

Eur J Neurosci. Author manuscript; available in PMC 2009 November 1.

Published in final edited form as:

Eur J Neurosci. 2009 May; 29(10): 2061–2073. doi:10.1111/j.1460-9568.2009.06743.x.

# Evaluating choices by single neurons in the frontal lobe: outcome value encoded across multiple decision variables

#### Steven W. Kennerley and Jonathan D. Wallis

Helen Wills Neuroscience Institute and Department of Psychology, University of California at Berkeley, 132 Barker Hall MC 3190, Berkeley, CA 94720-3190, USA

#### **Abstract**

Damage to the frontal lobe can cause severe decision-making impairments. A mechanism that may underlie this is that neurons in the frontal cortex encode many variables that contribute to the valuation of a choice, such as its costs, benefits and probability of success. However, optimal decision-making requires that one considers these variables, not only when faced with the choice, but also when evaluating the outcome of the choice, in order to adapt future behaviour appropriately. To examine the role of the frontal cortex in encoding the value of different choice outcomes, we simultaneously recorded the activity of multiple single neurons in the anterior cingulate cortex (ACC), orbitofrontal cortex (OFC) and lateral prefrontal cortex (LPFC) while subjects evaluated the outcome of choices involving manipulations of probability, payoff and cost. Frontal neurons encoded many of the parameters that enabled the calculation of the value of these variables, including the onset and offset of reward and the amount of work performed, and often encoded the value of outcomes across multiple decision variables. In addition, many neurons encoded both the predicted outcome during the choice phase of the task as well as the experienced outcome in the outcome phase of the task. These patterns of selectivity were more prevalent in ACC relative to OFC and LPFC. These results support a role for the frontal cortex, principally ACC, in selecting between choice alternatives and evaluating the outcome of that selection thereby ensuring that choices are optimal and adaptive.

# Keywords

anterior cingulate cortex; decision-making; orbitofrontal cortex; prefrontal cortex; reward

#### Introduction

Selecting the best course of action requires a consideration of multiple decision variables, in particular the magnitude of the action's payoff, its probability of success, and the costs in terms of time and effort (Kahneman & Tversky, 1979; Stephens & Krebs, 1986; Bautista *et al.*, 2001). We have shown that single neurons in the anterior cingulate cortex (ACC), lateral prefrontal cortex (LPFC) and orbitofrontal cortex (OFC) encode the predicted value of choices based on these three decision variables (Kennerley *et al.*, 2009). However, decision-making frameworks highlight that optimal behaviour not only rests upon the valuation of each choice alternative based on these variables (choice evaluation), but also on the evaluation of the experienced outcome in order to determine whether adaptations in future choice behaviour are necessary (Rangel *et al.*, 2008). Outcome-related activity is evident in the frontal cortex; unrewarded or unexpected outcomes modulate ACC, LPFC and OFC (Niki & Watanabe,

<sup>©</sup> The Authors (2009). Journal Compilation © Federation of European Neuroscience Societies and Blackwell Publishing Ltd *Correspondence*: Dr S. W. Kennerley, as above. E-mail: E-mail: skennerley@berkeley.edu.

1979; Hikosaka & Watanabe, 2000; Tremblay & Schultz, 2000; Walton *et al.*, 2004; Sallet *et al.*, 2007), and these areas are modulated by the experience of different magnitudes or types of outcomes (Hikosaka & Watanabe, 2000; Rolls, 2000; O'Doherty *et al.*, 2001; Amiez *et al.*, 2006; Roesch *et al.*, 2006). Moreover, damage to either OFC or ACC impairs the ability to utilize the value of an outcome to guide future decisions (Shima & Tanji, 1998; Kennerley *et al.*, 2006; Murray *et al.*, 2007).

Although ACC, LPFC and OFC all contribute to the decision-making process, little is known about the neural correlates of, and relationship between, choice evaluation and outcome evaluation in decision-making (Rangel et al., 2008). One possibility is that different populations of neurons are recruited for choice and outcome evaluation, as the two processes can involve very different types of information. For example, whereas choice evaluation may depend on predicted values associated with sensory stimuli (e.g. the sight of fruit), outcome evaluation may necessarily be encoded via a different sensory modality (e.g. does the fruit taste sweet?). Moreover, different frontal areas may be specialized for representing different aspects of an outcome. For example, ACC and LPFC have stronger connections with motor areas than OFC and so may be more involved in encoding the physical costs associated with an outcome (Dum & Strick, 1993; Carmichael & Price, 1995; Petrides & Pandya, 1999). In contrast, OFC has stronger connections with gustatory and olfactory areas, and so may be more involved in encoding the reward (Mesulam & Mufson, 1982; Carmichael & Price, 1995). Alternatively, given that some neurons encode choice value as a common or integrated value signal across decision variables (Amiez et al., 2006; Padoa-Schioppa & Assad, 2006; Kennerley et al., 2009), it may be that single neurons encode a common value signal during both choice and outcome evaluation. To address these issues, we trained two rhesus monkeys to make choices that varied along three physically different decision dimensions ('payoff', 'probability' and 'cost') and recorded the activity of single neurons simultaneously from ACC, LPFC and OFC while subjects made their choices.

#### Materials and methods

#### Subjects and neurophysiological procedures

The subjects were two rhesus macaques (Macaca mulatta) who were 5–6 years old and weighed 8–11 kg at the time of recording. We regulated the daily fluid intake of our subjects to maintain motivation on the task. Our methods for neurophysiological recording are reported in detail elsewhere (Wallis & Miller, 2003). Briefly, before surgery the animals were administered ketamine (10 mg/kg i.m.), and anaesthesia was subsequently maintained using isoflurane (2-4%) in balance with oxygen. We implanted both subjects with a titanium head positioner for restraint, and two recording chambers were positioned at an angle to allow access to ACC, LPFC and OFC, the positions of which were determined using a 1.5 T magnetic resonance imaging (MRI) scanner. Postoperative analgesia (buprenorphine; 0.01–0.03 mg / kg i.m.) was provided. We recorded simultaneously from ACC, LPFC and OFC using arrays of 10-24 tungsten microelectrodes (FHC Instruments, Bowdoin, ME, USA). In subject A, we recorded from LPFC and OFC in the left hemisphere, and ACC, OFC and LPFC from the right hemisphere. In subject B, we recorded from ACC and LPFC in the left hemisphere, and OFC and LPFC in the right hemisphere. We determined the approximate distance to lower the electrodes from the MRI images and advanced the electrodes using custom-built, manual microdrives until they were located just above the cell layer. We then slowly lowered the electrodes into the cell layer until we obtained a neuronal waveform. 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 different brain regions. Waveforms were digitized and analysed offline (Plexon Instruments, Dallas, TX, USA). All procedures were in accord with the National

Institute of Health guidelines and the recommendations of the U.C. Berkeley Animal Care and Use Committee.

We reconstructed our recording locations by measuring the position of the recording chambers using stereotactic methods. We plotted the positions onto the MRI sections using commercial graphics software (Adobe Illustrator, San Jose, CA, USA). We confirmed the correspondence between the MRI sections and our recording chambers by mapping the position of sulci and grey and white matter boundaries using neurophysiological recordings. We traced and measured the distance of each recording location along the cortical surface from the lip of the ventral bank of the principal sulcus. The positions of the other sulci, relative to the principal sulcus, were also measured in this way, allowing the construction of the unfolded cortical maps shown in Fig. 7.

#### Behavioural task

We used NIMH Cortex (http://www.cortex.salk.edu) to control the presentation of the stimuli and the task contingencies. We monitored eye position with an infrared system (ISCAN, Burlington, MA, USA). Each trial began with the subject fixating a central square cue 0.3° in width (Fig. 1A). If the subject maintained fixation within 1.8° of the cue for 1000 ms (fixation epoch), two pictures (2.5° in size) appeared at 5.0° to the left and right of fixation. Each picture was associated with either: (i) a specific probability of obtaining a fixed amount of juice (probability trials); (ii) a specific amount of juice (payoff trials); or (iii) a specific number of lever presses required to obtain a fixed amount of juice. After 1500 ms, the fixation cue changed colour, indicating that the subject was free to indicate its choice. Subject A made his choice by moving a lever in the direction of the chosen picture, while subject B made his choice by making a saccade to the chosen picture. On probability and payoff trials we delivered the reward immediately following the subject's choice and the chosen picture remained on the screen throughout reward delivery. On cost trials, the chosen picture remained on the screen and both subjects had to press a lever in order to earn the reward. Each time the subject made a lever press, the chosen picture was briefly extinguished (50 ms) until the required number of lever presses was made and reward was then delivered. A 1000-ms intertrial interval (ITI) separated each trial. We tailored the precise reward amounts and flow rate during juice delivery for each subject to ensure that they received their daily fluid aliquot over the course of the recording session and to ensure that they were sufficiently motivated to perform the task. Consequently, the reward epoch lengths and the magnitude of reward differed slightly in the two subjects. The reward epoch lengths in subject A were: P = 1400 ms; cost = 1700 ms; payoff = 150, 400, 750, 1150 and 1600 ms. The reward epoch lengths in subject B were: probability = 2200 ms; cost = 2800 ms; payoff = 400, 700, 1300, 1900 and 2500 ms.

