

Double dissociation of value computations in orbitofrontal and anterior cingulate neurons

Steven W Kennerley^{1,2}, Timothy E J Behrens^{1,3} & Jonathan D Wallis^{2,4}

Damage to prefrontal cortex (PFC) impairs decision-making, but the underlying value computations that might cause such impairments remain unclear. Here we report that value computations are doubly dissociable among PFC neurons. Although many PFC neurons encoded chosen value, they used opponent encoding schemes such that averaging the neuronal population extinguished value coding. However, a special population of neurons in anterior cingulate cortex (ACC), but not in orbitofrontal cortex (OFC), multiplexed chosen value across decision parameters using a unified encoding scheme and encoded reward prediction errors. In contrast, neurons in OFC, but not ACC, encoded chosen value relative to the recent history of choice values. Together, these results suggest complementary valuation processes across PFC areas: OFC neurons dynamically evaluate current choices relative to recent choice values, whereas ACC neurons encode choice predictions and prediction errors using a common valuation currency reflecting the integration of multiple decision parameters.

Prefrontal cortex (PFC) supports optimal and rational decision-making. Damage to areas such as orbitofrontal cortex (OFC), anterior cingulate cortex (ACC) and lateral prefrontal cortex (LPFC) is associated with severe decision-making impairments^{1–8}, impairments not typically found with damage outside PFC. This implies that PFC neurons must support value computations that are essential for decision-making. However, although anatomical subdivisions within the PFC appear specialized in their function when examined with circumscribed lesions^{2,4,5} or functional magnetic resonance imaging^{9–12}, such regional dissociations have rarely been described in the underlying neuronal activity. Similar determinants of an outcome's subjective value can be found in the firing rates of neurons in ACC, OFC and LPFC^{13–26}.

Although there is a technical reason for this apparent discrepancy—very few studies have simultaneously recorded single neurons from multiple PFC areas to identify regional specialization—there is also a more fundamental difficulty in uncovering such regional dissociations: the firing rates of PFC neurons show substantial heterogeneity even within a PFC subregion²⁷. This heterogeneity is evident in two forms. First, neurons recorded millimeters (or less) from each other may encode different features of the decision process²⁸: neurons encode different parameters that influence the value of the choice (for example, reward size or type; delay, risk, effort associated with obtaining reward)^{13,15,17,22–24}, or neuronal activity may be modulated by different trial events (for example, choice, movement, outcome)^{15,25}. Second, and potentially more problematic, neurons that encode the same value parameter often do so with opponent encoding schemes: while one neuron may increase firing rate as a function of value, its neighbor may increase firing rate as value decreases^{15,17,19,21,23–25,29–31}. This feature of PFC neurons is in stark contrast to activity measured, for example,

in dopamine neurons, which encode reward and reward prediction error at predictable times, using a unified coding scheme^{32–34}.

Such heterogeneity of prefrontal coding makes it particularly important to examine the activity of single neurons but has rendered it difficult to dissociate regional patterns of PFC neuronal firing according to qualitative, rather than simply quantitative, differences. To address this issue, we simultaneously recorded the activity of single neurons in ACC, LPFC and OFC while monkeys (*Macaca mulatta*) made choices that varied in both the cost and benefit of a decision. We report two clear functional dissociations between PFC areas. First, a subpopulation of ACC neurons showed two unique features that distinguish these neurons from other value coding neurons: (i) at choice, these ACC neurons encoded all decision variables using a unified coding scheme of positive valence, and (ii) these same neurons also encoded reward prediction errors. Such activity was absent in OFC and LPFC. Second, we show that neurons in OFC, but not ACC, encode the value of current choices relative to the recent history of choice values.

RESULTS

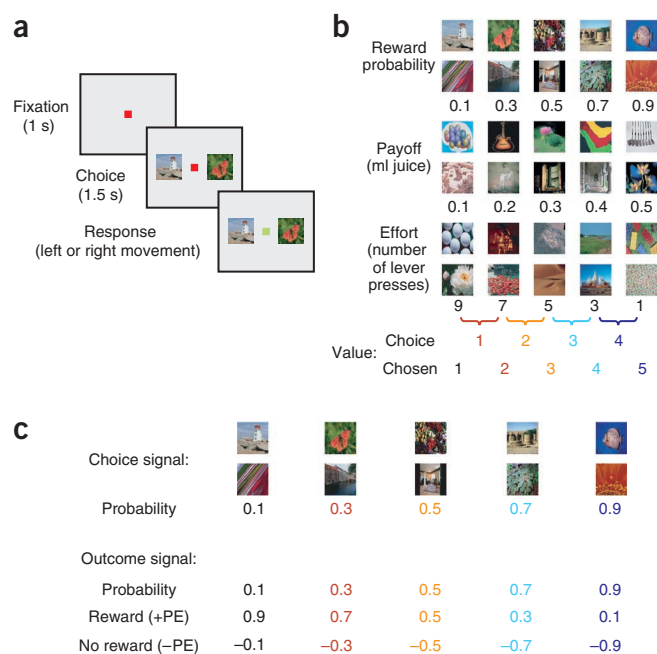
We trained two subjects to choose between pairs of pictures associated with different probabilities or sizes of reward, or different physical effort (cost) to obtain reward (**Fig. 1**)^{15,25}. In each trial the two stimuli varied along only one decision variable; hence, there was always a correct choice. Trials from each decision variable were intermixed and randomly selected. Behavioral analyses revealed¹⁵ that both subjects performed at a high level, choosing the more valuable outcome on 98% of the trials. We report neuronal activity from 257, 140 and 213 neurons located in LPFC, OFC and ACC, respectively (see Online Methods and **Supplementary Fig. 1** for recording locations).

¹Institute of Neurology, University College London, London, UK. ²Helen Wills Neuroscience Institute, University of California, Berkeley, Berkeley, California, USA.

³Oxford Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB), University of Oxford, John Radcliffe Hospital, Oxford, UK. ⁴Department of Psychology, University of California, Berkeley, Berkeley, California, USA. Correspondence should be addressed to S.W.K. (s.kennerley@ucl.ac.uk).

Received 22 June; accepted 19 September; published online 30 October 2011; doi:10.1038/nn.2961

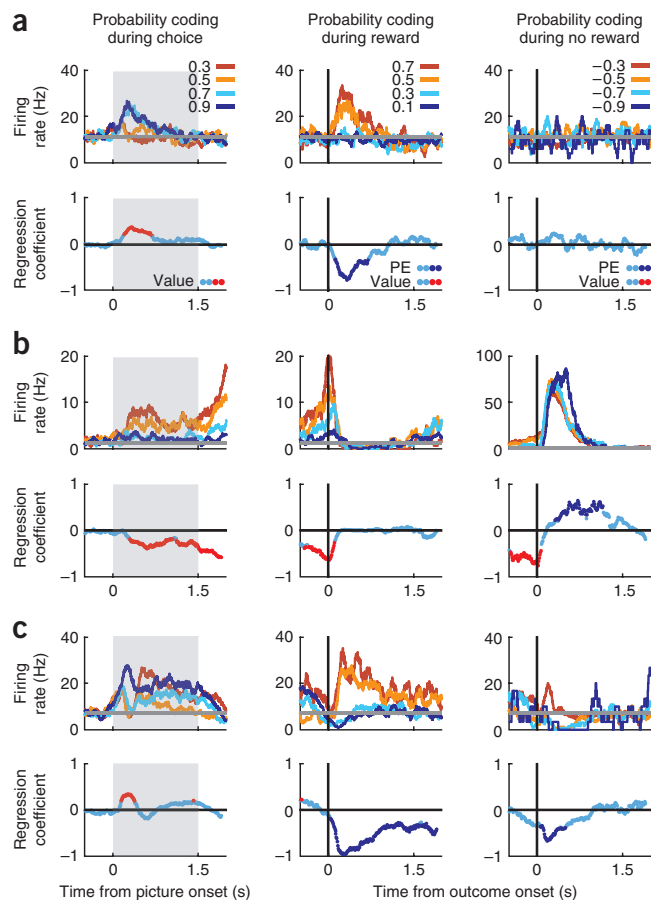
Figure 1 The behavioral task and experimental contingencies. **(a)** Subjects made choices between pairs of presented pictures. **(b)** There were six sets of pictures, each associated with a specific outcome. We varied the value of the outcome by manipulating either the amount of reward the subject would receive (payoff), the likelihood of receiving a reward (probability) or the number of times the subject had to press a lever to earn the reward (effort). We manipulated one parameter at a time, holding the other two fixed. Pictures presented were always adjacent to one another in terms of value; that is, choices were 1 versus 2, 2 versus 3, 3 versus 4 or 4 versus 5; hence, there were four choice values (1–4) per picture set and decision variable. All neuronal analyses were based on the chosen stimulus value (1–5). **(c)** Relationship between the probability of receiving a reward and various value-related parameters. During the choice phase, neurons could encode the value of the choice (the probability of receiving a reward). During the outcome phase, neurons could encode a value signal reflecting the probability of reward delivery, or they could encode a prediction error, which is the difference between the subject's expected value of the choice and the value of the outcome that was actually realized. On rewarded trials, these values would always be positive (+PE) whereas on unrewarded trials they would always be negative (–PE).



