Challenges of interpreting frontal neurons during value-based decision-making

Jonathan D. Wallis[1,2] and Erin L. Rich[1]*

[1] Helen Wills Neuroscience Institute, University of California Berkeley, Berkeley, CA, USA
[2] Department of Psychology, University of California Berkeley, Berkeley, CA, USA

The frontal cortex is crucial to sound decision-making, and the activity of frontal neurons correlates with many aspects of a choice, including the reward value of options and outcomes. However, rewards are of high motivational significance and have widespread effects on neural activity. As such, many neural signals not directly involved in the decision process can correlate with reward value. With correlative techniques such as electrophysiological recording or functional neuroimaging, it can be challenging to distinguish neural signals underlying value-based decision-making from other perceptual, cognitive, and motor processes. In the first part of the paper, we examine how different value-related computations can potentially be confused. In particular, error-related signals in the anterior cingulate cortex, generated when one discovers the consequences of an action, might actually represent violations of outcome expectation, rather than errors *per se*. Also, signals generated at the time of choice are typically interpreted as reflecting predictions regarding the outcomes associated with the different choice alternatives. However, these signals could instead reflect comparisons between the presented choice options and previously presented choice alternatives. In the second part of the paper, we examine how value signals have been successfully dissociated from saliency-related signals, such as attention, arousal, and motor preparation in studies employing outcomes with both positive and negative valence. We hope that highlighting these issues will prove useful for future studies aimed at disambiguating the contribution of different neuronal populations to choice behavior.

Keywords: value, reward, choice, decision-making, prediction error, valence, orbitofrontal, anterior cingulate

INTRODUCTION

Some of the first recordings of single neuron activity in frontal cortex noted the presence of neurons with various reward-related responses. Recordings in orbitofrontal cortex (OFC) and anterior cingulate cortex (ACC) found neurons that were active to cues that predicted reward, neurons that fired immediately preceding and during an expected reward and neurons that responded to the omission of an expected reward (Niki and Watanabe, 1976; Rosenkilde et al., 1981; Thorpe et al., 1983). Similar neurons were subsequently found throughout the frontal cortex (Ono et al., 1984; Watanabe, 1996; Leon and Shadlen, 1999; Amador et al., 2000; Roesch and Olson, 2003) and indeed in parietal (Platt and Glimcher, 1999), temporal (Liu and Richmond, 2000), and occipital cortex (Shuler and Bear, 2006). Due to the central role that reward plays in behavioral control, many cognitive processes can correlate with reward, so it is critical to define precisely the aspect of reward processing in which specific neuronal populations are engaged. Our review will focus on OFC and ACC. Even though reward-related responses are found throughout the frontal lobe, it is only damage to these two areas that produces a specific deficit in value-based decision-making (Bechara et al., 1998; Kennerley et al., 2006; Fellows and Farah, 2007).

Current theoretical models of value-based decision-making posit a series of distinct stages (Padoa-Schioppa, 2007; Rangel and Hare, 2010). First, the subject calculates the value of possible behavioral outcomes to derive a "goods space." This involves integrating the multiple parameters that go into making one outcome more preferable than another such as subjective preferences, magnitude of reward, or delay until reward delivery. Second, the subject calculates action values by subtracting the action costs involved in acquiring the goods from the value of the goods themselves. The separation of goods and actions makes sense from a computational perspective. The parameter space of potential goods and the space of possible actions are both vast and the complexity can be reduced by calculating the value of goods and actions independently. In addition to an argument from parsimony, neuroimaging findings support the notion that subjects can make choices in a goods space that is independent of action (Wunderlich et al., 2010), and these goods-based calculations appear to occur in OFC.

Within this framework it is evident that spurious correlations with goods values or action values might occur either upstream or downstream of the decision-making process. Calculating the value of a good requires the integration of its costs and benefits. For example, humans often calculate a good's value by integrating its desirability and price. The Porsche looks great; the price tag not so much. Similarly, animals often have to weigh the desirability of a good with relative availability in the environment (Stephens and Krebs, 1986). Performing these calculations

requires the integration of multiple sensory parameters that must be represented upstream of the calculation. Problematically for the interpretation of value signals, these sensory representations can easily correlate with the good's value. For example, a large piece of fruit is more rewarding to a hungry animal than a small piece of fruit. Thus, the firing rate of visual sensory neurons would correlate with the fruit's value, even though they are simply responding to the visual representation of the fruit's size rather than its value *per se*. Although our focus is on the frontal cortex, there is a good deal of sensory information encoded in this part of the brain that is relevant to the representation of rewards. For example, the ventral surface of the frontal lobe includes primary gustatory cortex and primary olfactory cortex (Cavada et al., 2000).

In order to dissociate value responses from the sensory responses that go into the value computation, it is important to show that the neuron responds to multiple dimensions of the value space. For example, if the same neuron increases its firing rate as the desirability of a good increases and decreases its firing rate as the price of that good increases, then we can reasonably conclude that the neuron encodes net value as a combination of costs and benefits. Neurons encoding multidimensional aspects of value have been identified throughout frontal cortex. For instance, we trained animals to perform a multidimensional choice task (**Figures 1A–C**) in which they had to make choices based on the magnitude of a juice, its probability of delivery and amount of work necessary to earn the juice (Kennerley et al., 2009). We found neurons encoding every decision variable and every combination of decision variables in ACC, OFC, and lateral prefrontal cortex (**Figure 1D**). Other groups have found prefrontal neurons that integrate the magnitude of a juice with its probability of delivery (Amiez et al., 2006), subjective preference for the juice (Padoa-Schioppa and Assad, 2006) and delay until its delivery (Hwang et al., 2009). Thus, there is ample evidence that frontal neurons do not solely respond to sensory dimensions of a good, but instead integrate multiple attributes that, in sum, determine the value of the good.

Downstream of the decision-making process, things are more complicated. Value signals can serve many different functions, including the reinforcement of behavior, the evaluation of alternative courses of action, and the prioritization of limited capacity behavioral and cognitive resources (Wallis and Kennerley, 2010). This means that neurons encoding other processes such as arousal or attention could correlate with expected value even though they are not encoding the value *per se* (Maunsell, 2004; Luk and Wallis, 2009). In the rest of this paper, we examine several places where these processes confuse the interpretation of value signals, and discuss attempts to disentangle them from the decision-making process.

PREDICTIONS, ERRORS, AND PREDICTION ERRORS

Value signals continue to be important even once a decision has been made and an action completed. Notably, the outcome of one's choices can be used to guide future decisions, thereby ensuring adaptive and efficient behavior. If the outcome of a choice was more valuable than expected then you should be more inclined to choose in a similar manner in future. In contrast, if the outcome was less valuable than expected, you should be less inclined to make

that choice again. The difference between the value of the expected outcome and the actual outcome is termed the prediction error, and was famously identified to be encoded by dopamine neurons in the ventral midbrain (Schultz et al., 1997). In this section, we examine to what extent value coding in frontal cortex relates to value predictions and prediction errors.

NEUROPHYSIOLOGICAL PROPERTIES OF ACC

Early single-unit recordings in ACC observed strong firing when a monkey made an error (Niki and Watanabe, 1979). Human neurophysiology studies later reported a negative potential over ACC when subjects made errors (Gehring et al., 1990; Falkenstein et al., 1991; Ito et al., 2003), which became known as the error-related negativity (ERN). Error signals were also observed using fMRI (Carter et al., 1998; Ullsperger and Von Cramon, 2003; Holroyd et al., 2004), and theories emerged suggesting that ACC was important for processing negative events, costs, or errors (Aston-Jones and Cohen, 2005) or monitoring for conflicts between competing types of information (Botvinick et al., 2004; Ridderinkhof et al., 2004). However, the picture emerging from single-unit studies soon became more complex.

In a task requiring a monkey to learn action–outcome associations using secondary reinforcers, ACC neurons were just as likely to respond to positive feedback as negative feedback and, furthermore, the response to positive feedback was strongest early in learning (Matsumoto et al., 2007). Thus, the neurons' response was strongest when the feedback was least expected, and weakest when it was fully predicted, exactly what one would expect from a prediction-error signal. ACC neuronal activity has also been recorded during the performance of a competitive game (Seo and Lee, 2007). On each trial, monkeys and a computer opponent chose one of two identical targets. Reward was delivered if both subject and computer chose the same target. Choice strategies were exploited by the computer opponent, so optimal behavior entailed choosing randomly. However, the monkeys did not behave completely randomly and their behavior indicated that they were estimating the long-term value of the response options. Many ACC neurons responded during the feedback period in a way that reflected both the value of the feedback (whether or not the animal received a reward) and the animal's estimation of the choice's long-term value, consistent with a reward prediction error.

