Triple Dissociation of Attention and Decision Computations across Prefrontal Cortex

LT Hunt[#1,2,3,4], **WMN Malalasekera**[#1], **AO de Berker**[1,2], **B Miranda**[1,5,6], **SF Farmer**[1], **TEJ Behrens**[2,3], and **SW Kennerley**[1]

[1]Sobell Department of Motor Neuroscience, University College London

[2]Wellcome Centre for Human Neuroimaging, University College London

[3]Wellcome Centre for Integrative Neuroimaging, University of Oxford

[4]Oxford Health NHS Foundation Trust

[5]International Neuroscience Doctoral Programme, Champalimaud Foundation

[6]Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa

[#] These authors contributed equally to this work.

Abstract

Naturalistic decision-making typically involves sequential deployment of attention to choice alternatives to gather information before a decision is made. Attention filters how information enters decision circuits, implying attentional control may shape how decision computations unfold. We recorded neuronal activity from three subregions of prefrontal cortex (PFC) while monkeys performed an attention-guided decision-making task. From the first saccade to decision-relevant information, a triple dissociation of decision- and attention-related computations emerged in parallel across PFC subregions. During subsequent saccades, orbitofrontal cortex activity reflected value comparison between currently and previously attended information. By contrast, anterior cingulate cortex carried several signals reflecting belief updating in light of newly attended information, integration of evidence to a decision bound, and an emerging plan for what action to choose. Our findings show how anatomically dissociable PFC representations evolve during attention-guided information search, supporting computations critical for value-guided choice.

Anatomical[1,2], neuroimaging[3,4] and lesion studies[5,6] indicate that prefrontal cortex (PFC) is central to value-guided choice. These techniques have functionally localized subcomponents of decision making to different subregions of PFC. However, explanations of

neuronal computations within these subregions vary widely across studies. Recent debates on the role of PFC subregions in value-guided decision making have been manifold. One debate centres on whether decision-related computations are performed in serial (certain subregions preceding others) or parallel (simultaneous, distributed activity across subregions)7,8. A second concerns whether stimulus valuation in orbitofrontal cortex (OFC) and adjacent ventromedial prefrontal cortex may be influenced by attention9–11. Further debate relates to whether anterior cingulate cortex (ACC) integrates evidence for different actions12–15, modifies behaviour in light of new evidence16–19, or evaluates evidence for alternative courses of action20,21. Resolving these debates demands a rich dataset that contrasts neuronal activity across multiple PFC subregions within a single paradigm, whilst experimentally controlling the order, duration and frequency with which choice options are attended and compared.

Real-world choices are typically guided by multiple shifts in attention between choice alternatives. Interactions between attention, information search and choice have been widely studied in the behavioral sciences22–28, and the order, duration and frequency of shifts in visual attention can strongly influence the eventual decision made24. The evolutionary expansion of primate PFC relative to other species may have been driven by primates' need to foveate, evaluate, remember and compare alternatives during visually-guided foraging29. However, it remains largely unknown how attentional reorienting affects PFC computations performed at a neuronal level during choice30. This is because decision paradigms in neuroscience have been predominantly conducted with central or uncontrolled fixation, meaning attentional focus is not placed under experimental control. By determining which information enters decision circuits, attention will affect the temporal dynamics of several decision-related computations, including stimulus identification, valuation, comparison to previously attended alternatives, and action selection. Dissociating the neural substrates of decision-related computations across PFC may therefore require synchronizing neural activity with attentional focus.

Here we contrast neuronal activity between macaque orbitofrontal (OFC), anterior cingulate (ACC) and dorsolateral prefrontal cortices (DLPFC) during sequential attention-guided information search and choice. When attention is first deployed to a choice alternative, a triple dissociation of attention and decision computations emerges in parallel across these three areas. As further information is sampled, OFC carries representations required for comparing currently and previously attended information. By contrast, multiple signals in ACC reflect belief updating in light of new evidence and relative valuation of different actions. These signals ramp towards final commitment to a choice. Our findings are consistent with models describing value comparison as an attention-guided bounded diffusion process24, but also more recent accounts that frame economic choice as a series of accept-reject decisions31.

Results

Experimental paradigm and subject behaviour

Our task design (Fig. 1a) mirrored established behavioural studies examining attention-guided information search during sequential, multi-attribute choice22,25,26,32. Each *option*,

presented on left and right sides of the screen, comprised two pre-learned picture cues, representing different *attributes* - the probability and magnitude of juice reward. Crucially, at trial start, all cues were hidden. Subjects made an instructed saccade towards a highlighted location to reveal cue 1. Following 300ms uninterrupted fixation, cue 1 was covered and another location highlighted, either vertically on the same *option*, or horizontally on the same *attribute*. Subjects saccaded here to reveal cue 2, again for 300ms. Hereafter, subjects could select either *option* using a manual left/right joystick movement. Alternatively, they could fixate one or both remaining highlighted cues in any order to reveal further information before making their decision. Following joystick choice, all four cues were revealed and juice reward was delivered with the chosen probability and magnitude. Picture cues, first/second highlighted location, and probability/magnitude attribute on top vs. bottom were pseudorandomly selected on each trial (with uniform distribution). Behavioural and neuronal data were collected from two macaque monkeys (*M. mulatta*).

Both monkeys used cue values appropriately to guide their choices (Fig. 1b). They chose the option with higher expected value on 76.6% and 79.8% of trials (monkeys F (n=25 sessions) and M (n=32 sessions) respectively), assigning approximately equal weight to reward probability and magnitude, and using all viewed cues to guide their choice (Fig. S1a/b). However, most choices were based upon partial information: subjects chose before all four cues had been evaluated on 85.5%/71.4% of trials (subjects F/M respectively). Choice accuracy based upon the pictures *observed*, rather than the true expected value, was substantially higher (86.3% and 87.4% for monkey F and M respectively). Surprisingly, choice accuracy was also higher on trials where subjects sampled *fewer* pieces of information (Fig. S1c). This was because such trials were associated with a higher value difference, and so subjects terminated these trials more quickly (Fig. S1d).

Subjects preferred to sample information from the option that they currently intended to choose[25,26]. This behavior revealed itself in two ways. Firstly, subjects were free to choose *where* to attend with their third saccade. On 'attribute' trials, this saccade was preferentially directed towards the option with the higher relative expected value between cue 1 and cue 2 (compare bottom-left versus top-right of Fig. 2a). Such behavior mirrors a recently identified bias towards 'sampling the favorite' in an equivalent experiment in human participants[25], and mirrors classic 'confirmation biases' in human hypothesis testing[33]. Secondly, once two cues had been presented, subjects were also free to decide *when* to stop sampling information and commit to a final choice. On 'option' trials, subjects sampled fewest pieces of information when cues 1 and 2 were highest in value, but most information when cues 1 and 2 were lowest in value (compare bottom-right versus top-left of Fig. 2b). This mirrors a (milder) 'positive evidence approach' bias in humans[25].

The latter bias (Fig. 2b) appears particularly surprising. Two low-valued cues in an 'option' trial provide conclusive evidence for choosing the option not yet attended. Yet monkeys nonetheless sampled from this option before committing to choosing it. This behavior is suboptimal in the context of two-alternative forced choice, yet more rational in the context of real-world decisions that comprise multiple, non-mutually exclusive alternatives. Here, evidence against one option does not provide evidence in favor of any particular alternative. A natural strategy for solving such choices is to consider one option to be the leading or

'foreground' candidate, and decide whether to *accept* or *reject* it31. This accept/reject decision might still rely upon value comparison, for example to the next best alternative23,34 or the average reward rate of the environment35.

Triply dissociable PFC population codes at first saccade

We recorded single unit activity from anterior cingulate (ACC), dorsolateral prefrontal (DLPFC) and orbitofrontal (OFC) cortices (n=189, 135 and 183 neurons respectively; see Fig. 3). ACC recordings were primarily from the dorsal bank of the cingulate sulcus (area 24); OFC recordings were primarily from the medial orbital gyrus (area 13); DLPFC recordings were primarily from dorsal and ventral banks of sulcus principalis (area 9/46).

A critical feature of our experiment is that following each saccade, the currently attended cue can be decomposed into multiple features: its associated *attribute* (magnitude or probability), *value* (level of reward probability/magnitude), *spatial position* (presented on top/bottom of screen), and *action* (left/right joystick response required to choose that option). In line with previous studies36, we found a degree of PFC subregion specificity in single neuron encoding of these features (see below). However, there was substantial between-neuron *heterogeneity* of decision-related computations encoded. This heterogeneity proved critical in robustly dissociating computations performed by each subregion.

We capitalised upon neuronal heterogeneity by assessing population-level encoding of decision computations. At the time when cue 1 was attended, we used representational similarity analysis (RSA). RSA correlates the normalised firing rate of the neural population between all conditions of interest37. This characterises task encoding across the neural population without strong prior assumptions on its structure. At cue 1 presentation, we performed RSA between 20 conditions: 5 probability cues and 5 magnitude cues, presented on either the left or right option. Here, as neurons were not all simultaneously recorded, we perform the analysis on 'pseudopopulations'. For each subregion, we collapse across recording sessions, and calculate the correlation matrix from the resulting [Neurons*Conditions] matrix of firing rates.

RSA revealed a striking triple dissociation of task-evoked neural codes across PFC subregions (Figs. 4a-c). This was consistent across subjects (Fig. S2). To formally compare subregion specificity and temporal evolution of population representations, we regressed templates onto RSA matrices to capture different features of the task design. DLPFC RSA reflected whether the subject was attending left or right (Fig. 4d). OFC representational similarity reflected the currently attended stimulus identity, irrespective of spatial position (Fig. 4e), and was also high for cues of similar attended value (Fig. 4f). ACC and DLPFC RSA showed a value code modulated by whether the subject was currently attending to the left or right option (Fig. 4g). ACC RSA also divided high-valued and low-valued items such that variance in ACC was best explained as a non-linear, categorical function of value (Fig. 4h; labelled 'accept/reject' coding for reasons explored below).

For intuition, we provide single neuron examples for these features in Fig. S3. We also present RSA matrices subdivided by top/bottom spatial position in Fig. S4. It is important to acknowledge that there is not 'pure selectivity' for any one feature in a given region; for

example, spatial attention is represented in both OFC and ACC (Fig. 4d), and other task features have some degree of representation in multiple regions. Nonetheless, there is strong regional specificity in the degree to which different subregions encode each feature.

Decision-related computations at cue 1 emerged in parallel across PFC rather than sequentially (Fig. 4d-h). Supplementary Video 1 reveals the temporal order and evolution of these different computations. We also plot the timecourses of coefficient of partial determination (CPD) sorted by region in Fig. S5. We quantified the time at which information relating to different factors was encoded in different subregions by analysing when CPD in Fig. 4d-h reached 75% of its maximum value (t_{75}, see Fig. S5b/c). Spatial attention affected representational similarity around the time of the saccadic eye-movement (t_{75}=24ms in ACC, 72ms in DLPFC, 67ms in OFC). The early rise time of this effect can be attributed to saccade generation as cue onset was timelocked to the saccade. Following this, coding of stimulus identity and attended value (t_{75}=269ms/241ms respectively in OFC) was comparable in latency to accept/reject coding (t_{75}=238ms in ACC, 224ms in OFC). By contrast, action value coding emerged significantly later (t_{75}=457ms in ACC, 369ms in DLPFC). In summary, attentional modulation occurred at time of the saccade; stimulus identification, valuation and accept/reject coding emerged in parallel across OFC and ACC; and this was subsequently translated into action value.

Value encoding also differed between ACC and OFC. RSA in OFC was consistent with a linear representation of cue 1 attended value (Figs. 4a/f). Additional analyses confirmed this result's robustness to the exact formulation of the 'attended value' template (see Supplementary Note, Figs. S11/S12). We hypothesised that this graded signal in OFC may be a critical substrate to support comparison of the *currently attended* cue value versus previously attended (*stored*) cue values during subsequent saccades[9,10,38].

By contrast, ACC value coding was more non-linear and categorical (Fig. 4h). Guided by a recent literature on ACC encoding expectancy violations and adapting behaviour in light of new evidence[14,16–18,20,21,34], as well as the pattern of information sampling in Fig. 2, we hypothesized that ACC activity might reflect whether to accept or reject the current 'foreground' option. In particular, a high-valued cue 1 might confirm the belief that the first attended option should be accepted, not rejected. This option would remain the 'foreground' candidate, from which subjects will likely sample further information[25,31,33] (Fig. 2). By contrast, low-valued cues would *disconfirm* this belief, leading to the item being rejected, and the alternative becoming the foreground option. This hypothesis can be more robustly evaluated during subsequent saccades, as subsequent cues might confirm or disconfirm the current 'foreground' candidate as the best choice[34]. This signal might become particularly prominent prior to choice, when confirmatory evidence becomes sufficient to commit to an action.