We used two sets of pictures for each decision variable to ensure that neuronal responses were not driven by the visual properties of the pictures. We used five different picture values for each decision variable and the two presented pictures were always adjacent in value (Fig. 1B). Thus, each picture set involved four distinct choice values, where choice value in this context refers to the value of the pair of stimuli available for choice on each particular trial. This design ensured that, aside from the pictures associated with the most or least valuable outcome, subjects chose or did not choose the pictures equally often. For example, a subject would choose picture C (Fig. 1B) on half of the trials (when it appeared with picture B) and not choose it on the other half of the trials (when it appeared with picture D). Thus, frequency with which a picture was selected could not account for differences in neuronal activity across decision value. Moreover, by only presenting adjacent valued pictures we were able to control for the difference in value for each of the choices and therefore the conflict or difficulty in making the choice. However, we note that with this task design, during the choice phase of the task, we cannot distinguish whether neurons are encoding the more or less valuable stimulus, or an

integrated value of the two stimuli. Equally, in the outcome phase, we assume that neurons are encoding the value of the experienced outcome, but it is possible that neuronal activity could instead reflect the value of the forgone outcome (Montague *et al.*, 2006).

We presented trials at random from a pool of 48 conditions: three decision variables; two picture sets; two responses (left / right); and four decision values. Subjects worked for ~600 trials per day. We defined correct choices as choosing the outcome associated with the largest amount of juice, most probable juice delivery and least amount of cost. We never punished the animal for choosing a less valuable outcome, for example, by using 'time-outs' or correction procedures. Nevertheless, the subjects rapidly learned to choose the more valuable outcomes consistently, typically taking just one or two sessions to learn a set of five picture–value associations during initial behavioural training. Once each subject had learned several picture sets for each decision variable, behavioural training was completed and two picture sets for each decision variable were chosen for each subject. Only these six picture sets for each subject were used during all recording sessions.

#### Statistical methods

We excluded trials in which a break fixation occurred and the repetition of the trial that followed such a break (19% of trials—subject B only) and trials where the subject chose the less valuable outcome (< 2% trials in both subjects). Although there were five possible outcomes for each decision variable, subjects virtually never chose the least valuable outcome. Thus, we excluded the least valuable outcome from our analysis. We constructed spike density histograms by averaging activity across the appropriate conditions using a sliding window of 100 ms.

We determined the time epochs over which to analyse neuronal activity by visually inspecting the spike density histograms in order to determine when neurons encoded parameters relating to the behavioural outcomes. For probability trials, we focused on two epochs: a 'reward' epoch, which characterized the neuronal response to the delivery of the reward; and a 'postreward' epoch, which characterized the neuronal response following the delivery of the reward. On rewarded trials, the reward epoch consisted of a 1000-ms period immediately after the subject made his choice, and the post-reward epoch consisted of a 1000-ms period following the offset of the reward (in fact, the 1000-ms ITI). On unrewarded trials, both epochs consisted of the 1000-ms ITI. For payoff trials, we defined the reward epoch as the time between the onset and offset of reward delivery, and the post-reward epoch as the 1000-ms period immediately after the offset of reward delivery. Because the reward epoch on payoff trials differed in length from trial-to-trial depending upon the amount of juice the subject received, we calculated the neuron's mean firing rate during the reward epoch by dividing the number of spikes during the reward epoch on each trial by the length of the reward epoch on that trial. The encoding of cost was evident in two epochs: while the subject was making the necessary movements of the lever; and during the subsequent delivery of the reward. Thus, we defined the 'movement' epoch as the time from the subject's choice until the delivery of the reward (i.e. once the subject had made all the necessary movements), and the reward epoch as the period from the onset of reward delivery until its offset. The movement epoch differed in length from trial-to-trial depending on how quickly the subject made the necessary lever presses, and so we calculated the neuron's mean firing rate during the movement epoch by dividing the number of spikes during the movement epoch on each trial by the length of the movement epoch on that trial.

To quantify neuronal selectivity we used an analysis of variance (ANOVA) with an alpha level of 0.01. For payoff and cost trials, the factors of the ANOVA were Value (from 1 to 4) and Response (left / right). For probability trials, we included an additional factor of Reward (whether the trial was rewarded or not). For neurons that showed a main effect of Value we also determined whether their firing rates had a positive or negative relationship with Value by fitting a linear

regression to the neuron's firing rate with the predictor variable being trial value (1–4). To asses whether the number of selective neurons in each area exceeded what would be expected by chance, we performed binomial tests (P < 0.05) using the alpha level from the ANOVA as the expected frequency. When more than one area contained selective neurons that exceeded chance levels, we used chi-squared tests (P < 0.05) to determine whether the prevalence of selective neurons differed significantly between areas.

For the probability trials, we determined the time course over which neurons encoded whether or not a reward had occurred by calculating a 'sliding' receiver operating characteristic (ROC). We calculated the ROC for overlapping 200-ms windows, beginning with the first 200 ms of the fixation period and then incrementing the window in 10-ms steps until we had analysed the entire 1000-ms reward epoch and calculated the ROC measure for each time point. Our measure of selectivity was the absolute difference between the ROC measure and the value expected by chance (0.5). We then used this analysis to determine the latency at which neurons encoded the presence or absence of the reward. We defined the latency for a given neuron as the time point at which its selectivity exceeded a criterion value (0.15) for three consecutive time bins. We determined the criterion value by calculating the 99th percentile of our selectivity values during the fixation period.

Unlike probability trials, where the subject only knew the actual outcome after they had made their choice, on payoff and cost trials the pictures indicated the outcome of the choice. Thus, neuronal selectivity encoding payoff or cost information often first appeared during the choice epoch and continued into the outcome epoch. Consequently, it was not possible to determine in a meaningful way the latency at which neurons encoded this information during the outcome epoch.

Finally, we contrasted neuronal selectivity during the outcome epoch with the selectivity that we previously observed during the choice epoch. Our goal was to determine whether neurons that encoded a given decision variable during the choice epoch also encoded that decision variable during the outcome epoch. A previous report describes in detail our methods for characterizing neuronal selectivity during the choice epoch (Kennerley *et al.*, 2009). In summary, for each neuron and each decision variable in turn, we performed a sliding regression analysis. We calculated a linear regression using the neuron's firing rate as the dependent variable and the choice's value as the predictive variable for overlapping 200-ms windows, beginning with the first 200 ms of the fixation period and then incrementing the window in 10-ms steps until we had analysed the entire 1500-ms choice epoch and calculated a linear regression for each time point. For each neuron, we then calculated for which decision variables its selectivity exceeded a criterion value during the choice epoch.

#### Results

We recorded the activity of 610 neurons from the frontal lobe. There were 257 neurons from LPFC defined as areas 9, 46, 45 and 47 / 121 (113 from subject A and 144 from subject B). There were 140 neurons from OFC defined as areas 11, 13 and 47 / 120 (58 from subject A and 82 from subject B). Finally, 213 neurons were in ACC, within the dorsal bank of the cingulate sulcus (70 from subject A and 143 from subject B). We previously reported the activity of these neurons during the choice phase of the task (Kennerley *et al.*, 2009). The current report focuses on characterizing their activity during the outcome phase of the task and determines how this relates to the neuron's activity during the choice phase. The previous report describes the subjects' behavioural performance in detail (Kennerley *et al.*, 2009). In summary, both subjects performed the task at a high level, choosing the more valuable outcome on 98% of the trials. We begin our description of the neuronal activity during the outcome phase by detailing how we characterized encoding for the outcomes associated with each of the three

different decision variables. We then describe how neuronal activity related across decision variables and how it related to activity during the choice phase.