However, subsequent studies revealed a more complex picture. For example, ACC neurons were recorded while an animal searched among four targets using trial and error to find the one associated with reward (Quilodran et al., 2008). Sometimes the animal would get lucky and discover the reward with the first target it selected. Sometimes it would not discover the rewarded target until the other three had been ruled out. Once the animal discovered the rewarded target, it was allowed to select it several more times and earn several more rewards. Similar to previous studies, many neurons responded to reinforcement. Responses were strongest when the correct target was first discovered and weaker when the target was reselected in order to receive more reward. However, the authors noted that the response did not resemble a prediction error. The response to the correct target was the same irrespective of whether it was selected on the first try (i.e., when there was a one in four chance of being correct) or whether it was

only selected after the other three targets had been ruled out (in which case the reward was certain). In other words, in this task the animal's prior expectancy of receiving a reward did not affect the neuronal response.

In order to clarify the role of ACC neurons in encoding reward prediction errors, we analyzed data from a task that minimized the effects of learning by using overlearned stimuli (**Figure 1**). The monkey had learned the probability with which these stimuli predicted reward delivery over many thousands of trials such

that their presentation would produce a specific expectancy of reward. A similar approach helped to delimit prediction errors in dopamine neurons (Fiorillo et al., 2003). We found a broad variety of responses in ACC neurons (**Figure 2**), including neurons that encoded whether something was better than expected (**Figure 2A**, positive prediction error) or worse than expected (**Figure 2B**, negative prediction error; Kennerley and Wallis, in press). Other groups, adopting a similar approach, have found that ACC neurons encode a saliency signal (i.e., whether an outcome
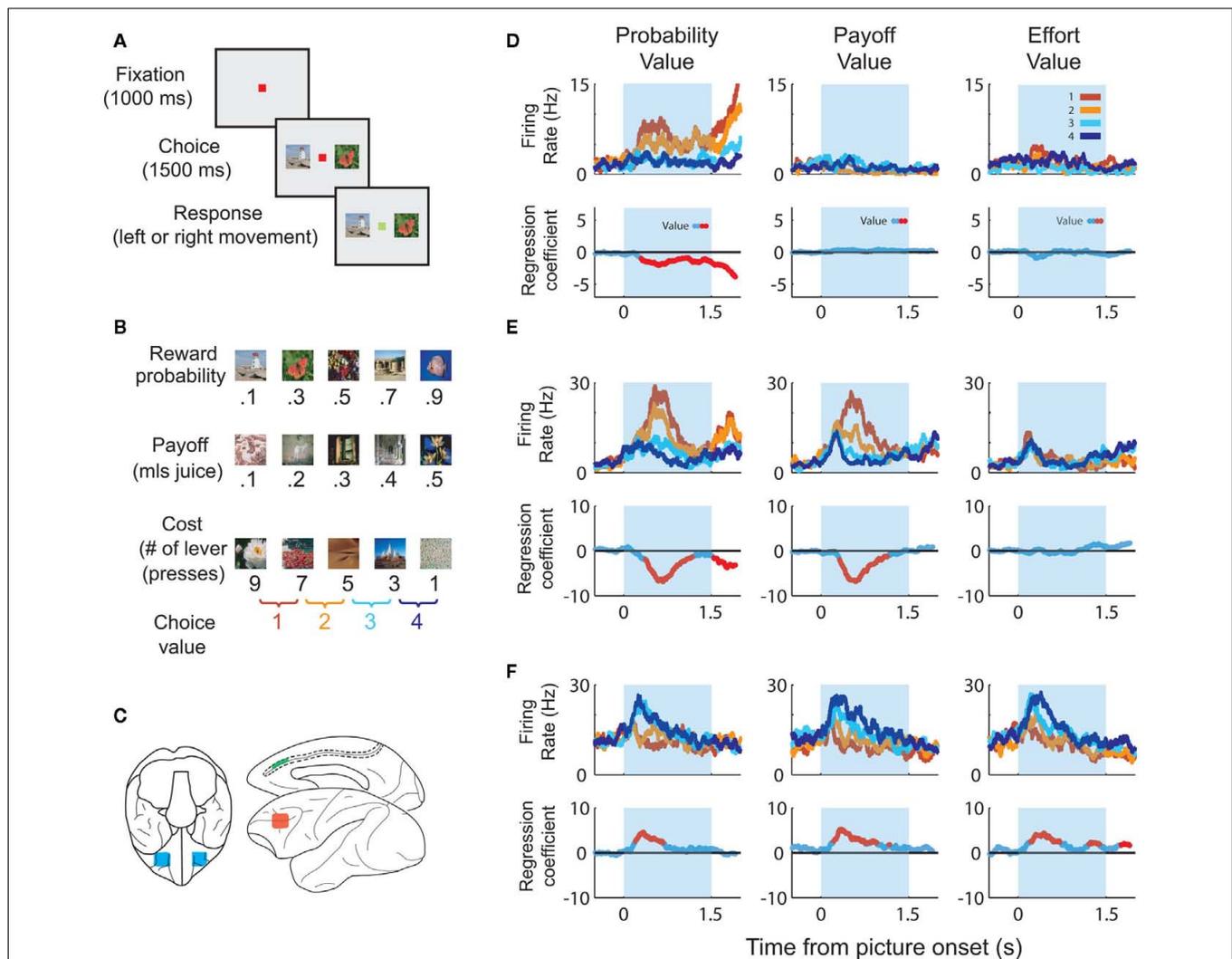

FIGURE 1 | Task parameters associated with the multidimensional choice task. (A) The task began with the subject fixating a central spot. Two pictures appeared, one on the left and one on the right. When the fixation spot changed color the subject selected one of the pictures and received the associated outcome. **(B)** Each picture was associated with a specific outcome. The "probability" pictures were associated with a set amount of juice, delivered on only a certain fraction of the trials. The "payoff" pictures were associated with different amounts of juice reward. The "cost" pictures were associated with a specific amount of juice, but the subject had to earn the juice by pressing a lever a different number of times. We only presented pairs of pictures that were from the same set and that were adjacent to one another in terms of value. Thus, for each set of pictures there were four

potential choices. **(C)** The approximate locations that we recorded in OFC (blue), ACC (green) and lateral prefrontal cortex (red). **(D)** The upper row of plots illustrates spike density histograms from a single ACC neuron sorted according to the value of the expected outcome of the choice. The lower row of plots illustrates a statistical measure of the extent to which the variance in the neuron's firing rate can be explained by the value of the choice. Portions of the curve shown in red indicate significant encoding of value at those time points. The neuron encodes value solely on probability trials with an increase in firing rate as the value of the choice decreases. **(E)** An ACC neuron that encodes value on probability and payoff trials, increasing its firing rate as value decreases. **(F)** An ACC neuron that encodes value for all decision variables, increasing its firing rate as value increases.
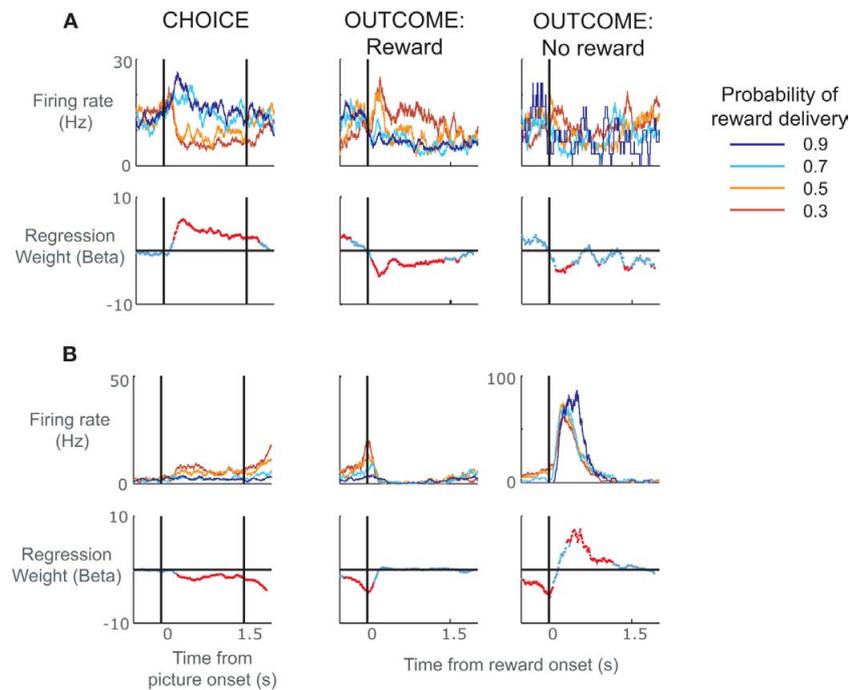
FIGURE 2 | Spike density histograms illustrating single neurons that encoded value information during the choice as well as the subsequent outcome of the choice. (A) The top row of plots consists of spike density histograms recorded from a single ACC neuron and sorted according to probability of reward delivery as indicated by the pictures. The three plots show activity during the choice phase, the outcome phase when a reward was delivered, and the outcome phase when a reward was not delivered. For the choice phase, the vertical lines relate to the onset of the pictures and the time at which the subject was allowed to make his choice. For the outcome phase, the vertical line indicates the onset of the juice reward. The lower row of plots indicates neuronal selectivity determined using regression to calculate the amount of variance in the neuron's firing rate at each time point that can be explained by the probability of reward delivery. Red data points