Attention-guided value comparison in OFC

To address our hypothesis concerning attention-guided value comparison in OFC, we used multiple linear regression to evaluate how strongly each neuron encoded the values of cues 1, 2, 3 and 4 across time on both option and attribute trials. In Fig. 5a, we plot the average coefficient of partial determination (CPD, a measure of variance explained by each

regressor; see Methods), timelocked to each of the first three cues. This shows that value encoding by OFC neurons peaked approximately 300ms after each cue was revealed but was then sustained above baseline as further cues were attended. Note that CPD in these single-neuron analyses is considerably lower than in Fig. 4f, but comparable to other studies of value-based decision making39,40. This is because these values reflect variance explained *across trials* in each neuron, whereas values in Fig. 4f reflect variance explained across the neural population (having *averaged* across trials for each condition).

Crucially, neuronal encoding of value was highly variable across the population. We again capitalised upon this heterogeneity to define population task-related 'subspaces' for value encoding. Task-related subspaces can be defined by using linear regression to define how sensitive each neuron is to experimental variables of interest, and then projecting the data into a space defined by these regression coefficients41. This analysis can again be performed on 'pseudopopulations' of non-simultaneously recorded neurons, as the regression is performed separately (within-session) for each neuron, before collapsing across sessions to define the (pseudo)population subspace.

For example, we found that single neuron T-statistics for the regression of cue 1 value when cue 1 was attended (ordinate in Fig. 5b) correlated positively with T-statistics for cue 2 value when cue 2 was attended (abscissa in Fig. 5b). These two regressors are orthogonal and defined at different task epochs (using a window of 150-350ms post-stimulus onset for each cue). This analysis therefore reveals a stable population subspace for the *currently attended* cue value.

We repeated this approach for different phases of the task, to ask how the *currently attended* cue value subspace (ordinates in Figs. 5c-e) correlates with subspaces encoding previously attended, or *stored*, cues across time (abscissae in Figs. 5c-e). This revealed a signature of value comparison in OFC between currently and previously attended cues10. For example, when cue 2 was attended on 'attribute' trials, the currently attended cue 2 value subspace correlated negatively with the stored cue 1 value subspace, representing the other option (Fig. 5c). This negative correlation indicates that neurons encoding the value of the currently attended option at cue do so *relative* to the value of the previously attended option, a key prediction of recent theories of economic choice24,31. Similarly, when cue 3 was attended on 'option' trials, the stored cue 1 and stored cue 2 values both represented the other option and were both negatively correlated with currently attended cue 3 value subspace (Figs. 5d/e). However, these two stored subspaces were themselves positively correlated at cue 3 on 'option' trials (Fig. 5f). This demonstrates that the two previously attended cues were combined at cue 3 to allow comparison with the currently attended cue.

A more complete description of the interaction between attention and value can be obtained by plotting the cross-correlation of these subspaces across time (Figs. 5g-j). This reveals how the same OFC population subspace would dynamically shift its encoding of values from positive to negative as the subject saccaded around the screen. The letters superimposed on these plots refer back to the correlations shown in Figs 5b-f.

Importantly, this signature of attention-guided value comparison was unique to OFC. Whilst the *currently attended* value subspace was present in DLPFC and ACC, value comparison with *stored* cues was absent in these regions (Fig 5k; Figs. S6-7). A formal comparison of each of the three correlations of interest (corresponding to those shown in Figs. 5c-e) across the three subregions (Fig. 5k) revealed significantly stronger population encoding in OFC than in DLPFC (attribute trials: VCue2 vs. VCue1, $p = 0.040$; option trials: VCue3 vs. VCue1, $p = 0.067$; VCue3 vs. VCue2, $p = 0.011$; Z-test after Fisher r-to-Z transformation) and in OFC than in ACC (attribute trials: VCue2 vs. VCue1, $p = 0.003$; option trials: VCue3 vs. VCue1, $p = 0.0003$; VCue3 vs. VCue2, $p = 0.00013$).

Parallel ACC signals for belief confirmation, choice commitment and action selection

We then evaluated ACC population activity across cues 2 and 3, based upon our earlier interpretation that Fig. 4h may represent a belief confirmation signal for accepting or rejecting the 'foreground' (current best) option. To test this hypothesis more rigorously, we included four regressors in our regression model that capture belief confirmation at subsequent cues, on both option and attribute trials. Whenever the evidence presented thus far suggests that the *currently attended* side should be chosen, we hypothesised that belief confirmation would scale *positively* with currently attended value. By contrast, when the evidence suggests that the *unattended* side should be chosen, belief confirmation would scale *negatively* with value (see Fig. S8). As a consequence, all four belief confirmation regressors were by definition orthogonal to *currently attended* value (Fig. S9).

We used these regressors to test whether ACC reliably encoded belief confirmation. We found that ACC population subspaces for each of these regressors were significantly correlated with each other and also to cue 1 belief confirmation (Fig. 6a/Fig. S10). As all five regressors are defined at different parts of the trial, this reveals a stable population code in ACC for accepting/rejecting the current belief, which was not present in OFC or DLPFC (Fig. 6a). We again formally compared the correlations between these regressors across subregions, using a Fisher r-to-Z transformation (Fig. 6b). Virtually all of these correlations were stronger in ACC than OFC/DLPFC (Fig. 6b, right panels), and the majority of individual comparisons were significant (Fig 6b, left panels).

We next asked whether this belief confirmation subspace in ACC might support commitment to a final decision[12,15]. To answer this, we examined the temporal evolution of belief confirmation subspace activity, using the regressors in Fig. S8/Fig. 6a. We used one half of all trials to define the subspace, and projected the data from the remaining half into this subspace to examine its evolution across time. To ensure statistical robustness, we repeated this procedure using 100 random splits of the data to obtain a distribution of these projection results, and then averaged across this distribution. Positive values on the ordinate of Fig. 6c/d thus indicates more activity in the subspace aligned with the 'belief confirmation' regressors in Fig. S8.

Time-varying ACC activity within this subspace showed distinct dynamics on trials of different reaction times (Fig. 6c). First, activity in this subspace separated short from long RT trials relatively early during the course of making a choice – even at the time of cue 1 presentation. One interpretation of this finding is that the first attended item is initially

referenced as the 'default' option to be accepted or rejected, and evidence is interpreted either in favour of or against this default34. Confirmatory evidence may lead to executing a final choice more rapidly (Fig. 2b), with faster RTs on these trials. Second, irrespective of reaction time, ACC activity ramped shortly prior to joystick movement (Fig. 6d). Activity within the belief confirmation subspace therefore became prominent immediately prior to commitment to action, on all trials.

Finally, Fig. 4g indicates that ACC contains a signal related to which action will be selected12,13,15,42. We defined a separate subspace for whether the subject would choose left or right on the current trial, adopting the same split-half approach as in Figs. 6c/d. Activity in the ACC action selection subspace also gradually ramped as evidence was revealed about which option to choose, and peaked immediately prior to action selection (Fig. 6e). Belief confirmation and action selection subspaces are orthogonal; the relationship between them can be seen in Supplementary Video 2.

Single neuron analyses recapitulate core findings at population level

The analyses in Figs. 4-6 explore how information is represented at the level of the neural ensemble rather than at the level of the single neuron. This exploits the known heterogeneity of PFC single neuron responses14,15,17,36,40,42 to study task representations distributed across a population of cells. There are strong theoretical and empirical reasons to motivate studying information representation at the population level43, which have motivated several recent studies of PFC neuronal responses10,18,21,41. However, much of the previous literature has emphasised information representation at the level of single neurons. To facilitate comparison with this literature, we examined whether there were differences between PFC subregions in the fraction of neurons selective for key variables at different stages of the task. These analyses recapitulated the core findings at the population level.

We first tested whether OFC had more neurons encoding value comparison between currently and previously attended stimuli than other subregions. We performed three analyses, analogous to Fig. 5c-e. At cue 2 of attribute trials (cf. Fig. 5c), we asked whether neurons encoded value difference between cue 2 and cue 1; at cue 3 of attribute trials, we asked whether they encoded value difference between cue 3 and cue 1 (cf. Fig. 5d); at the same timepoint, we asked whether they encoded value difference between cue 3 and cue 2 (cf. Fig. 5e). To consider a neuron as representing value difference, we required that the contrast of parameter estimates (i.e. (Value Attended) – (Value Unattended)) be significant, and also that (Value Attended) *and* (Value Unattended) be independently significant with opposing signs. At all three relevant timepoints, we found that a greater proportion of single neurons passed these criteria in OFC than in ACC/DLPFC (Fig. 7a). We collapsed across these three tests to show the fraction of single neurons passing these criteria at any of the three cues individually (Fig. 7b).

We next tested whether ACC had a larger proportion of neurons encoding belief confirmation than other subregions. Here, we asked whether each neuron significantly encoded the four regressors depicted in Fig. S8, corresponding to belief confirmation at cue 2 or 3 on 'option' or 'attribute' trials. These are the four regressors whose parameter estimates correlate with each other in ACC (Fig. S10) but not OFC/DLPFC (Fig. 6a) and

form the 'belief confirmation subspace' shown in Fig. 6c/d. At all four timepoints, there was a greater proportion of single neurons significantly encoding belief confirmation in ACC than in OFC or DLPFC (Fig. 7c).

Finally, we performed an additional regression analysis at cue 1 onset to examine how factors relating to the value of different task features was represented in PFC. We again capitalized upon the fact that each cue could be decomposed into multiple features: its associated action, attribute, spatial position, and value. These different values were entered into the same regression model, allowing us to test the unique contribution of each of these features in explaining variance in neuronal firing across different regions. Across all three regions, a significant fraction of neurons encoded Cue 1's value, irrespective of the cue's attribute, action or spatial position (Fig. 8; binomial test, all $p < 1*10^7$). We also found that single neurons encoded value in distinct frames of reference across PFC subregions.

First, a significant subset of ACC and DLPFC neurons (~18%) preferentially responded to the values of either left or right options (binomial test, both $p < 1*10^{-5}$). Both of these populations were significantly greater than OFC, which encoded action value at chance level (pairwise Chi^2 test, ACC vs. OFC: $p=0.002$, DLPFC vs. OFC: $p=0.003$). The timecourse of these signals (Fig. 8b) was similar to that identified in the population analysis of cue 1 activity using RSA (Fig. 4f/g).

As left and right options were spatially dissociated, ACC and DLPFC neurons might be encoding value with reference to various parts of space (as opposed to action). However, in a region tuned to spatial location rather than action, one would also expect to find neurons that differentiated value for cues on the top part of the visual display compared to the bottom part. In DLPFC, such a relationship held: an equally prevalent population of top-bottom 'spatial value' neurons was observed as left-right 'action value' neurons (binomial test, $p < 1*10^{-6}$). In ACC, this population was significantly smaller than the left-right value population (pairwise Chi^2 test $p=0.03$). Consistent with its strong modulation by attention in Fig. 4a, this suggests that DLPFC preferentially encodes value in the reference frame of spatial position, whereas ACC encodes it with respect to relevant choice actions.

Lastly, replicating previous results[44], we found a significant proportion of neurons in OFC (~24%) reflected attribute value (binomial test, $p < 1*10^7$): they preferentially responded to the value of cues for either probability or magnitude. This proportion was significantly greater than the representation of attribute value coding in either ACC or DLPFC (pairwise Chi^2 tests, OFC vs. ACC: $p=0.001$, OFC vs. DLPFC: $p=0.014$). We interpret this finding with a degree of caution, however. It is possible that OFC neurons could appear to reflect attribute value as an artifact of being particularly selective for individual stimulus identities (see RSA analysis, Fig. 4a/e, and example neuron Fig. S3b). There was not clear evidence for attribute-specific value coding in OFC using RSA (see Fig. 4a).

Discussion

In real-world decision tasks, value-guided decision making is shaped heavily by visual attention. Information gathering strategies of both human consumers[22–24,27] and foraging

animals45 are well characterised as consecutive consideration of each choice option and its component attributes. Our findings demonstrate that as attention is first deployed to a choice option, population codes for decision-related processes emerge simultaneously in ACC, OFC and DLPFC rather than sequentially (Fig. 4/8), lending strong support for distributed and parallel models of value-based choice7,8. As attention was redeployed to sample further information, OFC activity reflected attention-guided value comparison (Fig. 5/7), whereas ACC activity reflected belief updating in light of new evidence and commitment to a final action (Fig. 6/7).

In addition to providing functional dissociations across PFC subregions, our paradigm allowed us to explore subjects' information sampling behavior, which suggested how subjects might be solving the task. In particular, subjects were biased towards sampling information from an option that they currently intended to choose (Fig. 2a), even when that information would yield little or no information about the choice (Fig. 2b). This mirrors biases that we have recently observed in a human version of the same experiment25, and is consistent with monkeys' willingness to sacrifice reward to obtain information about reward delivery46. Such behaviors could be interpreted of in terms of a mechanism for solving value-based choice of having a 'foreground' option in mind, and deciding sequentially whether to accept or reject this option relative to alternatives8.