#### Probability trials: encoding the presence or absence of reward

The subjects experienced four different probabilities of receiving reward: specifically, 0.3, 0.5, 0.7 and 0.9 (as they never chose the 0.1 outcome). The first three of these outcomes (0.3, 0.5 and 0.7) are balanced in terms of the probability of receiving or not receiving a reward, so we focused on these conditions for our statistical analysis. The predominant selectivity was encoding of the presence or absence of reward. Figure 2 illustrates examples of single neurons encoding this information. A variety of temporal profiles of selectivity were evident, including neurons that were selective throughout reward delivery (Fig. 2A and B) and neurons that showed responses that were more phasic (Fig. 2C and D). Neurons fired more strongly to either the presence (Fig. 2A and D) or absence of reward (Fig. 2B and C).

We performed a three-way ANOVA with the factors of Value (the probability of receiving reward), Reward (whether the subject received the reward) and Response (whether the behavioural response was to the left or right). During the reward epoch, 244 / 610 or 40% of the neurons encoded whether the subject received the reward, showing a significant main effect of Reward. These neurons were significantly more prevalent in ACC (58%) than either LPFC (32%) or OFC (27%,  $\chi^2 = 42$ ,  $P < 1 \times 10^{-9}$ ). There was a weak bias for neurons to show a higher firing rate to the absence of reward rather than its presence (136 / 244 or 56% of reward-selective neurons, binomial, P < 0.05) and the bias was consistent across all three areas (LPFC 60%, OFC 55%, ACC 53%). A main effect of Response was displayed by 55 / 610 or 9% of neurons, with significantly more Response neurons in LPFC (13%) relative to OFC (6%) and ACC (6%,  $\chi^2 = 8.4, P < 0.05$ ). Approximately 10% (58/610) of neurons showed a significant main effect of Value, but the distribution of these neurons did not differ across areas (LPFC = 8%, OFC = 12%, ACC = 9%,  $\chi^2$  = 1.2, P > 0.1). Less than 1% of all neurons exhibited significant two-way or three-way interactions, which did not exceed the number of neurons expected by chance (binomial, P > 0.1). Thus, encoding of the presence or absence of reward dominated neuronal selectivity.

The pattern of results during the post-reward epoch was very similar. The presence or absence of reward dominated neuronal selectivity: 235 / 610 or 39% of the neurons encoded whether or not the reward had been delivered, showing a significant main effect of Reward. These neurons were significantly more prevalent in ACC (58%) compared with OFC (36%), and significantly more prevalent in OFC relative to LPFC (24%,  $\chi^2 = 58$ ,  $P < 1 \times 10^{-12}$ ). A significant main effect of Value was shown by 29 / 610 or 5% of neurons (ACC = 7%, OFC = 6%, LPFC = 2%), and 34/610 or 6% of neurons showed a significant main effect of Response (ACC = 4%, OFC = 5%, LPFC = 7%). In both cases the number of selective neurons did not differ across areas ( $\chi^2 < 5.0$ , P > 0.05 in both cases). Interactions were encoded by less than 2% of the neurons. In either epoch, 337/610 neurons were reward-selective, with 30% of these neurons selective only in the reward epoch, 28% selective only in the post-reward epoch and 42% selective in both epochs. Of these neurons, a clear majority (87%, binomial,  $P < 1 \times$  $10^{-19}$ ) displayed the same preference in both epochs (i.e. if they showed a higher firing rate to the presence of reward during the reward epoch, they would also show a higher firing rate to the presence of reward during the post-reward epoch). However, a minority of the neurons switched their preference. Figure 2C illustrates such a neuron. This neuron shows a dramatic increase in firing rate on unrewarded trials when the reward does not occur, and a smaller response on rewarded trials immediately after the end of reward delivery.

For each neuron, we calculated the latency at which it first encoded information regarding the presence or absence of reward using a sliding ROC analysis (see Materials and methods, and Fig 3). We then performed a Kruskal–Wallis ANOVA with the latency values as the dependent

variable and the brain area as the independent variable. There was a significant main effect of brain area ( $\chi^2 = 41_{2,349}$ ,  $P < 1 \times 10^{-8}$ ). Post hoc analysis revealed that ACC (median = 230 ms) encoded whether or not reward occurred significantly earlier than either LPFC (median = 430 ms) or OFC (median = 320 ms), while the latencies in LPFC and OFC did not significantly differ from one another.

# Payoff trials: encoding the amount of juice received

Figure 4 illustrates examples of neurons that encode the size of the delivered reward during either the reward epoch (Fig. 4A) or the post-reward epoch (Fig. 4B). The neuron in Fig. 4A modulates its firing rate as a function of reward magnitude immediately after reward onset, which is sustained through the time of reward offset and during the post-reward epoch. Thus, this neuron appears to encode both the expectation and experience of reward magnitude. In contrast, Fig. 4B illustrates a neuron that shows a phasic increase in activity, which appears synched to the time of reward offset. However, this response is not simply related to reward offset as the neuron's response increases as a function of reward magnitude. To quantify the incidence of such neurons we performed a two-way ANOVA with the factors of Value (the amount of reward received) and Response (whether the behavioural response was to the left or right). During the reward epoch, the majority of the selective neurons encoded the amount of reward: 204/610 or 33% of the neurons showed a significant main effect of Value. This encoding was significantly more prevalent in ACC (46%) than in OFC (30%) or LPFC (25%,  $\chi^2 = 21$ , P <0.0001). In contrast, only 10% (58 / 610) of the neurons showed a significant main effect of Response and the distribution of these neurons did not significantly differ between the areas  $(\chi^2 = 1.6, P > 0.1)$ . Only 2% (11 / 610) of all neurons showed a Value × Response interaction, which did not exceed the number of neurons expected by chance in any area (binomial, P >0.1). To determine the nature of the Value encoding, for each selective neuron (as determined from the ANOVA) we fit a slope to the neuron's firing rate across the four different reward amounts. There was an approximately even split between neurons that had a positive relationship between their firing rate and the reward amount (96/204 or 47%), and those that had a negative relationship (108 / 204 or 53%). These values did not significantly differ from chance (binomial, P > 0.1) and the pattern was consistent for all three areas.

We saw a similar pattern of results during the post-reward epoch. A significant main effect of Value was shown by 128 / 610 or 21% of the neurons. The encoding was more prevalent in OFC (27%) and ACC (24%) than LPFC (15%,  $\chi^2 = 10$ , P < 0.05). There was an even split between neurons with a positive relationship between firing rate and reward amount (65 / 128 or 51%), and those with a negative relationship (63 / 122 or 49%). However, within LPFC, there were significantly more neurons that increased their firing rate as reward amount decreased (25 / 38 or 66%; binomial, P < 0.05). In contrast, neurons in OFC (20 / 38, 53%) and ACC (32 / 52, 62%) tended to increase their firing rate as reward amount increased but this did not reach significance (binomial, P > 0.05). Only 40/610 (7%) of the neurons showed a significant main effect of Response and the distribution of these neurons did not significantly differ between areas ( $\chi^2$  < 4.6, P > 0.1). Only 4 / 610 (1%) of the neurons showed a Response × Value interaction, which did not exceed the number of neurons expected by chance in any area (binomial, P > 0.1). In either epoch, 258 / 610 neurons were reward-selective, with 50% of these neurons selective only in the reward epoch, 21% selective only in the post-reward epoch and 29% selective in both epochs. Of these neurons, a clear majority (82%, binomial,  $P < 1 \times 10^{-8}$ ) showed the same relationship between firing rate and reward amount in both epochs.

We also performed a *post hoc* trend analysis on each neuron that exhibited a main effect of Value in the reward and post-reward epochs to examine the degree to which the relationship between firing rate and payoff value was linear. In the reward and post-reward epochs, in 72

and 68% of the respective cases, the variance in the data was fit by a linear function with no residual variance left to explain. The addition of a quadratic function to the linear function significantly reduced the residual variance in 11 and 9% of the remaining cases, respectively. In sum, for the majority of neurons the encoding of payoff value was best described by a linear function.

# Cost trials: encoding the amount of work performed

Encoding of value on cost trials refers to the number of lever presses required to yield a fixed amount of reward, with fewer lever presses equating to more valuable trials. Figure 5A illustrates a neuron that encodes the value of the outcome as soon as the subject makes its choice, showing a more rapid increase in firing rate when the subject is required to make fewer lever presses. Such a neuron may also be described as anticipating the expected onset of reward delivery. Nevertheless, it is this parameter that differentiates low value from high value on cost trials. Figure 5B illustrates a neuron that shows differential firing during the reward delivery once the subject has completed all the movements even though reward magnitude is the same for all cost trials. The neuron shows a higher firing rate on low value trials where the subject has completed many lever presses to earn the reward.