indicate time points where the probability of reward delivery significantly predicted the neuron's firing rate. The neuron responded during the choice phase when pictures appeared that predicted reward delivery with high probability. It also responded during reward delivery, but only when the subject was least expecting to receive the reward. It shows little response when the subject did not receive a reward. In other words, the neuron encoded a positive prediction error, i.e., it responded when either choice offerings or outcomes were better than expected. **(B)** An ACC neuron that encoded a negative prediction error, i.e., it responded when events occurred that were worse than expected. The neuron responded when pictures appeared that predicted reward delivery with low probability, showed little response to the delivery of reward, and responded when reward was not delivered, particularly when the subject was expecting to receive a reward.

was unexpected irrespective of whether it was better or worse than expected; Hayden et al., 2011) as well as "fictive" error signals, neuronal responses to outcomes for actions that one did not take (Hayden et al., 2009).

Thus, the picture that is emerging from ACC is of an area that encodes a variety of signals that would be useful for learning, with a common thread being that they integrate information about the outcome of actions and their relationship to prior expectancies. This contrasts with activity recorded from dopamine neurons, where the vast majority of signals correlate with reward prediction errors (Fiorillo et al., 2003; Bayer and Glimcher, 2005; Bayer et al., 2007), encoding positive prediction errors with increased firing rates and negative prediction errors with decreased firing rates. ACC neurons also encode prediction errors, but with a good deal more heterogeneity. This heterogeneity should not be surprising. Cortex is responsible for performing multiple computations in parallel and integrating a diversity of information. Even neurons in primary sensory areas encode multiple parameters of a stimulus space (Carandini et al., 2005). In contrast, signals from neurotransmitter systems appear more uniform, performing

a single computation and broadcasting it to a large portion of the brain (Schultz, 1998; Yu and Dayan, 2005), although we note that recent studies have challenged whether signals from neurotransmitter systems are quite as homogenous as originally thought (Bromberg-Martin et al., 2010).

In our study, we found that the activity of ACC neurons during the feedback period tended to match that in the choice period. If a neuron responded to rewards that were better than expected, it also tended to respond when the choice was between better than average alternatives. For example the neuron shown in **Figure 2A** had the largest responses during the outcome period when the picture associated with a low probability of receiving reward unexpectedly resulted in reward, that is, it responded when the outcome was better than expected. During the choice period, the very same neuron had highest firing rates when the monkey was presented with the option to choose a picture with high reward probabilities, and lower firing rates when his best option was a lower probability picture. Because we always presented pictures that were adjacent in value (**Figure 1B**) and the subjects virtually always chose the best option, we cannot determine whether this activity is related to the

average value of the two stimuli or the higher-valued chosen stimulus. In either circumstance, however, the neuron could be viewed as responding to a situation that is better than expected, which, during the choice phase, was the presentation of stimuli indicating the highest probability of receiving a reward. This raises the question of whether we should reinterpret our original conclusions regarding ACC activity during the choice phase as encoding differences between the value of the present options and the average choice value (i.e., a choice prediction error) rather than the value of the choice *per se.*

Studying decision-making requires presenting subjects with choices. This is typically done in such a way as to minimize other cognitive processes, such as learning, that might confound the interpretation of neural activity related to decision-making. Choices are randomized, independent of one another and, in humans, frequently trial-unique. However, even with these precautions in place, subjects are still able to learn. They are learning the range and average value of the choices that the experimenter might present. Consequently, although activity during the choice may reflect predictions about the outcome of the choice alternatives, it could equally reflect a prediction error, encoding the current options relative to the other potential choices the subject might have expected. For instance, if a subject has extensive experience with three equally probable choices valued at 0, 1, and 2, the average value of a choice in this experiment is 1. An offered choice of 2 is better than expected and could produce a prediction error at the time of the offer. Furthermore, in the typical decision-making experiment, outcome values and prediction errors are often strongly correlated. That is, a highly valued outcome is likely to be better than average and generate a large prediction-error relative to a second option where both value and prediction error might be smaller.

Figure 3 illustrates this point graphically. In this example task, the subject is presented with two stimuli on each trial (S_R and S_L), indicates his choice with a response (right or left arrows) and receives an outcome (O; **Figure 3A**). The value of S_R and S_L, the value of the outcome, and the prediction errors are represented by the height of colored bars (**Figure 3B**). Note that we have drawn prediction-error activity as it would appear in neuronal activity: computationally, events that are worse than expected should generate a negative prediction error, but neurons cannot have a negative firing rate. Thus, prediction errors are encoded relative to the neuron's mean or baseline firing rate (although it is not trivial to determine such a baseline in a high-level cognitive area such as prefrontal cortex). On trial N, the subject is presented with two low value stimuli and its choice (S_L) unexpectedly yields a large reward (green). The outcome generates a large positive prediction error (orange) because it was better than expected. However, neurons encoding prediction-errors could respond to the presentation of the reward-predicting stimuli as well as the receipt of the rewarding outcome. This is illustrated as a prediction error during the choice phase. In trial N, the choice prediction error is low because the values of the stimuli are low relative to other stimuli in the set. On trial $N + 1$, the subject is presented with two higher value stimuli, chooses S_R and unexpectedly receives a small reward. In this case, the choice prediction error is high, because S_R was expected to yield a large reward, however the outcome was

worse than expected, leading to a negative prediction error during the outcome phase. Finally, on trial $N + 2$, the subject is presented with a high and a low value stimulus, chooses S_R and receives a large reward as expected. Here the choice prediction error is high, because the subject is given the option of a high valued stimulus, and there is no outcome prediction error (or the height of the bar is about at "baseline") because the outcome was fully predicted. From this illustration, it is evident that if you were to focus solely on neuronal activity during the choice phase you would not be able to differentiate neurons encoding prediction errors from those encoding the value of the chosen stimulus. Yet these signals have very different implications for the larger question of how the brain computes value-based decisions.

A prominent role for ACC in the encoding of learning signals is consistent with the dopaminergic input that this region receives. All areas of frontal cortex receive dopaminergic input, but it is particularly heavy in ACC (Williams and Goldman-Rakic, 1993). Furthermore, while dopamine signals have a very short latency (typically <100-ms from the onset of a reward or reward-predictive stimulus) ACC signals evolve over a longer timeframe. We found that the median time of ACC neurons to encode the amount of reward predicted by a stimulus was 230-ms (Kennerley and Wallis, 2009a). This would also be consistent with the dopaminergic input into ACC being responsible for the prediction errors that are observed there. However, we also note that the theory that dopamine neurons encode prediction errors is not without controversy. For example, there is debate as to whether prediction-error activity in dopamine neurons is a cause or consequence of value learning (Berridge, 2007). Further, some authors have suggested that responses of dopamine neurons are too rapid to encode prediction errors and instead respond to sensory properties of unexpected stimuli, and true prediction-error encoding should occur with longer latencies, in higher cortical areas (Redgrave and Gurney, 2006). These ideas raise the possibility that prediction errors could initially be computed in ACC during learning and then used to train up the short latency dopamine responses.