We focus here on understanding the relationship between value encoding at both choice and outcome, and determining whether the encoding of choice value is sensitive to the previous history of choices and experienced outcomes.

Encoding of outcomes relative to choice expectancies

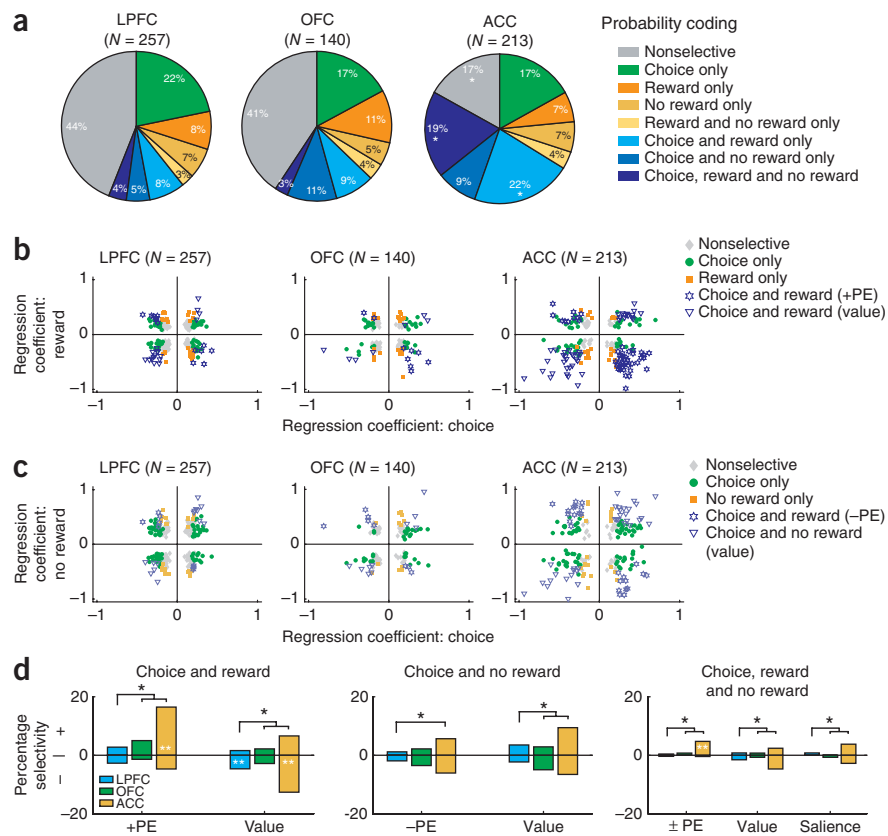
A hallmark of a prediction error neuron is that it encodes the discrepancy between expected and experienced outcomes. Thus, its firing rate should significantly correlate with probability at both choice and outcome, but should do so with an opposite relationship (Fig. 1c).



For example, dopamine neurons show a positive relationship between firing rate and reward probability when the stimulus is evaluated, and a negative relationship during the experienced outcome³². In contrast, if a neuron were simply encoding reward probability at choice and outcome (that is, a value rather than prediction error signal), then its firing rate should have the same relationship with probability during both epochs. Many ACC neurons showed increased activity when outcomes were better than expected (Fig. 2a), consistent with encoding a positive prediction error (+PE). Other neurons encoded a negative prediction error (–PE; Fig. 2b), indicating the outcome was worse than expected. In addition, some ACC neurons encoded probability at choice and encoded both a +PE and –PE at outcome (Fig. 2c), an activity pattern very similar to dopamine neurons³². Finally, some neurons encoded reward probability at both choice and outcome with the same pattern of activity, indicative of an outcome value (rather than prediction error) signal (Supplementary Fig. 2a,b).

Figure 2 Three single-neuron examples showing choice and outcome activity on probability trials during rewarded and unrewarded trials. For each neuron, the upper row of plots illustrates spike density histograms sorted according to the chosen stimulus value (choice epoch) or the size of the prediction error (outcome epochs). Data are not plotted for the least valuable choice because it was rarely chosen by the subjects. The horizontal gray line indicates the neuron's baseline firing rate as determined from the 1-s fixation epoch immediately before the onset of the pictures. Gray shading indicates choice epoch. The trace in the lower row of plots illustrates the change in magnitude of regression coefficients for reward probability coding across the course of the trial. Red data points indicate significant value encoding at choice (GLM-1 t -statistic ≥ 3.1 for two consecutive time bins; see Online Methods); dark blue data points indicate significant prediction error encoding at outcome (GLM-2 t -statistic ≥ 3.1 for two consecutive time bins). **(a)** An ACC neuron that encoded +PE. It increased firing rate during choice as chosen probability increased, and it increased firing rate during rewarded outcomes when the subject was least expecting to receive a reward (low-probability trials). **(b)** An ACC neuron that encoded –PE. It increased firing rate during choice as chosen probability decreased, and it increased firing rate during unrewarded outcomes when the subject was expecting to receive a reward (high-probability trials). **(c)** An ACC neuron that encoded chosen probability and at outcome encoded +PE and –PE.

Figure 3 Relationship between neuronal encoding during the choice and outcome epoch. **(a)** Prevalence of neurons encoding probability during different task events (choice, rewarded outcomes and/or unrewarded outcomes). * $P < 0.05$ (chi-squared tests), significant difference between proportion of selective neurons in ACC and proportion in LPFC and OFC. More ACC neurons encoded reward probability during both choice and outcome phases (blue shading). **(b,c)** Scatter plots of the regression coefficients for probability coding during choice and rewarded outcomes **(b)** and choice and unrewarded outcomes **(c)** for all neurons in each brain area. The point of maximum selectivity for each neuron and epoch is plotted, so data are biased away from zero. **(b)** ACC neurons that encoded probability positively during the choice encoded probability negatively during rewarded outcomes (star), consistent with a +PE. Neurons that encoded probability positively during rewarded outcomes, consistent with encoding a value signal at outcome (triangle). **(c)** There was no consistent relationship between the encoding of probability during the choice and unrewarded outcomes. **(d)** Bar plots summarizing the different value representations, based on the sign of the choice probability regression coefficient (+ or -). Single black asterisks, * $P < 0.05$ for differences between PFC areas (chi-squared tests); double white asterisks, ** $P < 0.05$ for the proportion of neurons with positive or negative regression coefficients different from the chance 50%-50% split (binomial test). Position of white asterisks indicates the larger population.



To distinguish whether neurons encoded a prediction error or a value signal at outcome, we compared the signs of the regression coefficients for encoding reward probability during the choice and outcome epochs (general linear models (GLMs) 1 and 2, respectively; see Online Methods). ACC neurons were significantly more likely than LPFC and OFC neurons to encode reward probability during both the choice and outcome epochs (Fig. 3a). Furthermore, many ACC neurons encoded reward probability information at both choice and outcome with oppositely signed regression coefficients, indicative of a prediction error signal (Fig. 3b,c). Indeed, ACC neurons were significantly more likely to encode +PE ($\chi^2 = 33$, $P < 9 \times 10^{-8}$), -PE ($\chi^2 = 13$, $P = 0.0016$) or both +PE and -PE ($\chi^2 = 10$, $P = 0.0085$) compared to both LPFC and OFC (Fig. 3d). ACC neurons were also significantly more likely to encode +PE than -PE ($\chi^2 = 6.2$, $P = 0.013$). Finally, neurons that encoded reward probability in the same direction during the choice and outcome—indicative of an outcome value signal—were also most commonly found in ACC (Fig. 3d; $\chi^2 = 24$, $P < 7 \times 10^{-6}$).