An accept/reject strategy for value-based choice might be considered quite natural in the wider context of the real-world foraging decisions faced by our evolutionary ancestors31. These are inherently sequential in nature and involve decisions such as whether to accept or reject a current patch. Such patch-leaving decisions rely upon ramping signals in ACC prior to action selection20,47. In our task, ACC categorised the first attended option non-linearly into cues that might be accepted or rejected (Fig. 4c/h); had a stable 'belief confirmation' code for accepting/rejecting the foreground option in light of new evidence (Fig. 6a/S8/S10); and integrated that evidence towards a decision bound (Fig. 6c/d) while signalling which action would be chosen (Fig. 6e). There are a number of seemingly discrepant accounts of ACC function such as its role in value updating3,5,14, action-outcome prediction13, information seeking21, behavioural adaptation18,20,47 and action selection5,15,42. The computations we identified in ACC are consistent with these accounts but occurred at distinct time points or within orthogonal subspaces (Supplementary Video 2), thereby reconciling some of the outstanding debate concerning ACC function.

By contrast, OFC initially carried a representation of the first attended stimulus identity and its value, consistent with its anatomical projections from regions of inferotemporal cortex48 representing highly processed visual information such as object identity49. As further information was attended, simultaneous coding of attended and stored information emerged uniquely in OFC populations (Fig. 5) and provided a relative value coding mechanism for how choice options are compared. Such attention-guided relative value coding mechanisms form a central component of value comparison in recent accept/reject models of economic choice31, as well as other decision models24. Our findings are consistent with similar findings in adjacent ventromedial prefrontal cortex (VMPFC)9,10 but extend these results in important ways. For example, on 'option' trials we reveal how the OFC neural ensemble combines the value associated with multiple components of an option. Attribute integration

did not occur immediately upon attending to the second cue, but instead only once the third cue was attended on the alternative option (Fig. 5f/h) and value comparison could take place (Fig. 5d/e/i/j). Our findings also indicate that with respect to attended stimuli, attention-guided value comparison is neuroanatomically specific (cf. Figs. S6/S7). We note, however, that value comparison may still be supported in other structures in complementary frames of reference[7,8], for instance in the space of action value in ACC (Fig. 4f).

An important caveat of our study is that analyses were based on pseudo-ensembles, not large ensembles of simultaneously recorded neurons. We typically isolated between 5 and 25 single units per recording session (see Supplementary Table S1). Further insight into the relationship between different PFC subregions' decision dynamics might be obtained with higher-yield simultaneous population recording techniques.

In summary, theoretical models propose decision-making requires several computations including stimulus identification, valuation and integration with other attributes, comparison to previous choice options and action selection[7,8,50]. More recently it has been suggested that value-based decision making may be linked to other forms of choice such as sequential foraging decisions[8,31]. Although value-related signals are commonly found in PFC, a mechanistic account describing how distinct PFC subregions contribute to these computations has been lacking. Using a naturalistic information search and decision task that afforded exploration of how decision-related computations evolve as evidence informing choice is attended, we isolated these computations dissociably across PFC subregions. Our results therefore provide a unifying account of how PFC subregions support value-guided choice.

Online Methods

Subjects

Two adult male rhesus monkeys (*Macaca mulatta*), M and F, were used as subjects and weighed 7-10kg at the time of neuronal data collection. Both were ~4 years old at the start of the experiment. We regulated their daily fluid intake to maintain motivation on the task. All experimental procedures were approved by the UCL Local Ethical Procedures Committee and the UK Home Office, and carried out in accordance with the UK Animals (Scientific Procedures) Act.

Behavioral Protocol

Subjects sat head restrained in a behavioral chair facing a 19" computer monitor placed approximately 57cm away from the subjects' eyes. The height of the screen was adjusted so that the center of the screen aligned with neutral eye level for the subject. A voltage-gating joystick (APEM Components, UK) was placed in front of the subject out of his line of sight and was used to make manual responses during the task. Eye position and pupil tracking was achieved using an infrared camera (ISCAN ETL-200) sampled at 240Hz. The behavioral paradigm was run using the MATLAB based toolbox MonkeyLogic (http://www.monkeylogic.net/, Brown University, USA)[51–53]. All joystick and eye position was relayed to MonkeyLogic and for use online during the task and also recorded by

MonkeyLogic at 1000Hz. Juice delivery was achieved by using a precision peristaltic (ISMATEC IPC) to pump juice to a spout placed at the lips of the subject. Subject M was given dilute (50%) apple juice while Subject F drank dilute (50%) mango juice.

Subjects were taught the value of a set of 10 isoluminant pictures cues pertaining to either magnitude or probability value (see *Task* for further details) using secondary conditioning on a separate day preceding data acquisition. This set of cues was then used for the following 1-4 recording sessions at which point a new set of cues would be taught to the subject. In total Subject M learnt 13 separate sets of cues, while Subject F learnt 11 sets.

Task

A representation of the task structure is shown in main Figure 1A. Subjects initiated the trial by maintaining saccadic fixation on the center of the screen and central fixation of the joystick for 500ms. Once this was achieved two options were presented on the screen (left and right of center). Each option consisted of two pre-learned picture cues assigned to two different value attributes, probability of reward (10%, 30%, 50%, 70%, 90%) and magnitude of juice reward (0.15AU, 0.35AU, 0.55AU, 0.75AU, 0.95AU). The cues were uniformly sampled (with replacement, i.e. it was sometimes the case that the same cue would appear on both options). Reward magnitude (volume) was varied by manipulating the length of time a reward pump was driven, and the absolute values (i.e. reward time) associated with each stimulus varied slightly between subjects. Importantly, all four picture cues were covered up by grey squares with the exception of one, which was covered by a blue square. The blue square informed the subject the location of a required saccade. Once the subject made a saccade and fixated the blue square, the blue square was replaced by the picture cue, which the subject was required to continuously fixate for 300ms. If continuous fixation was not achieved within 1200ms the trial was aborted and the subject received a short timeout. Once this fixation period was achieved, the cue was covered and a second blue square was presented, indicating the location of the required second saccade. The position of this blue square indicated to the subject the type of trial being experienced. If the blue square was for the second cue of the same option subjects were in an 'Option trial', whereas if the blue square was for the same attribute cue of the second option then this was an 'Attribute trial'. Selection of trial types was pseudorandom. Once the subject made a saccade to the blue square, the blue square was replaced by the picture cue, and the subject was again required to maintain fixation of the second cue for 300ms. After this point, the subjects were relatively unconstrained. The two remaining unexplored locations were now replaced by blue squares. The subject could either choose an option using a joystick movement (left/ right) based on the value of the currently known information, or saccade to one or both of the remaining cues (in any order) as they wanted (with the 3rd cue requiring 300ms of fixation before subjects could saccade and uncover the information of the 4th cue) before making a choice. Importantly, however, they were prevented from viewing any cue that they had already seen. Once a response was made all four cues were uncovered (for 500ms for Subject F and 1000ms for Subject M), after which juice reward feedback was given with the probability and reward magnitude chosen by the subject.

Note that the position of the probability/magnitude cues were counterbalanced across trials – i.e. on half of all 'Option'/'Attribute' trials, the probability cues would appear on the top row, and on the other half of trials the magnitude cues would appear on the top row. Attribute locations also corresponded across options (i.e. if probability was on the top row for the left option, it would also be on the top row for the right option). Additionally, the location of the first cue was counterbalanced across all four possible spatial locations across trials.

'Option' and 'Attribute' trials were pseudorandomly interleaved during blocks of 50 trials. Between each of these blocks subjects were presented with a block of 25 trials, where all four picture cues were presented immediately (so called 'Simultaneous' trials). Data from these trials will be discussed in a separate publication.

Neuronal Recordings

Subjects were implanted with a titanium head positioner for restraint, and then subsequently implanted with two recording chambers that were located using pre-operative 3T MRI and stereotactic measurements. Post-operatively we used gadolinium attenuated MRI imaging and electrophysiological mapping of gyri and sulci to confirm chamber placement. The center of each chamber along the anterior-posterior (AP) coordinate plane was as follows; Subject M: left: AP 30.5, right: AP 33, Subject F: left: AP 34, right: AP 32.5. The chambers were angled along the medial-lateral plane to target different frontal regions (see Figure 3). Craniotomies were then performed inside each chamber to allow neuronal recordings.

During each recording session, neuronal activity was measured using tungsten microelectrodes (FHC Instruments, Bowdoin, USA) that were lowered into the brain through a grid using using custom-built manual microdrives or chamber-mounted motorized microdrives (FlexMT; AlphaOmega Inc.). During a typical recording session, 8-24 electrodes were lowered bilaterally into multiple target regions until well-isolated neurons were found. Neuronal data was recorded at 40kHz using a Plexon Omniplex system (Dallas, USA). Single unit isolation was achieved with manual spike sorting, using Plexon Offline Sorter (Dallas, USA). We randomly sampled neurons; no attempt was made to select neurons based on responsiveness. This procedure ensured an unbiased estimate of neuronal activity thereby allowing a fair comparison of neuronal properties between the different brain regions. Note that neural populations used to perform analyses in Figs. 4-6 are therefore not all simultaneously recorded, but they are pseudopopulations constructed across multiple recording sessions. Each neuron is first averaged across conditions (Fig. 4) or regressed across trials (Fig. 5/6) to identify the neuron's response to experimental variables, allowing us to then collapse across sessions, as in previous studies using similar approaches[10,39,41]. We excluded neurons with an average firing rate of <1Hz from further analysis (9 units in ACC, 21 in DLPFC, 12 in OFC; *n* reported in main text are after these units have been removed). No statistical methods were used to pre-determine sample sizes but our sample sizes are similar to those reported in previous publications (refs 10,11,14,15,20,38–40,42,46).

We recorded neuronal data from three target regions: anterior cingulate cortex (ACC), dorsolateral prefrontal cortex (DLPFC), and orbitofrontal cortex (OFC). We considered ACC

to be the entire dorsal bank of the anterior cingulate sulcus from AP 27-37mm. Our LPFC recordings spanned both dorsal and ventral banks of the principal sulcus but were concentrated towards the former. All neurons recorded lateral to the medial orbital sulcus and medial to the lateral orbital sulcus was considered OFC. In some sessions, all three regions were recorded simultaneously, whereas in other sessions only two were targeted. The total number of units with firing rate >1Hz in each brain region for each recording session is shown in Supplementary Table 1. Some recordings were also made in ventromedial prefrontal cortex (VMPFC), but these will be discussed in a separate publication. We used the gadolinium-enhanced MRI along with electrophysiological observations during the process of lowering each electrode to estimate the location of each recorded neuron and produce a histological map of the neuronal population (see Figure 3). All data was subsequently analysed using custom-written MATLAB code, using MATLAB 2017a (MathWorks, Natick, USA). Data collection and analysis were not performed blind to the conditions of the experiments, as neuroanatomical location of recording sites had to be known when lowering electrodes for recording.

Representational similarity analysis (RSA) at Cue 1 presentation (Figures 4/S2/S4, Supplementary Video 1)

To calculate the representational similarity matrices shown in Figure 4A-C, we first calculated the average firing rate for each neuron for each of the 20 conditions of interest: when the lowest to highest probability cue was presented on the left at Cue 1, when the lowest to highest magnitude cue was presented on the left at Cue 1, lowest to highest probability cue on right, and lowest to highest magnitude cue on right. This firing rate was computed between 100ms and 500ms after Cue 1 onset. We then normalized across these 20 conditions, subtracting the mean and dividing by the standard deviation.

Repeating this for every neuron yielded a matrix with dimensions (neurons*20). For two conditions (i, j), we computed the correlation coefficient across neurons between row i and row j of this matrix, which is plotted in element (i,j) of the representational similarity matrix. For Supplementary Video 1, we repeated the same procedure on sliding windows of +/- 100ms from the timepoint of interest.