To quantify the incidence of cost selectivity, we performed a two-way ANOVA with factors of Value (number of lever presses) and Response (whether the behavioural response was to the left or right). During the movement epoch, the most common neuronal selectivity was encoding of Value. A significant main effect of Value was shown by 141 / 610 or 23% of the neurons, with no significant interaction. This encoding was significantly more prevalent in ACC (33%) than either OFC (21%) or LPFC (16%,  $\chi^2 = 18$ , P < 0.001). To determine the nature of the Value encoding, for each selective neuron (as determined from the ANOVA) we fit a slope to the neuron's firing rate across the four different cost conditions. There was an approximately even split between neurons that had a positive relationship between firing rate and lever presses (65/ 141 or 46%) as had a negative relationship (76 / 147 or 54%). These values did not significantly differ from chance (binomial, P > 0.05) and the pattern was consistent for all three areas. Other types of neuronal selectivity were also evident. A main effect of Response was shown by 83 / 610 or 14% of the neurons, with no significant interaction. This encoding was significantly more prevalent in LPFC (17%) than OFC (8%,  $\chi^2 = 5.3$ , P < 0.05), while the prevalence in ACC (14%) did not differ from the other two areas. Finally, only 1% of the neurons showed a significant Value × Response interaction, which did not exceed the number of neurons expected by chance in any area (binomial, P > 0.1).

We saw a similar pattern of results during the reward epoch. Encoding of Value was again the most prevalent neuronal selectivity. A significant main effect of Value was shown by 89 / 610 or 15% of the neurons, with no significant interaction. These neurons were significantly more common in ACC (21%) than either OFC (11%) or LPFC (12%,  $\chi^2 = 8.7$ , P < 0.05). There was an even split between those neurons that increased their firing rate as cost increased (51%) and those that increased their firing rate as cost decreased (49%, binomial, P > 0.1). Encoding of Response was significantly more common in LPFC (14%) than OFC (3%,  $\chi^2 = 11$ , P < 0.01), while the prevalence in ACC (11%) did not differ from the other two areas. Less than 1% of the neurons showed a significant Value × Response interaction, which did not exceed the number of neurons expected by chance in any area (binomial, P > 0.1). In either epoch, 198 / 610 neurons were Value selective, with 54% selective only during the movement epoch, 23% selective only during the reward epoch, and 24% selective in both epochs. Of these neurons, 74% showed the same relationship between firing rate and lever presses in both epochs.

We also performed a trend analysis on each neuron that exhibited a main effect of Value in the movement and reward epochs to examine the degree to which the relationship between firing rate and cost value was linear. In the movement and reward epochs, in 70 and 67% of the

respective cases, the variance in the data was fit by a linear function with no residual variance left to explain. The addition of a quadratic function to the linear function significantly reduced the residual variance in 13 and 5% of the remaining cases, respectively. In sum, as was the case for payoff, the activity of the majority of neurons encoding value on cost trials was best described by a linear function. This is consistent with our previous finding that the majority of value selectivity during the choice phase could also be fit by a linear function (Kennerley *et al.*, 2009).

#### **Encoding schemes across epochs**

Having characterized how neurons encode the different outcomes of the choice, we next sought to determine how selectivity in the outcome epoch related to selectivity in the choice epoch. In particular, we wanted to determine whether neurons that encoded particular decision variables during the choice epoch encoded those same decision variables during the outcome epoch. Our methods for classifying neurons during the choice epoch are summarized in the Materials and methods, and described in detail in our previous report (Kennerley et al., 2009). Our analyses of neuronal activity during the outcome, described above, focused on the most prevalent type of encoding in the two different epochs for each decision variable (e.g. reward epoch and the post-reward epoch). Thus, following probability choices, we defined selective neurons as those that encoded whether the choice was rewarded or not, payoff neurons as those encoding the magnitude of the reward on payoff trials and cost neurons as those that encoded the number of lever presses on cost trials. We considered neurons to encode a decision variable during the outcome phase if they showed selectivity in either of the two outcome epochs. The results of this analysis revealed that neurons that encoded the value of the trial during both the choice and outcome epochs of the task were most common in ACC and relatively rare in LPFC, and this was true for all three decision variables (Table 1). We also examined to what extent neurons switched which decision variable they encoded in each epoch. Of the 320 neurons that encoded at least one decision variable in both the choice and outcome epochs, 244 (76%) encoded the same variable in both epochs, while 76 (24%) encoded a different decision variable in each epoch. In summary, frontal neurons, particularly those in ACC and OFC, appear to play an important role in bridging the period between making one's choice and experiencing the results of that choice by encoding both anticipated and actual outcomes.

#### **Encoding schemes across decision variables**

Another important issue is the degree to which encoding of outcome value is specific to a particular decision variable as opposed to encoding a more abstract representation of outcome value across variables. Seven possible encoding schemes are possible with three decision variables: neurons that are selective only to one decision variable; neurons that encode two of the three decision variables (e.g. payoff and cost only); and neurons that encode all three decision variables. We determined whether neurons in the outcome epoch encoded reward presence, payoff or cost information using the same method as the above analysis. Many neurons encoded the outcome of the choice across multiple decision variables. Figure 6A-C shows one such example. The encoding scheme is very different depending on the specific decision variable. For probability trials, the neuron shows a sharp increase in firing rate when no reward occurs. For payoff trials, the neuron shows a gradual increase in firing rate, which is more rapid for smaller amounts of reward. For cost trials, the neuron shows a depression in firing rate that is most pronounced when the fewest lever presses are required, and the neuron does not respond to the actual delivery of the reward, which occurs immediately after the final movement. In other words, for this neuron, the firing rate is consistently highest across decision variables when the outcome is least valuable, irrespective of the physical characteristics of the outcome. Figure 6D and E shows the proportion of neurons that encoded the outcome associated with either a single decision variable, or two or more decision variables. We

observed every possible combination of decision variables encoded. However, neurons that encoded the outcome for two or more decision variables were significantly more prevalent in ACC than LPFC or OFC.

Many neurons showed a linear relationship between their firing rate and the value of the trial during the delivery of the reward on payoff and cost trials. (In contrast, on probability trials neurons tended to encode simply the presence or absence of reward irrespective of the value of the choice.) This raised the possibility that on payoff and cost trials neurons were encoding an integrated cost and payoff signal, or in other words the amount of juice discounted by the amount of work necessary to earn the juice. If this was the case, then the neuron's firing rate should show the same relationship to increasing payoff as to decreasing cost. Of the 51 neurons that were selective during the reward epoch of both payoff and cost trials, 31 / 51 (61%) had the same relationship (i.e. slope) to increasing payoff as to decreasing cost. However, the remaining 20 / 51 (39%) of the neurons had the opposite relationship to increasing payoff as to decreasing cost. The proportion of neurons having the same slope in both epochs did not exceed the number of neurons expected by chance (binomial, P > 0.05). We note that this contrasts with the selectivity during the choice phase of the task, where the clear majority of neurons that were selective on both payoff and cost trials showed the same relationship to increasing payoff as to decreasing cost (99 / 134 or 74%; binomial, P < 0.05). In summary, unlike the choice phase of the task, at the outcome phase our neuronal data did not unequivocally support the notion that neurons were integrating costs and payoffs.

#### **Recording locations**

We plotted the position of outcome-selective neurons on reconstructed cortical maps (Fig. 7). There was no evidence that neurons encoding a particular decision variable were clustered within a brain area.

# **Discussion**

During a period where monkeys evaluated the outcome of their choices, we found many neurons throughout the frontal cortex that encoded various parameters relating to the outcome. On payoff and cost trials, these parameters related directly to the value of the outcome, such as the onset and offset of reward and the number of lever presses required to earn the reward. In contrast, on probability trials, neurons predominantly encoded the presence or absence of reward on a specific trial, but showed little encoding that discriminated the value of the trial, specifically the frequency across trials that a given choice was rewarded. Many frontal neurons encoded a decision variable during both the choice and outcome phases of the task, suggesting that they play an important role in selecting between alternative behaviours and evaluating the outcome associated with that selection. There was no evidence that specific areas of the frontal cortex were specialized for encoding specific types of outcomes, but ACC was significantly more likely, relative to OFC and LPFC, to encode outcomes related to any of the outcome variables manipulated. Moreover, neurons that encoded outcomes for a single decision variable were equally prevalent throughout the frontal cortex, while neurons that encoded two or more decision variables were significantly more common in ACC.

