, and lower NVEs predicted a greater likelihood of choosing
193 the second target. Overall, classification accuracy significantly greater than chance: across 32
194 sessions, the mean area under the ROC curve (AUC) was 0.528 (SEM 0.008, $p = 0.003$, Fig.
195 2i.)

While low in absolute terms, this AUC is comparable to choice probabilities reported for single MT neurons during motion discrimination[1,7]. Moreover, it is notable that this choice-predictive activity is detectable well before the decision reaction time, and before the value of the other target is represented in OFC firing (Fig. 2d).

These choice probability effects – derived from value representations measured at the population-level – are also substantially higher than those reported for single OFC cells in prior studies[22–24]. To directly compare population-based choice probabilities to single cells in our own study, we focused on the cells with non-zero weights in Eqns. 3 and 4, i.e. the cells that contribute to the decoded population value signal. Using the sign of each cell's weight to specify the positive class (see Methods), the AUC was calculated individually for every cell (n = 515 total, mean 16.1 per session). The average AUC was not significantly different from chance (0.497, SEM 0.002, p = 0.14), and was significantly smaller than the session-wise average AUCs derived from the population (p = 0.0009 by rank sum test, Fig. 2i). Choice probabilities were also not distinguishable from chance when using the cells' weights to calculate a weighted mean AUC over the 515 cells (0.493, bootstrapped 99% confidence interval [0.483 0.501]); or when finding the mean AUC of the highest-weighted cell in a given session (mean AUC 0.496, SEM 0.013, p = 0.94 by t-test) These latter two results rule out the possibility that significant population-derived choice probabilities are due only to the weighting applied to each cell, or due to the decoded value signal being dominated by one highly-predictive cell. Consistent with prior work[34], both the population-derived and single-cell choice probabilities were slightly larger when AUCs were calculated within conditions using non-normalized data (see Methods). However, with this method, population-based choice probabilities were still significantly larger than in single cells (Fig. S4). Thus, the neural representation of the value of the first target, decoded from ensembles of primate OFC neurons, provided robust choice probability effects that were not evident at the level of the individual neurons.

222

223 To evaluate choice probability over the course of the trial, the analysis above was performed in
224 bins of 200ms at increments of 25ms, and was performed separately for the variables '1st
225 value', '2nd value', and '1st chosen'. (For '1st chosen', spiking variance attributable to the two
226 target values was removed before finding the cells weights; see Methods for details.) As shown
227 in Fig. 3a-b, significant choice probability effects for '1st value' and '2nd value' were evident
228 starting ~200ms after viewing the respective targets, mirroring the timing pattern for the fraction
229 of cells encoding these variables (Fig. 2d-e). The opposite signs of the '1st value' and '2nd
230 value' effects indicate that higher '1st value' NVEs predict more choices in favor of the first
231 target, whereas higher '2nd value' NVEs predict the opposite.

232

233 Notably, the choice probability effect for '1st value' was maintained during the 'late' epoch
234 indicated by the shading in Fig. 3b, appearing to exceed the magnitude of choice probability
235 effects in the 'early' epoch indicated in Fig. 3a. We therefore focused on this late epoch for
236 follow-up analyses for all three variables of interest.

237

238 First, as shown in the black and blue bars in Fig. 3c, the population-based choice probabilities
239 for '1st value' and '2nd value' were larger than the average choice probabilities in the single
240 cells with non-zero weights (a total of 512 or 16.0 cells per session for '1st value', and 621 or
241 19.4 cells per session for '2nd value'). This indicates that within this epoch, the population-
242 based NVEs depended upon the summed activity of many cells that individually had only weak
243 choice-predictive activity. In contrast, for '1st chosen', the population-based and single-cell
244 choice probabilities were not different (gray bars). This can be explained by the fact that only 31
245 cells (mean 0.97 per session) had non-zero weights for '1st chosen' in this epoch; as a result, in
246 most sessions the NVEs for '1st chosen' reflected the activity of only a single cell, and
247 consequently that the single-cells and NVEs yielded virtually the same level of classifier

performance. The results were similar when each monkeys' data was analyzed individually (Fig. S3), and when using non-condition-normalized data to calculate AUCs (Fig. S4). Results were also similar when using the decision RT as the temporal reference (lever lift in Fig. 1a), and computing choice probability using the final 100ms of neural activity prior to the lever lift (Fig. S5).

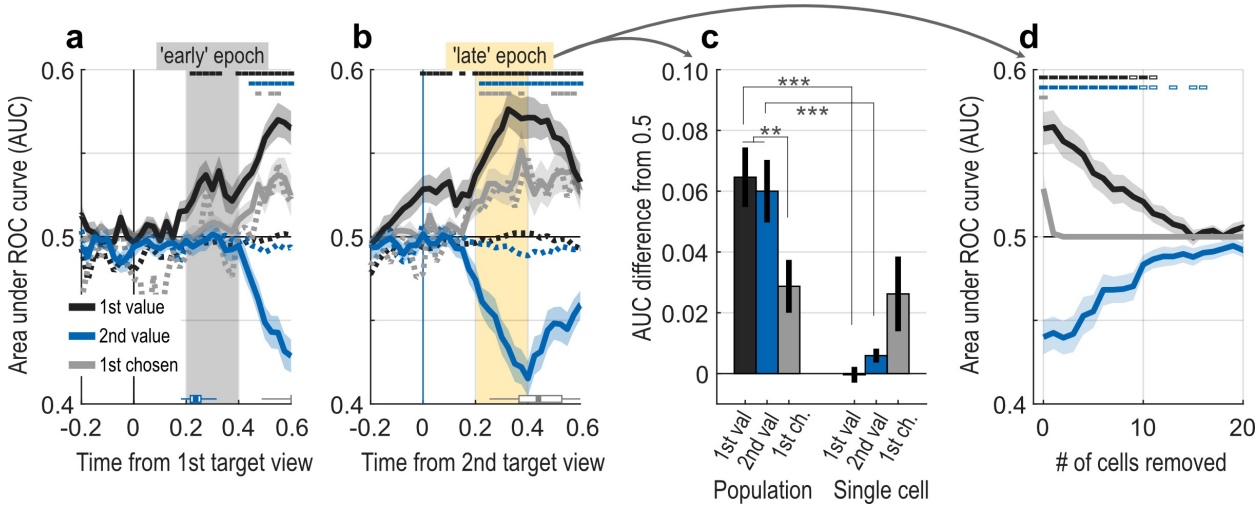



**Figure 3**: **Choice probability over time and neuron-dropping analysis (a-b)** Choice probability effects are time-locked to target viewing times. Population-based choice probabilities are given by solid lines with shading (means and SEMs, aggregated over 32 sessions and both monkeys). Small rectangles at top indicate bins for which the population effects are significantly different from 0.5 at a threshold of $p < 0.005$ (uncorrected) using a one-sample t-test. Dotted lines give the means of choice probabilities calculated from single cells individually (no SEM shown, numbers of cells vary by bin, see Methods). The blue and gray horizontal box plots give the distribution of second target viewing times and choice RTs, respectively. The shaded epochs in panels (a) and (b) indicate epochs used in subsequent analyses (see main text). **(c)** Comparison of population and single-cell choice probabilities during the late epoch in panel (b). To facilitate comparison across variables, AUCs were transformed to the same positive scale; this was done by either subtracting 0.5 from the AUCs (for '1st value' and '1st chosen') or by multiplying the AUCs by -1 and then adding 0.5 ('2nd value'). ** and *** indicate significant differences by t-test at uncorrected thresholds of $p < 0.01$ and $p < 0.001$, respectively. The population data show the mean and SEM for 32 sessions; the single cell data show the mean and SEM of 512 cells for '1st value', 621 cells for '2nd value', and 31 cells for '1st chosen'. **(d)** Neuron dropping analysis for choice probability effects in the late epoch. Line colors indicate variables as in panel (a). Lines and shading give mean and SEM over 32 sessions. Open and closed rectangles indicate significant difference from 0.5 at uncorrected thresholds of $p < 0.05$ and $p < 0.005$ (respectively) by t-test. See Fig. S3 for data from individual monkeys.