RSA template-based regression (Figure 4D-H)

We used multiple linear regression to assess the contribution of several potential 'template' neural codes to the RSA matrices within each region. Each of the 400 elements of each region's RSA template was explained using the following regression model:

$$r_{(i,j)} = \beta_0 + \sum_{n=1}^{6} \beta_n Template_{n(i,j)} + \epsilon_{(i,j)}$$

Where r denotes the correlation coefficient matrix computed using RSA, and there are six 'template' matrices onto which the RSA matrix is regressed. We estimated β_{0-6} using ordinary least squares, minimizing the sum of squared residuals ϵ. The six template matrices were as follows:

Template 1: Identity matrix – accounting for all RSA matrices being 1 when element *i*=element *j* (note that this is a regressor of no interest, to model out the unity correlation between a condition and itself)

Template 2 (Figure 4D): 'Spatial attention' – accounting for representational similarity between cues presented on the same side, but dissimilarity between cues on opposite sides (1 if *i*<=10 and *j*<=10, 1 if *i*>=11 and *j*>=11, -1 elsewhere)

Template 3 (Figure 4E): 'Stimulus identity' – accounting for representational similarity between the same stimulus being presented on left/right options (1 where |*i*-*j*|=10, 0 elsewhere)

Template 4 (Figure 4F): 'Attended value' – accounting for representational similarity between similarly valued items and representational dissimilarity between dissimilarly valued items (*ranked value(i)** *ranked value(j)*, where ranked value is -2 for the lowest ranked stimulus within an attribute (i.e. 10% probability, 15% maximal reward magnitude), -1 for the 2nd lowest ranked (30% probability, 35% maximal reward magnitude), 0 for the median ranked (50 % probability, 55% maximal reward magnitude), 1 for the 2nd highest ranked (70% probability, 75% maximal reward magnitude), 2 for the highest ranked (90% probability, 95% maximal reward magnitude)) – see supplementary note for further justification of the structure of this regressor

Template 5 (Figure 4G): 'Left/right value' – interaction of template 4 with spatial attention - i.e. set to the same value as template 4 for cues presented on the same side, and set to 0 for cues presented on opposite sides

Template 6 (Figure 4H): 'Accept/reject' - accounting for representational similarity between cues that might lead to ultimately accepting the current alternative (good items similar to other good items; bad items similar to other bad items), and representational dissimilarity between dissimilar items in terms of acceptance/rejection (*sign* of *attended value* template)

For the middle panels in Figure 4D-H this model was estimated on RSA matrices from 100-500ms post-stimulus, as in Figure 4A-C; for the bottom panels of Figure 4D-H it was performed on sliding windows of +/- 100ms from the timepoint of interest, as in Supplementary Video 1. In these panels we plot the coefficient of partial determination (CPD) for each regressor across time, which is defined for EV X_i as follows:

$$CPD(X_i) = \left[SSE(X_{\sim i}) - SSE(X_{\sim i}, X_i) \right] / SSE(X_{\sim i})$$

where *SSE(X)* refers to the sum of squared errors in a GLM that includes a set of EVs *X*, and $X_{\sim i}$ is a set of all the EVs included in the full model except X_i [14,40].

Prior to running the regression model, each template was normalized by dividing by its maximum absolute value (so that the minimum possible value of each template was -1, and the maximum value of each template was +1). This normalization was simply to place the regressors on a common scale, so that when plotted in Figure 4, the same color axis could be

used to describe all regressors. Importantly, this normalization has no bearing on either the CPD or T-statistics, as both of these measures are scale-free.

To quantify the latency at which different factors were represented across time, we calculated the timepoint at which the CPD reached 75% of its maximal value over time, a statistic we label t_{75} in the paper. To simulate how different instantiations of the noise might affect our estimate of t_{75}, we permuted the residuals from the original GLM and added these permuted residuals to $X\beta$ (where X is the design matrix and β are the parameter estimates). We then recalculated the time-varying measure of CPD and re-estimate t_{75} for each instantiation of the noise. The resulting distribution of values of t_{75} from this analysis are shown in Figure S5c (100 permutations were performed).

Statistical inference on RSA template-based regression model

We tested the significance of each template within each region by computing the T-statistic for each β coefficient (i.e. $\hat{\beta}_n/\hat{\sigma}_n$, where $\hat{\sigma}_n$ denotes the standard errors of each coefficient estimate). We compared differences between regions by computing F-statistics equivalent to a one-way ANOVA (see https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FEAT/UserGuide#ANOVA:_1-factor_4-levels for example). Importantly, however, when calculating these statistics on a correlation matrix, they may not be parametrically distributed in the null distribution (due to observations not being independently and identically distributed). To overcome this, we built a non-parametric null distribution for each test of interest, by permuting the identities of the 20 cues (i.e. values 1-5 on probability/magnitude, left/right), recomputing the RSA matrix, and rerunning the regression. We then computed the T-statistics and F-statistics on this permuted data, and compared the true statistics to the permuted null distribution to obtain p-values 54. We performed 10,000 permutations.

General linear model (GLM), underlying analyses in Figures 5-7

For the analyses shown in Figures 5-7, we first estimated a general linear model on the firing rate of each individual neuron, timelocked with respect to Cue 1 presentation, Cue 2 presentation, Cue 3 presentation, and joystick movement (response). Each neuron's firing rate was explained using a GLM containing 18 explanatory variables (EVs), detailed below, estimated using ordinary least squares. Note that *EVs 1-6* are critical for the analyses shown in Figure 5 and Figures S6/S7, *EVs 13-16* are critical for the analyses shown in Figure 6A-C and Figure S10, and *EVs 17-18* are critical for the analyses shown in Figure 6D.

EV 1 captured the linear effect of changing the first attended cue's value from the lowest value to highest value, collapsing across probability and magnitude cues, selectively on 'option trials'. Specifically, if the lowest ranked probability/magnitude item was presented they were valued -2; if the second lowest ranked item was presented -1; third lowest, 0; second highest, 1; highest, 2.

EVs 2-4 were similar to EV 1, but reflected the second, third and fourth attended cue's value respectively (for option trials only). On trials where the third or fourth cue was not attended on an option trial (because the subject responded without sampling all cues), the corresponding EVs were valued 0.

EVs 5-6 were similar to EVs 1-2, but reflected the first and second attended cue's value respectively for 'attribute trials' only. EVs 7-8 were similar to EVs 3-4, but reflected the third and fourth attended cue's value on attribute trials where the subject saccaded diagonally back to the first side of the screen (0 on vertical saccade trials), whereas EVs 9-10 reflected the third and fourth attended cue's value on attribute trials where the subject saccaded vertically to the second side of the screen (i.e. 0 on diagonal saccade trials). Note that there is no need to split option trials by third saccade direction, as unlike in option trials the third saccade is always to the second side of the screen.

EV 11 was an indicator variable for option trials (1 on option trials, 0 otherwise); EV 12 was an indicator variable for attribute trials (1 on attribute trials, 0 otherwise). Note that EVs 11 and 12 sum to produce a constant term, thereby capturing variation in the mean firing rate of the cell across time.

EVs 13-16 were variables that all captured the extent to which the Cue value observed at Cue 2 and Cue 3 were *consistent* (belief confirmation) or *inconsistent* (belief disconfirmation) with the currently held belief as to which option would be rewarded. They are described below, but for clarity, they are also depicted in Figure S8. Two key points are pertinent: (a) by design, all four EVs were largely *orthogonal to the value of Cue 1, Cue 2 and Cue 3* (although see note on EV 16 below); (b) they each rely upon different cues and different trials, and so are *orthogonal to each other* by design.

EV 13 (Figure S8A) was the same as EV 2 – i.e. the value of cue 2 on option trials – but crucially, it was multiplied by 1 whenever the value of the first cue was *greater* than the average value (i.e. best or second best picture cues), multiplied by -1 whenever the value of the first cue was *lower* than the average value (i.e. worst or second worst picture cue), and multiplied by 0 whenever it was of average value (middle picture cue). EV 13 therefore was positively signed whenever Cue 2 was consistent with Cue 1 (e.g. low-valued cue followed by another low-valued cue, or high-valued cue followed by another high-valued cue).

EV 14 (Figure S8B) was the same as EV 6 – i.e. the value of cue 2 on attribute trials – but was multiplied by 1 when the first cue's value was *lower* than average, by -1 whenever the first cue's value was *higher* than average, and by 0 when cue 1 was of average value. Again, this meant that EV 14 was positively signed whenever it was consistent with Cue 1 (e.g. low-valued cue on the left followed by high-valued cue right both favor the right action, or high-valued cue on the left followed by low-valued cue on the right both favor a left action).

EV 15 (Figure S8C) was the same as EV 3 – i.e. the value of cue 3 on option trials – but was multiplied by 1 whenever the first and second cue were lower than average value (when EV 1 + EV 2 was negative), by -1 whenever the first and second cue were higher than average value (when EV 1 + EV 2 was positive), and by 0 when the first and second cue were of average value(when EV 1 + EV 2 equalled 0).

EV 16 (Figure S8D) was similarly defined to EVs 7 and 9 – i.e. the value of cue 3 on attribute trials – but crucially relies upon an interaction of the relative value of the first and second cue, and which side the subject decided to attend with the third saccade. On trials where the subject's third saccade was diagonal back to option 1, it was EV 7 multiplied by 1

when (EV 5> EV 6), multiplied by -1 when (EV 6>EV 5), and multiplied by 0 when (EV 5=EV 6). On trials where the subject's third saccade was vertical within option 2, it was EV 9 multiplied by 1 when (EV 6>EV 5), multiplied by -1 when (EV 5>EV 6), and multiplied by 0 when (EV 5=EV 6). Note that because subjects' decision whether to make a third saccade to the same side as option 1 relied upon the relative value of Cue 1 and Cue 2, there existed some positive correlation between EV16 and EVs 7 and 9 (mean r^2 of 0.167 and 0.194 respectively, see Figure S9). Nevertheless, including all three EVs together in the GLM directly controls for this correlation with value, by partialling out any variance that can be attributed to EVs 7 or 9 from the parameter estimate for EV 16.

EV 17 was defined in terms of action selectivity on option trials. It was valued 1 on option trials where the subject chose left, -1 on option trials where the subject chose right, and 0 on attribute trials.

EV 18 was defined in terms of action selectivity on attribute trials. It was valued 1 on attribute trials where the subject chose left, -1 on attribute trials where the subject chose right.

We estimated this multiple regression model on neuronal firing rate in sliding 200ms bins, stepped in 10ms time-windows, from 100ms pre-cue to 500ms post-cue (when stimulus-locked), or from 500ms pre-response to 100ms post-response (when response-locked). We excluded trials where subjects viewed fewer than 3 cues from this analysis.

Peri-stimulus correlation and cross-correlation of parameter estimates from GLM (Figure 5/ Figures S6/S7)

Once the model in the previous section was estimated for each neuron, we then correlated, *across neurons*, T-statistics associated with parameter estimates for different EVs. This allowed us to examine how population subspaces encoding different variables related to each other, at various timepoints through the trial. Note that in one case (Figure 5B) we collapse across parameter analyses from option and attribute trials for clarity. Parameter estimates in Figure 5B-F/K were taken from 250ms post-stimulus, whereas in Figure 5G-J they were repeated on all possible combinations of time-points to produce cross-correlation matrices of parameter estimates. In figure 5K, we performed a Fisher r-to-Z transformation to test the differences between these correlations between subregions.

Statistical inference on cross-correlation of parameter estimates

To test whether areas of high/low correlation between parameter estimates were significantly larger than would be expected by chance, we used a cluster-based permutation test[54]. We identified clusters in the cross-correlation map that were larger than a cluster-forming threshold (set at $|r|>0.2$; similar results could be obtained with other cluster-forming thresholds). We then permuted (across neurons) one of the two sets of parameter estimates used to compute the cross-correlation matrix, and identified clusters that exceeded the cluster-forming threshold in the permuted data. For each of the 1,000 permutations, we stored the size of the largest cluster. This provided a null distribution of maximum cluster sizes that would be expected by chance. We used the 99.9[th] percentile of this null

distribution as a threshold for deeming whether cluster sizes observed in the data were significant, at a p-value of $p<0.001$ (corrected for multiple comparisons).

Projection of ACC activity onto belief confirmation/chosen response subspaces (Figure 6 and Figure S10, Supplementary Video 2)

To identify whether there was a stable subspace representing 'belief confirmation' in each brain region (Figure 6A), we investigated whether the parameter estimates for all four regressors that captured belief confirmation in our GLM were correlated (Figure S9). The parameter estimates used were EV 13, 300ms after Cue 2 onset; EV 14, 300ms after Cue 2 onset; EV 15, 300ms after Cue 3 onset; EV 16, 300ms after Cue 3 onset. We also asked whether this subspace was similar to the subspace for Cue 1 value (i.e. EV1 + EV5, 300ms after Cue 1 onset), based on the idea that Cue 1 'value' responses in ACC are better conceived in terms of belief confirmation about accepting or rejecting the first attended cue (cf. results in Figure 4C, 4H). We again performed a Fisher r-to-Z transformation to test the differences across subregions between the correlations in these subspaces.

This approach uniquely identified a stable subspace for belief confirmation in ACC. Once this stable subspace was identified (see Figure 6A/S10), we asked how activity in this subspace evolved in trials where the subject took different lengths of time to make his final choice response (Figure 6 and Supplementary Video 2). For each neuron, we split trials into five separate bins depending upon response time from Cue 1 onset, and averaged neuronal firing for these different trial types. For each bin, this yielded a matrix with dimensions *time*neurons*.