A notable feature of ACC activity was that the sign of the regression coefficient at choice distinguished the type of signal that would be encoded at outcome—prediction error or value representation (Fig. 3d). ACC neurons that encoded +PE (35/45; z-score = 3.7, $P < 0.0002$, binomial test) or both +PE and -PE (10/11; z-score = 2.7, $P = 0.0059$, binomial test) were significantly more likely to show a positive relationship between firing rate and probability during choice. In contrast, neurons that encoded a value signal on rewarded trials were significantly more likely to do so when the choice regression coefficient was negative (27/41; z-score = -2.2, $P = 0.03$, binomial test).

Thus, if an outcome-selective neuron encoded reward probability positively at choice, then it flipped sign and encoded a prediction error at outcome; but if it encoded probability negatively at choice, then it encoded a value signal at outcome (Fig. 3b-d). Only 17 of 610 neurons encoded a saliency signal (see Online Methods and Supplementary Fig. 2c).

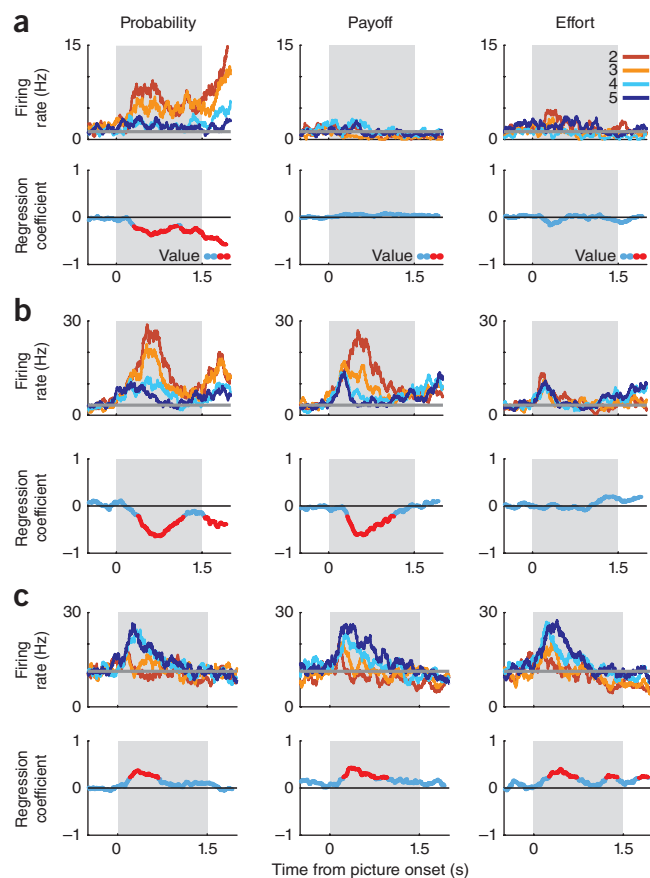
ACC multiplexes choice value with a positive valence

The previous analysis indicated that value coding during the choice phase predicted different value representations. Neurons that encoded chosen probability with a positive valence subsequently encoded a prediction error at outcome, whereas neurons that encoded chosen probability with a negative valence encoded a value signal at outcome. To determine whether the sign of the regression coefficient at choice could predict other types of value representations, we compared the relationship between firing rate and choice value for all three decision variables (GLM-1). PFC neurons showed considerable heterogeneity in the representation of choice value; whereas some PFC neurons specifically encoded one decision variable (Fig. 4a), other neurons flexibly encoded multiple decision variables (Fig. 4b,c). Value coding for each decision variable was present in all PFC areas, but each was significantly more prevalent in ACC than in LPFC or OFC (Fig. 5a, $\chi^2 > 23$, $P < 2 \times 10^{-6}$ in all cases). Neurons encoding all three decision variables (Fig. 4c) were also more common in ACC than in LPFC or OFC (Fig. 5a, $\chi^2 = 42$, $P < 7 \times 10^{-10}$). Notably, across all three decision variables and all three PFC areas, neurons were equally likely to increase firing rate as value either increased or decreased (Fig. 5a).

Figure 4 Three neurons that encode the value of different choice variables. For each neuron, the upper row of plots illustrates spike density histograms sorted according to the value (2–5) of the expected outcome of the choice. Data are not plotted for the least valuable choice because it was rarely chosen by the subjects. The horizontal gray line indicates the neuron's baseline firing rate as determined from the 1-s fixation epoch immediately before the onset of the pictures. The trace in the lower row of plots illustrates the change in the magnitude of regression coefficients for value coding across the course of the choice epoch. Red indicates time bins with significant value coding (GLM-1 t -statistic ≥ 3.1 for two consecutive time bins; see Online Methods). (a) An ACC neuron that encoded value solely on probability trials with an increase in firing rate as the value of the choice decreased. (b) An ACC neuron that encoded value on probability and payoff trials, increasing its firing rate as value decreased. (c) An ACC neuron that encoded value for all decision variables, increasing its firing rate as value increased.

The noteworthy prevalence of opponent encoding of choice value may have important implications for interpreting brain activity in human studies, which typically use measures of average neuronal activity (for example, electroencephalography, functional magnetic resonance imaging) that may cancel out opposing value signals. To look at this issue directly, we calculated the coefficient of partial determination³⁵ (CPD; see Online Methods), which measures how much variance in each neuron's firing rate is explained by each GLM regressor (Fig. 5b–d). We determined the absolute mean CPD for all neurons in each PFC area, and found that chosen value consistently explained more of the variance in firing rate in ACC neurons than in LPFC or OFC neurons (Fig. 5b). We then sorted neurons by the sign of their regression coefficient and calculated the CPD for each population (Fig. 5c). As suggested by the single-neuron analysis, choice value coding averages to zero in all three PFC regions when both the positive and negative neuronal populations were averaged together (Fig. 5d). This is particularly noteworthy given the prevalence of value selectivity (Fig. 5a): the lack of a net value signal at the population level is the consequence of many highly selective neurons encoding the same value parameter but with opposite coding schemes.

However, there was a clear exception to this averaging effect: at the time of choice, ACC neurons that encoded all three decision variables showed a strong bias to encode value positively at



both the single-neuron (Fig. 5a, right column; 49/70; z -score = 3.3, $P = 0.0005$, binomial test) and population level (Fig. 5d, right column). Thus, both in number and in total variance explained, these ACC neurons are a specific subpopulation that multiplex value using a unified coding scheme of increasing firing rate as chosen value increases.