NEUROPHYSIOLOGICAL PROPERTIES OF OFC
With regard to OFC, a broad consensus seems to be emerging that OFC neurons encode value predictions rather than prediction errors (Roesch et al., 2010). We found no evidence of prediction-error signals in OFC in monkeys using the same task in which we detected prediction errors in ACC (Kennerley and Wallis, in press). Although OFC neurons encoded whether or not a reward occurred, there was little evidence that this signal was influenced by the animal's prior expectancy of receiving the reward. In humans, subjects have been required to bid on food items while simultaneously winning or losing money, thereby enabling prediction errors to be uncorrelated from value signals (Hare et al., 2008). fMRI revealed that OFC activity correlated with value while ventral striatal activity correlated with prediction errors. In rats performing an odor-guided task, OFC neurons encoded predictions but not prediction errors (Roesch et al., 2006), while the opposite was true for dopamine neurons (Roesch et al., 2007).

Anatomically, OFC is in an ideal position to encode the reward value of sensory stimuli. It receives input from high-level sensory
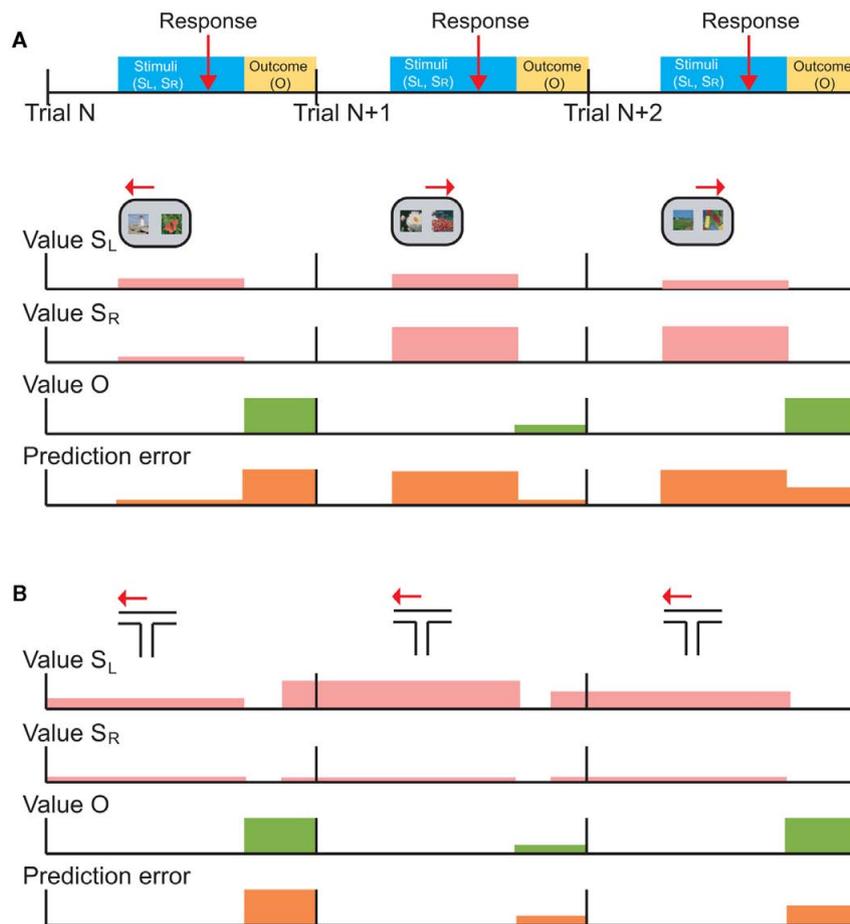
FIGURE 3 | Schematic depictions of typical choice tasks for primates and rodents are shown with the putative neuronal signals that those tasks should generate. (A) The temporal occurrence of behavioral events common to both tasks. On each trial (starting at the vertical black bars), the subject is presented with a choice between left (S_L) and right (S_R) stimuli, makes a response (red arrow), and receives an outcome (yellow shading). (B) Choice task typical in monkeys or humans, in which the subject is presented with the choice of two visual stimuli, S_L and S_R, selects one (left or right arrows), and then receives an outcome. The rows with pink bars show hypothetical learned values of S_L and S_R. A typical choice task conducted in humans and primates uses well learned stimuli from a larger set of reward-predictive stimuli, so the depicted values are shown as if they are well-known. The height of the bars indicates the degree of value, so that S_L has a slightly higher value than S_R on

trial N, and so forth. The following two rows show the value of the actual outcome (Value O) and prediction-errors generated throughout the trial. A prediction error can be generated at the time of the choice, since the subject does not know specifically which choice will be presented. This choice prediction error is the difference between the value of the presented options, and the average value of the complete set of possible options. (C) A typical choice task conducted in rodents, in which the animal chooses between one of two arms in a T-maze, and receives an outcome. In this case, the same choice is effectively presented on every trial, so value predictions for S_L and S_R can be updated at the time of outcome receipt (green). This is shown as value predictions (pink) updating prior to the start of the next trial (i.e., shaded bars are shifted to the left). Furthermore, there is no choice prediction error because each trial consists of the same two choice options.

areas (Carmichael and Price, 1995b) as well as limbic structures responsible for processing rewards, such as the amygdala and hypothalamus (Carmichael and Price, 1995a). In addition, posterior OFC is responsible for the integration of taste and smell (Rolls and Baylis, 1994). Finally, OFC neurons encode the amount of reward predicted by a stimulus quickly, typically within 200-ms of the presentation of the stimulus. This is significantly quicker than neurons in ACC (Kennerley and Wallis, 2009a) or lateral prefrontal cortex (Wallis and Miller, 2003). This suggests that OFC neurons could serve as a source of reward information for the rest of the frontal cortex and perhaps even for subcortical structures. Indeed, a recent study has examined the relationship between OFC

and dopamine. The authors disconnected OFC from the dopaminergic system using a crossed inactivation procedure, and showed that the two needed to interact in order for rats to learn from unexpected outcomes (Takahashi et al., 2009). They suggested that OFC provides dopamine neurons with a prediction as to the expected outcome. The dopamine neurons can then use this information, along with information about the actual outcome, in order to calculate a prediction error.

An exception to the consensus that OFC neurons encode predictions is a study examining the ability of rats to learn a probabilistically rewarded T-maze, which found encoding of prediction errors in OFC at the time of reward delivery (Sul et al., 2010). It

is possible that this difference reflects a difference in functional anatomy between rodents and primates. Most recordings from OFC in primates focus on areas 11 and 13, dysgranular cortex that may not have a homolog in rodents (Wise, 2008). It is the posterior, agranular OFC in primates that is the likely homolog of rodent OFC, yet this OFC region is frequently neglected by primate neurophysiologists. Thus, it is possible that if primate neurophysiologists were to record from this posterior OFC region they would see prediction errors.

However, it is also possible that differences in the way in which choice behavior is tested between primates and rats may contribute to observed neurophysiological differences. In primates and humans, each trial typically involves a two-alternative choice between reward-predictive stimuli whose outcome contingencies have been previously learned and that are drawn from a larger set of possible reward-predictive stimuli (e.g., **Figure 3B**). Thus, at the time of reward delivery a prediction error can be calculated, but the subject cannot make any specific predictions about the next trial, since it does not know which choice will be presented next. In contrast, rodents are typically tested in a learning situation involving the same two-alternative choice on every trial. For example, in the T-maze task the two alternatives are the right or left goal arms; in a task requiring nose-pokes, alternatives might be right or left ports. This means that at the time of reward delivery, not only can the rat calculate a reward prediction error, but it can also represent value predictions about the next trial.

The T-maze task is illustrated in **Figure 3C**. On the first trial, the subject selects a low value arm (S_L) and unexpectedly receives a large reward. The subject can immediately update the value of this arm for the next trial. On trial $N + 1$, the subject repeats this choice (S_L) and unexpectedly receives a small reward. The subject again updates its estimate of the value of this arm for the next trial, in this case decreasing its estimate of the value. Note that in this situation, prediction-errors generated by the outcome can overlap with the predictions for the next trial. These signals should be distinguishable during learning: as performance improves, predictions will begin to accurately reflect the value of the choice while prediction errors will tend toward zero. However, neuronal representations of stimulus values might change for reasons other than learning, for example due to adaptation (Padoa-Schioppa, 2009) or satiation (Bouret and Richmond, 2010).