To examine activity within different subspaces, we then regressed this matrix onto a projection matrix composed of two key 'weights' per neuron, i.e. T-statistics of contrasts of parameter estimates of interest, estimated from the GLM. This projection matrix therefore had dimensions *neurons*(2 PEs)*. The two contrasts of interest were:

1. The average parameter estimates for belief confirmation, i.e. EV 13, 300ms after Cue 2 onset; EV 14, 300ms after Cue 2 onset; EV 15, 300ms after Cue 3 onset; EV 16, 300ms after Cue 3 onset;

2. The average parameter estimates for left vs. right action selection, i.e. EV 17 and EV 18, 200ms prior to response onset;

Regressing the time*neurons matrix onto the neurons*(2 PEs) gives rise to the sliding analysis that is shown in Figure 6. In Figure 6B/C, we plot the stimulus-locked and response-locked parameter estimates for contrast 1 respectively, reflecting the population activity in the belief confirmation subspace for trials of different length. In Figure 6D, we plot the response-locked parameter estimates for contrast 2, reflecting population activity in the left/right action selection subspace in trials of different length. In both cases, we baseline corrected subspace activity to the time of Cue 1 onset +/- 50ms. Supplementary Video 2 provides a representation of how activity in both of these subspaces progresses during the course of the trial.

Crucially, we avoided using the same data for estimating different neurons' weights in the projection matrix as for plotting population activity. To achieve this, we first split the data into odd and even trials; we estimated the projection matrix weights using the GLM on the odd trials, and projected these weights onto firing rates on the even trials; we then repeated the same process with even trials for GLM estimation and odd trials for projection; finally, we averaged subspace activity together across odd and even-trial analyses.

Further detail on methods is available online in the Life Sciences Reporting Summary.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Yeterian EH, Pandya DN, Tomaiuolo F, Petrides M. The cortical connectivity of the prefrontal cortex in the monkey brain. Cortex; a journal devoted to the study of the nervous system and behavior. 2012; 48:58–81. DOI: 10.1016/j.cortex.2011.03.004 [PubMed: 21481342]

2. Haber, Suzanne N; Behrens, Timothy E J. The neural network underlying incentive-based learning: implications for interpreting circuit disruptions in psychiatric disorders. Neuron. 2014; 83:1019–1039. DOI: 10.1016/j.neuron.2014.08.031 [PubMed: 25189208]

3. Rushworth, Matthew FS; Noonan, MaryAnn P; Boorman, Erie D; Walton, Mark E; Behrens, Timothy E. Frontal cortex and reward-guided learning and decision-making. Neuron. 2011; 70:1054–1069. DOI: 10.1016/j.neuron.2011.05.014 [PubMed: 21689594]

4. Donoso M, Collins AG, Koechlin E. Foundations of human reasoning in the prefrontal cortex. Science. 2014; 344:1481–1486. DOI: 10.1126/science.1252254 [PubMed: 24876345]

5. Rudebeck PH, et al. Frontal cortex subregions play distinct roles in choices between actions and stimuli. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2008; 28:13775–13785. DOI: 10.1523/JNEUROSCI.3541-08.2008 [PubMed: 19091968]

6. Vaidya AR, Fellows LK. Testing necessary regional frontal contributions to value assessment and fixation-based updating. Nature Communications. 2015; 6:10120.doi: 10.1038/ncomms10120

7. Cisek P. Making decisions through a distributed consensus. Current opinion in neurobiology. 2012; 22:927–936. DOI: 10.1016/j.conb.2012.05.007 [PubMed: 22683275]

8. Hunt LT, Hayden BY. A distributed, hierarchical and recurrent framework for reward-based choice. Nature Reviews Neuroscience. 2017; 18:172–182. DOI: 10.1038/nrn.2017.7 [PubMed: 28209978]

9. Lim SL, O'Doherty JP, Rangel A. The decision value computations in the vmPFC and striatum use a relative value code that is guided by visual attention. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2011; 31:13214–13223. DOI: 10.1523/JNEUROSCI.1246-11.2011 [PubMed: 21917804]

10. Strait CE, Blanchard TC, Hayden BY. Reward value comparison via mutual inhibition in ventromedial prefrontal cortex. Neuron. 2014; 82:1357–1366. DOI: 10.1016/j.neuron.2014.04.032 [PubMed: 24881835]

11. McGinty, Vincent B; Rangel, A; Newsome, William T. Orbitofrontal cortex value signals depend on fixation location during free viewing. Neuron. 2016; 90:1299–1311. DOI: 10.1016/j.neuron. 2016.04.045 [PubMed: 27263972]

12. Hare TA, Schultz W, Camerer CF, O'Doherty JP, Rangel A. Transformation of stimulus value signals into motor commands during simple choice. Proceedings of the National Academy of Sciences. 2011; 108:18120–18125. DOI: 10.1073/pnas.1109322108

13. Alexander WH, Brown JW. Medial prefrontal cortex as an action-outcome predictor. Nature Neuroscience. 2011; 14:1338–1344. DOI: 10.1038/nn.2921 [PubMed: 21926982]

14. Kennerley SW, Behrens TEJ, Wallis JD. Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. Nature Neuroscience. 2011; 14:1581–1589. DOI: 10.1038/nn.2961 [PubMed: 22037498]

15. Cai X, Padoa-Schioppa C. Neuronal encoding of subjective value in dorsal and ventral anterior cingulate cortex. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2012; 32:3791–3808. DOI: 10.1523/JNEUROSCI.3864-11.2012 [PubMed: 22423100]

16. Behrens TE, Woolrich MW, Walton ME, Rushworth MF. Learning the value of information in an uncertain world. Nat Neurosci. 2007; 10:1214–1221. DOI: 10.1038/nn1954 [PubMed: 17676057]

17. Bernacchia A, Seo H, Lee D, Wang XJ. A reservoir of time constants for memory traces in cortical neurons. Nat Neurosci. 2011; 14:366–372. DOI: 10.1038/nn.2752 [PubMed: 21317906]

18. Karlsson MP, Tervo DGR, Karpova AY. Network resets in medial prefrontal cortex mark the onset of behavioral uncertainty. Science. 2012; 338:135–139. DOI: 10.1126/science.1226518 [PubMed: 23042898]

19. Bryden DW, Johnson EE, Tobia SC, Kashtelyan V, Roesch MR. Attention for Learning Signals in Anterior Cingulate Cortex. Journal of Neuroscience. 2011; 31:18266–18274. DOI: 10.1523/jneurosci.4715-11.2011 [PubMed: 22171031]

20. Hayden BY, Pearson JM, Platt ML. Neuronal basis of sequential foraging decisions in a patchy environment. Nat Neurosci. 2011; 14:933–939. DOI: 10.1038/nn.2856 [PubMed: 21642973]

21. Stoll FM, Fontanier V, Procyk E. Specific frontal neural dynamics contribute to decisions to check. Nature Communications. 2016; 7:11990.doi: 10.1038/ncomms11990

22. Payne JW. Task complexity and contingent processing in decision-making - information search and protocol analysis. Organ Behav Hum Perf. 1976; 16:366–387.

23. Bettman JR, Luce MF, Payne JW. Constructive consumer choice processes. Journal of Consumer Research. 1998; 25:187–217.

24. Krajbich I, Armel C, Rangel A. Visual fixations and the computation and comparison of value in simple choice. Nat Neurosci. 2010; 13:1292–1298. DOI: 10.1038/nn.2635 [PubMed: 20835253]

25. Hunt LT, Rutledge RB, Malalasekera WMN, Kennerley SW, Dolan RJ. Approach-induced biases in human information sampling. PLOS Biology. 2016; 14:e2000638.doi: 10.1371/journal.pbio. 2000638 [PubMed: 27832071]

26. Stewart N, Hermens F, Matthews WJ. Eye Movements in Risky Choice. Journal of Behavioral Decision Making. 2016; 29:116–136. DOI: 10.1002/bdm.1854 [PubMed: 27522985]

27. Gidlof K, Wallin A, Dewhurst R, Holmqvist K. Using Eye Tracking to Trace a Cognitive Process: Gaze Behaviour During Decision Making in a Natural Environment. J Eye Movement Res. 2013; 6

28. Daddaoua N, Lopes M, Gottlieb J. Intrinsically motivated oculomotor exploration guided by uncertainty reduction and conditioned reinforcement in non-human primates. Sci Rep. 2016; 6:20202.doi: 10.1038/srep20202 [PubMed: 26838344]

29. Passingham, RE, Wise, SP. The Neurobiology of the Prefrontal Cortex. Oxford University Press; 2012. 26–64. Ch 2

30. Gottlieb J, Hayhoe M, Hikosaka O, Rangel A. Attention, reward, and information seeking. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2014; 34:15497–15504. DOI: 10.1523/JNEUROSCI.3270-14.2014 [PubMed: 25392517]

31. Hayden BY. Economic choice: the foraging perspective. Curr Opin Behav Sci. 2018; 24:1–6. DOI: 10.1016/j.cobeha.2017.12.002

32. Fellows LK. Deciding how to decide: ventromedial frontal lobe damage affects information acquisition in multi-attribute decision making. Brain. 2006; 129:944–952. DOI: 10.1093/brain/awl017 [PubMed: 16455794]

33. Nickerson RS. Confirmation bias: A ubiquitous phenomenon in many guises. Review of General Psychology. 1998; 2:175–220. DOI: 10.1037/1089-2680.2.2.175

34. Boorman ED, Rushworth MF, Behrens TE. Ventromedial prefrontal and anterior cingulate cortex adopt choice and default reference frames during sequential multi-alternative choice. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2013; 33:2242–2253. [PubMed: 23392656]

35. Wittmann MK, et al. Predictive decision making driven by multiple time-linked reward representations in the anterior cingulate cortex. Nat Commun. 2016; 7:12327.doi: 10.1038/ncomms12327 [PubMed: 27477632]

36. Kaping D, Vinck M, Hutchison RM, Everling S, Womelsdorf T. Specific contributions of ventromedial, anterior cingulate, and lateral prefrontal cortex for attentional selection and stimulus valuation. PLoS biology. 2011; 9:e1001224.doi: 10.1371/journal.pbio.1001224 [PubMed: 22215982]

37. Kriegeskorte N. Representational similarity analysis – connecting the branches of systems neuroscience. Frontiers in Systems Neuroscience. 2008; doi: 10.3389/neuro.06.004.2008

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

39. Hunt LT, Behrens TE, Hosokawa T, Wallis JD, Kennerley SW. Capturing the temporal evolution of choice across prefrontal cortex. Elife. 2015; 4doi: 10.7554/eLife.11945

40. Cai X, Kim S, Lee D. Heterogeneous coding of temporally discounted values in the dorsal and ventral striatum during intertemporal choice. Neuron. 2011; 69:170–182. DOI: 10.1016/j.neuron.2010.11.041 [PubMed: 21220107]

41. Mante V, Sussillo D, Shenoy KV, Newsome WT. Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature. 2013; 503:78–84. DOI: 10.1038/nature12742 [PubMed: 24201281]

42. Kennerley SW, Dahmubed AF, Lara AH, Wallis JD. Neurons in the frontal lobe encode the value of multiple decision variables. Journal of cognitive neuroscience. 2009; 21:1162–1178. DOI: 10.1162/jocn.2009.21100 [PubMed: 18752411]

43. Yuste R. From the neuron doctrine to neural networks. Nat Rev Neurosci. 2015; 16:487–497. DOI: 10.1038/nrn3962 [PubMed: 26152865]

44. O'Neill M, Schultz W. Coding of Reward Risk by Orbitofrontal Neurons Is Mostly Distinct from Coding of Reward Value. Neuron. 2010; 68:789–800. DOI: 10.1016/j.neuron.2010.09.031 [PubMed: 21092866]

45. O'Brien WJ, Browman HI, Evans BI. Search strategies of foraging animals. American Scientist. 1990; 78:152–160.

46. Blanchard, Tommy C; Hayden, Benjamin Y; Bromberg-Martin, Ethan S. Orbitofrontal Cortex Uses Distinct Codes for Different Choice Attributes in Decisions Motivated by Curiosity. Neuron. 2015; 85:602–614. DOI: 10.1016/j.neuron.2014.12.050 [PubMed: 25619657]

47. Kolling N, Behrens T, Wittmann MK, Rushworth M. Multiple signals in anterior cingulate cortex. Current opinion in neurobiology. 2016; 37:36–43. DOI: 10.1016/j.conb.2015.12.007 [PubMed: 26774693]

48. Cavada C, Company T, Tejedor J, Cruz-Rizzolo RJ, Reinoso-Suarez F. The anatomical connections of the macaque monkey orbitofrontal cortex. A review. Cereb Cortex. 2000; 10:220–242. [PubMed: 10731218]

49. DiCarlo JJ, Zoccolan D, Rust NC. How does the brain solve visual object recognition? Neuron. 2012; 73:415–434. DOI: 10.1016/j.neuron.2012.01.010 [PubMed: 22325196]

50. Padoa-Schioppa C. Neurobiology of Economic Choice: A Good-Based Model. Annual Review of Neuroscience. 2011; 34:333–359. DOI: 10.1146/annurev-neuro-061010-113648