#### Relationship to other outcome-related signals: prediction errors and error monitoring

There are two prominent neurophysiological signals related to the processing of behavioural outcomes: the prediction error signal encoded by dopamine neurons; and error-related activity, which is thought to arise in ACC. The outcome signals that we observed were qualitatively distinct from either of these signals.

Dopamine neurons encode a prediction error, which is the discrepancy between a predicted and actual outcome (Schultz, 1998; Fiorillo et al., 2003; Satoh et al., 2003; Bayer & Glimcher, 2005; Tobler et al., 2005; Roesch et al., 2007). For example, in a task where different stimuli predicted the delivery of juice with different probabilities, dopamine neurons showed a phasic response to the onset of reward delivery whose magnitude linearly increased as the probability of reward decreased (Fiorillo et al., 2003). When the stimulus indicated a low probability of reward, dopamine neurons fired strongly to the occurrence of reward. When the stimulus predicted a high probability of reward, dopamine neurons only fired weakly to the occurrence of reward. This pattern of activity was consistent with a prediction error. In contrast, the activity of frontal neurons in our task showed little influence of the predicted likelihood of reward on neuronal activity. Instead, neuronal selectivity during the outcome of probability trials largely related to encoding whether a reward occurred. This is consistent with recent findings from functional neuroimaging that indicate outcome value is correlated with activity in frontal areas, whereas prediction errors are correlated with activity in the ventral striatum (McClure et al., 2003; O'Doherty et al., 2003; Knutson et al., 2005; Tobler et al., 2006; Hare et al., 2008; Rolls et al., 2008). The dopamine signal is also more restricted with regard to the direction of response. Dopamine neurons consistently increase their firing rate to cues that predict more valuable outcomes or when outcomes are more valuable than expected (Fiorillo et al., 2003; Tobler et al., 2005; Roesch et al., 2007). In contrast, as outcome value increased an approximately equal number of frontal neurons increased their firing rate as decreased their firing rate.

A second prominent outcome-related neurophysiological signal is present when subjects detect an erroneous or unrewarded event. This signal is thought to arise in ACC, as ACC activity is modulated by errors or violations in reward expectancy (Niki & Watanabe, 1979; Falkenstein et al., 1991; Gehring et al., 1993; Ullsperger & von Cramon, 2001; Ito et al., 2003; Debener et al., 2005), although LPFC (Gehring & Knight, 2000; Ullsperger & von Cramon, 2006) and OFC (Turken & Swick, 2008) may contribute to the process. Our results on probability trials support this framework: neurons throughout the frontal cortex responded to the absence of reward but these responses were most prevalent and occurred earliest in ACC. However, our results also suggest that an explanation of outcome activity solely in terms of expectancy violation or error monitoring is overly simplistic (Rushworth et al., 2004; Walton et al., 2004; Rushworth & Behrens, 2008). Many of the neurons that responded to the absence of the reward encoded other aspects of the outcome, such as the magnitude of the reward and the physical effort necessary to earn the reward. This suggests that the role of the frontal cortex in outcome evaluation generalizes across different types of outcomes rather than being specific to monitoring for the presence of reward or errors. Indeed, the decision-making impairments following frontal lobe damage reflect an inability to evaluate a range of decision variables (Bechara et al., 1994; Fellows, 2006; Kennerley et al., 2006; Rudebeck et al., 2006a, b, 2008; Walton et al., 2006; Murray et al., 2007). This suggests that for many frontal neurons, especially in ACC, outcomes are encoded based on their intrinsic value, irrespective of the physical modality in which value is manipulated.

In sum, frontal neurons provide a much richer representation of the outcome than can be accommodated within a prediction error or error monitoring framework. Frontal neurons showed a great deal of diversity in the nature and timing of their responses. For many frontal neurons the outcome activity appeared to be time-locked to the choice and its immediate outcome. However, in other neurons, the selectivity peaked just prior to the offset of the outcome, or during the ITI after the outcome epoch terminated. On payoff trials such encoding could indicate the magnitude of the received reward, while on cost trials it could encode the amount of effort necessary to earn the reward. Thus, frontal neurons contain a particularly rich representation of the behavioural outcome that is not necessarily simply related to the received reward, but also why the reward is valuable. In sum, across the population, especially within

ACC, there appears to be a representation of all of the relevant information necessary to determine the value of the outcome.

### A conceptual framework for valuation

A recent framework for decision-making has proposed four distinct processes to valuation (Rangel *et al.*, 2008). First, the brain needs to represent what alternatives are available and determine the internal (e.g. hunger level, subjective preference) and external states (e.g. terrain, weather) that may inform the valuation of those alternatives. Second, given these states, the brain must then compute the variables that will enable efficient selection between those alternatives, such as the benefits, costs and risks associated with each alternative. These variables, along with the internal and external states, may then get integrated into an overall value signal. The third component in decision-making is selecting the action that will most efficiently realise the most valuable alternative. Finally, once the choice has been made the brain must compute the value of the obtained outcome. Depending on whether the obtained outcome matched the predicted value of that alternative, a prediction error signal can be generated and modify the value of the alternatives thereby ensuring that future choices are adaptive.

The role of the frontal cortex in this process appears most consistent with integrating relevant information in order to calculate values. Our previous results during choice evaluation highlighted a role for ACC and OFC in representing both an integrated value signal as well as the individual components that go into calculating that value signal, such as the size of the payoff or the amount of work that is necessary to earn the payoff (Kennerley et al., 2009). Our current results show that at the outcome phase of the task ACC and OFC neurons are encoding many of the parameters necessary for determining the value of the outcome (Wallis, 2007). In addition, our results highlight aspects of the valuation process that do not appear to involve frontal cortex. We found little evidence that frontal neurons encode reward prediction errors on probability trials. This is perhaps surprising, as we know that such contingencies produce prediction error signals in dopamine neurons (Fiorillo et al., 2003; Tobler et al., 2005; Morris et al., 2006) and frontal cortex is a major recipient of dopaminergic input (Williams & Goldman-Rakic, 1993). In addition, frontal neurons encode a signal similar to a prediction error during the performance of tasks that require subjects to rapidly adapt their behaviour in response to changes in reward contingencies (Amiez et al., 2005; Matsumoto et al., 2007; Seo & Lee, 2007). One possible explanation is that dopaminergic prediction errors encode expectancy violations whenever they occur, whereas frontal areas arbitrate the influence of this signal. In learning the probability trials our subjects have to learn to ignore the absence of reward; if the subject selects the picture that is rewarded on 90% of the trials but does not get a reward then the optimal strategy is to select that picture again when it appears, not switch and choose the other picture. In this context, frontal cortex may dampen the influence of dopamine such that modifications of the original valuations are less likely when dopamine signals a prediction error. Consistent with this idea, neuroimaging studies show that volatility (the probability that reward contingencies will change) modulates ACC activity, suggesting ACC may dictate how much influence individual outcomes should be given in guiding adaptive behaviour (Behrens et al., 2007).

#### **Encoding of the behavioural response**

Although many neurons encoded the value of the different outcomes, few neurons encoded responses and even fewer exhibited response by value interactions. Several studies have emphasized the role of frontal cortex in encoding links between actions and outcomes (Shima & Tanji, 1998; Hadland *et al.*, 2003; Matsumoto *et al.*, 2003; Kennerley *et al.*, 2006; Rudebeck *et al.*, 2008). However, in these studies optimal behaviour was learned by assigning outcome value to different actions. In contrast, our task could be learned via associations between stimuli

and outcome values. Many ACC neurons were, however, modulated by the number of actions the subjects performed in order to acquire the reward (c.f. Procyk *et al.*, 2000; Shidara & Richmond, 2002). These results suggest that while frontal neurons can encode motor parameters during choice evaluation or when action value is being manipulated, unless different actions are consistently associated with different values, outcome evaluation may occur largely independently of the previous motor response.