This raises the question as to why other rodent studies did not see the same type of activity in OFC, since they also used the same two-alternative choice (Roesch et al., 2006). A key difference is that in this study rats were only given a free choice on about one-third of the trials. On the other two-thirds of the trials they were forced to choose one of the alternatives. This makes the task more similar to primate choice tasks, in that the rat is unable to predict the value of the next trial. In this situation, rat OFC neurons did not encode any signal that looked like a prediction error at the time of reward delivery, raising the possibility that these signals are actually related to predictions about the next trial.

SUMMARY

Across a broad range of studies OFC activity appears most consistent with encoding value predictions, and ACC activity appears most consistent with value prediction errors. In theory, there should be little problem in separating these two types of signal in the choice situation. The subject is faced with a choice, makes its selection and receives an outcome. At the time of choice, neurons should encode a prediction regarding the value of the potential outcomes. At the time of the outcome, neurons should encode a prediction-error reflecting the discrepancy between the actual outcome and the prediction. In practice, however, things are more problematic. Prediction errors can be generated at the time of the choice, because the subject is comparing the choice with other potential choices that may have occurred, and predictions can be generated at the time of the outcome if the subject is going to experience the same choice on the next trial. It is important to recognize that trials in behavioral tasks do not take place in isolation and computational processes occurring within the temporal limits of one trial could reflect the influence of past or upcoming trials.

SALIENCY AND ITS EFFECTS ON ATTENTION, AROUSAL, AND MOTOR PREPARATION

Valuable items are salient. Even under experimental conditions, a high value item can trigger a variety of processes linked to its saliency, including an increase in attention, arousal, and motor preparation (Maunsell, 2004; Luk and Wallis, 2009). These processes, in turn, can have clear behavioral consequences. For example, offers of larger or more immediate rewards increase motivation and attentiveness to tasks, resulting in fewer incorrect responses and fewer errors in task execution, such as breaks in visual fixation (Kennerley and Wallis, 2009b). Larger rewards enhance preparation for response execution, so that motor responses are faster when more desirable outcomes are at stake (Kawagoe et al., 1998; Leon and Shadlen, 1999; Hassani et al., 2001; Kobayashi et al., 2002; Roesch and Olson, 2003). Rewards generate psychophysiological measures of arousal such as changes in galvanic skin conductance (Bechara et al., 1996) and heart rate and blood pressure (Braesicke et al., 2005). Finally, value can even correlate with muscle tone in the neck and jaw as recorded by EMG, likely a result of arousal or preparation for ingestive behaviors (Roesch and Olson, 2003, 2005). These are potent demonstrations that behavioral and physiological responses can be tightly coupled to value, and neural encoding of these effects could be indistinguishable from value encoding. Indeed, at the neural level salient items have widespread effects. There are stronger representations of more valuable cues in nearly all cortical areas considered, even primary sensory areas (Pantoja et al., 2007; Pleger et al., 2008; Serences, 2008), likely because of the heightened attentional and motivational salience of valuable items. However, interpretations become difficult when multiple signals correlating with value are found in frontal regions. For instance, neural representations of cognitive processes like working memory in ventrolateral PFC are influenced by reward magnitude (Kennerley and Wallis, 2009b) even though a lesion of lateral PFC has no effect on value-based decisions (Baxter et al., 2009). Analysis of latencies to respond to valuable stimuli found that OFC encodes value earlier than lateral PFC, suggesting that value information is passed from OFC to other PFC regions (Wallis and Miller, 2003). While these lateral PFC signals likely serve important functions, such as allocating cognitive resources appropriately, it is important to distinguish these downstream effects from value calculations themselves.

The only way to dissociate putative value signals from signals relating to saliency is to use stimuli or events that are aversive. Aversive stimuli (e.g., electric shock) have a negative value in that they negatively reinforce actions and can motivate avoidance behavior. True value signals should distinguish appetitive stimuli (rewards) and aversive stimuli (punishments). In contrast, saliency is associated with expectation of either punishment or reward, so that neural responses correlating with salience should be similar under rewarding and punishing conditions (Lin and Nicolelis, 2008).

Before we consider how these ideas have been applied to the interpretation of neuronal data, there are two additional issues we should consider. First, it is not necessarily the case that rewards and punishments will be encoded on the same scale. One neuronal population could encode the value of appetitive stimuli while a separate population could encode the value of aversive stimuli. Indeed, two prominent theories regarding the organization of value information have posited separate representations of appetitive and aversive information. One theory suggests that positive and negative outcomes are encoded by medial and lateral OFC respectively (Kringelbach and Rolls, 2004; Frank and Claus, 2006), while another argues that this distinction lies between OFC and ACC respectively (Aston-Jones and Cohen, 2005). Though there are data to support and contradict both theories, in principle, it is possible that different neural circuits represent different valences. In contrast, saliency signals by definition cannot discriminate the valence of the stimulus: if they did, they would be operationally identical to value signals.

There are also psychological reasons why rewards and punishments might not be encoded by the same neuronal population. Whereas subjects work to obtain rewards, they work to avoid aversive outcomes. This introduces a key paradox of avoidance learning: as learning progresses, there is less and less exposure to the reinforcing stimulus. By standard reinforcement learning theory, this situation should produce extinction, yet robust avoidance learning is readily obtained (Solomon et al., 1953). The influential two-process theory (Mowrer, 1947) suggests that aversively conditioned cues come to elicit a negative emotional state through Pavlovian conditioning. Responses that terminate the cue are then reinforced by the reduction of the negative emotional state. A similar two-process theory has been postulated to underlie learning about rewards (Rescorla and Solomon, 1967). In this case, the cues activate positive emotional states, which in turn elicit responses toward the desired outcome. Thus, if learning requires the activation of specific emotional states, it is possible that different neuronal populations will be responsible for the representation of different emotional states rather than a single neuronal population encoding value along a common scale.
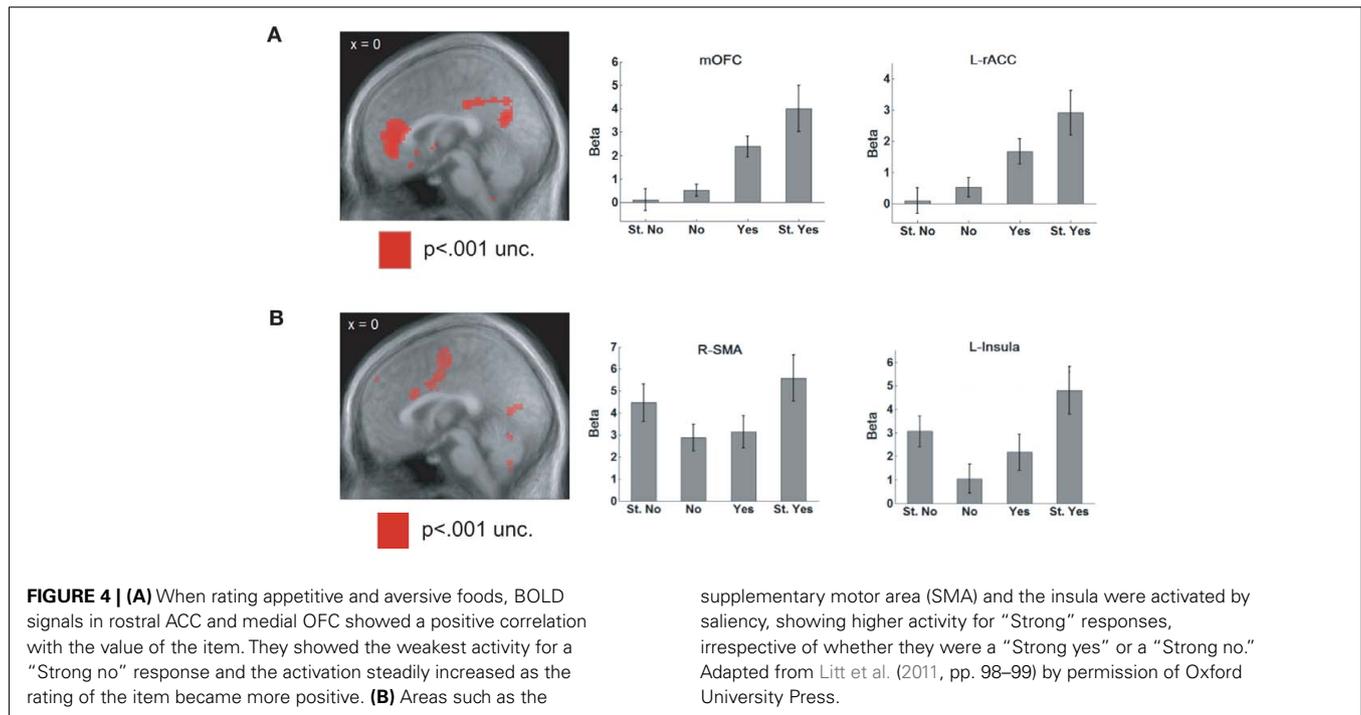
A second issue relates to the conflation of costs with aversive stimuli. Motivated behavior typically accrues certain costs, such as the time and effort involved in acquiring a desired outcome, or the risk that the desired outcome will not be obtained. Although costs influence behavior (e.g., all other things being equal the subject will choose the outcome whose acquisition involves the lowest costs), the desired outcome, not the cost, provides motivation for behavior. The subject's goal is to acquire an appetitive stimulus or avoid an aversive stimulus, and the cost is a necessary evil in obtaining that goal.