51. Asaad WF, Eskandar EN. A flexible software tool for temporally-precise behavioral control in Matlab. J Neurosci Methods. 2008; 174:245–258. DOI: 10.1016/j.jneumeth.2008.07.014 [PubMed: 18706928]

52. Asaad WF, Eskandar EN. Achieving behavioral control with millisecond resolution in a high-level programming environment. J Neurosci Methods. 2008; 173:235–240. DOI: 10.1016/j.jneumeth.2008.06.003 [PubMed: 18606188]

53. Asaad WF, Santhanam N, McClellan S, Freedman DJ. High-performance execution of psychophysical tasks with complex visual stimuli in MATLAB. J Neurophysiol. 2013; 109:249–260. DOI: 10.1152/jn.00527.2012 [PubMed: 23034363]

54. Nichols TE, Holmes AP. Nonparametric permutation tests for functional neuroimaging: a primer with examples. Human brain mapping. 2002; 15:1–25. [PubMed: 11747097]

55. Hunt LT, Malalasekera WMN, Kennerley SW. Recordings from three subregions of macaque prefrontal cortex during an information search and choice task. CRCNS.org. 2018; doi: 10.6080/K0PZ5712
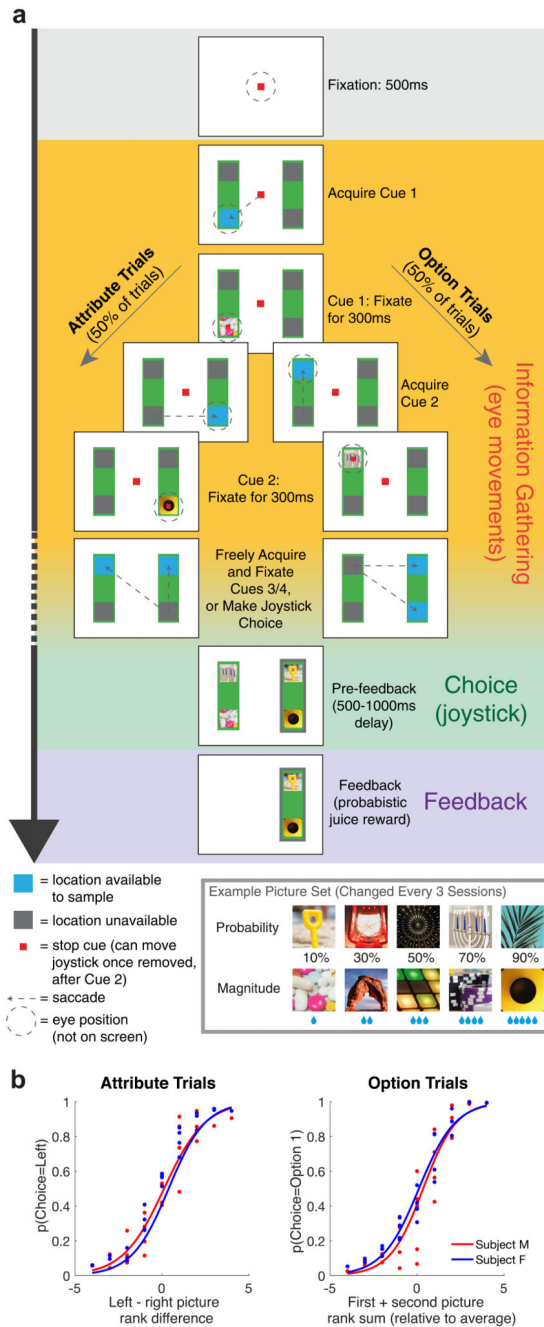
Fig. 1. Experimental paradigm and basic subject behavior.
(A) Task design. Subjects chose between a left and right option (green rectangles) using a manual joystick movement, after sequentially sampling 2, 3, or 4 cues that revealed reward probability and magnitude to the subject. Blue squares indicated locations available for information sampling. On attribute trials (left panels), cues 1 and 2 were on opposing options but the same attribute; on option trials (right panels), cues 1 and 2 were different attributes of the same option. **(B)** Choice behavior as function of cue 1 and 2 value. **(i)** Probability of choosing left option on attribute trials, as function of left-right picture rank

difference (where 1 is lowest rank picture on each attribute, and 5 is highest rank picture). **(ii)** Probability of choosing option 1 on option trials, as function of first plus second picture value (relative to middle value picture 3). See also Fig. S1 for additional analyses of choices.
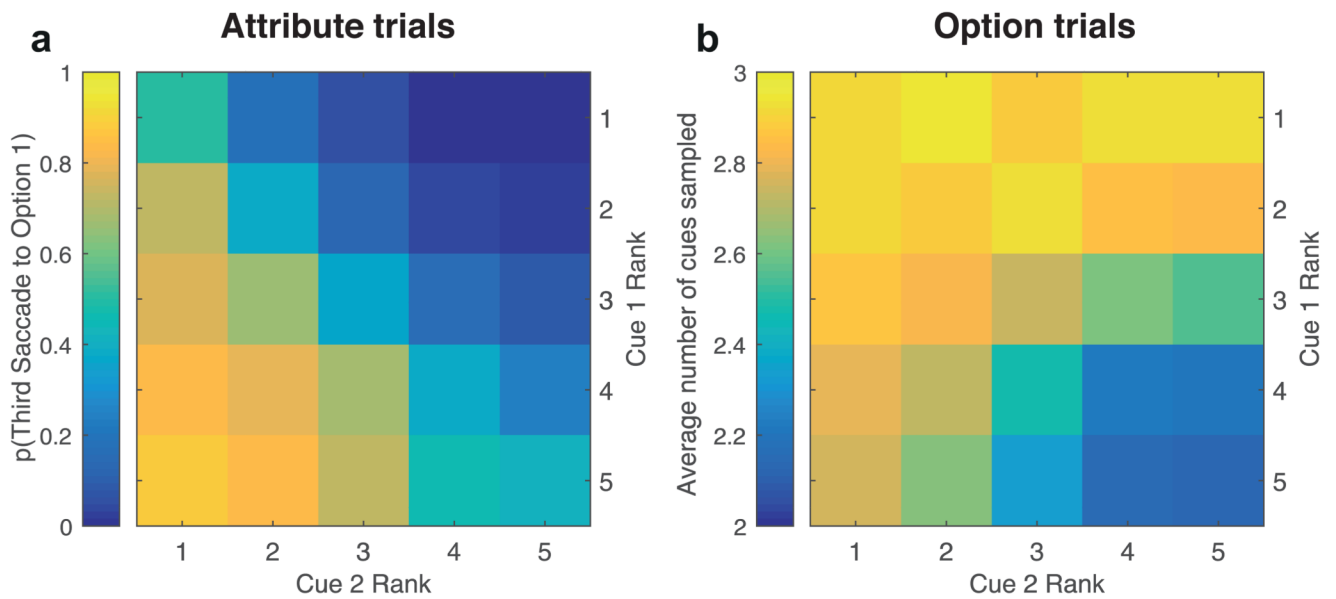
Fig. 2. Information sampling behaviour on Attribute and Option trials reveals preference for sampling current best alternatiive.

(**A**) On attribute trials where a third cue was sampled, the probability of third saccade being directed diagonally back towards option 1, rather than vertically to remain on option 2, is high when Cue 1 is high value and Cue 2 is low value. (**B**) On option trials, the number of cues sampled is lowest when Cues 1 and 2 are high value, but surprisingly it is highest when Cues 1 and 2 are low value, even though this provides complete information about which option to choose.
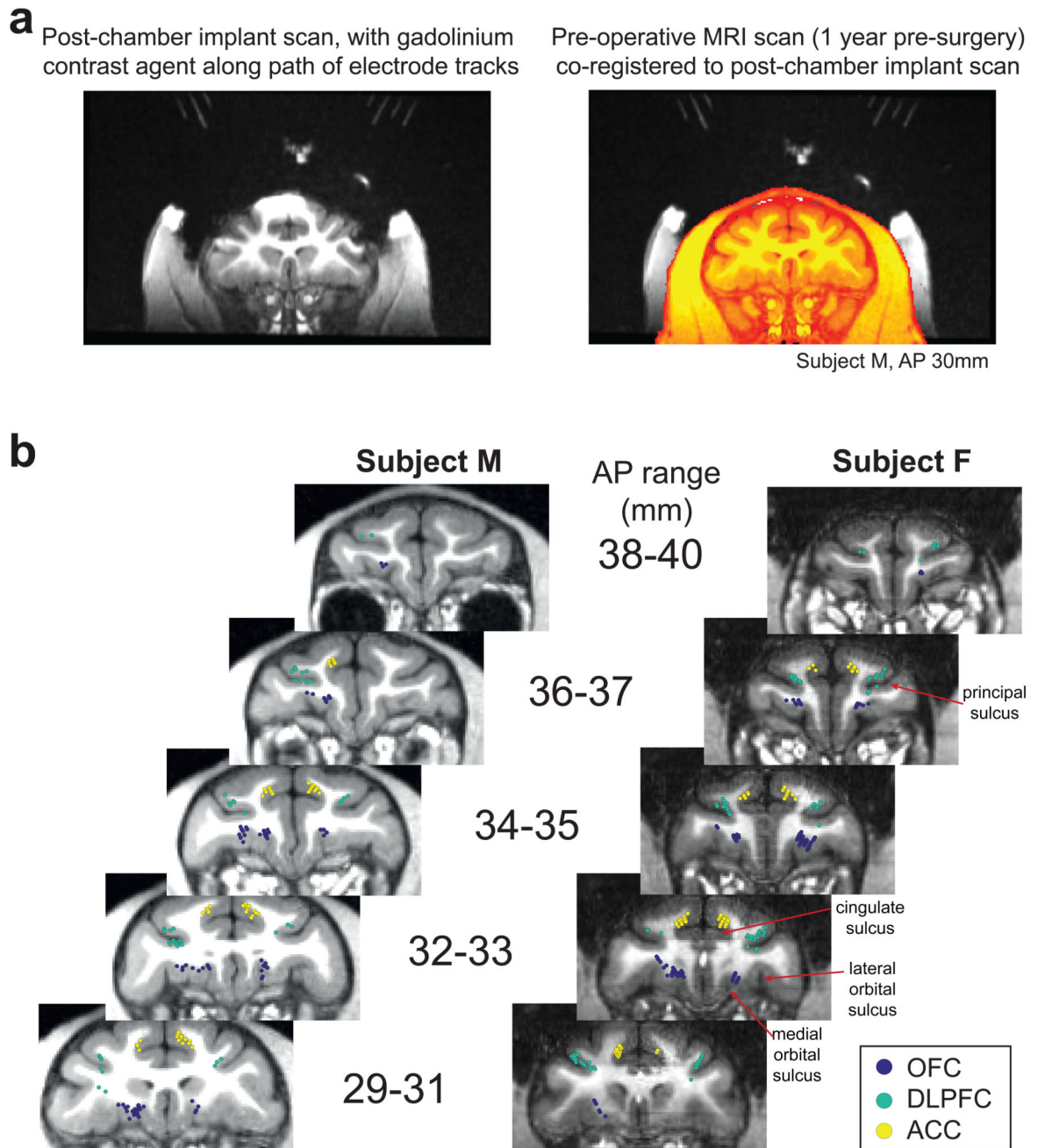
Fig. 3. Recording locations.