255 We also used the late epoch data to estimate the fraction of cells contributing to the population-
256 based choice probability effects, by iteratively removing the cells with the highest absolute
257 encoder weights in each session. Classification performance for the two value variables
258 degraded in approximately linear fashion, falling to near-chance levels after the removal of
259 about ~10 cells on average (Fig. 3d, black and blue lines). In contrast, choice probabilities for
260 '1st chosen' fell to chance after removal of just one cell (Fig. 3d, gray line), consistent with the
261 fact that on average only a single cell per session was assigned a non-zero weight. These
262 findings were consistent in both monkeys (Fig. S3), and were similar when using non-condition-
263 normalized data to calculate AUCs (Fig. S4). Choice probabilities did not depend upon the total
264 number of cells recorded in each session (Fig. S6a), and did not drastically change when using
265 either a smaller or larger pool of cells isolated under stricter or more liberal sorting quality
266 criteria, respectively (Fig. S6b-c). We conclude that, within the sampling regime of this study,
267 choice probability effects depend upon value signals distributed relatively evenly over a modest
268 fraction of recorded cells (~20% on average), and that sampling from a significantly larger pool
269 of neurons in the same regions of OFC would not necessarily produce large gains in classifier
270 performance.
271
272 Within the late epoch in Fig. 3b, the representations of '1st value' and '2nd value' were nearly
273 orthogonal, both with respect to the subspaces encoding the value variables, and, critically, with
274 respect to the trial-wise fluctuations in value signals. Specifically, there was no correlation
275 between the cell weights for '1st value' and '2nd value' ($r = -0.006$, $p = 0.81$, for $n = 1{,}450$ cells
276 pooled across 32 sessions, Fig. S7a), and only a small negative correlation between their NVEs
277 ($r = -0.075$, $p < 1e-10$, for $n = 9{,}920$ test trials pooled across 32 sessions, Fig. S7b). We also
278 compared these late epoch value representations to representations of '1st value' in the early
279 epoch in Fig. 3a; the results are summarized in Fig. S7c-f. The most notable finding was that the
280 '1st value' NVEs in the early epoch were only weakly correlated with the two late epoch NVEs

281 ($R^2 < 0.008$ for both). The overall lack of correlation between the NVEs suggests that they are
282 separable, and that they may therefore make an independent contribution towards the trial-to-
283 trial variability in choice outcomes.
284
285 To quantify the net choice-predictive effects of the NVEs at both early and late epochs, we fit a
286 model that explained choice outcomes through a combination of both behavioral and neural
287 predictors. Using logistic regression, the fraction of first target choices was explained as a
288 function of the difference in target values, and, critically, the NVEs for '1st value' at the early
289 epoch, as well as NVEs for both '1st value' and '2nd value' at the late epoch (Eqn. 5). As
290 expected, the mean regression estimate for the effect of relative target value was significantly
291 positive (mean 1.50, SEM 0.09, 32 sessions, $p < 1e-10$), indicating that the greater the relative
292 value of the first-viewed target, the greater the fraction of first target choices. For the '1st value'
293 NVEs, the session-wise average regression estimates were significantly greater than zero, both
294 at the early (0.343, SEM 0.100, n = 32 sessions, $p = 0.001$ by t-test) and the late epochs (0.660,
295 SEM 0.141, $p = 0.0001$), with the late estimates significantly larger than the early estimates
296 (mean difference 0.317, SEM 0.154, $p = 0.049$ by paired t-test). For '2nd value' NVEs,
297 regression estimates were significantly less than zero (-0.555, SEM 0.100, $p < 1e-5$). Thus,
298 when considered within the same neural-behavioral model, variability in neural value signals at
299 both time points explain variability in choice outcome, consistent with their making independent
300 contributions to behavior.
301
302 **DISCUSSION**
303 In summary, we show for the first time a trial-to-trial relationship between variability in OFC
304 value representations and variability in value-based choices in primates performing at
305 behavioral threshold. This newly identified neural-behavioral link complements important prior
306 work that links OFC value representations to decision reaction times[11] and to choice behavior at

307 the session level[10], as well as detailed analyses of single-cell choice probability effects in
308 different OFC sub-populations[22,24]. Thus, our findings are significant because they fill a critical
309 missing piece of a long-standing puzzle in neuroeconomics – i.e. they show how cell-level
310 representations of value may contribute to the outcome of value-based choices at the resolution
311 of single trials.
312
313 Beyond the use of ensemble recording and analysis, two design elements in this study likely
314 contributed to the robust choice probability effects we identified. First, equal-value choice
315 conditions provided maximal choice variability, and therefore maximal leverage for detecting a
316 correspondence between choice outcome and neural value signals[30]. Moreover, when the net
317 evidence for a given decision is weak, behavior is presumably more sensitive to small
318 fluctuations in neural activity – again providing ideal conditions for identifying neural-behavioral
319 correlations[1]. Second, the fact that the target values were revealed serially (based on the
320 monkey's gaze sequence) permitted us to isolate neural activity related to specific choice
321 targets. Together, our approach permitted single-trial estimates of subjective value to be
322 obtained for specific choice targets, and correlated with the variable choices of animals
323 performing at threshold levels.
324
325 With respect to decision mechanisms, the results give several direct insights into how OFC
326 value signals may be read out by downstream circuits that are more proximal to the ultimate
327 motor output. First, they suggest a primary role for cells encoding value signals ('1st value' and
328 '2nd value'), and a smaller role for cells encoding choice behavior in binary fashion ('1st
329 chosen'). Second, they suggest a significant contribution from monotonically value-coding cells
330 that are read out via a simple weighted sum of their activity – although this does not preclude an
331 additional contribution from cells with non-monotonic or non-linear value encoing[35,36]. Third,
332 when both targets values are represented simultaneously (late epoch in Fig. 3b), they are

encoded by orthogonal neural populations whose trial-wise value representations are only minimally correlated. This suggests that value representations are nearly independent within OFC, and therefore linearly separable by downstream circuits[37].

Finally, variability in the '1st value' signals measured after viewing the first target were not strongly correlated with variability in '1st value' signals measured after viewing the second target Thus, even though a representation of the first target persists in OFC well after it is initially viewed, the stochastic variation present in the initial representation is not carried over into the late representation. Importantly, the value signals at both time points significantly explained variance in choice outcome. On the one hand, this suggests that downstream circuits may accumulate value evidence over time as targets are viewed, rather than instantaneously once the value of both targets is known[38,39]. However, late value signals explain more variance than early value signals (see Eqn. 5 and related results), even though a similar fraction of cells encodes '1st value' at each time point (Fig. 2d-e). In other words, the fraction of cells encoding value at a given time point did not necessarily correspond to the degree of choice-predictive activity. The mechanisms underlying this discrepancy are unclear, and will require further study. Possible explanations include the existence of output-potent and output-null subspaces within the OFC[40]; gated or time-varying evidence accumulation in downstream circuits[41]; or the presence of a 'leaky' evidence accumulation process, such that the influence of early evidence decays substantially over time[42].

The apparently linear read-out of value signals for value-based choice indicates potential similarities to mechanisms for perceptual motion discrimination, in which a linear read-out of motion-sensitive MT cells is sufficient to explain motion judgements[3,7]. The study of perceptual decision mechanisms may therefore provide a fruitful roadmap for future experiments. These include, for example: further examination of correlated noise among value-coding neurons,

359 especially with respect to cells that encode different choice targets[24,26]; the identification of
360 circuits downstream from OFC that may represent accumulated evidence [38,39]; and the use of
361 microstimulation to selectively target value-coding cells with choice-predictive activity, to enact a
362 direct causal test of their role in behavior[2]. In summary, the identification of robust value-based
363 choice probability has the potential to open many new avenues for understanding the
364 computational and neuronal mechanisms of economic choice.
365
366 **METHODS**
367 SUBJECTS AND APPARATUS
368 All procedures were performed in accordance with the NIH Guide for the Care and Use of
369 Laboratory Animals, and were approved by the Animal Care and Use Committees of Stanford
370 University, and Rutgers University – Newark. The subjects were two adult male rhesus
371 macaques designated K and C, weighing ~14kg each at the time of the study. Data from
372 Monkey K were acquired at Stanford University, and from Monkey C were acquired at Rutgers
373 University – Newark. They were implanted with MR-compatible head holders and recording
374 chamber (Crist Instruments, Hagerstown, MD); craniotomies were also performed to allow
375 access to the OFC. Monkey K had bilateral chambers and craniotomies, permitting bilateral
376 recordings. Monkey C had a single chamber and craniotomy overlying the left OFC. All
377 procedures were performed under full surgical anesthesia using aseptic techniques and
378 instruments, with analgesics and antibiotics given pre-, intra- and/or post-operatively as
379 appropriate. Data were collected while the monkeys were head-restrained and seated ~57 cm
380 from a fronto-parallel CRT monitor displaying the task stimuli (120Hz refresh rate, 1024x768
381 resolution). Three response levers (ENV-612M, Med Associates, Inc., St. Albans, VT) were
382 placed in front of the subjects, within their reach. The “center” lever was located approximately
383 21cm below the display center, and the other two levers were located approximately 8.5cm to
384 the left and right of the center lever. Stimulus presentation, reward delivery, and monitoring of

behavioral responses was controlled through MATLAB (Mathworks, Inc., Natick, MA), using the ROME toolbox (R. Kiani, 2012). Horizontal and vertical eye position was recorded noninvasively at 250Hz (Eyelink, SR Research, Mississauga, Ontario, Canada). Fluid rewards were delivered via a gravity-fed reservoir and solenoid valve. Neural activity, eye position, and task event data were acquired and stored using a Plexon Omniplex system (Plexon, Inc., Dallas, TX). Analysis was performed using custom routines in MATLAB 2019b and the R computing environment (version 4.0.2) with the *glmnet* package (version 4.0-2).