DISSOCIATING VALUE AND SALIENCY SIGNALS

The goal of dissociating value and saliency signals motivated an experiment in which hungry humans were shown a variety of food items and asked whether they would like to eat them (Litt et al., 2011). They provided ratings of "Strong no," "No," "Yes," or "Strong yes." The food items were chosen to be appetitive (e.g., potato chips) or aversive (e.g., baby food). BOLD signals in rostral ACC and medial OFC showed a positive correlation with the value of the item, lowest for items rated "Strong no" and highest for "Strong yes" (**Figure 4A**). In contrast, areas such as the supplementary motor area (SMA) and the insula consistently showed higher activity for "Strong" responses, irrespective of whether they were a "Strong yes" or a "Strong no," consistent with a saliency-related signal (**Figure 4B**).

However, there is an important caveat to the interpretation of neuroimaging results. Neuroimaging studies tend to largely localize value signals to the ventral part of the medial wall of prefrontal cortex, yet single neurons encoding value are found throughout frontal cortex (Wallis and Kennerley, 2010). This suggests that neuroimaging methods are underestimating the extent of frontal cortex involved in valuation processes. A possible explanation for this lies in the fact among value encoding neurons, those that increase their firing rate as values increase are found in approximately equal numbers as those that increase their firing rate as values decrease (Kennerley and Wallis, 2009a; Kennerley et al., 2009; Padoa-Schioppa, 2009). It is possible that such signals could cancel one another out when averaged together in the BOLD signal. Support for this idea has come from recent studies that have directly compared the BOLD response to underlying neuronal activity in area MT (Lippert et al., 2010). Neuronal activity (whether measuring action potentials or local field potentials) shows parametric tuning related to the direction of motion of a visual stimulus. However, the BOLD response is evoked by stimuli moving in any direction, precisely as though tuned populations were being added together thereby masking the tuning. Consequently, it is important to dissociate value and saliency signals at the single-unit level.

The first study that attempted to systematically dissociate these two signals required a monkey to choose between stimuli that were associated either with different amounts of juice or different lengths of a "time-out" (the monkey had to simply sit and wait a designated amount of time until the next trial started and did not receive any juice; Roesch and Olson, 2004). Both OFC and premotor neurons tended to increase firing rate as expected rewards increased, appearing to code the value of different reward magnitudes. However, only OFC neurons decreased firing as expected punishment increased, and thus scaled with the value of both positive and negative outcomes. Premotor neurons, in contrast, increased firing to increasing penalties, suggesting that they code information related to motivation, arousal, or motor preparation. Although this study suggests that OFC neurons encode rewards and punishments along a single scale, there is an alternative explanation. It is not clear that a "time-out" is necessarily a punisher and could instead be construed as a cost that must be overcome in order to obtain reward on subsequent trials. Indeed, several studies suggest that OFC may be responsible for integrating reward information with temporal costs (Roesch and Olson, 2005; Rudebeck et al., 2006; Kable and Glimcher, 2007).

FIGURE 4 | (A) When rating appetitive and aversive foods, BOLD signals in rostral ACC and medial OFC showed a positive correlation with the value of the item. They showed the weakest activity for a "Strong no" response and the activation steadily increased as the rating of the item became more positive. (B) Areas such as the supplementary motor area (SMA) and the insula were activated by saliency, showing higher activity for "Strong" responses, irrespective of whether they were a "Strong yes" or a "Strong no." Adapted from Litt et al. (2011, pp. 98–99) by permission of Oxford University Press.

Subsequent studies have explored OFC responses to cues that predict more unambiguous punishers such as electric shock (Hosokawa et al., 2007) or an air puff to the face (Morrison and Salzman, 2009). Notably, there was evidence that single OFC neurons encoded the appetitive and aversive outcomes along a single scale. For example, they would show a strong response to a large reward, a weaker response to a small reward and an even weaker response to the aversive stimulus. Importantly, there was no evidence of a functional topography of responses. Neurons that showed stronger responses to aversive events appeared to be randomly interspersed with neurons that showed stronger responses to appetitive events (Morrison and Salzman, 2009), casting doubt on the theory that appetitive encoding is located more medially than aversive coding (Kringelbach and Rolls, 2004; Frank and Claus, 2006). In this study, cue–outcome associations were conditioned in a Pavlovian manner, eliminating decision-making from the task design. It will be interesting in future studies to examine how these findings extend to choice behavior.

GAINS AND LOSSES

Our discussion so far has focused on positive punishment: punishing behavior by presenting an aversive stimulus. However, there is a second class of punishment, negative punishment, in which a subject is punished by the removal of an appetitive stimulus. Most studies of valuation in humans involve winning and losing money (Breiter et al., 2001; Knutson et al., 2005), which is a form of negative punishment, in that losing money consists of the removal of an appetitive stimulus. Critically, negative punishment requires the use of secondary reinforcement. This is because once a subject has received a primary reinforcement, such as a shock or a food reward, there is no way to take it back. In contrast, a secondary reinforcer, such as money, can be removed before the subject has had the ability to consume it.

Few studies in animals have involved negative punishment. One exception is a study that examined the ability of animals to play a competitive game for tokens (Seo and Lee, 2009). Monkeys played against a computer opponent, trying to guess which of two targets the computer would choose on each trial. If both subject and computer chose the same target, the subject had a high probability of gaining a token; if they chose different targets there was the risk of losing a token. For every six tokens won, the animal received juice as a primary reward. Although the optimal approach to the task was to choose randomly, monkeys tended to modify their behavior based on their previous choices, demonstrating that gains and losses affected choice behavior. As such, individual neurons in multiple cortical areas, including dorsomedial PFC, dorsolateral PFC, and dorsal ACC, had differential responses to gains and losses, the first time the effects of negative punishment have been seen at the single neuron level. Furthermore, individual neurons showed opposing responses to gains and losses relative to neutral outcomes. For example, they might show a strong response to a gain, a weaker response to a neutral outcome, and little or no response to a loss indicating that reward and punishment may be coded along a single value dimension.

Nevertheless, this use of gains and losses of conditioned reinforcers remains an exception in the animal literature. Most animal studies do not include punishment, and if they do it is typically positive punishment. The precise implications of this disconnect between human and animal studies remains unclear, but recent findings suggest that different regions of OFC may be responsible for the encoding of primary and secondary reinforcement. For example, monetary reward, a secondary reinforcer, activates more anterior regions of OFC than erotic pictures, a primary reinforcer (Sescousse et al., 2010). Furthermore, aversive conditioning based on monetary loss does not activate the amygdala, which is highly interconnected with OFC, while the same conditioning based on

electric shock does (Delgado et al., 2011), suggesting that OFC may also respond differentially to positive and negative punishment. Consequently, different conclusions may be reached by investigators studying decision-making in animals or humans, not because of a genuine species difference, but rather because of a difference in the way the species are tested behaviorally. In future neurophysiology studies it will be important to bring clarity to these issues by comparing single neuron responses in both OFC and ACC to primary and secondary reinforcement.

SUMMARY

In sum, measures of value and saliency signals are highly correlated unless tasks employ both rewarding and punishing outcomes. Aversive events can include either primary punishment, such as electric shock, or negative punishment, such as the loss of a valuable item. In either case they should be distinguished from a cost that accompanies reward, since it is unknown whether costs and punishments are coded similarly at the neural level. A number of studies have now successfully disambiguated value from saliency signals, and found that ACC and OFC activity correlates with value, not saliency. It is important to keep pursuing these types of distinctions, since they have significant implications for our interpretation of neuronal activity.