(A) Strategy for reconstructing path of electrode tracks. Left: After surgery for chamber implantation, prior to craniotomy, subjects underwent MRI scan with a custom-built implant placed within the chamber. This contained the MRI contrast agent gadolinium along the trajectory of potential recording paths at regular 4mm intervals. Note the prominent susceptibility artifact due to titanium chambers does not affect gadolinium trajectories, which were intentionally located away from the chamber. Right: This scan could be co-registered to a pre-operative scan without susceptibility artifact (in orange; note that head

appears smaller due to muscle growth between scans), to reliably reconstruct recording locations. This technique was further verified by identification of grey/white matter boundaries during lowering of electrodes along different trajectories. **(B)** Recording locations of orbitofrontal neurons (OFC), dorsolateral prefrontal neurons (DLPFC), and anterior cingulate neurons (ACC), shown on coronal sections. AP range denotes position anterior to interaural plane in stereotactic coordinates.
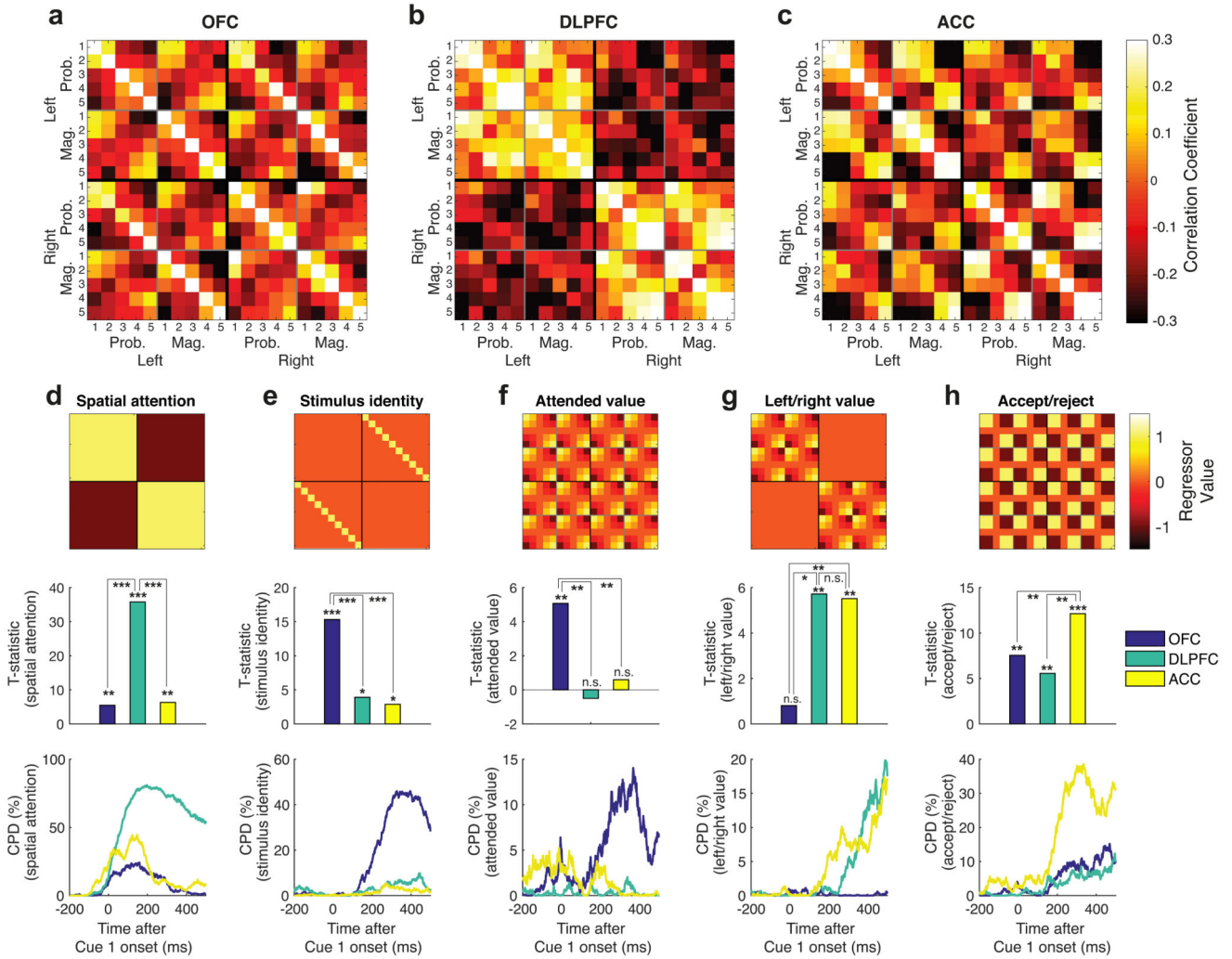
Fig. 4. Triple dissociation of task-evoked neural codes across OFC (n=183 units), DLPFC (n=135 units) and ACC (n=189 units) at Cue 1 presentation.
(**A**) OFC representaional similarity analysis between the 10 different cue identities – probability and magnitude, sorted from lowest (1) to highest (5) ranked picture – when presented on left and right options. For matrix element (*i,j*), color denotes Pearson's correlation coefficient of Z-scored firing rate between condition *i* and condition *j* across the OFC neuronal population. Firing rate is averaged from 100ms to 500ms post-stimulus (see also Supplementary Video 1). Note this matrix collapses across top/bottom spatial position; see Fig. S4 for 40x40 matrix, splitting conditions by spatial position. (**B**)/(**C**) As panel (**A**), but for DLPFC and ACC populations respectively. (**D**)-(**H**) Results from multiple linear regression of 'templates' onto RSA matrices. In each column, top subpanel shows template matrix, middle subpanel shows corresponding regression T-statistics for each region, and bottom subpanel shows coefficient of partial determination from sliding regression of templates onto RSA matrices, using sliding window of +/- 100ms. See Methods for full description of template matrices, regression model and statistical inference via non-parametric permutation test. (**D**) 'Spatial attention' template (differentiating cues on left vs.

right-hand option) was particularly prominent in DLPFC (DLPFC: T_{399}=35.782, p<$1*10^{-4}$), and significantly more so than other regions (one-way ANOVA $F_{2,1179}$ = 311.18, p<$1*10^{-4}$; post-hoc comparison for DLPFC>ACC: p<$1*10^{-4}$; post-hoc comparison for DLPFC>OFC: p<$1*10^{-4}$); **(E)** 'stimulus identity' template (responding similarly to the same cue irrespective of side) was particularly strong in OFC (T_{399}=15.3173, p<$1*10^{-4}$), again more so than other regions ($F_{2,1179}$ = 32.77, p<$1*10^{-4}$; post-hoc comparisons: OFC>ACC: p<$1*10^{-4}$; OFC>ACC: p<$1*10^{-4}$); **(F)** 'Attended value' template (representing cue 1 value irrespective of stimulus location or attribute) was prominent in OFC (T_{399}= 5.0697, p=0.0036), and significantly more so than in other regions ($F_{2,1179}$ = 6.73, p=0.0126; post-hoc comparisons: OFC>DLPFC: p=0.0017; OFC>ACC: p=0.0002); **(G) '**Left/right value' template was prominent principally within ACC (T_{399}=5.5151, p=0.0096) and DLPFC (T_{399}=5.7156, p=0.0062), and significantly more so than in OFC ($F_{2,1179}$ = 9.08, p=0.011; post-hoc comparisons: ACC>OFC: p=0.0003; DLPFC>OFC: p=0.011); **(H)** 'Accept/reject' template (reflecting whether Cue 1 had high (rank 4 or 5) versus low value (rank 1 or 2)) was strongest in ACC (T_{399}=12.1217, p<$1*10^{-4}$) and significantly more so than other regions ($F_{2,1179}$ = 17.20, p=0.0006; post-hoc comparisons: ACC>OFC: p<$1*10^{-4}$; ACC>DLPFC: p=0.0004). In panels **(D)-(F)**, * denotes p<0.05, ** denotes p<0.005, *** denotes p<0.0005. See also Fig. S5.
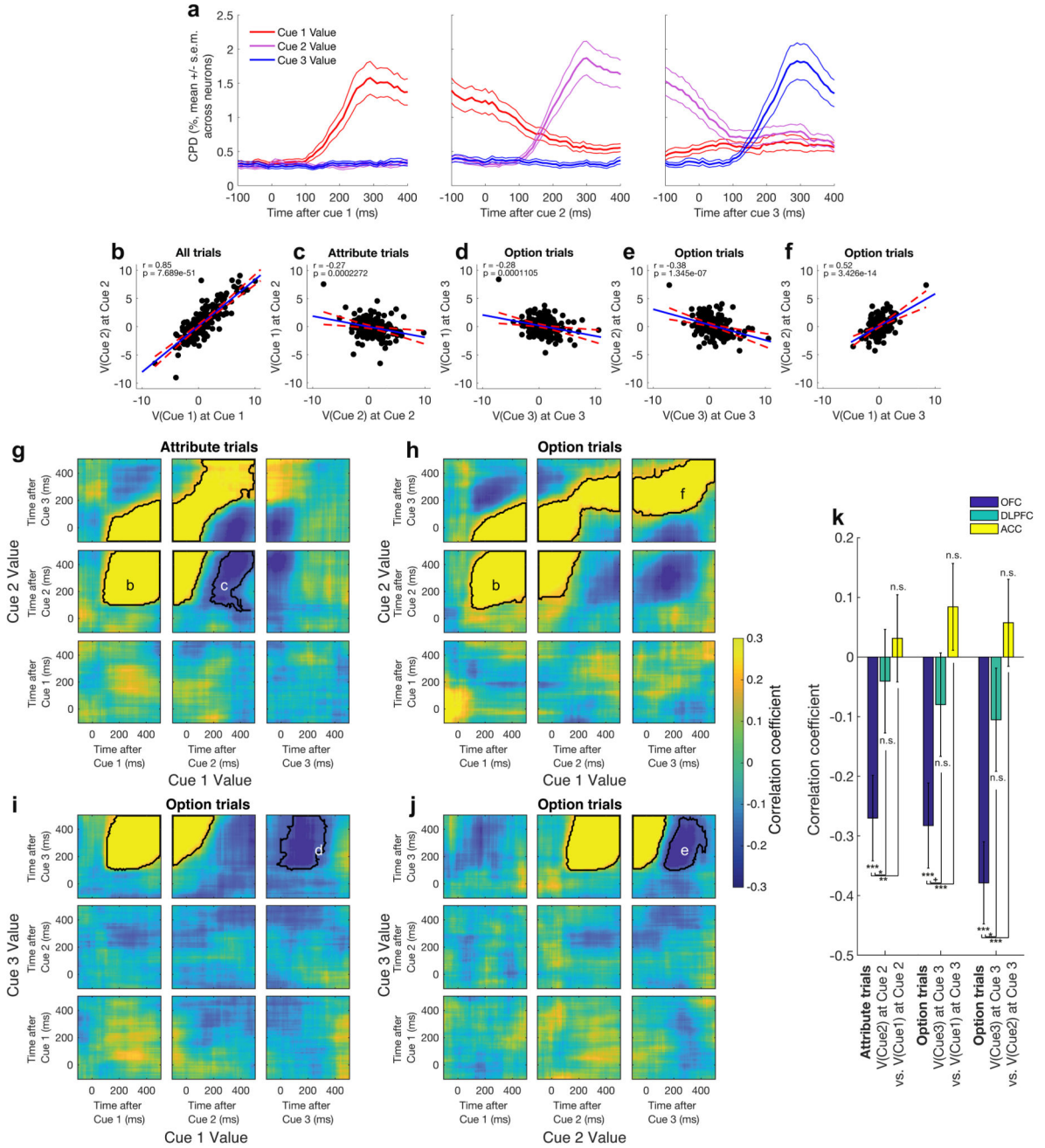
Fig. 5. Valuation subspaces for attended ('online') and stored cues, supporting attention-guided value comparison in OFC (n=183 units).
(A) Coefficient of partial determination for Cue 1 Value, Cue 2 Value and Cue 3 value, timelocked to each cue's presentation using a sliding 200ms window. Lines denote mean +/- s.e.m. across neurons. **(B)** Positive relationship between T-statistics from regression of value of cue 1 when cue 1 is presented (ordinate), and value of cue 2 when cue 2 is presented (abscissa), implying a stable subspace reflecting the currently attended value. **(C)** When cue 2 is being attended on attribute trials, the online value subspace (ordinate) correlates