DECISION TASK

*Task structure*: Fig. 1a shows an abbreviated task sequence; Fig. S1a shows the full sequence as described here: Monkeys initiated a trial by fixating a central point on the display, and manually depressing the center lever (Fig. S1a). After holding fixation and the center lever for a variable duration of 1-1.5s, the fixation point disappeared, and two ‘target arrays’ appeared, centered 7.5 degrees of visual angle to the left and right of the display center (Fig. 1a). As soon as the arrays appeared, the monkeys were allowed to shift their gaze as they desired to view the targets, in any order and for any amount of time, until they initiated a choice. Eye position was monitored throughout, but had no programmatic effect on the trial outcome (with the exception of enforced central fixation necessary to initiate a trial). The monkeys were encouraged to view each target at least once, through the use of visual crowding and gaze contingent programming, described in detail below. The monkeys were free to choose at any time, by first lifting the center lever (extinguishing the targets), and then pressing either the left or right lever to choose the left or right target, respectively. After the left/right press, a juice reward of 1-5 drops was delivered, depending on the value of the chosen target (Fig. 1b). To discourage the monkey from engaging in significant deliberation (e.g. changes of mind) after the center lever lift, the interval between center lift and left/right press was limited to 400ms

410 (Monkey K) or 500ms (Monkey C); failure to respond within this deadline resulted in an aborted
411 trial.
412
413 *Choice targets*: The choice targets were colored glyphs, each a unique combination of one of
414 four colors and one of four shapes. A full example set of 12 targets is shown in Fig. 1b. New
415 targets sets were generated every 1-5 sessions, to minimize over-learning of the stimuli.
416 The available shape primitives and the set generation procedure is given in Fig. S1b.
417
418 Target values ranged from 1-5 drops of juice. Importantly, every value stratum contained at
419 least two unique targets, permitting us to offer choices between two distinct but equal-value
420 targets. Note that the term 'target value' refers to a target's reward association (1-5 juice drops),
421 and the term 'target identity' refers to a target's unique color/shape combination, of which there
422 are a total of 12 in every set (Fig. 1b). Targets and crowders (see below) were all 0.75x0.75
423 degrees of visual angle. The patch colors used for targets and crowders were selected to be
424 isoluminant. The nominal patch luminance for all colors used in Monkey C's sessions was ~2.7
425 cd/m^2 with negligible background luminance; and for Monkey K's session was ~22 cd/m^2 with
426 background luminance of ~4cd/m^2, as measured by a Tektronix luminance J17 photometer
427 with a J1820 head. Note that during data collection, the luminance of the target stimuli was
428 reduced by up ~50% of these nominal values (with no change to the crowders), in order to
429 better obscure the targets within the crowder array (Fig. 1c). Note that the luminance reduction
430 did not prevent the monkeys from identifying the targets, because choice performance was
431 nearly perfect for the easiest trials (Fig. 1d, value differences of 3 and 4).
432
433 *Target arrays, visual crowding, and gaze-contingent mask*: Two methods were used to
434 encourage the monkeys to directly fixate both targets prior to making a choice. First, each target
435 was surrounded by six 'crowders', which were randomly generated, multi-colored '+' or '×'

436 shapes that had no direct association with reward. (Fig. 1c). Crowders reduce the effectiveness
437 of peripheral vision[28,29], making it difficult to identify a target without using high-acuity foveal
438 vision. Second, we used gaze-contingent programming to initially obscure the targets until the
439 first eye movement was made: At the moment the target arrays first appeared on the display,
440 the two targets were masked with a randomly generated crowder (Fig. S1A). Therefore, no
441 information about target identity or value was initially available. At the moment the monkey's eye
442 position breached the central fixation window (diameter 3.5 degrees), both masks were
443 removed, and the choice targets were shown in their place, one each at the center of the two
444 target arrays. Because the display frame rate interval (~8ms) was shorter than the time needed
445 to complete a saccade (~20-40ms), the masks were usually replaced with the targets while the
446 eyes were still moving. Thus, the task design ensured that the monkeys could gain no
447 information about the target values until they initiated a saccade towards one of the two targets.
448 After the targets were shown, they remained in place until the monkey initiated a choice by
449 lifting the center lever, at which point both targets and all crowders were extinguished.
450
451 NEURAL RECORDINGS
452 Linear recording arrays (Plexon V-Probes) were introduced into the brain through a sharpened
453 guide tube whose tip was inserted 1-3mm below the dura. Probes had 16, 24, or 32 channels,
454 spaced either 50 or 100μm apart; up to 2 probes were used per hemisphere (32-80 channels,
455 mean 50 for monkey C, 75 for Monkey K). OFC was identified on the basis of gray/white matter
456 transitions, and by consulting a high-resolution MRI acquired from each animal. We targeted the
457 fundus and lateral bank of the medial orbital sulcus and the laterally adjacent gyrus (Fig. S2),
458 over anterior-posterior coordinates ranging from +32 to +38mm anterior to the intra-aural
459 landmark[43].
460

Neural unit signals were sorted by automated spike detection and sorting routines (Kilosort2, default parameters, except for ops.th = [4 12] and ops.lam = 15) followed by manual curation (Phy2 GUI). Sorted units were considered eligible for analysis if they had less than 2% of inter-spike intervals (ISIs) below 2ms, and also had less than 25% of spikes estimated to have come from 'rogue' units, according to the methods of Hill et al[44]. Based upon follow-up manual inspection (Plexon Offline Sorter), these criteria selected all well-isolated single units, as well as less-well isolated units lightly contaminated with rogue spikes. We collectively refer to units that meet these thresholds as 'cells', recognizing that they represent a mix of units of varying isolation quality. The key findings in this study were confirmed when using a more stringent 1% ISIs threshold, and 10% rogue spike threshold (Fig. S6b-c).

ANALYSIS

*Behavior and key task variables*: Choice performance was quantified by a logistic regression in which the fraction of left choices was explained by the left-minus-right target value (in drops of juice, Fig. 1d). The number of fixations per trial and the decision RTs were averaged by task difficulty, defined as the best-minus-worst target value (Fig. 1e,f). For neural data analyses, only the trials in which monkeys viewed both targets were used (99.3% of trials for Monkey C, and 93.2% of trials for Monkey K). Neural data analysis was based on three task variables defined in a viewing order-based reference frame: '1st value' and '2nd value' were defined as the number of juice drops associated with the first and second target (respectively) viewed in every trial; and '1st chosen' was defined as *true* when the monkey chose the first target he viewed in the trial, and *false* otherwise.

*Single cell generalized linear model (GLM)*: To quantify the fraction of cells encoding the variables above, we fit for each cell a stepwise linear model: First, we fit the following GLM, assuming a Poisson error function:

$$SPIKES = \beta_0 + \beta_1 * val1 + \beta_2 * val2 + RESID \quad (1)$$

Where *SPIKES* is the spike count of a single cell in a single time bin, *val1* and *val2* are the variables '1st value' and '2nd value' as defined above, and *RESID* indicates the model residuals. We then fit a second model:

$$RESID = \beta_0 + \beta_1 * ch1 \quad (2)$$

Where RESID is the residuals from the model fit in Eqn. 1, and *ch1* is the Boolean variable '1st chosen' defined above. As a result, all of the variance in spiking due to *val1* and *val2* was accounted for before assessing the effects of *ch1*; this ensured that the effects of *ch1* could only be attributed to this variable, and could not attributed in any way to collinearity between *ch1* and the two value variables. This approach was chosen because it mirrors that of the encoder/decoder models used for choice probability calculations (Eqns. 3 and 4 below).

The fraction of cells with a significant effect of each variable was quantified at an uncorrected threshold of $p < 0.001$, chosen to produce minimal false positives before target viewing. The GLM used spike counts observed in 100ms bins, moved at 25ms increments, using the viewing of the first and second targets as temporal reference points.