#### Conclusion

In conclusion, neurons throughout frontal cortex encode a rich, diverse representation of the events that follow a behavioural choice, particularly in comparison to previously observed outcome-related neurophysiological signals. Encoding of the outcome was particularly evident in ACC relative to OFC and LPFC. Neurons encoded the various parameters that would enable an animal to determine the magnitude of the received payoff, the amount of work that was necessary to earn payoff and the likelihood of a successful outcome. Many neurons encoded this information in a multiplexed fashion and encoded the information across both the choice and outcome phases of the task. Taken together these results support a role for ACC in evaluating alternative choices and evaluating the outcome of the choice, thereby enabling future behaviour to select between alternatives in the most advantageous manner.

# Acknowledgements

The project was funded by NIDA grant R01DA19028 and NINDS grant P01NS040813 to J.D.W., and NIMH training grant F32MH081521 to S.W.K. Both authors contributed equally to all aspects of the project. The authors have no competing interests.

#### **Abbreviations**

ACC, anterior cingulate cortex; ITI, intertrial interval; LPFC, lateral prefrontal cortex; MRI, magnetic resonance imaging; OFC, orbitofrontal cortex; ROC, receiver operating characteristic.

#### References

- Amiez C, Joseph JP, Procyk E. Anterior cingulate error-related activity is modulated by predicted reward. Eur. J. Neurosci 2005;21:3447–3452. [PubMed: 16026482]
- Amiez C, Joseph JP, Procyk E. Reward encoding in the monkey anterior cingulate cortex. Cereb. Cortex 2006;16:1040–1055. [PubMed: 16207931]
- Bautista LM, Tinbergen J, Kacelnik A. To walk or to fly? How birds choose among foraging modes. Proc. Natl Acad. Sci. USA 2001;98:1089–1094. [PubMed: 11158599]
- Bayer HM, Glimcher PW. Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron 2005;47:129–141. [PubMed: 15996553]
- Bechara A, Damasio AR, Damasio H, Anderson SW. Insensitivity to future consequences following damage to human prefrontal cortex. Cognition 1994;50:7–15. [PubMed: 8039375]
- Behrens TE, Woolrich MW, Walton ME, Rushworth MF. Learning the value of information in an uncertain world. Nat. Neurosci 2007;10:1214–1221. [PubMed: 17676057]
- Carmichael ST, Price JL. Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. J. Comp. Neurol 1995;363:642–664. [PubMed: 8847422]
- Debener S, Ullsperger M, Siegel M, Fiehler K, von Cramon DY, Engel AK. Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. J. Neurosci 2005;25:11730–11737. [PubMed: 16354931]
- Dum, RP.; Strick, PL. Cingulate Motor Areas. In: Vogt, BA.; Gabriel, M., editors. Neurobiology of Cingulate Cortex and Limbic Thalamus: A Comprehensive Handbook. Cambridge, MA: Birkhaeuser; 1993. p. 415-441.

Falkenstein M, Hohnsbein J, Hoormann J, Blanke L. Effects of crossmodal divided attention on late ERP components II. Error processing in choice reaction tasks. Electroencephalogr. Clin. Neurophysiol 1991;78:447–455. [PubMed: 1712280]

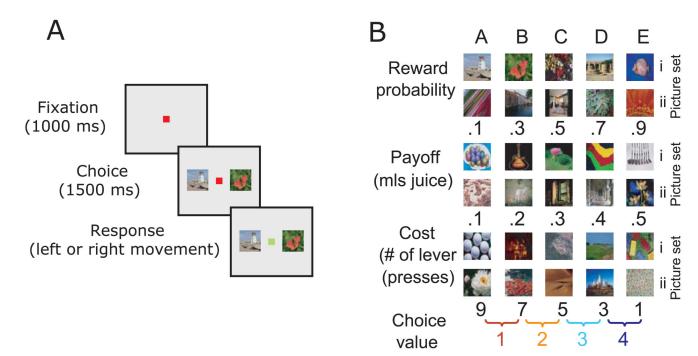
- Fellows LK. Deciding how to decide: ventromedial frontal lobe damage affects information acquisition in multi-attribute decision making. Brain 2006;129:944–952. [PubMed: 16455794]
- Fiorillo CD, Tobler PN, Schultz W. Discrete coding of reward probability and uncertainty by dopamine neurons. Science 2003;299:1898–1902. [PubMed: 12649484]
- Gehring WJ, Knight RT. Prefrontal-cingulate interactions in action monitoring. Nat. Neurosci 2000;3:516–520. [PubMed: 10769394]
- Gehring WJ, Goss B, Coles MG, Meyer DE. A neural system for error detection and compensation. Psychol. Sci 1993;4:385–390.
- Hadland KA, Rushworth MF, Gaffan D, Passingham RE. The anterior cingulate and reward-guided selection of actions. J. Neurophysiol 2003;89:1161–1164. [PubMed: 12574489]
- Hare TA, O'Doherty J, Camerer CF, Schultz W, Rangel A. Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. J. Neurosci 2008;28:5623–5630. [PubMed: 18509023]
- Hikosaka K, Watanabe M. Delay activity of orbital and lateral prefrontal neurons of the monkey varying with different rewards. Cereb. Cortex 2000;10:263–271. [PubMed: 10731221]
- Ito S, Stuphorn V, Brown JW, Schall JD. Performance monitoring by the anterior cingulate cortex during saccade countermanding. Science 2003;302:120–122. [PubMed: 14526085]
- Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. Econometrica 1979;47:263–291
- Kennerley SW, Walton ME, Behrens TE, Buckley MJ, Rushworth MF. Optimal decision making and the anterior cingulate cortex. Nat. Neurosci 2006;9:940–947. [PubMed: 16783368]
- Kennerley SW, Dahmubed AF, Lara AH, Wallis JD. Neurons in the frontal lobe encode the value of multiple decision vairables. J. Cogn. Neurosci 2009;21:1162–1178. [PubMed: 18752411]
- Knutson B, Taylor J, Kaufman M, Peterson R, Glover G. Distributed neural representation of expected value. J. Neurosci 2005;25:4806–4812. [PubMed: 15888656]
- Matsumoto K, Suzuki W, Tanaka K. Neuronal correlates of goal-based motor selection in the prefrontal cortex. Science 2003;301:229–232. [PubMed: 12855813]
- Matsumoto M, Matsumoto K, Abe H, Tanaka K. Medial prefrontal cell activity signaling prediction errors of action values. Nat. Neurosci 2007;10:647–656. [PubMed: 17450137]
- McClure SM, Berns GS, Montague PR. Temporal prediction errors in a passive learning task activate human striatum. Neuron 2003;38:339–346. [PubMed: 12718866]
- Mesulam MM, Mufson EJ. Insula of the old world monkey III: efferent cortical output and comments on function. J. Comp. Neurol 1982;212:38–52. [PubMed: 7174907]
- Montague PR, King-Casas B, Cohen JD. Imaging valuation models in human choice. Annu. Rev. Neurosci 2006;29:417–448. [PubMed: 16776592]
- Morris G, Nevet A, Arkadir D, Vaadia E, Bergman H. Midbrain dopamine neurons encode decisions for future action. Nat. Neurosci 2006;9:1057–1063. [PubMed: 16862149]
- Murray EA, O'Doherty JP, Schoenbaum G. What we know and do not know about the functions of the orbitofrontal cortex after 20 years of cross-species studies. J. Neurosci 2007;27:8166–8169. [PubMed: 17670960]
- Niki H, Watanabe M. Prefrontal and cingulate unit activity during timing behavior in the monkey. Brain Res 1979;171:213–224. [PubMed: 111772]
- O'Doherty J, Kringelbach ML, Rolls ET, Hornak J, Andrews C. Abstract reward and punishment representations in the human orbitofrontal cortex. Nat. Neurosci 2001;4:95–102. [PubMed: 11135651]
- O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ. Temporal difference models and reward-related learning in the human brain. Neuron 2003;38:329–337. [PubMed: 12718865]
- Padoa-Schioppa C, Assad JA. Neurons in the orbitofrontal cortex encode economic value. Nature 2006;441:223–226. [PubMed: 16633341]

Petrides M, Pandya DN. Dorsolateral prefrontal cortex: comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. Eur. J. Neurosci 1999;11:1011–1036. [PubMed: 10103094]