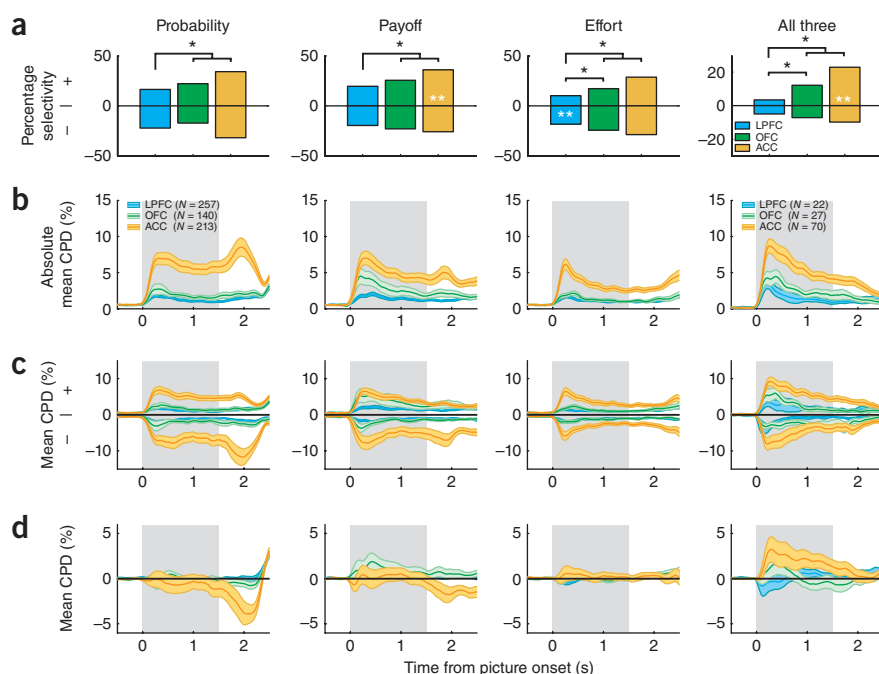
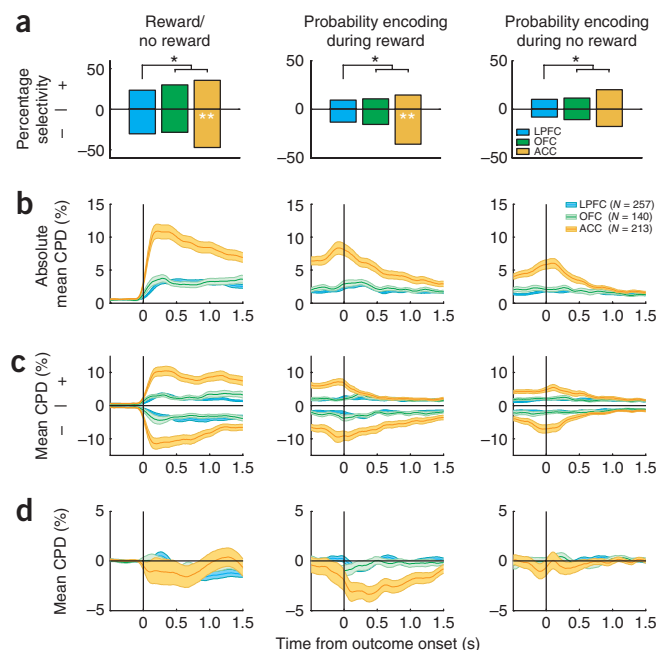
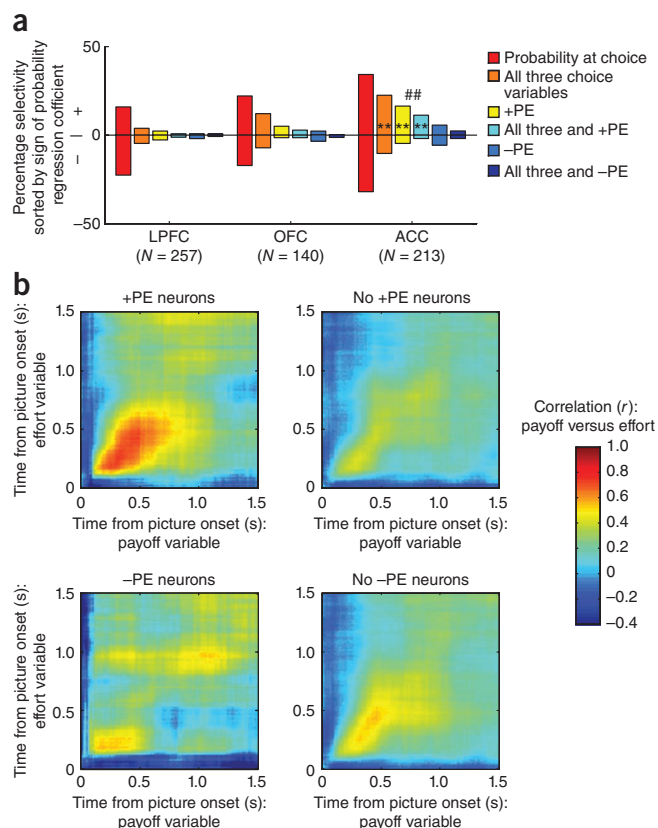


Figure 5 Population encoding of value during the choice epoch. (a) Prevalence of neurons encoding choice value with a positive or negative regression coefficient. Conventions as in Figure 3d. (b–d) The mean (solid line) and standard error (shading) of value selectivity as determined by the CPD for each value regressor. For the first three columns (probability, payoff, effort), the CPD is calculated from GLM-1 and averaged across all recorded neurons. The final column (all three) is restricted to neurons that encoded all three decision variables, and the CPD is calculated from GLM-3. (b) The absolute mean CPD across all neurons in each brain area. (c) The mean CPD plotted separately for those neurons that encoded choice value with a positive (+) or negative (–) regression coefficient. The two populations encoded approximately the same amount of choice value information. (d) The population mean CPD, averaged across neurons that encoded choice value with either positive or negative regression coefficients. The mirrored plots from panel (c) average so that approximately zero information about choice value remains at the population level, with the exception that the subpopulation of ACC neurons that encoded all three decision variables showed a positive bias (that is, spiking increased as choice value increased).

Figure 6 Population analyses of neuronal activity during the outcome epoch of probability trials. Conventions as in **Figure 5**. (a) Left: prevalence of neurons with positive or negative regression coefficients for encoding a categorical signal about reward presence or absence (for example, neurons with a positive regression coefficient increase firing rate more on rewarded than on unrewarded outcomes). Middle and right: prevalence of neurons with positive or negative regression coefficients for encoding reward probability on rewarded (middle) or unrewarded (right) trials (for example, neurons with a negative regression coefficient show a linear increase in firing rate as reward probability decreases). (b) The absolute mean CPD for the three different regressors from GLM-2. (c) The mean CPD plotted separately for those neurons that showed a positive or negative regression coefficient for the three different regressors. The two populations encoded approximately the same amount of outcome information. (d) The population mean CPD, averaged across neurons that showed both positive and negative regression coefficients for each of the three different regressors. The mirrored plots from panel (c) average so that approximately zero information about each regressor remains at the population level, with the exception that the ACC population retains information about rewarded outcomes with a negative bias (that is, spiking on rewarded outcomes increased as probability decreased, akin to a +PE).

Outcome negativity in ACC explained by prediction errors

A second exception to the averaging effect occurred during the outcome of probability trials: ACC showed a negative net value bias (**Fig. 5d**, left column). There are two possible sources of this bias. First, the negative bias could arise from signaling errors, as has been previously reported in ACC activity^{36,37}. If neurons signal an error, then their firing rate would increase as the chosen probability decreases because the frequency of unrewarded trials (errors) increases. Second, the negative bias could arise from the many single ACC neurons that encoded reward prediction errors (**Fig. 3b**). If neurons encode a +PE, then firing rate would increase as chosen probability decreases because a reward on a low-probability trial is 'better than expected'.



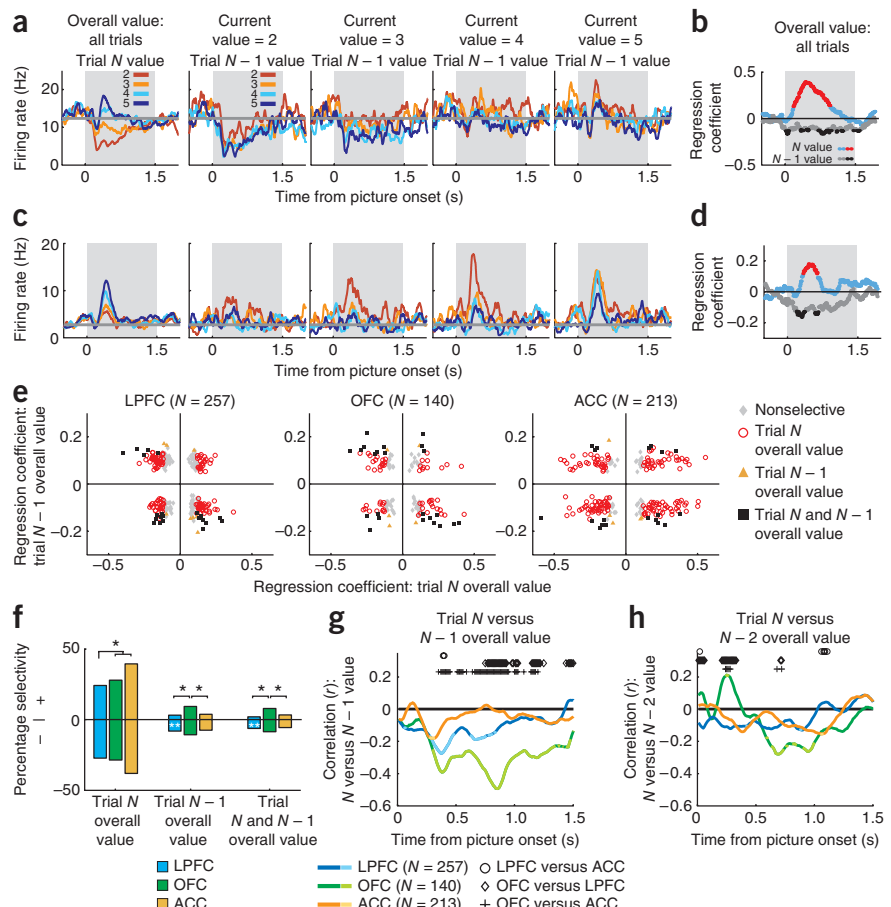
To determine the source of this negative bias, we examined whether neurons were sensitive to reward/no reward, or showed a linear relationship between firing rate and chosen probability on rewarded or unrewarded trials (GLM-2; see Online Methods). Neurons that encode reward/no reward reflect a categorical outcome; a positive or negative regression coefficient indicates that neurons are more responsive to reward presence or absence, respectively (the latter is commonly considered an error signal). Although ACC neurons tended to encode categorical reward delivery with a negative regression coefficient (**Fig. 6a**, left), the amount of outcome information (**Fig. 6b**, left) encoded by the populations with positive and negative regression coefficients was approximately equal (**Fig. 6c**, left) such that it averaged to zero and thus did not contribute to the negative bias at the population level (**Fig. 6d**, left). Similarly, although many ACC neurons were modulated by chosen probability on unrewarded outcomes (**Fig. 6a**, right), this information was absent at the population level (**Fig. 6d**, right). In contrast, ACC neurons showed a strong bias for a negative relationship with reward probability on rewarded trials (**Fig. 6a**, middle), leading to a large net negativity in the population response (**Fig. 6d**, middle).