*Choice probability measured from population activity*: At a high level, our approach entails the classification of choice outcomes using population-level representations of value. While there are many potential approaches for binary classification, we chose choice probability to facilitate comparison to prior studies[1,22–24]. With regards to the calculation of population value signals, we chose a LASSO regression model[33], also known as L1-regularized linear regression. LASSO

513 was chosen because it enforces a sparse set of regression estimates – meaning in this case
514 that it clearly defines a subset of cells that contribute to the population value signal. As
515 described below, having a clearly defined subset of contributing cells facilitates a direct
516 comparison between population-based and single-cell choice probability effects.
517
518 Choice probability analysis proceeded in three steps: fitting a model to explain an objective
519 variable as a function of neural activity (encoding); using the model to estimate that objective
520 variable in held-out trials (decoding); and classification of choice outcomes using the decoded
521 estimates of the objective variable (classification). Note that the encoding and decoding steps
522 can be interpreted as a form of dimensionality reduction and projection: the fitted parameters of
523 the encoding model define an axis (a 1-D sub-space) within the neural state space that encodes
524 the objective variable, and the decoded estimates quantify the projection of neural activity onto
525 that axis.
526
527 Three objective variables were tested in separate analyses: ‘1st value’, ‘2nd value’, and ‘1st
528 chosen’, defined above. To show the time course of choice probability effects, the analysis was
529 performed using spike counts observed in 200ms bins, moved at 25ms increments, using two
530 different temporal reference points: the viewing of the first target in each trial, and the viewing of
531 the second target in each trial. Results were similar when using 100ms bins (not shown). These
532 steps were performed independently within each bin, meaning that fitted parameters could vary
533 across bins – in particular, that from one bin to the next, different cells could be assigned non-
534 zero weights in the LASSO model described in Eqn. 3 below. For some follow-up analyses we
535 defined an ‘early’ epoch, 0.2 to 0.4s after first target viewing; and a ‘late’ epoch, 0.2 to 0.4s after
536 second target viewing (Fig. 3a-b).
537

First, in the encoding step, the data were stratified into a training and test set, maintaining whenever possible equal numbers of trial types, defined by the values of the first- and second-viewed target in each trial. Then we fit the following L1-regularized linear model (LASSO)[33] to the training trials, using the `glmnet` package in R:

$$Y = \beta_0 + \beta_1 S_1 + \beta_2 S_2 \ldots \beta_N S_N \,, \qquad (3)$$

where $Y$ is one of the three objective variables described above ('1st value', '2nd value', or '1st chosen'), $S_i$ is the spike count observed from cell *i* in a given bin, and *N* is the number of simultaneously recorded cells in a session. The resulting fitted parameters, $\{\beta_1, \beta_2, \ldots \beta_N\}$, can be interpreted as a set of weights that indicate the degree to which the firing of each cell uniquely contributes to explaining variance in the dependent variable $Y$. The model was fit by minimizing the negative log likelihood plus a penalty $\lambda|\{\beta_1, \beta_2, \ldots \beta_N\}|_1$, where $||_1$ indicates the L1 norm. The scaling parameter $\lambda$ (lower-bounded by zero) was selected to minimize cross-validation error (10-folds) within the training trials.

This form of regularization results in a sparse set of weights[33], such that the weights for cells that explain negligible variance are set to zero. In this way, the model unambiguously selects a subset of cells (those with non-zero weights) whose activity contributes to the encoding of the objective variable. Having a clearly defined subset of cells is important, because it allows the population-based choice probability effects to be directly compared to single-cell choice probabilities over the exact same cells (see below). Note that the cell weights were fit independently in every 200ms time bin, meaning that different subsets of cells could be defined at each bin. While the weights in Eqn. 3 were generally consistent with the single-cell GLM

562 estimates obtained from Eqn. 1, they were not identical, due to the fact that in Eqn. 3, a given
563 cells' weight depends on the activity of other cells fit in the same model. For details, see Fig. S8.
564
565 Before fitting Eqn. 3, the effects of non-relevant variables were removed from the spiking data
566 (training trials only), so that non-relevant variables would not influence the weights pertaining to
567 the objective variable. This was most important for the epochs where more than once variable
568 was being represented in neural activity (e.g. following fixations onto the second target, Fig. 2e),
569 but was performed at all time bins for the sake of consistency. When $Y$ was the variable '1st
570 value', our goal was to remove from the spiking data all variance associated with the value of
571 the (non-relevant) second-viewed target. This was necessary because the distribution of the
572 variable '2nd value' was not guaranteed to be identical across every stratum of the variable '1st
573 value'. As a result, changes in firing driven by '2nd value' could be inappropriately attributed to
574 '1st value', leading to inaccurate regression weights. Therefore, before model fitting, the spiking
575 data were normalized with respect to '2nd value', by grouping trials according to the identity of
576 the second-viewed target, and then mean-centering each of those groups. Likewise, when $Y$
577 was the variable '2nd value', the spiking data were normalized according to '1st value' in similar
578 fashion. When $Y$ was the variable '1st chosen', the goal was to remove variance associated with
579 either target value, in order to isolate spiking activity related only to choice outcome. To do this,
580 spiking data were normalized according to *both* '1st value' *and* '2nd value' simultaneously, by
581 obtaining the residuals from a two-way factorial ANOVA, in which spiking is explained by the
582 both identity of the first-viewed target and the identity of the second-viewed target, with no
583 interaction term.
584
585 Next, in the decoding step, we compute the following weighted sum in a set of held-out test
586 trials:
587

588 $$\hat{Y} = \beta_0 + \sum_{i=1}^{N} \beta_i S'_i \quad (4)$$

589

590 where $\beta_0$ and $\beta_i$ are the intercept and fitted parameters from Eqn. 3, and $S'_i$ is the spike count
591 from the *i*-th cell, and *N* is the total number of cells. This yields for every held-out trial the
592 weighted sum of spike counts across all simultaneously observed cells. In their raw form, these
593 weighted sums estimate the objective variable $Y$ in Eqn. 3; for example, when the objective
594 variable is '1st value', $\hat{Y}$ is on average, a monotonically increasing function of the first-viewed
595 target value (Fig. 2f.) However, the goal of this analysis is to consider the *trial-to-trial variability*
596 in $\hat{Y}$, i.e. the degree to which the decoded estimates exceed or fall short of their expected value
597 from one trial to the next. To isolate this variability, we normalize $\hat{Y}$ by taking the residuals
598 obtained from a two-way factorial ANOVA, in which $\hat{Y}$ is explained by the both identity of the
599 first-viewed target and the identity of the second-viewed target (with no interaction term).
600 This normalized $\hat{Y}$ is referred to as the Normalized Variable Estimates (NVEs) for the objective
601 variable (Fig. 2g). Note that normalization is performed with respect to the 12 unique stimulus
602 *identities*, rather than to the 5 levels of reward value (Fig. 1b.); this is necessary to control for
603 the possibility that the monkeys may idiosyncratically assign different subjective values to
604 targets that have equal reward associations.

605

606 The NVEs therefore reflect trial-to-trial variability in the population-level representations of the
607 objective variable, and serve as the input data for classification, as follows: The NVEs are used
608 to classify held-out trials according to whether the first target was chosen, using standard
609 receiver-operating characteristic (ROC) analyses. While encoding and decoding were
610 performed over all of the training and test trials (respectively), classification was only performed
611 over test trials for which the two targets differed by 0 or 1 drop of juice, because these were the
612 only trials with sufficient choice variability (mean 145 trials per session for Monkey K; 175 trials

613 per session for Monkey C). This produced for every session one area under the ROC curve
614 (AUC). For all three objective variables, the positive class was defined as choices in favor of the
615 first-viewed target. Therefore, the AUC was above 0.5 when higher NVEs corresponded to a
616 greater tendency to select the first target, and was below 0.5 when higher NVEs were
617 associated with a tendency to select the second target.
618
619 *Choice probability in individual cells*: This analysis used the same binned spiking data as the
620 population-based choice probability analysis described above, except that it was performed
621 individually for each cell. Before analysis, spike counts that were first normalized with respect to
622 the target identities, using a two-way factorial ANOVA in which spiking is explained by both the
623 identity of the first-viewed target and the identity of the second-viewed target, with no interaction
624 term. This is the same normalization used to transform decoded value signals (obtained from
625 Eqn. 4) into NVEs. ROC calculations were based upon those trials in the test set in which the
626 target values differed by 0 or 1 were used.
627
628 To directly compare the single cell with population results, average choice probabilities were
629 calculated for three groups of cells: cells with non-zero weights in Eqn. 3 when the objective
630 variable was ‘1st value’; cells with non-zero weights for the objective variable ‘2nd value’; and
631 cells with non-zero weights for the objective variable ‘1st chosen’. In other words, mean choice
632 probabilities were calculated for the cells that contributed to the NVE calculations for each
633 respective variable. The positive class of each cell was determined by the sign of the weight
634 from Eqn. 3: when the weight was above zero, the positive class was defined as choices in
635 favor of the first-viewed target; and when the weight was below zero, the positive class was
636 defined as second target choices. This ensures that the direction of the single-cell ROC results
637 (i.e. whether the AUC is above or below 0.5) have the same interpretation as the population-
638 based ROC results. Note that the three groups of cells were not mutually exclusive: an

individual cell could be used to calculate mean choice probabilities for more than one variable, if it was assigned non-zero weights for more than one variable.