- Procyk E, Tanaka YL, Joseph JP. Anterior cingulate activity during routine and non-routine sequential behaviors in macaques. Nat. Neurosci 2000;3:502–508. [PubMed: 10769392]
- Rangel A, Camerer C, Montague PR. A framework for studying the neurobiology of value-based decision making. Nat. Rev. Neurosci 2008;9:545–556. [PubMed: 18545266]
- Roesch MR, Taylor AR, Schoenbaum G. Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. Neuron 2006;51:509–520. [PubMed: 16908415]
- Roesch MR, Calu DJ, Schoenbaum G. Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. Nat. Neurosci 2007;10:1615–1624. [PubMed: 18026098]
- Rolls ET. The orbitofrontal cortex and reward. Cereb. Cortex 2000;10:284-294. [PubMed: 10731223]
- Rolls ET, McCabe C, Redoute J. Expected value, reward outcome, and temporal difference error representations in a probabilistic decision task. Cereb. Cortex 2008;18:652–663. [PubMed: 17586603]
- Rudebeck PH, Buckley MJ, Walton ME, Rushworth MF. A role for the macaque anterior cingulate gyrus in social valuation. Science 2006a;313:1310–1312. [PubMed: 16946075]
- Rudebeck PH, Walton ME, Smyth AN, Bannerman DM, Rushworth MF. Separate neural pathways process different decision costs. Nat. Neurosci 2006b;9:1161–1168. [PubMed: 16921368]
- Rudebeck PH, Behrens TE, Kennerley SW, Baxter MG, Buckley MJ, Walton ME, Rushworth MF. Frontal cortex subregions play distinct roles in choices between actions and stimuli. J. Neurosci 2008;28:13775–13785. [PubMed: 19091968]
- Rushworth MF, Behrens TE. Choice, uncertainty and value in prefrontal and cingulate cortex. Nat. Neurosci 2008;11:389–397. [PubMed: 18368045]
- Rushworth MF, Walton ME, Kennerley SW, Bannerman DM. Action sets and decisions in the medial frontal cortex. Trends Cogn. Sci 2004;8:410–417. [PubMed: 15350242]
- Sallet J, Quilodran R, Rothe M, Vezoli J, Joseph JP, Procyk E. Expectations, gains, and losses in the anterior cingulate cortex. Cogn. Affect. Behav. Neurosci 2007;7:327–336. [PubMed: 18189006]
- Satoh T, Nakai S, Sato T, Kimura M. Correlated coding of motivation and outcome of decision by dopamine neurons. J. Neurosci 2003;23:9913–9923. [PubMed: 14586021]
- Schultz W. Predictive reward signal of dopamine neurons. J. Neurophysiol 1998;80:1–27. [PubMed: 9658025]
- Seo H, Lee D. Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. J. Neurosci 2007;27:8366–8377. [PubMed: 17670983]
- Shidara M, Richmond BJ. Anterior cingulate: single neuronal signals related to degree of reward expectancy. Science 2002;296:1709–1711. [PubMed: 12040201]
- Shima K, Tanji J. Role for cingulate motor area cells in voluntary movement selection based on reward. Science 1998;282:1335–1338. [PubMed: 9812901]
- Stephens, DW.; Krebs, JR. Foraging Theory. Princeton: Princeton University Press; 1986.
- Tobler PN, Fiorillo CD, Schultz W. Adaptive coding of reward value by dopamine neurons. Science 2005;307:1642–1645. [PubMed: 15761155]
- Tobler PN, O'Doherty JP, Dolan RJ, Schultz W. Human neural learning depends on reward prediction errors in the blocking paradigm. J. Neurophysiol 2006;95:301–310. [PubMed: 16192329]
- Tremblay L, Schultz W. Reward-related neuronal activity during go-nogo task performance in primate orbitofrontal cortex. J. Neurophysiol 2000;83:1864–1876. [PubMed: 10758098]
- Turken AU, Swick D. The effect of orbitofrontal lesions on the error-related negativity. Neurosci. Lett 2008;441:7–10. [PubMed: 18597940]
- Ullsperger M, von Cramon DY. Subprocesses of performance monitoring: a dissociation of error processing and response competition revealed by event-related fMRI and ERPs. Neuroimage 2001;14:1387–1401. [PubMed: 11707094]
- Ullsperger M, von Cramon DY. The role of intact frontostriatal circuits in error processing. J. Cogn. Neurosci 2006;18:651–664. [PubMed: 16768367]

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

- Wallis JD, Miller EK. From rule to response: neuronal processes in the premotor and prefrontal cortex. J. Neurophysiol 2003;90:1790–1806. [PubMed: 12736235]
- Walton ME, Devlin JT, Rushworth MF. Interactions between decision making and performance monitoring within prefrontal cortex. Nat. Neurosci 2004;7:1259–1265. [PubMed: 15494729]
- Walton ME, Kennerley SW, Bannerman DM, Phillips PE, Rushworth MF. Weighing up the benefits of work: behavioral and neural analyses of effort-related decision making. Neural Netw 2006;19:1302–1314. [PubMed: 16949252]
- Williams SM, Goldman-Rakic PS. Characterization of the dopaminergic innervation of the primate frontal cortex using a dopamine-specific antibody. Cereb. Cortex 1993;3:199–222. [PubMed: 8100725]



**Fig. 1.**(A) Sequence of events in the behavioural task. A trial began with the subject fixating a central cue for 1000 ms, after which two pictures appear to either side at fixation. After 1500 ms, the fixation cue changed colour, indicating that the subject should indicate his choice. (B) The pictures and their associated outcomes.

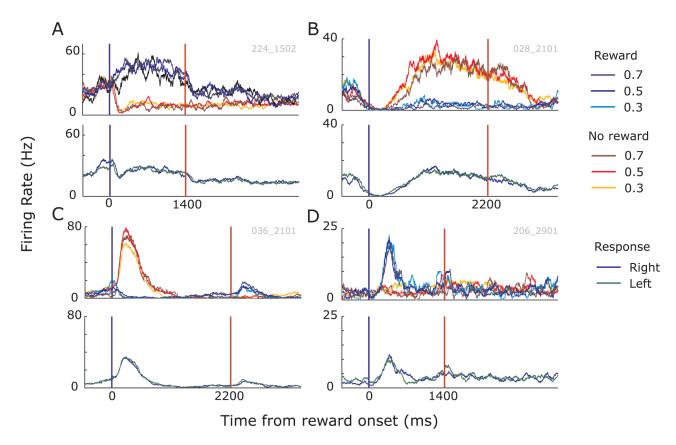


Fig. 2. Spike density histograms illustrating the activity of neurons during the outcome phase of probability trials. In each case, the top panel illustrates activity sorted according to the expected probability of reward delivery (signalled by the pictures during the choice epoch) and whether reward did or did not occur, while the bottom panel illustrates activity sorted by the subject's behavioural response. The blue vertical line indicates the onset of the reward and the red vertical line indicates its offset. (The two subjects received different amounts of reward to ensure that they received their daily aliquot of fluid during a single recording session.) (A) An ACC neuron that shows a sustained elevation of firing during the delivery of reward. (B) An ACC neuron that shows a phasic response when reward does not occur. (C) An ACC neuron that shows a phasic response when reward does not occur. This neuron also shows a phasic response to the offset of the reward. (D) An LPFC neuron that shows a phasic response to the delivery of reward. None of these neurons discriminated between the different behavioural responses.

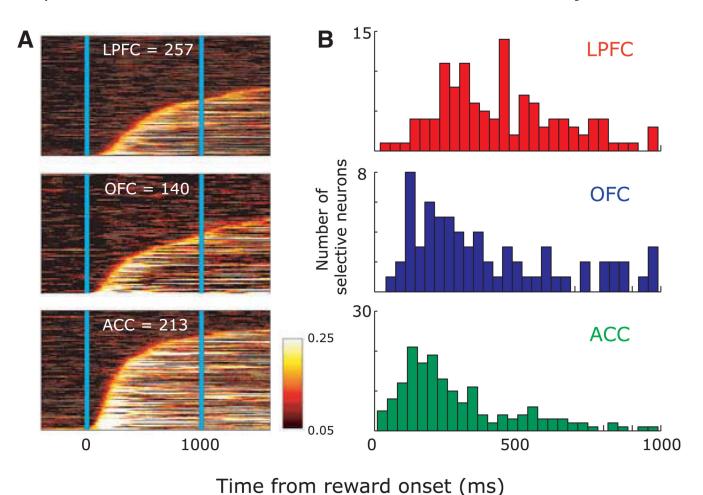


Fig. 3. (A) The time-course of selectivity for encoding whether a reward occurred on probability trials. For each brain area, the entire neuronal population from which we recorded is shown. On each plot, each horizontal line indicates the data from a single neuron, and the colour code illustrates how the ROC value changes across the course of the trial. The ROC value can be thought of as the probability that you could correctly identify whether or not a reward had occurred solely by looking at the neuron's firing rate. It thus ranges from 0.5 (the neuron provides no information as to reward occurrence) to 1.0 (the neuron perfectly indicates reward occurrence). The colour scale for ROC values refers to the deviation from an ROC value of 0.5, thus neurons shaded in white (0.25 on the colour scale) have ROC values of 0.75 or greater. We have sorted the neurons on the y-axis according to the latency at which their selectivity exceeds criterion during the 1000-ms reward epoch. The vertical blue lines indicate the extent of this epoch. The dark area at the top of each plot consists of neurons that did not reach the criterion during the outcome epoch. (B) The histograms illustrate the distribution of neuronal latencies at which selectivity appeared in each brain area. Selectivity appeared significantly earlier in anterior cingulate cortex (ACC) relative to either lateral prefrontal cortex (LPFC) or orbitofrontal cortex (OFC).