Figure 7 Relationship between encoding common value and +PE.

(a) Proportion of ACC neurons encoding probability at choice with a positive or negative regression coefficient. ACC neurons that encoded all three decision variables during the choice and/or +PE at outcome were significantly more likely to encode probability during the choice phase with a positive (as opposed to negative) regression coefficient (** $P < 0.05$, binomial test). ACC neurons that encoded +PE at outcome were the same subpopulation of neurons that also encoded all three decision variables at choice (69% of +PE neurons also encoded all three decision variables, exceeding the expected frequency based on the odds of encoding payoff and effort information with a positive regression coefficient; ## $P < 0.01$, binomial test). (b) For those neurons that encoded probability during the choice, plots of the correlation of their value encoding on payoff trials with the value encoding on effort trials (GLM-1) separated according to whether they did or did not code either +PE or -PE at outcome. This establishes whether the likelihood of encoding payoff and effort value information depends on encoding +PE or -PE. Only neurons that encoded probability during the choice phase and +PE at outcome showed a significant correlation with value encoding on payoff and effort trials (Pearson's correlation, $P < 0.05$ corrected for multiple comparisons).

Figure 8 Neuronal encoding of value history.

(a) An OFC neuron encoding current and past trial choice value. This neuron increased firing rate as current choice value increased (leftmost plot). Additionally, with trials sorted by the $N - 1$ trial value, firing rate increased both as the current trial value increased and as the previous trial value decreased (four rightmost plots): this neuron is modulated by the difference in the current and past trial value. (b) Dynamics of the encoding shown in **a** as determined from the regression coefficients of GLM-3. Significant bins for current and $N - 1$ value in red and black, respectively (GLM-3 t -statistic ≥ 3.1 for two consecutive time bins; see Online Methods). (c,d) Data presented as in **a,b** for an LPFC neuron that encoded current trial value relative to previous trial value. (e) Scatter plot of regression coefficients for current (N) and past trial ($N - 1$) value for all neurons per brain area. Differently colored symbols indicate different selectivity patterns. (f) Proportion of neurons encoding current trial N and/or $N - 1$ trial chosen value at the time of the current trial choice, sorted by the sign of the regression coefficient for current trial value (left) or the sign of the regression coefficient for $N - 1$ trial value (middle, right). Conventions as in **Figure 3d**. (g,h) Mean correlation (r) between regression coefficients for encoding value of the current and previous trial (g) or current and two trials ago (h). Lighter colors signify the bins where the correlation was significant ($P < 0.01$). Symbols indicate significant differences in r values between areas (Fisher's Z transformation, $P < 0.01$).



In other words, the ACC negativity at outcome (**Fig. 5d**, left column) is driven primarily by low-probability trials being rewarded (that is, +PE response) rather than unrewarded outcomes (that is, error response).

Prediction errors use an abstract value scale

So far we identified two populations of ACC neurons that were biased toward encoding probability positively during choice: neurons that encoded value across all three decision variables (**Fig. 5**, right column) and neurons that encoded +PEs (**Fig. 3d**). Our results suggest these two selectivity patterns are found in the same population of neurons, but only in ACC (**Fig. 7a**). Neurons that encoded all three decision variables and +PE were significantly more likely to be found in ACC, but only when chosen probability was encoded with a positive regression coefficient. For example, the neuron in **Figure 4c** (choice activity) is the same neuron as in **Figure 2a** (outcome activity).

To test whether +PE neurons were also likely to encode information about all three choice variables, we sorted neurons that encoded probability during choice into two groups, depending on whether these neurons did or did not encode a prediction error during outcome. There is no a priori reason why one group should be more likely to encode information about payoff and effort than the other group. However, we found a very strong correlation between the encoding of payoff and effort information for the group of neurons that encoded +PEs, but not for the other groups (**Fig. 7b**). This correlation began as early as 200 ms after the start of the choice epoch, suggesting that these neurons are multiplexing choice value early enough to reflect a role in the decision process. Thus, there seems to be a special population of neurons in ACC that multiplex choice value across decision variables, do so with a positive valence and encode +PEs. This suggests that ACC encodes decision value and +PEs using a common value currency.

Relative value coding: the effect of value history

To facilitate optimal and adaptive choice, the brain must represent the reinforcement history of one's choices. We examined whether PFC neurons encoded the overall value history of choices (that is, value based on probability, payoff and effort) while controlling for reward history (actual reward amount received) (GLM-3; see Online Methods). Many OFC neurons were sensitive to both the current and previous trial choice value, showing increasing firing rate as the current trial value increased and as the past trial value decreased (**Fig. 8a,b**). In other words, such neurons encoded the current value relative to the value of the last choice. Some neurons in LPFC showed similar activity profiles (**Fig. 8c,d**).

We examined the distribution of regression coefficients for encoding current (N) and past ($N - 1$) trial value for all neurons (**Fig. 8e**) to determine the number of neurons in each PFC area that encoded the current and/or past trial value at the time of the current choice (**Fig. 8f**). Whereas 77% of ACC neurons encoded current choice value (significantly more than in LPFC and OFC, $\chi^2 = 34$, $P < 4 \times 10^{-8}$), only 11% of ACC neurons encoded past trial value. In fact, OFC was significantly more likely to encode past trial value ($\chi^2 = 6.2$, $P = 0.04$) or both current and past trial value (**Fig. 8f**; $\chi^2 = 6.4$, $P = 0.04$) compared to either ACC or LPFC. Further, the value history representation in OFC was not simply persistent activity from the previous trial, but rather it emerged at the time of the current choice as if it was used as a reference for the calculation of the current choice value (**Supplementary Fig. 3a–c**). Across several studies, this is the first time we have seen a value signal encoded more prevalently by OFC than ACC^{15,21,25}. In contrast, at the time of the current choice,

ACC neurons encoded the actual reward amount received from the previous trial (**Supplementary Fig. 3**; $\chi^2 = 6.0$, $P = 0.05$).

If neurons encode current choice value relative to the past choice values, we might expect neurons to encode both value signals simultaneously with a negative correlation. Across the entire population, the correlation between current and past trial value was significantly stronger in the OFC population than in either the LPFC or ACC population, and it was negative (**Fig. 8g**; Fisher's Z transformation, $P < 0.01$). This correlation emerged 300 ms into the current choice, consistent with when the current choice would be evaluated. The current choice value was even modulated by the value of choices two trials into the past (**Fig. 8h**), but only in OFC. This effect was not driven by a subpopulation of very selective neurons, as this correlation remained significant (**Supplementary Fig. 3d,e**; Fisher's Z transformation, $P < 0.01$) across the OFC population even when the most selective neurons were excluded (see **Supplementary Results**). Thus, the OFC neuronal population not only tracks the recent history of choice value, but uses this history to encode current choice value on a relative scale.

DISCUSSION

Decision-making frameworks highlight the importance of multiple value computations necessary for decision-making, including representing decision variables that influence current choices and encoding choice outcomes to guide future choices²⁸. This might explain why neuronal correlates of value have been identified across the entire brain, from PFC to sensorimotor areas^{16,27,38,39}. To understand the relationship between value signals and decision-making, we must identify what functions different value representations support. Recent comparative lesion studies suggest that despite the ubiquity of value representations across the brain, the computations supported by PFC areas make fundamental and specialized contributions to decision-making^{2,4–6}. Here we describe our findings that different PFC neuronal populations encode different value computations.