*Choice probability calculated with non-condition-normalized data:* We confirmed the main findings of the study using alternative method for calculating AUCs (Fig. S4). In this method, trials were grouped according to condition, and the non-condition-normalized data (either $\hat{Y}$ in Eqn. 4 or the spike counts for single neurons) were used to calculate a separate AUC within each condition. Then, within each session the AUCs were averaged across conditions to arrive at the session-wise AUC. As expected from prior work[34], with this method the resulting AUCs were slightly higher than AUCs calculated from normalized data as described above. Note that using non-condition-normalized data typically requires that every condition have a sufficient number of trials to calculate a per-condition AUCs, and that this was generally *not* the case in our study: Each condition was defined by the identities of the first- and second-viewed targets. With 12 unique targets, there were 66 conditions (12 choose 2), and therefore a relatively small number of trials per condition. For conditions with too few trials, no AUC calculation could be performed, and as a result these trials were essentially discarded, and made no contribution towards to calculation of the session-wise average AUC. Specifically, only 34.6% of trials were used to calculate AUCs using non-condition-normalized data; for this reason, the main analyses use normalized data.

*Neuron dropping*: The population choice probability analysis above was repeated over a successively smaller pool of cells. Within a session the *N* recorded cells were ranked according to the absolute value of the weights obtained in Eqn. 3. Then, the cell with the highest absolute weight was removed from the pool, and decoding and classification steps were performed, yielding an AUC based upon the *N-1* remaining cells. This procedure was repeated until no cells remained. Results are plotted for the first 20 removed cells.

*Logistic regression with behavioral and neural regressors*: The fraction of first target choices in the test trials was explained by the following logistic regression:

$$logit(ch1) = \beta_0 + \beta_1(val1 - val2) + \beta_2 NVE1_{early} + \beta_3 NVE1_{late} + \beta_4 NVE2_{late} \quad (5)$$

where logit() is the logit function, *ch1* is a Boolean variable indicating a choice in favor of the first-viewed target, *val1* and *val2* are variables ‘1st value’ and ‘2nd value’ defined above, *NVE1* and *NVE2* refer to the NVEs for ‘1st value’ and ‘2nd value’ defined above, and the subscripts indicate data measured at the ‘early’ and ‘late’ epochs indicated in Fig. 3a-b. The model was fit for each session, and then the regression estimates were averaged across sessions. Similar results (not shown) were obtained from a regression that also included the NVEs for ‘1st chosen’ measured at the ‘late’ epoch; and also from a regression that includes a Boolean variable that takes ‘true’ when the monkey selected the first-viewed target on the *previous* trial, and ‘false’ otherwise. This latter term accounts for any potential effects of choice hysteresis, i.e. the potential for the monkey’s choice to be influenced by the choices in the trial immediately prior.

## SUPPLEMENTARY FIGURES

### FIGURE S1

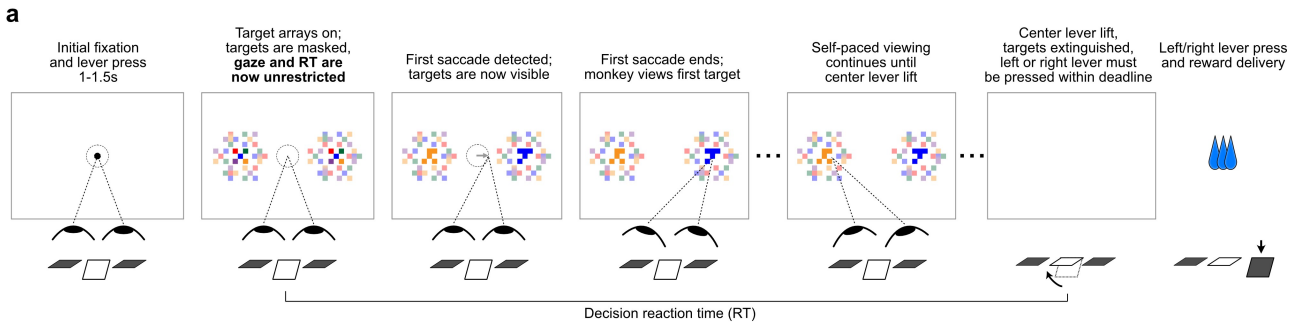


**Figure S1**: **(a)** Illustration of the complete task sequence. Note that for illustration purposes, in this panel the background is white, and the surrounding crowders are shown at reduced saturation relative to the centrally located targets. However, on the actual task display, the background was dark, and target luminance was reduced relative to the crowders (see Fig. 1c), in order to better obscure the targets when they are not being viewed directly. **(b)** The 7 shape primitives used to generate choice stimuli. To create a new stimulus set, four shapes are selected without replacement from the 7 primitives; then each is rotated randomly by 0, 90, 180, 270 degrees. The four rotated shapes are then combined with four colors selected randomly from equally-spaced points on a color wheel defined within the RGB gamut in the CIELUV color space. A sub-set of 12 targets is then selected, and assigned to reward values ranging from 1-5 drops (Fig. 1b).

**FIGURE S2**

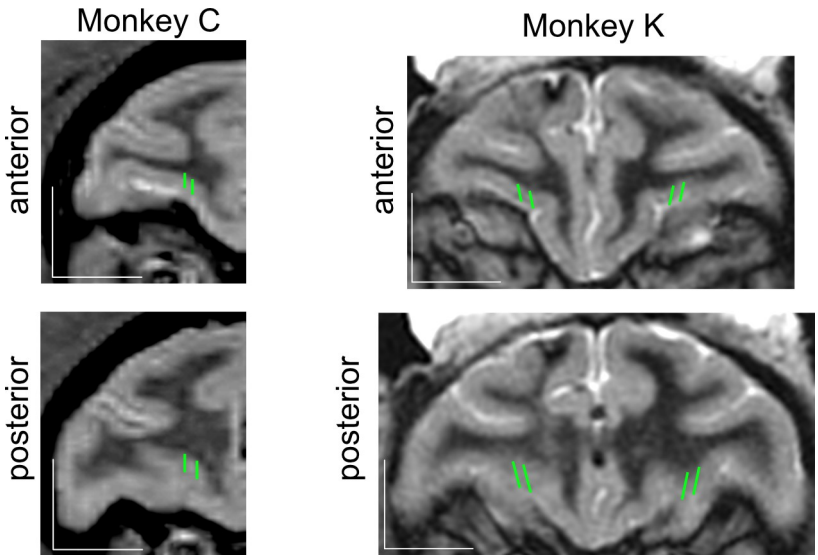
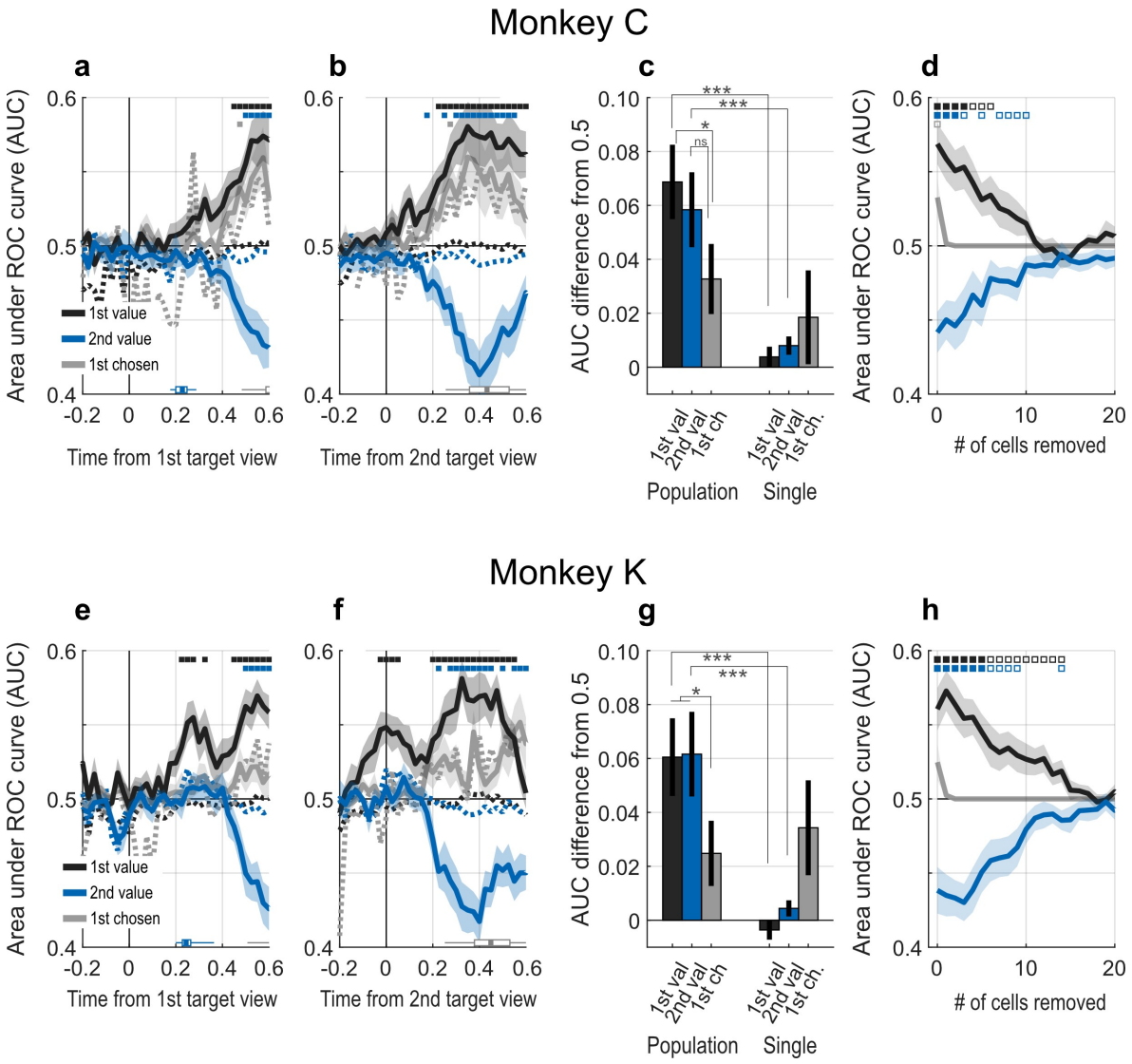