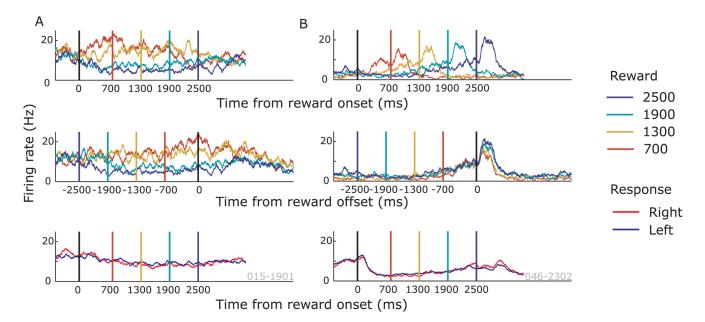


Fig. 4.

Spike density histograms illustrating the activity of neurons during the outcome phase of payoff trials. The top panel indicates neuronal selectivity synchronized to the onset of reward delivery. Each coloured vertical line illustrates the offset of reward delivery with the different colours indicating the different reward magnitudes. The middle panel illustrates neuronal selectivity synchronized to the offset of reward delivery, with each coloured vertical line showing the onset of reward delivery. The bottom panel illustrates neuronal activity synchronized to reward delivery and sorted according to the behavioural response. (A) An ACC neuron that differentiates between the different reward amounts from the beginning of reward delivery. It shows its highest firing rate to the smallest reward amount, and the lowest firing rate to the largest reward amount. (B) An ACC neuron that differentiates between the different reward amounts at the offset of the reward. The neuron's firing rate increases as the size of the reward that the subject receives increases. For both neurons, there is no difference in firing rate between the left and right behavioural responses.

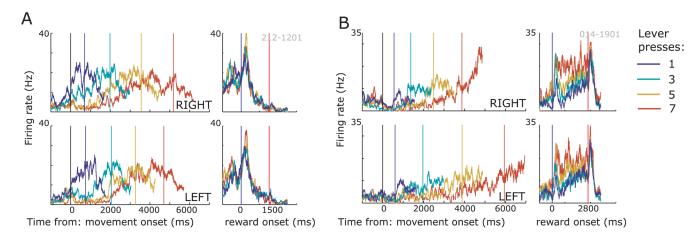


Fig. 5.

Spike density histograms illustrating the activity of neurons during the outcome phase of cost trials. Neuronal activity is synchronized to the subject's choice, at which point the subject is required to begin making the movements that are necessary to earn the reward (left panel), or to the onset of reward delivery (right panel). Neuronal activity is shown separately for rightward and leftward behavioural responses. For the left panels the vertical lines illustrate the median time at which the subject completed the necessary number of lever presses to earn the reward for each of the conditions. For the right panels the vertical lines illustrate the onset and offset of reward delivery. (A) An OFC neuron that shows differential activity during the movement epoch, showing higher levels of firing when fewer lever presses are necessary. (B) An ACC neuron that shows differential activity during the reward epoch, showing a higher firing rate if the subject had made more lever presses to earn the reward. For both neurons, there is no difference in firing rate between the left and right behavioural responses.

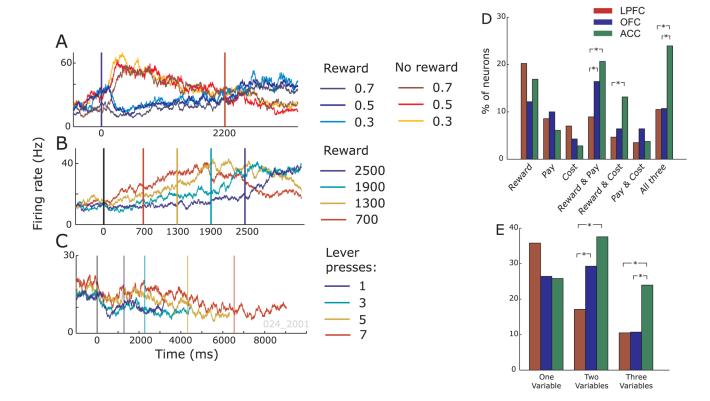


Fig. 6. Spike density histograms illustrating the activity of an anterior cingulate cortex (ACC) neuron on (A) probability, (B) payoff and (C) cost trials. The figures are illustrated according to the conventions of Fig 2, Fig 4 and Fig 5. On probability trials the neuron shows a sustained elevation of firing when reward does not occur, on payoff trials it shows a gradual increase in firing rate that is steeper for smaller reward amounts, and on cost trials the neuron shows a gradual depression in firing rate as the subject completes the movements that is steepest for the conditions requiring the fewest movements. In other words, in all cases the neuron shows its highest firing rate to the least valuable outcome. (D) Percentage of neurons that encode the value of the outcome depending on which decision variable was being manipulated. Reward refers to neurons encoding the presence on absence of reward on probability trials. (E) Percentage of neurons that encode outcome value across one, two or three decision variables. Asterisks in (D) and (E) indicate the proportions that are significantly different from one another (chi-squared test, P < 0.05).

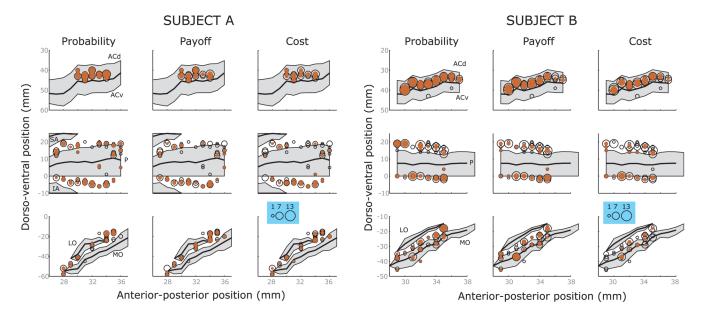


Fig. 7.
Locations of all recorded neurons (open circles) and neurons selective for the different decision variables (filled circles) in subject A and subject B. The size of the circles indicates the number of neurons at that location. We measured the anterior—posterior position from the interaural line (*x*-axis), and the dorso-ventral position relative to the lip of the ventral bank of the principal sulcus (0 point on *y*-axis). Grey shading indicates unfolded sulci. See Materials and methods for details regarding the reconstruction of the recording locations. ACd, dorsal bank of the anterior cingulate sulcus; ACv, ventral bank of the anterior cingulate sulcus; IA, inferior arcuate sulcus; LO, lateral orbital sulcus; MO, medial orbital sulcus; P, principal sulcus; SA, superior arcuate sulcus.

# NIH-PA Author Manuscript

**NIH-PA Author Manuscript** 

NIH-PA Author Manuscript

Percentages of neurons in LPFC, OFC and ACC that encode the value of different decision variables during either the choice epoch, the outcome epoch, or both epochs

	Probability /	Probability / Reward presence		Payoff			Cost		
	LPFC	OFC	ACC	LPFC	OFC	ACC	LPFC	OFC	ACC
Choice only	10	13	14	17	14	21	16	18	23
Outcome only	30	26	27	21	16	20	19	16	16
Both epochs	14	19	*84	*11	28	35	9	111	<sub>28</sub> *

Note that selectivity on the probability trials during the outcome phase refers to neurons selective for the presence or absence of reward rather than probability per se. ACC, anterior cingulate cortex; LPFC, lateral prefrontal cortex; OFC, orbitofrontal cortex.

\* Proportion of selective neurons that was significantly different from the other two areas (chi-square, P < 0.05).