Prefrontal encoding of prediction errors

Although previous studies have reported an outcome signal that correlates with prediction error in both ACC^{19,40,41} and OFC⁴² (but see ref. 26), our study is, to our knowledge, the first to report single PFC neuron activity time-locked to both stimulus presentation and outcome in a probabilistic context, analogous to the conditions used to test for these signals in dopamine neurons^{32,34}. We identified three main findings. First, only ACC neurons encoded probability during the stimulus and outcome phases with an opposite relationship consistent with a prediction error signal. Second, ACC neurons more commonly encoded +PEs compared to –PEs, consistent with previous reports in both ACC⁴⁰ and dopamine neurons⁴³. We have also shown that focal ACC lesions cause reward-based, but not error-based, learning and decision-making impairments¹, consistent with a functional link between dopamine and ACC for learning from +PEs. Third, just like dopamine neurons³², prediction error neurons in ACC encoded reward probability with a positive valence during stimulus presentation and a negative valence during reward delivery.

An important issue for future consideration is whether the value encoding scheme is of relevance. Many PFC neurons encode value with opposing encoding schemes, whereby neurons are equally likely to encode value positively or negatively^{15,17,19,21,23–25,29,30}. Although opponent encoding schemes might be a fundamental feature of a decision-making network³¹, such encoding schemes introduce potential interpretation difficulties for techniques that average across neuronal

populations, such as those in human neuroimaging. Averaging across opponent encoding schemes could average away value signals, as it did in many of our population analyses. However, despite an even distribution of ACC neurons that showed a positive or negative relationship between choice probability and firing rate, the former and latter groups were predictive of a prediction error and a value signal at outcome, respectively, suggesting that these two populations have functionally distinct roles.

ACC has been associated with error-related activity^{36,37}. However, in many studies, errors occur less frequently than rewards, suggesting ACC outcome activity might instead reflect violations in expectancy^{44,45}, or how informative an outcome is for guiding adaptive behavior^{9,46,47}. We counterbalanced the likelihood of a rewarded or unrewarded (error) outcome on probability trials to unconfound this issue. A comparison of ACC outcome activity suggested that an approximately equal number of neurons were sensitive to reward and unrewarded (error) outcomes, but that error-related neurons were more likely to encode a categorical signal, whereas reward-related neurons were often modulated by prior reward expectancy, consistent with a +PE. Such results are consistent with our previous findings: ACC sulcus lesions did not impair error-related performance but impaired optimal choice after a new action was rewarded¹. This suggests that when choices are rewarded in the context of a novel, volatile or low reward expectation environment, a critical function of ACC may be to signal a +PE to indicate that the outcome is better than expected and hence promote future selection of that choice. In sum, rather than simply error detection, ACC encodes a rich outcome representation that incorporates prior expectancies and can inform future behavior²⁵.

Choice value coded in a common neuronal currency

Formal decision models suggest that determining an action's value requires the integration of the costs and benefits of a decision, thus generating a single value estimate for each decision alternative⁴⁸. This net utility signal can be thought of as a type of neuronal currency that can facilitate comparison of very disparate choice outcomes⁴⁸. However, such a valuation system operating in isolation would make learning about individual decision variables problematic, as it would not be possible to update the value estimate of individual variables (for example, effort costs in isolation from reward benefits) if the only signal the valuation system received was in terms of overall utility. It may be optimal for the brain to encode two types of choice value signals: a variable-specific value signal that represents why an option is valuable, and a net utility signal that reflects the integrated net value of the option to bias action selection.

Our results suggested that variable-specific value signals were equally prevalent across all three PFC areas¹⁵, whereas neurons that encoded all three decision variables occurred primarily in ACC. Furthermore, ACC neurons that encoded all three decision variables during choice used a unified coding scheme of positive valence and encoded +PE on rewarded outcomes. This suggests that both choice value and prediction error coding, at least in many ACC neurons, is based on a common valuation scale reflecting the integration of individual decision variables.

Reinforcement history and fixed versus adaptive coding

Despite the abundance of value signals encoded by ACC neurons during both choice and outcome epochs, the encoding of current choice value in ACC was insensitive to the overall value of previous choices. Instead, ACC neurons encoded the actual reward history of choices, consistent with previous reports^{1,19,42,49}. The encoding of reward history by ACC might reflect its role in tracking the reward rate, or volatility of possible outcomes, to promote adaptive behavior⁴⁷.

Previous studies have shown that some ACC and OFC neurons adapt their firing rate when the scale of choice values is altered^{14,24,29,30}, whereas other neurons encode a fixed value representation invariant to changes in the set of offers available^{30,50}. Adaptive value coding is efficient because it provides flexibility to encode value across decision contexts that differ substantially in value, whereas fixed value coding is important because it provides value transitivity, a key characteristic of economic choice⁵⁰.

Our results indicate that adaptive coding is specific to OFC and is extremely dynamic, operating on a trial-by-trial basis. OFC neurons showed a stronger response when the current trial value differed by a large amount from the previous trial. Yet OFC neurons did not encode a prediction error at outcome²⁶, so this cannot be characterized as a general discrepancy signal. Adaptive coding in OFC also differed from ACC prediction error coding in its distributed nature. The OFC neurons that had weak or nonsignificant value coding still showed a robust anticorrelation between current and past value, even for outcomes more than one trial into the past. Thus, adaptive value coding is a general feature of the OFC population, whereas prediction error computations are performed by a subpopulation of ACC neurons.

Interpretational issues

There are some caveats to our results. First, single-unit neurophysiology has sampling bias. We strove to reduce this bias (for example, by not prescreening neurons for response properties), but it is impossible to eliminate. The most easily detectable electrical signals are generated by the largest neurons, and different areas may have subtly different sampling biases. Nevertheless, simultaneous recording eliminates many confounds associated with comparison of neurophysiological responses in different brain areas. Such comparisons frequently must be made across different studies from different labs using different behavioral and statistical methods. Second, the precise behavioral contingencies and training history may have influenced the direction of our effects. Indeed, neuroimaging studies suggest ACC activations can be stronger to errors or rewards depending on which outcome is more frequent as a result of task design^{44,45}. Future research can test the generalizability of our findings to other behavioral contexts.

In conclusion, our results show complementary encoding of value between ACC and OFC. ACC encodes value on a fixed scale using a common value currency and uses this value signal to encode prediction errors: how choice outcomes relate to prior expectancies. In contrast, OFC encodes an adaptive value signal that is dynamically adjusted on the basis of the recent value history of choice environments.

METHODS

Methods and any associated references are available in the online version of the paper at <http://www.nature.com/natureneuroscience/>.

Note: Supplementary information is available on the Nature Neuroscience website.

ACKNOWLEDGMENTS

The project was funded by US National Institute on Drug Abuse grant R01DA19028, US National Institute of Neurological Disorders and Stroke grant P01NS040813 to J.W. and US National Institute on Mental Health training grant F32MH081521 to S.W.K. T.E.J.B. and S.W.K. are supported by the Wellcome Trust.

AUTHOR CONTRIBUTIONS

S.W.K. designed the experiment, collected and analyzed the data and wrote the manuscript. T.E.J.B. analyzed the data and edited the manuscript. J.D.W. designed the experiment, supervised the project and edited the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at <http://www.nature.com/natureneuroscience/>.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