negatively with the subspace reflecting the stored value of cue 1 (abscissa). **(D)-(F)** When cue 3 is being attended on option trials, the attended value subspace (ordinate of **(D)** and **(E)**) correlates negatively with the subspace for stored values of both cue 1 (abscissa of **(D)**) and cue 2 (abscissa of **(E)**). The two stored subspaces are positively correlated **(F)**. **(G)-(J)** Cross-correlation matrices reflecting the time-varying relationship between different value subspaces on attribute trials (in **(G)**) and option trials (in **(H)-(J)**). Heat maps reflect the correlation coefficient between T-statistics from regression for each cue's value across the OFC population. Superimposed letters refer back to the correlations that are plotted above, in panels **(B)-(F)**. Black lines denote significant clusters (p<0.001, cluster-based permutation test, corrected for multiple comparisons). **(K)** Negative correlations in parts **(C)-(E)** are only present in OFC, and significantly stronger in OFC than in DLPFC (n=135 units) and ACC (n=189 units); bars show correlation coefficient +/- s.e. (*** = p<0.001, ** = p<0.01, * = p<0.05 (two-tailed), '+' = p<0.05 (one-tailed), Fisher's r-to-Z transform). Correlations in parts **B-F** and **K** are defined using a window of 150-350ms post-stimulus onset. All correlation coefficients are Pearson's correlation. See also Figs. S6/S7.
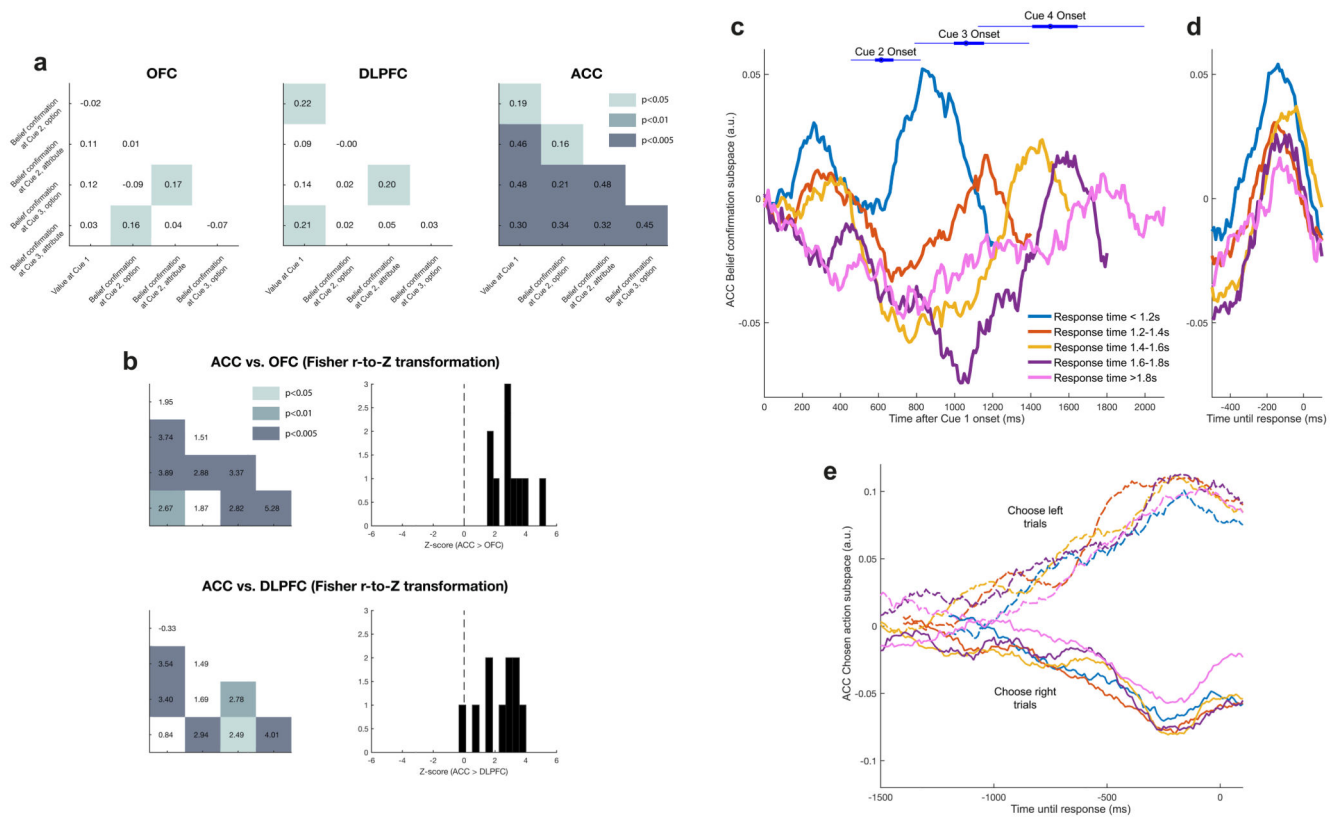
Fig. 6. Multiple signals in ACC reflect belief confirmation, commitment to a course of action, and accumulation of evidence for left/right movement (n=189 units).

(A) Positive population correlations between four orthogonal regressors that reflect 'belief confirmation' at cue 2/cue 3 on both option and attribute trials (cf. Fig. S8), and initial value population response at cue 1 (cf. Fig. 4H), demonstrating a stable belief confirmation subspace in ACC across multiple cues. The value inside each box denotes the Pearson correlation of parameter estimates across the neural population in each region; shading denotes significance. See Fig. S10 for individual ACC correlations. (**B**) Formal comparison of correlations shown in part (**A**) using Fisher's r-to-Z transformation against OFC (n=183 units) and DLPFC (n=135 units); left panels show the individual Z-scores/significance for each comparison (same layout as in part (**A**)), right panels show distribution of Z-scores. **(C)/(D)** Projecting ACC population activity into belief confirmation subspace reveals ramping immediately prior to the commitment to joystick movement. In part **(C)**, trials are sorted by response time and time-locked to cue 1 onset (blue bars above are box plots of cue onset times for other cues); in part **(D)**, trials are time-locked to response onset. **(E)** An orthogonal subspace in ACC reflects the emergence of an action plan to choose the left or right option. Activity in this subspace progressively favors one action over the other across time. See Supplementary Video 2 for relationship between belief confirmation and action selection subspaces.
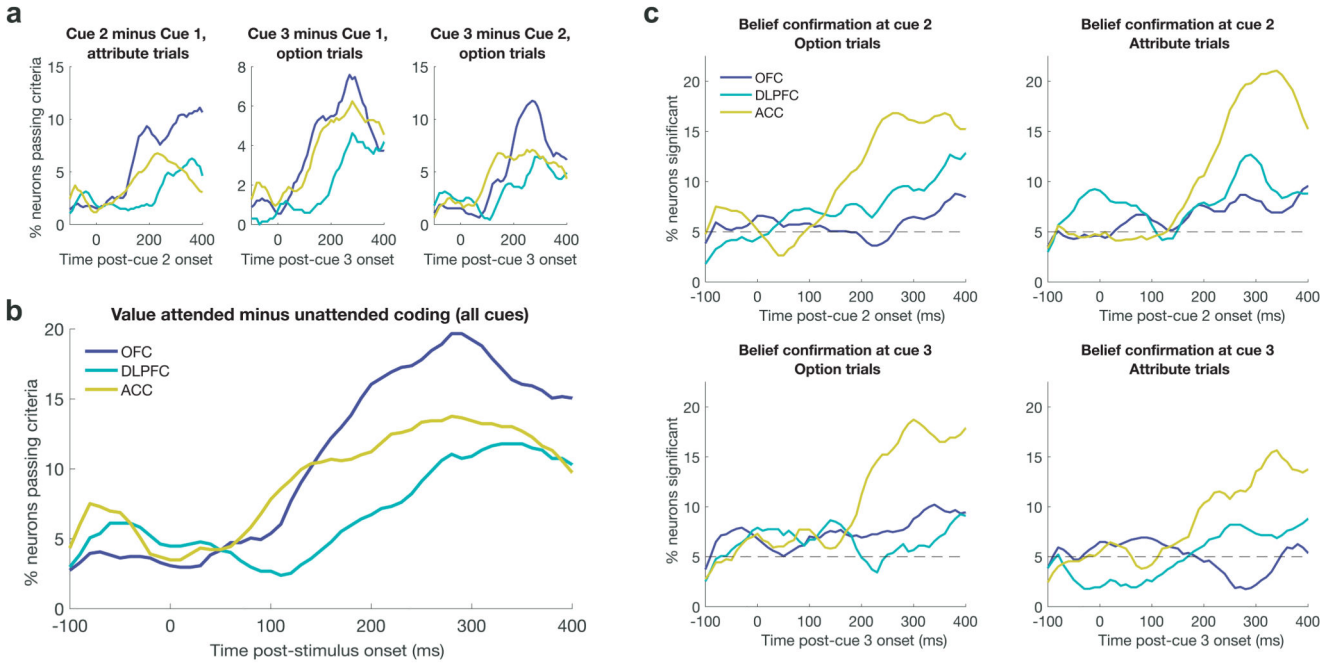
Fig 7. Single neuron analysis of attention-guided value comparison and belief confirmation recapitulates findings of population analysis.

(**A**) Attention-guided value comparison can be defined using three separate epochs and sets of regressors (see text for details). All three definitions show a greater proportion on neurons significantly encoding attention-guided value comparison in OFC (n=183 units) than in DLPFC (n=135 units) or ACC (n=189 units). (**B**) This figure shows the percentage of neurons that encoded value attended minus value unattended across any of the three different epochs in part (**A**). (**C**) There are four separate ways of defining 'belief confirmation' (see text and Fig. S8 for details). All four definitions show a greater proportion of neurons significantly encoding belief confirmation in ACC than in OFC or DLPFC.
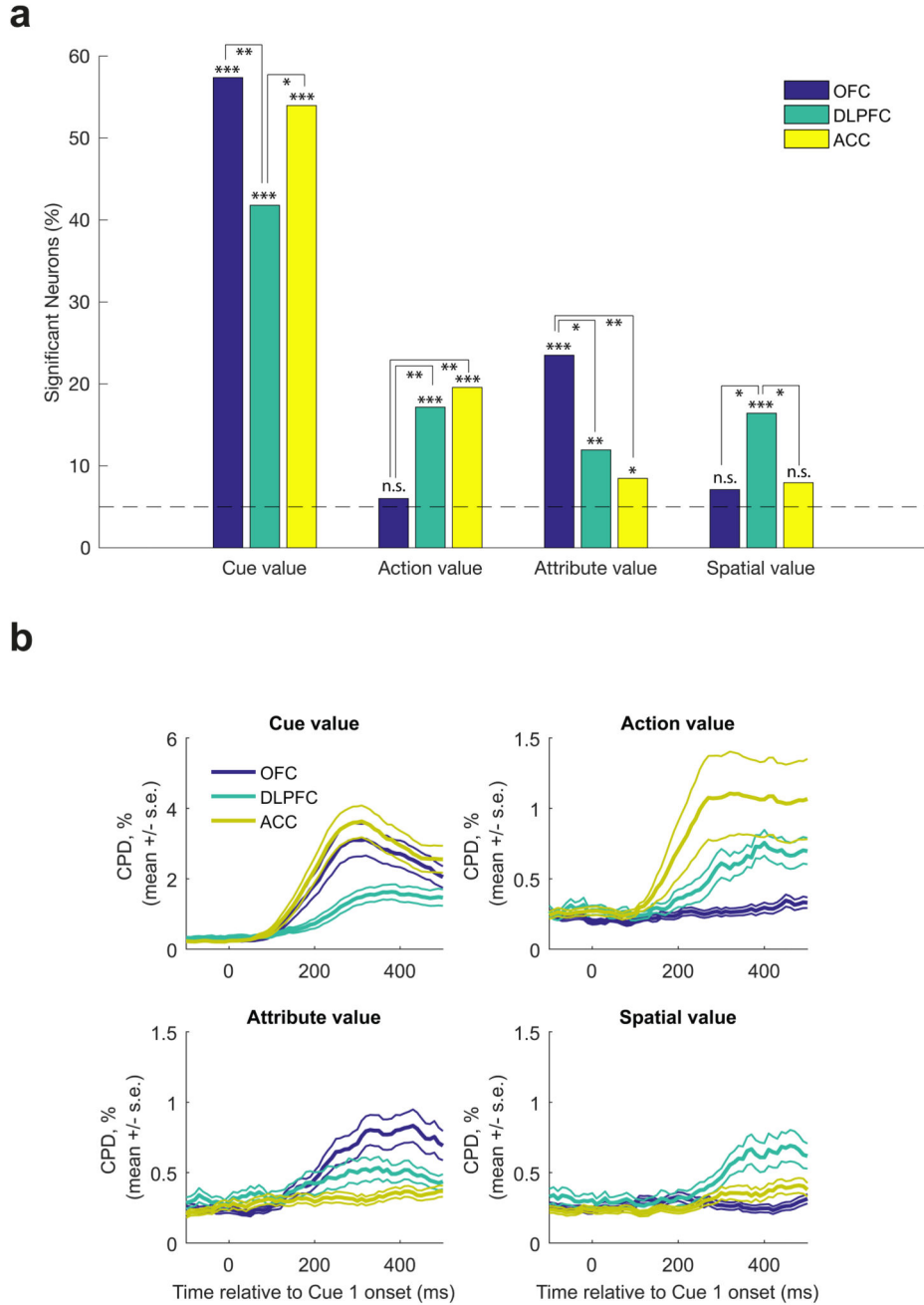
Fig 8. Single neuron correlates of task variables at Cue 1 presentation.

(**A**) Fractions of ACC (n=189 units), DLPFC (n=135 units) and OFC (n=183 units) neurons significantly encoding cue value, action value (left minus right), attribute value (probability minus magnitude), and spatial value (top minus bottom) in firing rates 100-500ms post-Cue 1 onset. Significance denotes two-sided binomial test against chance encoding (5%) for each bar (Cue value: ACC $p<1*10^{-5}$, DLPFC $p<1*10^{-5}$, OFC $p<1*10^{-5}$; Action value: ACC $p<1*10^{-5}$, DLPFC $p<1*10^{-5}$, OFC p=0.4139; Attribute value: ACC p=0.029, DLPFC p=0.0023, OFC $p<1*10^{-5}$; Spatial value: ACC p=0.0572, DLPFC $p<1*10^{-5}$, OFC

p=0.0845), and pairwise Chi^2 tests for comparisons between regions; * denotes p<0.05, ** denotes p<0.01, *** denotes p<0.001. (**B**) Coefficient of partial determination estimated in sliding 200ms bins for cue, action, attribute and spatial value in each region (mean +/- s.e. across neurons (ACC, n=189 units; DLPFC, n=135 units; OFC, n=183 units)).