**Figure S2**: OFC recordings targeted the medial orbital gyrus, as well as the lateral bank of the medial orbital sulcus, over an anterior-posterior range of approximately +32 to +38mm relative to the intra-aural landmark. A representative anterior and posterior MR section (T2-weighted) is shown for each animal, with representative array trajectories shown in green. In Monkey C, recordings were from the left hemisphere only; in Monkey K, recordings were typically obtained bi-laterally. Scale bars are 10mm.

713

714 **FIGURE S3**




715

716 **Figure S3**: Choice probability and neuron dropping analyses for individual monkeys. Same

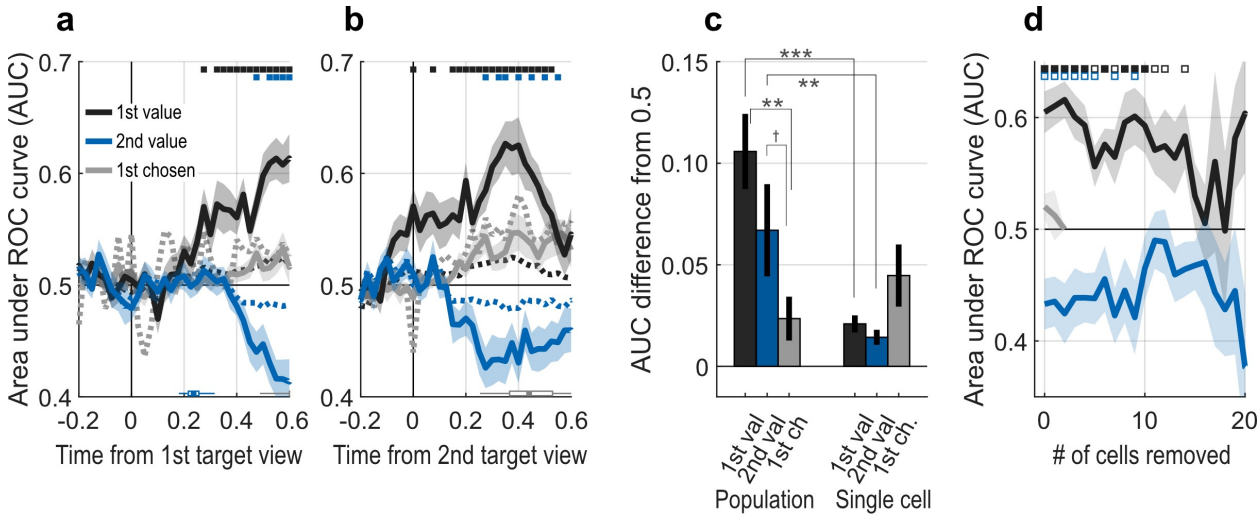
717 format and conventions as Fig. 3a-d. In panels (c) and (g), * and *** indicate significant

718 differences at uncorrected thresholds of $p < 0.05$ and $p < 0.001$, respectively. 'ns' indicates

719 $p>0.05$.

720

721 **FIGURE S4**




722

723 **Figure S4:** Choice probability and neuron dropping analyses when using non-condition-
724 normalized data to calculate the area under the ROC curve (see Methods). Note the overall
725 larger effects in comparison to Fig. 3. In panel (c), † indicates p = 0.07. Otherwise, all
726 conventions are the same as in Fig. 3a-d. Note that the single-cell choice probabilities for the
727 '1st value' and '2nd value are comparable to those reported in prior studies that calculate AUCs
728 with non-condition-normalized data: Specifically, Conen and Padoa-Schioppa[24] report mean
729 single-cell choice probabilities of 0.013 greater than chance, with a standard error of 0.007 in a
730 sample of 229 cells (their Figure 7A), implying a standard deviation of ~0.11 in their sample. In
731 panel (c), the mean choice probability for '1st value' is 0.021 greater than chance, with a
732 standard deviation of 0.096 in sample of 512 cells; and for '2nd value' the mean choice
733 probability is 0.014 above chance, standard deviation 0.091, in a sample of 621 cells. Thus, in
734 our study, single OFC cells exhibit choice probabilities that do not appear drastically different
735 from previous reports both in terms of the mean effects and standard deviation across cells.
736

**FIGURE S5**

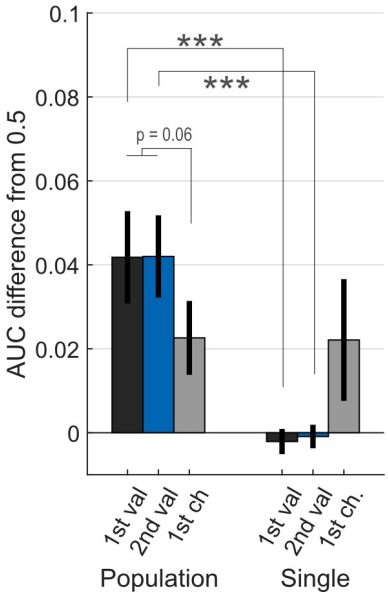



**Figure S5** Choice probability effects, placed on the same positive scale, using the data from the final 100ms before the decision RT in each trial. Conventions are the same as in Fig. 3c.

**FIGURE S6**

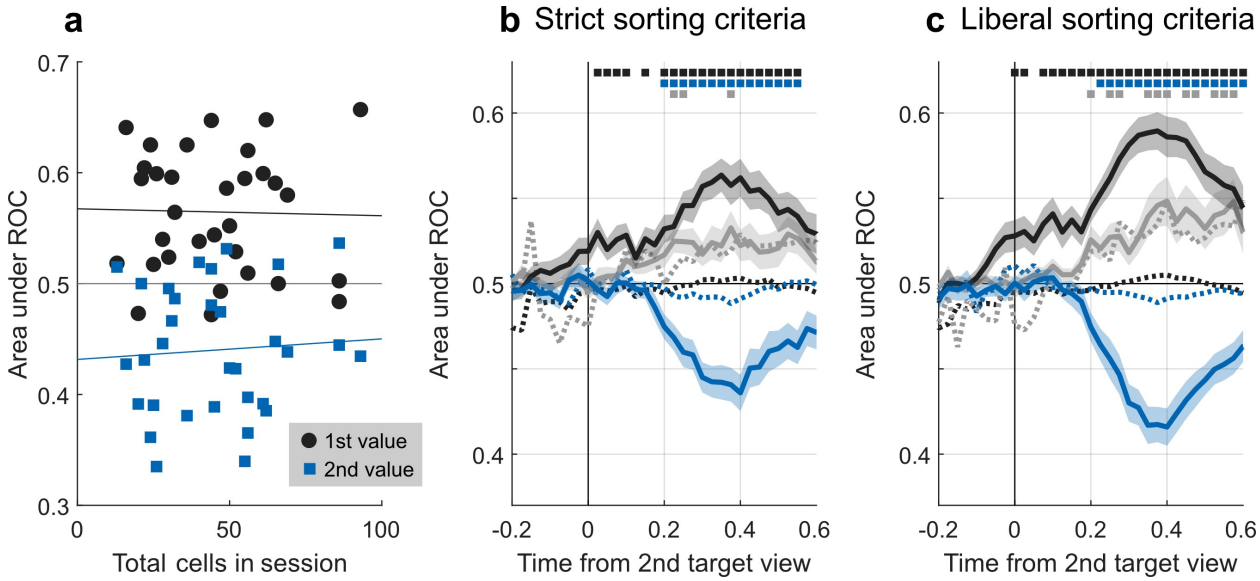



**Figure S6**: **(a)** AUCs as a function of the number of cells recorded in each session, using the data from the 'late' epoch in Fig. 3b. Solid lines give linear fit. For '1st value' the correlation was -0.024 (p = 0.90), and for '2nd value' was 0.067 (p = 0.71). **(b-c)** Choice probability effects referenced to second target viewing, using cells meeting 'strict' sort quality criteria (<1% of ISIs shorter than 2ms, and <10% of ISIs estimated to come from rogue spikes) or 'liberal' sorting criteria (<50% of ISIs from rogue spikes, no restriction on ISIs shorter than 2ms). Mean cells per session for strict criteria were 23.9, and for liberal criteria were 77.6. Conventions are the same as in Fig. 3b.