1. Kennerley, S.W., Walton, M.E., Behrens, T.E., Buckley, M.J. & Rushworth, M.F. Optimal decision making and the anterior cingulate cortex. *Nat. Neurosci.* **9**, 940–947 (2006).
2. Rudebeck, P.H. *et al.* Frontal cortex subregions play distinct roles in choices between actions and stimuli. *J. Neurosci.* **28**, 13775–13785 (2008).
3. Walton, M.E., Behrens, T.E., Buckley, M.J., Rudebeck, P.H. & Rushworth, M.F. Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. *Neuron* **65**, 927–939 (2010).
4. Buckley, M.J. *et al.* Dissociable components of rule-guided behavior depend on distinct medial and prefrontal regions. *Science* **325**, 52–58 (2009).
5. Noonan, M.P. *et al.* Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. *Proc. Natl. Acad. Sci. USA* **107**, 20547–20552 (2010).
6. Fellows, L.K. Deciding how to decide: ventromedial frontal lobe damage affects information acquisition in multi-attribute decision making. *Brain* **129**, 944–952 (2006).
7. Williams, Z.M., Bush, G., Rauch, S.L., Cosgrove, G.R. & Eskandar, E.N. Human anterior cingulate neurons and the integration of monetary reward with motor responses. *Nat. Neurosci.* **7**, 1370–1375 (2004).
8. Bechara, A., Damasio, A.R., Damasio, H. & Anderson, S.W. Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition* **50**, 7–15 (1994).
9. Walton, M.E., Devlin, J.T. & Rushworth, M.F. Interactions between decision making and performance monitoring within prefrontal cortex. *Nat. Neurosci.* **7**, 1259–1265 (2004).
10. Daw, N.D., O'Doherty, J.P., Dayan, P., Seymour, B. & Dolan, R.J. Cortical substrates for exploratory decisions in humans. *Nature* **441**, 876–879 (2006).
11. Prévost, C., Pessiglione, M., Metereau, E., Cléry-Melin, M.L. & Dreher, J.C. Separate valuation subsystems for delay and effort decision costs. *J. Neurosci.* **30**, 14080–14090 (2010).
12. Rushworth, M.F. & Behrens, T.E. Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat. Neurosci.* **11**, 389–397 (2008).
13. Padoa-Schioppa, C. & Assad, J.A. Neurons in the orbitofrontal cortex encode economic value. *Nature* **441**, 223–226 (2006).
14. Tremblay, L. & Schultz, W. Relative reward preference in primate orbitofrontal cortex. *Nature* **398**, 704–708 (1999).
15. 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 (2009).
16. Roesch, M.R. & Olson, C.R. Impact of expected reward on neuronal activity in prefrontal cortex, frontal and supplementary eye fields and premotor cortex. *J. Neurophysiol.* **90**, 1766–1789 (2003).
17. O'Neill, M. & Schultz, W. Coding of reward risk by orbitofrontal neurons is mostly distinct from coding of reward value. *Neuron* **68**, 789–800 (2010).
18. Seo, H. & Lee, D. Behavioral and neural changes after gains and losses of conditioned reinforcers. *J. Neurosci.* **29**, 3627–3641 (2009).
19. Seo, H. & Lee, D. Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *J. Neurosci.* **27**, 8366–8377 (2007).
20. Wallis, J.D. & Miller, E.K. Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *Eur. J. Neurosci.* **18**, 2069–2081 (2003).
21. Kennerley, S.W. & Wallis, J.D. Encoding of reward and space during a working memory task in the orbitofrontal cortex and anterior cingulate sulcus. *J. Neurophysiol.* **102**, 3352–3364 (2009).
22. Kim, S., Hwang, J. & Lee, D. Prefrontal coding of temporally discounted values during intertemporal choice. *Neuron* **59**, 161–172 (2008).
23. Morrison, S.E. & Salzman, C.D. The convergence of information about rewarding and aversive stimuli in single neurons. *J. Neurosci.* **29**, 11471–11483 (2009).
24. Sallet, J. *et al.* Expectations, gains, and losses in the anterior cingulate cortex. *Cogn. Affect. Behav. Neurosci.* **7**, 327–336 (2007).
25. Kennerley, S.W. & Wallis, J.D. Evaluating choices by single neurons in the frontal lobe: outcome value encoded across multiple decision variables. *Eur. J. Neurosci.* **29**, 2061–2073 (2009).
26. Takahashi, Y.K. *et al.* The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. *Neuron* **62**, 269–280 (2009).
27. Wallis, J.D. & Kennerley, S.W. Heterogeneous reward signals in prefrontal cortex. *Curr. Opin. Neurobiol.* **20**, 191–198 (2010).
28. Rangel, A., Camerer, C. & Montague, P.R. A framework for studying the neurobiology of value-based decision making. *Nat. Rev. Neurosci.* **9**, 545–556 (2008).
29. Padoa-Schioppa, C. Range-adapting representation of economic value in the orbitofrontal cortex. *J. Neurosci.* **29**, 14004–14014 (2009).
30. Kobayashi, S., Pinto de Carvalho, O. & Schultz, W. Adaptation of reward sensitivity in orbitofrontal neurons. *J. Neurosci.* **30**, 534–544 (2010).
31. Machens, C.K., Romo, R. & Brody, C.D. Flexible control of mutual inhibition: a neural model of two-interval discrimination. *Science* **307**, 1121–1124 (2005).
32. Fiorillo, C.D., Tobler, P.N. & Schultz, W. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* **299**, 1898–1902 (2003).
33. Matsumoto, M. & Hikosaka, O. Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* **459**, 837–841 (2009).

34. Tobler, P.N., Fiorillo, C.D. & Schultz, W. Adaptive coding of reward value by dopamine neurons. *Science* **307**, 1642–1645 (2005).
35. Cai, X., Kim, S. & Lee, D. Heterogeneous coding of temporally discounted values in the dorsal and ventral striatum during intertemporal choice. *Neuron* **69**, 170–182 (2011).
36. Holroyd, C.B. & Coles, M.G. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol. Rev.* **109**, 679–709 (2002).
37. Debener, S. *et al.* Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *J. Neurosci.* **25**, 11730–11737 (2005).
38. Platt, M.L. & Glimcher, P.W. Neural correlates of decision variables in parietal cortex. *Nature* **400**, 233–238 (1999).
39. Shuler, M.G. & Bear, M.F. Reward timing in the primary visual cortex. *Science* **311**, 1606–1609 (2006).
40. Matsumoto, M., Matsumoto, K., Abe, H. & Tanaka, K. Medial prefrontal cell activity signaling prediction errors of action values. *Nat. Neurosci.* **10**, 647–656 (2007).
41. Amiez, C., Joseph, J.P. & Procyk, E. Anterior cingulate error-related activity is modulated by predicted reward. *Eur. J. Neurosci.* **21**, 3447–3452 (2005).
42. Sul, J.H., Kim, H., Huh, N., Lee, D. & Jung, M.W. Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. *Neuron* **66**, 449–460 (2010).
43. Bayer, H.M. & Glimcher, P.W. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47**, 129–141 (2005).
44. Oliveira, F.T., McDonald, J.J. & Goodman, D. Performance monitoring in the anterior cingulate is not all error related: expectancy deviation and the representation of action-outcome associations. *J. Cogn. Neurosci.* **19**, 1994–2004 (2007).
45. Jessup, R.K., Busemeyer, J.R. & Brown, J.W. Error effects in anterior cingulate cortex reverse when error likelihood is high. *J. Neurosci.* **30**, 3467–3472 (2010).
46. Jocham, G., Neumann, J., Klein, T.A., Danielmeier, C. & Ullsperger, M. Adaptive coding of action values in the human rostral cingulate zone. *J. Neurosci.* **29**, 7489–7496 (2009).
47. Behrens, T.E., Woolrich, M.W., Walton, M.E. & Rushworth, M.F. Learning the value of information in an uncertain world. *Nat. Neurosci.* **10**, 1214–1221 (2007).
48. Montague, P.R. & Berns, G.S. Neural economics and the biological substrates of valuation. *Neuron* **36**, 265–284 (2002).
49. Bernacchia, A., Seo, H., Lee, D. & Wang, X.J. A reservoir of time constants for memory traces in cortical neurons. *Nat. Neurosci.* **14**, 366–372 (2011).
50. Padoa-Schioppa, C. & Assad, J.A. The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nat. Neurosci.* **11**, 95–102 (2008).

ONLINE METHODS

Subjects. Two male rhesus monkeys (*M. mulatta*), ages 5 and 6, were subjects. Our methods for neurophysiological recording are reported in detail elsewhere^{15,20}. We recorded simultaneously from ACC, LPFC and OFC using arrays of 10–24 tungsten microelectrodes (**Supplementary Fig. 1**). We recorded 257 LPFC neurons from areas 9, 46, 45 and 47/12l (subject A, 113; subject B, 144), 140 OFC neurons from areas 11, 13 and 47/12o (A, 58; B, 82), and 213 ACC neurons from area 24c in the dorsal bank of the cingulate sulcus (A, 70; B, 143). We randomly sampled neurons to enable a fair comparison of neuronal properties between different brain regions. There were no significant differences in the proportion of selective neurons between subjects ($P > 0.05$, chi-squared tests), so we pooled neurons across subjects. All procedures complied with guidelines from the US National Institutes of Health and the University of California Berkeley Animal Care and Use Committee.

Task. Task details have been described previously¹⁵. We used NIMH Cortex (<http://www.cortex.salk.edu/>) to control the presentation of the stimuli and the task contingencies. We monitored eye position and pupil dilation using an infra-red system sampling at 125 Hz (ISCAN). Briefly, subjects made choices between two pictures associated with either (i) a specific number of lever presses required to obtain a fixed magnitude of juice reward, with the probability and magnitude of reward held constant (effort trials), (ii) a specific amount of juice with probability and effort held constant (payoff trials) or (iii) a specific probability of obtaining a juice reward with effort and payoff held constant (probability trials, **Fig. 1**). All trials were randomly intermixed. We used five different picture values for each decision variable, and the two presented pictures were always adjacent in value. Thus, there were four different choice values (1 versus 2, 2 versus 3, 3 versus 4, 4 versus 5) per decision variable. This ensured that, aside from pictures associated with the most or least valuable outcome, subjects chose or did not choose the pictures equally often. Moreover, by only presenting adjacently valued pictures, we controlled for the difference in value for each of the choices and therefore the conflict or difficulty in making the choice. We used two sets of pictures for each decision variable to ensure that neurons were not encoding the visual properties of the pictures, but very few neurons encoded the picture set¹⁵, so for all analyses here we collapsed across picture set. We defined correct choices as choosing the outcome associated with the larger reward probability, larger payoff and least effort. We defined chosen value as the value of the chosen stimulus, even when the less valuable stimulus was chosen. Subjects rapidly learned to choose the more valuable outcomes consistently within a behavioral session. Once each subject had learned several picture sets for each decision variable, behavioral training was completed. Two picture sets for each decision variable were chosen for each subject and used during all recording sessions.