CONCLUSION

It has been over 30 years since the first studies determined that frontal neurons showed responses that predicted reward outcomes (Niki and Watanabe, 1976; Rosenkilde et al., 1981). In the ensuing decades, researchers have made a great deal of progress in understanding how positive and negative outcomes can influence choices. Formal behavioral models have been adopted, which, in turn, have allowed for a more quantitative analysis of neuronal responses. In this paper, we have outlined a number of challenges confronted when assessing the neural correlates of these behavioral models. We hope that this will prove useful for disambiguating the contribution of different neuronal populations to choice behavior.

ACKNOWLEDGMENTS

REFERENCES

Amador, N., Schlag-Rey, M., and Schlag, J. (2000). Reward-predicting and reward-detecting neuronal activity in the primate supplementary eye field. *J. Neurophysiol.* 84, 2166–2170.

Amiez, C., Joseph, J. P., and Procyk, E. (2006). Reward encoding in the monkey anterior cingulate cortex. *Cereb. Cortex* 16, 1040–1055.

Aston-Jones, G., and Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annu. Rev. Neurosci.* 28, 403–450.

Baxter, M. G., Gaffan, D., Kyriazis, D. A., and Mitchell, A. S. (2009). Ventrolateral prefrontal cortex is required for performance of a strategy implementation task but not reinforcer devaluation effects in rhesus monkeys. *Eur. J. Neurosci.* 29, 2049–2059.

Bayer, H. M., and Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47, 129–141.

Bayer, H. M., Lau, B., and Glimcher, P. W. (2007). Statistics of midbrain dopamine neuron spike trains in the awake primate. *J. Neurophysiol.* 98, 1428–1439.

Bechara, A., Damasio, H., Tranel, D., and Anderson, S. W. (1998). Dissociation of working memory from decision making within the human prefrontal cortex. *J. Neurosci.* 18, 428–437.

Bechara, A., Tranel, D., Damasio, H., and Damasio, A. R. (1996). Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex. *Cereb. Cortex* 6, 215–225.

Berridge, K. C. (2007). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology (Berl.)* 191, 391–431.

Botvinick, M. M., Cohen, J. D., and Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends Cogn. Sci.* 8, 539–546.

Bouret, S., and Richmond, B. J. (2010). Ventromedial and orbital prefrontal neurons differentially encode internally and externally driven motivational values in monkeys. *J. Neurosci.* 30, 8591–8601.

Braesicke, K., Parkinson, J. A., Reekie, Y., Man, M. S., Hopewell, L., Pears, A., Crofts, H., Schnell, C. R., and Roberts, A. C. (2005). Autonomic arousal in an appetitive context in primates: a behavioural and neural analysis. *Eur. J. Neurosci.* 21, 1733–1740.

Breiter, H. C., Aharon, I., Kahneman, D., Dale, A., and Shizgal, P. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron* 30, 619–639.

Bromberg-Martin, E. S., Matsumoto, M., and Hikosaka, O. (2010). Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron* 68, 815–834.

Carandini, M., Demb, J. B., Mante, V., Tolhurst, D. J., Dan, Y., Olshausen, B. A., Gallant, J. L., and Rust, N. C. (2005). Do we know what the early visual system does? *J. Neurosci.* 25, 10577–10597.

Carmichael, S. T., and Price, J. L. (1995a). Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. *J. Comp. Neurol.* 363, 615–641.

Carmichael, S. T., and Price, J. L. (1995b). Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. *J. Comp. Neurol.* 363, 642–664.

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., and Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science* 280, 747–749.

Cavada, C., Company, T., Tejedor, J., Cruz-Rizzolo, R. J., and Reinoso-Suarez, F. (2000). The anatomical connections of the macaque monkey orbitofrontal cortex. A review. *Cereb. Cortex* 10, 220–242.

Delgado, M. R., Jou, R. L., and Phelps, E. A. (2011). Neural systems underlying aversive conditioning in humans with primary and secondary reinforcers. *Front. Neurosci.* 5:71. doi:10.3389/fnins.2011.00071

Falkenstein, M., Hohnsbein, J., Hoorman, J., and Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalogr. Clin. Neurophysiol.* 78, 447–455.

Fellows, L. K., and Farah, M. J. (2007). The role of ventromedial prefrontal cortex in decision making: judgment under uncertainty or judgment per se? *Cereb. Cortex* 17, 2669–2674.

Fiorillo, C. D., Tobler, P. N., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299, 1898–1902.

Frank, M. J., and Claus, E. D. (2006). Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychol. Rev.* 113, 300–326.

Gehring, W. J., Coles, M. G., Meyer, D. E., and Donchin, E. (1990). The error-related negativity: an event-related potential accompanying errors. *Psychophysiology* 27, S34.

Hare, T. A., O'Doherty, J., Camerer, C. F., Schultz, W., and Rangel, A. (2008). Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *J. Neurosci.* 28, 5623–5630.

Hassani, O. K., Cromwell, H. C., and Schultz, W. (2001). Influence of expectation of different rewards on behavior-related neuronal activity in the striatum. *J. Neurophysiol.* 85, 2477–2489.

Hayden, B. Y., Heilbronner, S. R., Pearson, J. M., and Platt, M. L. (2011). Surprise signals in anterior cingulate cortex: neuronal encoding of unsigned reward prediction errors driving adjustment in behavior. *J. Neurosci.* 31, 4178–4187.

Hayden, B. Y., Pearson, J. M., and Platt, M. L. (2009). Fictive reward signals in the anterior cingulate cortex. *Science* 324, 948–950.

Holroyd, C. B., Nieuwenhuis, S., Yeung, N., Nystrom, L., Mars, R. B., Coles, M. G., and Cohen, J. D. (2004). Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. *Nat. Neurosci.* 7, 497–498.

Hosokawa, T., Kato, K., Inoue, M., and Mikami, A. (2007). Neurons in the macaque orbitofrontal cortex code relative preference of both rewarding and aversive outcomes. *Neurosci. Res.* 57, 434–445.

Hwang, J., Kim, S., and Lee, D. (2009). Temporal discounting and intertemporal choice in rhesus monkeys. *Front. Behav. Neurosci.* 3:9. doi:10.3389/neuro.08.009.2009

Ito, S., Stuphorn, V., Brown, J. W., and Schall, J. D. (2003). Performance monitoring by the anterior cingulate cortex during saccade countermanding. *Science* 302, 120–122.

Kable, J. W., and Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* 10, 1625–1633.

Kawagoe, R., Takikawa, Y., and Hikosaka, O. (1998). Expectation of reward modulates cognitive signals in the basal ganglia. *Nat. Neurosci.* 1, 411–416.

Kennerley, S. W., Dahmubed, A. F., Lara, A. H., and Wallis, J. D. (2009). Neurons in the frontal lobe encode the value of multiple decision variables. *J. Cogn. Neurosci.* 21, 1162–1178.

Kennerley, S. W., and Wallis, J. D. (2009a). Encoding of reward and space during a working memory task in the orbitofrontal cortex and anterior cingulate sulcus. *J. Neurophysiol.* 102, 3352–3364.

Kennerley, S. W., and Wallis, J. D. (2009b). Reward-dependent modulation of working memory in lateral prefrontal cortex. *J. Neurosci.* 29, 3259–3270.

Kennerley, S. W., and Wallis, J. D. (in press). Double dissociation of value computations in orbitofrontal cortex and anterior cingulate cortex. *Nat Neurosci.* doi: 10.1038/nn.2961. [Epub ahead of print].

Kennerley, S. W., Walton, M. E., Behrens, T. E., Buckley, M. J., and Rushworth, M. F. (2006). Optimal decision making and the anterior cingulate cortex. *Nat. Neurosci.* 9, 940–947.

Knutson, B., Taylor, J., Kaufman, M., Peterson, R., and Glover, G. (2005). Distributed neural representation of expected value. *J. Neurosci.* 25, 4806–4812.

Kobayashi, S., Lauwereyns, J., Koizumi, M., Sakagami, M., and Hikosaka, O. (2002). Influence of reward expectation on visuospatial processing in macaque lateral prefrontal cortex. *J. Neurophysiol.* 87, 1488–1498.

Kringelbach, M. L., and Rolls, E. T. (2004). The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. *Prog. Neurobiol.* 72, 341–372.

Leon, M. I., and Shadlen, M. N. (1999). Effect of expected reward magnitude on the response of neurons in the dorsolateral prefrontal cortex of the macaque. *Neuron* 24, 415–425.