**FIGURE S7**

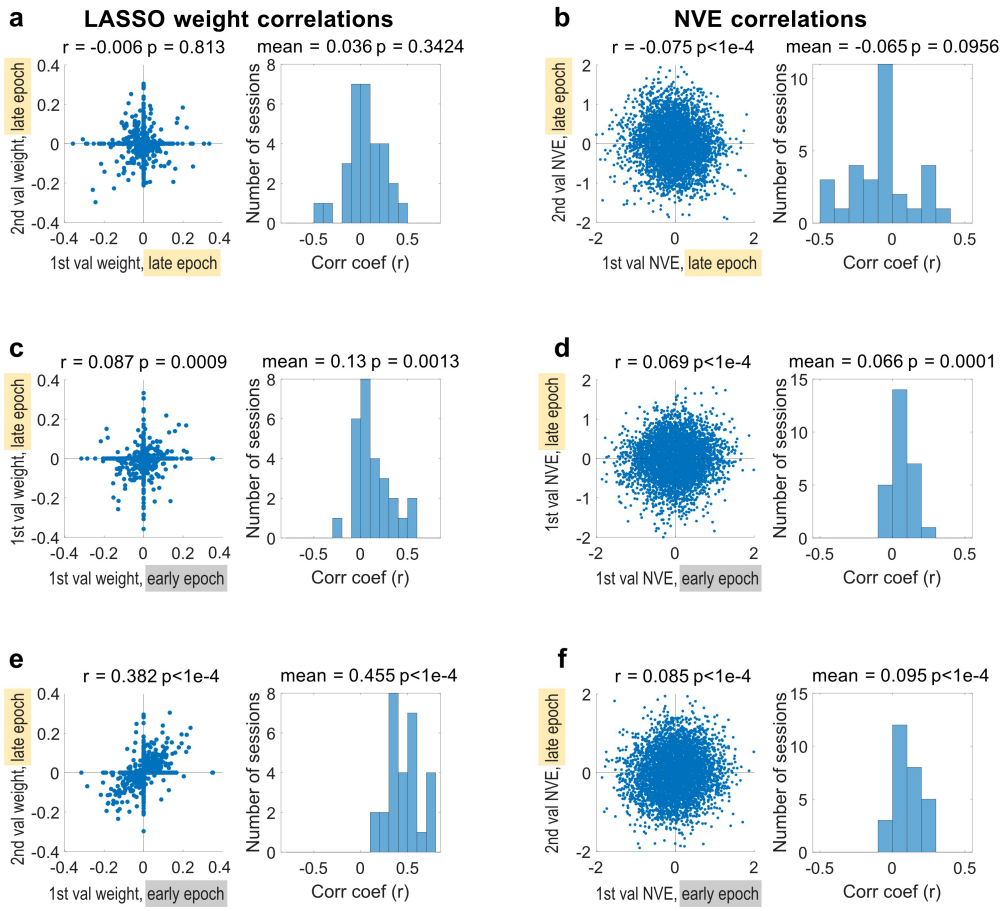



**Figure S7**: Correlations between LASSO regression weights from Eqn. 3 (panels a, c, e); and between NVEs (panels b, d, f). Each panel shows a scatter plot in which data are collapsed across all 32 sessions, and a histogram of correlations calculated within each session. All data are from either the 'late' epoch in Fig. 3b, or the 'early' epoch in Fig. 3a. **(a)** Correlation between weights from '1st value' and '2nd value' LASSO models, using data from the late epoch. On the scatter plot, each dot indicates a single cell (n = 1,450 total); note that the many cells have zero weights due to the L1 regularization used by the LASSO model. **(b)** Correlation between NVEs for '1st value' and '2nd value', from the late epoch. On the scatter plot, each dot indicates a single test trial (n = 9,920 total). **(c-d)** Correlations between '1st-value' LASSO weights (c) and between NVEs (d) measured at two different time points. The first time point is the early epoch (x-axis), and the second time point is the late epoch (y-axis). Conventions are otherwise the same as in panels (a) and (b). **(e-f)** Correlations between LASSO weights (e) and between NVEs (f), for '1st-value' cells during the early epoch (x-axis), and '2nd-value' cells during the late epoch (y-axis). Conventions are the same as in panels (a) and (b).

**FIGURE S8**

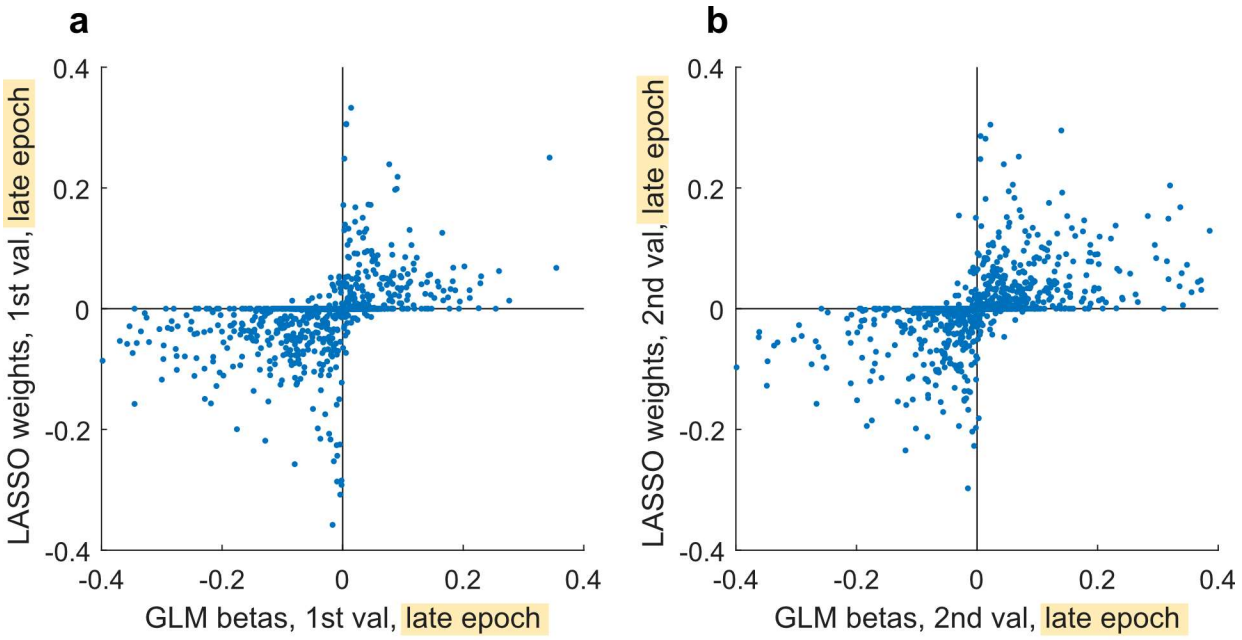



**Figure S8**: Comparison of regression estimates obtained from single-cell GLMs (Eqn. 1) with the weights assigned by the LASSO model (Eqn. 3), for all 1,450 cells recorded in the study. **(a)** GLM estimates and LASSO weights for the variable '1st value' at the late epoch defined in Fig. 3b. **(b)** GLM estimates and LASSO weights for the variable '2nd value' at the late epoch defined in Fig. 3b. Each dot indicates a single cell. While there is general consistency between these measures, especially in terms of their signs, the correlations were modest: for '1st value r = 0.384, and for '2nd value' r = 0.416. Correlations were also modest between the absolute values (not illustrated): for '1st value' r = 0.222, and for '2nd value' r = 0.249. In other words, cells identified as strongly value coding by the single-cell GLM were not guaranteed to be identified as strongly value coding in the LASSO model. This discrepancy is explained by the fact that the LASSO model considers the *joint* contribution of all of the cells in a session, and uses a penalty term to shrink the weights of cells that don't contribute uniquely to explaining variance in the objective variable. The GLM, in contrast, evaluates cells one at a time, and returns an estimate that does not depend on the activity of other cells recorded simultaneously. Thus, a cell that is strongly modulated by value, but that is also strongly correlated with other value-coding cells in the same session, may have a smaller weight in the LASSO than in the GLM. Likewise, a cell that is only weakly modulated by value, but that explains a portion of the model variance in Eqn. 3 not accounted for by other cells may have a larger LASSO weight than GLM estimate.