We tailored the precise reward amounts during juice delivery for each subject to ensure that they received their daily fluid allotment over the course of the recording session. Consequently, the duration for which the reward pump was active (and hence the magnitude of delivered rewards) differed slightly between the two subjects. The reward durations in subject A were as follows: probability, 1,400 ms; effort, 1,700 ms; payoff, 150, 400, 750, 1,150 or 1,600 ms. The reward durations in subject B were as follows: probability, 2,200 ms; effort, 2,800 ms; payoff, 400, 700, 1,300, 1,900 or 2,500 ms. We also individually adjusted the reward magnitudes so that the five levels of value for each decision variable were approximately subjectively equivalent to both subjects as confirmed by choice preference tests outside the recording experiment. For example, subject A was indifferent in choices between the level 3 probability condition (0.5 probability of 1,400-ms reward, no effort), the level 3 payoff condition (1.0 probability of 750-ms reward, no effort) and the level 3 effort condition (1.0 probability of 1,600-ms reward, effort of five presses) when faced with choice preference tests between these conditions, but he consistently preferred any of the level 3 conditions over lower probability, lower payoff or higher effort conditions. Although the task was not designed to pit the different decision variables against each other to determine subjective equivalence, the subjective indifference to each level of value for each decision variable suggested that each level of value of each decision variable had approximately equal value. We used this measure of value in GLM-3 below to assess the effects of choice value history on current trial neuronal activity.

Data analysis. Behavioral data have been reported elsewhere¹⁵. For all neuronal analyses, firing rate was regressed against the chosen stimulus value, even when the

less valuable stimulus was chosen (<2% of all choices; see **Supplementary Fig. 4**). Analyses using average (instead of chosen) value produced nearly identical results. We constructed spike density histograms by averaging neuronal activity across appropriate conditions using a sliding window of 100 ms. To calculate neuronal selectivity, we fit three sliding GLMs (see below) to each neuron's standardized firing rate observed during a 200-ms time window, starting in the first 200 ms of the fixation epoch and then shifted in 10-ms steps until we had analyzed the entire trial. This enabled us to determine neuronal selectivity in a way that was independent of absolute neuronal firing rates, which is useful when comparing neuronal populations that can differ in the baseline and dynamic range of their firing rates. For each neuron, we also calculated the mean firing rate during the 1,000-ms fixation epoch to serve as a baseline.

We used three different GLMs. In GLM-1, we modeled the relationship between the neuron's standardized firing rate in the choice epoch and the value of the stimulus that would subsequently be chosen as a function of each decision variable (namely, probability, payoff or effort). Three regressors defined the relevant decision variable for the trial (probability, payoff or effort) to account for possible differences in mean firing rates across the different conditions. Three regressors defined the interaction of chosen value and decision variable (that is, value in probability trials, value in payoff trials and value in effort trials).

GLM-2 focused solely on the probability trials and examined how each neuron's standardized firing rate in the outcome epoch was influenced by whether the trial was rewarded and by the probability of getting a reward. The outcome epoch was defined as the 1,000-ms epoch between reward onset (or when reward would have been delivered on unrewarded trials) to 1,000 ms after reward onset. We included four regressors in the design matrix. The first was a constant, to model the mean firing rate across trials. The second differentiated rewarded from unrewarded trials (that is, the encoding of a categorical outcome). The third modeled a linear relationship between firing rate and reward probability on rewarded trials only. The fourth modeled a linear relationship between firing rate and reward probability on unrewarded trials only. The third and fourth regressors were orthogonalized with respect to the first and second to ensure that parameter estimates for the second regressor would reflect differences in mean firing rate between rewarded and unrewarded trials.

In GLM-3 we examined how each neuron's standardized firing rate in the choice epoch of all trials was influenced by either the choice value of the preceding trial or the actual reward history. The design matrix included six regressors. The first was a constant, to model the mean firing rate across trials. The second through fifth encoded the value of trials N (current value) through $N - 3$ (value of the choice made 1, 2 or 3 trials in the past) and the sixth was the actual experienced amount of reward (zero for unrewarded trials) from trial $N - 1$ to ensure that any satiety effect from a reward on the last trial did not influence encoding of value on the current trial.

All regressors in all GLMs were normalized to a scale of 1–5 (except for reward/no reward, which was normalized to a scale of 0–1), so that regression coefficients would be of approximately equal weight across analyses. Although we used standardized firing rates for the GLMs, this becomes an issue when correlating regression coefficients because it tends to under-weight neurons with large dynamic firing rate ranges (that tend to be selective) and over-weight neurons with low dynamic ranges (that tend to be nonselective), potentially leading to type II (false negative) errors. Therefore, for these correlation analyses (**Figs. 7b and 8g,h** and **Supplementary Fig. 3d,e**) we used actual neuronal firing rates.

We defined a selective neuron as any neuron for which a regressor from any of the time bins in the sliding GLMs had a t -statistic ≥ 3.1 for two consecutive time bins. We chose this criterion to produce acceptable type I (false positive) error levels, quantified by examining how many neurons reached the criterion for encoding chosen value during the 1,000-ms fixation epoch (before the choice stimuli appear; thus the number of neurons reaching criterion for this variable should not exceed chance levels). Our type I error for this analysis was acceptable: crossings of our criterion by chance occurred less than 3% of the time for a 1,000-ms epoch. The point of maximum selectivity for each neuron was defined as the time bin in each sliding GLM that had the largest t -statistic for each regressor. The sign of the regression coefficient at this bin was used to define the relationship between firing rate and value and was also used in the scatter plots of **Figures 3b,c** and **8e**. For neurons that encoded all three decision variables at choice, we determined the sign of the regression coefficient at peak selectivity for each decision variable. The majority of neurons (76%) that encoded all three



decision variables at choice did so with the same signed regression coefficient across all variables. Excluding neurons that had opposing signed regression coefficients across choice decision variables did not alter any of the results.

To determine the contribution of each regressor in explaining the variance in a neuron's firing rate, we calculated the CPD and then averaged this value across all neurons in each brain area to obtain an estimate of the amount of variance explained for each regressor at the population level. The CPD for regressor X_i is defined by

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

where $\text{SSE}(X)$ refers to the sum of squared errors in a regression model that includes a set of regressors X , and X_{-i} is a set of all the regressors included in the full model except X_i (ref. 35).

The signs of the regression coefficients become important when examining whether PFC neurons encode prediction errors. The relationship between encoding value and encoding the discrepancy from expected value (that is, prediction error) at the time of outcome are opposite (**Fig. 1c**): the larger the probability at

choice, the smaller the +PE if that trial is rewarded and the smaller (or more negative) the -PE if that trial is not rewarded. In other words, if a neuron has the same relationship between probability and firing rate during both the choice and the outcome, then this neuron is not encoding a prediction error but rather encoding the value (or reward probability) of the trial. In contrast, if a neuron encodes probability at choice and outcome with an opposite relationship, then this would reflect a prediction error signal. Thus, to be defined as a prediction error neuron, the neuron was required to meet three conditions: it had to significantly encode probability at the time of choice (probability regressor from GLM-1, t -statistic ≥ 3.1 for two consecutive time bins), it had to encode the discrepancy between actual and expected reward during the outcome (rewarded and/or non-rewarded outcome regressor(s) from GLM-2, t -statistic ≥ 3.1 for two consecutive time bins) and it had to do so with an opposite relationship (that is, oppositely signed regression coefficients). We further examined whether neurons encoded a salience signal. To be defined as a salience signal, neurons had to significantly encode probability during both rewarded and unrewarded outcomes (rewarded and/or non-rewarded outcome regressor(s) from GLM-2, t -statistic ≥ 3.1 for two consecutive time bins) but with oppositely signed regression coefficients (**Supplementary Fig. 2c**).