Lin, S. C., and Nicolelis, M. A. (2008). Neuronal ensemble bursting in the basal forebrain encodes salience irrespective of valence. *Neuron* 59, 138–149.

Lippert, M. T., Steudel, T., Ohl, F., Logothetis, N. K., and Kayser, C. (2010). Coupling of neural activity and fMRI-BOLD in the motion area MT. *Magn. Reson. Imaging* 28, 1087–1094.

Litt, A., Plassmann, H., Shiv, B., and Rangel, A. (2011). Dissociating valuation and saliency signals during decision-making. *Cereb. Cortex* 21, 95–102.

Liu, Z., and Richmond, B. J. (2000). Response differences in monkey TE and perirhinal cortex: stimulus association related to reward schedules. *J. Neurophysiol.* 83, 1677–1692.

Luk, C. H., and Wallis, J. D. (2009). Dynamic encoding of responses and outcomes by neurons in medial prefrontal cortex. *J. Neurosci.* 29, 7526–7539.

Matsumoto, M., Matsumoto, K., Abe, H., and Tanaka, K. (2007). Medial prefrontal cell activity signaling prediction errors of action values. *Nat. Neurosci.* 10, 647–656.

Maunsell, J. H. (2004). Neuronal representations of cognitive state: reward or attention? *Trends Cogn. Sci. (Regul. Ed.)* 8, 261–265.

Morrison, S. E., and Salzman, C. D. (2009). The convergence of information about rewarding and aversive stimuli in single neurons. *J. Neurosci.* 29, 11471–11483.

Mowrer, O. H. (1947). On the dual nature of learning a reinterpretation of "conditioning" and "problem-solving". *Harv. Educ. Rev.* 17, 102–150.

Niki, H., and Watanabe, M. (1976). Cingulate unit activity and delayed response. *Brain Res.* 110, 381–386.

Niki, H., and Watanabe, M. (1979). Prefrontal and cingulate unit activity during timing behavior in the monkey. *Brain Res.* 171, 213–224.

Ono, T., Nishino, H., Fukuda, M., Sasaki, K., and Nishijo, H. (1984). Single neuron activity in dorsolateral prefrontal cortex of monkey during operant behavior sustained by food reward. *Brain Res.* 311, 323–332.

Padoa-Schioppa, C. (2007). Orbitofrontal cortex and the computation of economic value. *Ann. N. Y. Acad. Sci.* 1121, 232–253.

Padoa-Schioppa, C. (2009). Range-adapting representation of economic value in the orbitofrontal cortex. *J. Neurosci.* 29, 14004–14014.

Padoa-Schioppa, C., and Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature* 441, 223–226.

Pantoja, J., Ribeiro, S., Wiest, M., Soares, E., Gervasoni, D., Lemos, N. A., and Nicolelis, M. A. (2007). Neuronal activity in the primary somatosensory thalamocortical loop is modulated by reward contingency during tactile discrimination. *J. Neurosci.* 27, 10608–10620.

Platt, M. L., and Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature* 400, 233–238.

Pleger, B., Blankenburg, F., Ruff, C. C., Driver, J., and Dolan, R. J. (2008). Reward facilitates tactile judgments and modulates hemodynamic responses in human primary somatosensory cortex. *J. Neurosci.* 28, 8161–8168.

Quilodran, R., Rothe, M., and Procyk, E. (2008). Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron* 57, 314–325.

Rangel, A., and Hare, T. (2010). Neural computations associated with goal-directed choice. *Curr. Opin. Neurobiol.* 20, 262–270.

Redgrave, P., and Gurney, K. (2006). The short-latency dopamine signal: a role in discovering novel actions? *Nat. Rev. Neurosci.* 7, 967–975.

Rescorla, R. A., and Solomon, R. L. (1967). Two-process learning theory: relationships between Pavlovian conditioning and instrumental learning. *Psychol. Rev.* 74, 151–182.

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., and Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science* 306, 443–447.

Roesch, M. R., Calu, D. J., Esber, G. R., and Schoenbaum, G. (2010). All that glitters . . . dissociating attention and outcome expectancy from prediction errors signals. *J. Neurophysiol.* 104, 587–595.

Roesch, M. R., Calu, D. J., and Schoenbaum, G. (2007). Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat. Neurosci.* 10, 1615–1624.

Roesch, M. R., and Olson, C. R. (2003). Impact of expected reward on neuronal activity in prefrontal cortex, frontal and supplementary eye fields and premotor cortex. *J. Neurophysiol.* 90, 1766–1789.

Roesch, M. R., and Olson, C. R. (2004). Neuronal activity related to reward value and motivation in primate frontal cortex. *Science* 304, 307–310.

Roesch, M. R., and Olson, C. R. (2005). Neuronal activity in primate orbitofrontal cortex reflects the value of time. *J. Neurophysiol.* 94, 2457–2471.

Roesch, M. R., Taylor, A. R., and Schoenbaum, G. (2006). Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. *Neuron* 51, 509–520.

Rolls, E. T., and Baylis, L. L. (1994). Gustatory, olfactory, and visual convergence within the primate orbitofrontal cortex. *J. Neurosci.* 14, 5437–5452.

Rosenkilde, C. E., Bauer, R. H., and Fuster, J. M. (1981). Single cell activity in ventral prefrontal cortex of behaving monkeys. *Brain Res.* 209, 375–394.

Rudebeck, P. H., Walton, M. E., Smyth, A. N., Bannerman, D. M., and Rushworth, M. F. (2006). Separate neural pathways process different decision costs. *Nat. Neurosci.* 9, 1161–1168.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *J. Neurophysiol.* 80, 1–27.

Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.

Seo, H., and Lee, D. (2007). Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *J. Neurosci.* 27, 8366–8377.

Seo, H., and Lee, D. (2009). Behavioral and neural changes after gains and losses of conditioned reinforcers. *J. Neurosci.* 29, 3627–3641.

Serences, J. T. (2008). Value-based modulations in human visual cortex. *Neuron* 60, 1169–1181.

Sescousse, G., Redoute, J., and Dreher, J. C. (2010). The architecture of reward value coding in the human orbitofrontal cortex. *J. Neurosci.* 30, 13095–13104.

Shuler, M. G., and Bear, M. F. (2006). Reward timing in the primary visual cortex. *Science* 311, 1606–1609.

Solomon, R. L., Kamin, L. J., and Wynne, L. C. (1953). Traumatic avoidance learning: the outcomes of several extinction procedures with dogs. *J. Abnorm. Psychol.* 48, 291–302.

Stephens, D. W., and Krebs, J. R. (1986). *Foraging Theory*. Princeton: Princeton University Press.

Sul, J. H., Kim, H., Huh, N., Lee, D., and Jung, M. W. (2010). Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. *Neuron* 66, 449–460.

Takahashi, Y. K., Roesch, M. R., Stalnaker, T. A., Haney, R. Z., Calu, D. J., Taylor, A. R., Burke, K. A., and Schoenbaum, G. (2009). The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. *Neuron* 62, 269–280.

Thorpe, S. J., Rolls, E. T., and Maddison, S. (1983). The orbitofrontal cortex: neuronal activity in the behaving monkey. *Exp. Brain Res.* 49, 93–115.

Ullsperger, M., and Von Cramon, D. Y. (2003). Error monitoring using external feedback: specific roles of the habenular complex, the reward system, and the cingulate motor area revealed by functional magnetic resonance imaging. *J. Neurosci.* 23, 4308–4314.

Wallis, J. D., and Kennerley, S. W. (2010). Heterogeneous reward signals in prefrontal cortex. *Curr. Opin. Neurobiol.* 20, 191–198.

Wallis, J. D., and Miller, E. K. (2003). Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *Eur. J. Neurosci.* 18, 2069–2081.

Watanabe, M. (1996). Reward expectancy in primate prefrontal neurons. *Nature* 382, 629–632.

Williams, S. M., and Goldman-Rakic, P. S. (1993). Characterization of the dopaminergic innervation of the primate frontal cortex using a dopamine-specific antibody. *Cereb. Cortex* 3, 199–222.

Wise, S. P. (2008). Forward frontal fields: phylogeny and fundamental function. *Trends Neurosci.* 31, 599–608.

Wunderlich, K., Rangel, A., and O'Doherty, J. P. (2010). Economic choices can be made using only stimulus values. *Proc. Natl. Acad. Sci. U.S.A.* 107, 15005–15010.

Yu, A. J., and Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron* 46, 681–692.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.