## REFERENCES


1. Britten, K. H., Newsome, W. T., Shadlen, M. N., Celebrini, S. & Movshon, J. A. A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis. Neurosci.* **13**, 87–100 (1996).
2. Salzman, C. D., Britten, K. H. & Newsome, W. T. Cortical microstimulation influences perceptual judgements of motion direction. *Nature* **346**, 174–177 (1990).
3. Yates, J. L., Katz, L. N., Levi, A. J., Pillow, J. W. & Huk, A. C. A simple linear readout of MT supports motion direction-discrimination performance. *J. Neurophysiol.* **123**, 682–694 (2019).
4. Katz, L. N., Yates, J. L., Pillow, J. W. & Huk, A. C. Dissociated functional significance of decision-related activity in the primate dorsal stream. *Nature* **535**, 285–288 (2016).
5. Gold, J. I. & Shadlen, M. N. The Neural Basis of Decision Making. *Annu. Rev. Neurosci.* **30**, 535–574 (2007).
6. Shadlen, M. N. & Newsome, W. T. Neural Basis of a Perceptual Decision in the Parietal Cortex (Area LIP) of the Rhesus Monkey. *J. Neurophysiol.* **86**, 1916–1936 (2001).
7. Cohen, M. R. & Newsome, W. T. Estimates of the Contribution of Single Neurons to Perception Depend on Timescale and Noise Correlation. *J. Neurosci.* **29**, 6635–6648 (2009).
8. Thorpe, S. J., Rolls, D. E. T. & Maddison, S. The orbitofrontal cortex: Neuronal activity in the behaving monkey. *Exp. Brain Res.* **49**, 93–115 (1983).
9. Tremblay, L. & Schultz, W. Relative reward preference in primate orbitofrontal cortex. *Nature* **398**, 704–708 (1999).
10. Padoa-Schioppa, C. & Assad, J. A. Neurons in the orbitofrontal cortex encode economic value. *Nature* **441**, 223–226 (2006).
11. Rich, E. L. & Wallis, J. D. Decoding subjective decisions from orbitofrontal cortex. *Nat. Neurosci.* **advance online publication**, (2016).
12. Yamada, H., Louie, K., Tymula, A. & Glimcher, P. W. Free choice shapes normalized value signals in medial orbitofrontal cortex. *Nat. Commun.* **9**, 162 (2018).
13. Ballesta, S., Shi, W., Conen, K. E. & Padoa-Schioppa, C. *Values Encoded in Orbitofrontal Cortex Are Causally Related to Economic Choices*. http://biorxiv.org/lookup/doi/10.1101/2020.03.10.984021 (2020) doi:10.1101/2020.03.10.984021.
14. Knudsen, E. B. & Wallis, J. D. Closed-Loop Theta Stimulation in the Orbitofrontal Cortex Prevents Reward-Based Learning. *Neuron* (2020) doi:10.1016/j.neuron.2020.02.003.
15. Murray, E. A., Moylan, E. J., Saleem, K. S., Basile, B. M. & Turchi, J. Specialized areas for value updating and goal selection in the primate orbitofrontal cortex. *eLife* **4**, e11695 (2015).
16. Rudebeck, P. H., Saunders, R. C., Prescott, A. T., Chau, L. S. & Murray, E. A. Prefrontal mechanisms of behavioral flexibility, emotion regulation and value updating. *Nat. Neurosci.* **16**, 1140–1145 (2013).
17. Setogawa, T. *et al.* Neurons in the monkey orbitofrontal cortex mediate reward value computation and decision-making. *Commun. Biol.* **2**, 126 (2019).

18. Eldridge, M. A. G. *et al.* Disruption of relative reward value by reversible disconnection of orbitofrontal and rhinal cortex using DREADDs in rhesus monkeys. *Nat. Neurosci.* **19**, 37–39 (2016).
19. Kennerley, S. W., Dahmubed, A. F., Lara, A. H. & Wallis, J. D. Neurons in the Frontal Lobe Encode the Value of Multiple Decision Variables. *J. Cogn. Neurosci.* **21**, 1162–1178 (2008).
20. McGinty, V. B., Rangel, A. & Newsome, W. T. Orbitofrontal Cortex Value Signals Depend on Fixation Location during Free Viewing. *Neuron* **90**, 1299–1311 (2016).
21. Hunt, L. T. *et al.* Triple dissociation of attention and decision computations across prefrontal cortex. *Nat. Neurosci.* **21**, 1471–1481 (2018).
22. Padoa-Schioppa, C. Neuronal Origins of Choice Variability in Economic Decisions. *Neuron* **80**, 1322–1336 (2013).
23. Kimmel, D. L., Elsayed, G. F., Cunningham, J. P. & Newsome, W. T. Value and choice as separable and stable representations in orbitofrontal cortex. *Nat. Commun.* **11**, 3466 (2020).
24. Conen, K. E. & Padoa-Schioppa, C. Neuronal variability in orbitofrontal cortex during economic decisions. *J. Neurophysiol.* **114**, 1367–1381 (2015).
25. Yoo, S. B. M., Sleezer, B. J. & Hayden, B. Y. Robust Encoding of Spatial Information in Orbitofrontal Cortex and Striatum. *J. Cogn. Neurosci.* **30**, 898–913 (2018).
26. Averbeck, B. B., Latham, P. E. & Pouget, A. Neural correlations, population coding and computation. *Nat. Rev. Neurosci.* **7**, 358–366 (2006).
27. Enel, P., Wallis, J. D. & Rich, E. L. Stable and dynamic representations of value in the prefrontal cortex. *eLife* **9**, e54313 (2020).
28. Crowder, E. A. & Olson, C. R. Macaque monkeys experience visual crowding. *J. Vis.* **15**, 14 (2015).
29. Whitney, D. & Levi, D. M. Visual crowding: a fundamental limit on conscious perception and object recognition. *Trends Cogn. Sci.* **15**, 160–168 (2011).
30. Fetsch, C. R. The importance of task design and behavioral control for understanding the neural basis of cognitive functions. *Curr. Opin. Neurobiol.* **37**, 16–22 (2016).
31. Strait, C. E., Blanchard, T. C. & Hayden, B. Y. Reward Value Comparison via Mutual Inhibition in Ventromedial Prefrontal Cortex. *Neuron* doi:10.1016/j.neuron.2014.04.032.
32. Rudebeck, P. H., Ripple, J. A., Mitz, A. R., Averbeck, B. B. & Murray, E. A. Amygdala contributions to stimulus–reward encoding in the macaque medial and orbital frontal cortex during learning. *J. Neurosci.* 0933–16 (2017) doi:10.1523/JNEUROSCI.0933-16.2017.
33. Tibshirani, R. Regression Shrinkage and Selection via the Lasso. *J. R. Stat. Soc. Ser. B Methodol.* **58**, 267–288 (1996).
34. Kang, I. & Maunsell, J. H. R. Potential confounds in estimating trial-to-trial correlations between neuronal response and behavior using choice probabilities. *J. Neurophysiol.* **108**, 3403–3415 (2012).
35. Rigotti, M. *et al.* The importance of mixed selectivity in complex cognitive tasks. *Nature* **497**, 585–590 (2013).
36. Saez, A., Rigotti, M., Ostojic, S., Fusi, S. & Salzman, C. D. Abstract Context Representations in Primate Amygdala and Prefrontal Cortex. *Neuron* **87**, 869–881 (2015).
37. Yoo, S. B. M. & Hayden, B. Y. The Transition from Evaluation to Selection Involves Neural Subspace Reorganization in Core Reward Regions. *Neuron* **0**, (2019).

875 38. Krajbich, I., Armel, C. & Rangel, A. Visual fixations and the computation and comparison of
876 value in simple choice. *Nat Neurosci* **13**, 1292–1298 (2010).
877 39. Shadlen, M. N. & Shohamy, D. Decision Making and Sequential Sampling from Memory.
878 *Neuron* **90**, 927–939 (2016).
879 40. Kaufman, M. T., Churchland, M. M., Ryu, S. I. & Shenoy, K. V. Cortical activity in the null
880 space: permitting preparation without movement. *Nat. Neurosci.* **17**, 440–448 (2014).
881 41. Okazawa, G., Sha, L., Purcell, B. A. & Kiani, R. Psychophysical reverse correlation reflects
882 both sensory and decision-making processes. *Nat. Commun.* **9**, 3479 (2018).
883 42. Usher, M. & McClelland, J. L. On the Time Course of Perceptual Choice: The Leaky
884 Competing Accumulator Model. 72.
885 43. Öngür, D. & Price, J. L. The Organization of Networks within the Orbital and Medial
886 Prefrontal Cortex of Rats, Monkeys and Humans. *Cereb. Cortex* **10**, 206–219 (2000).
887 44. Hill, D. N., Mehta, S. B. & Kleinfeld, D. Quality Metrics to Accompany Spike Sorting of
888 Extracellular Signals. *J. Neurosci.* **31**, 8699–8705 (2011).
889