Published in final edited form as:

J Cogn Neurosci. 2014 July ; 26(7): 1347–1362. doi:10.1162/jocn_a_00573.

Functional Organization of the Orbitofrontal Cortex

Erin L. Rich and Jonathan D. Wallis

Department of Psychology and Helen Wills Neuroscience Institute, University of California at Berkeley, Berkeley, California, USA

Abstract

Emerging evidence suggests that specific cognitive functions localize to different subregions of orbitofrontal cortex (OFC), but the nature of these functional distinctions remains unclear. One prominent theory, derived from human neuroimaging, proposes that different stimulus valences are processed in separate orbital regions, with medial and lateral OFC processing positive and negative stimuli respectively. Thus far, neurophysiology data have not supported this theory. We attempted to reconcile these accounts by recording neural activity from the full medial-lateral extent of the orbital surface in monkeys receiving rewards and punishments via gain or loss of secondary reinforcement. We found no convincing evidence for valence selectivity in any orbital region. Instead, we report differences between neurons in central OFC and those on the inferior-lateral orbital convexity (IC), in that they encoded different sources of value information provided by the behavioral task. Neurons in IC encoded the value of external stimuli, whereas those in OFC encoded value information derived from the structure of the behavioral task. We interpret these results in light of recent theories of OFC function and propose that these distinctions, not valence selectivity, may shed light on a fundamental organizing principle for value processing in orbital cortex.

Introduction

Converging evidence shows that the prefrontal cortex (PFC) encodes the value of stimuli and events, and these signals are thought to play a crucial role in guiding behavior toward optimal choices (Rushworth, Noonan, Boorman, Walton, & Behrens, 2011; Wallis & Kennerley, 2010). The OFC in particular appears important for predicting choice outcomes and learning to make adaptive decisions (Gläscher et al., 2012; Padoa-Schioppa, 2011; Rangel & Hare, 2010; Schoenbaum, Takahashi, Liu, & McDannald, 2011; Wallis, 2012). However, the orbital cortex is a large and heterogeneous area (Carmichael & Price, 1994), and it remains unclear how decision-related information is organized within OFC.

One prominent theory, based on results in the human functional neuroimaging literature, suggests that different orbital regions specialize in evaluating rewards and punishments (Kringelbach & Rolls, 2004). This and more recent studies (e.g. (Elliott, Agnew, & Deakin, 2010; Hayes & Northoff, 2012; Liu, Hairston, Schrier, & Fan, 2011) find that rewards and affectively positive stimuli increase fMRI BOLD responses in medial OFC (mOFC; Chib,

Rangel, Shimojo, & O'Doherty, 2009; Kim, Shimojo, & O'Doherty, 2011), while punishment and affectively negative stimuli increase BOLD responses on the inferior-lateral orbital convexity (IC) and anterior insula (Elliott et al., 2010; Fujiwara, Tobler, Taira, Iijima, & Tsutsui, 2008; Hayes & Northoff, 2012; Seymour et al., 2005) (Figure 1A). Thus, a medial-lateral gradient of valence processing has been proposed across the orbital surface (Kringelbach & Rolls, 2004; Liu et al., 2011; O'Doherty, Kringelbach, Rolls, Hornak, & Andrews, 2001). It is debated whether these effects are driven by hedonic properties of reward, reinforcement value or behaviorally-relevant information content (Elliott et al., 2010). However, the general notion of segregated valence processing in OFC has contributed to theories of the functional underpinnings of disorders such as addiction (Crunelle, Veltman, Booij, Emmerik-van Oortmerssen, & van den Brink, 2012; Ma et al., 2010), anxiety disorder (Milad & Rauch, 2007), and depression (McCabe, Cowen, & Harmer, 2009; McCabe, Mishor, Cowen, & Harmer, 2010). Therefore, it is crucial to understand at the single neuron level whether distinct circuits mediate valence processing within the orbital cortex.

Studies of single OFC neurons thus far have not supported this theory of organization. While some OFC neurons respond more strongly to rewards and others to punishment, these appear to be anatomically intermingled (Morrison & Salzman, 2009). However, there are experimental differences that may account for this discrepancy. First, these neurons were recorded primarily from the central region of OFC (areas 11 and 13). One possibility is that valence-selective responses occur in the more medial (mOFC) and lateral (IC) orbital areas that were not evaluated by neurophysiology. To address this, the present study was designed to assess positive and negative valence encoding by single neurons across the full extent of the orbital surface.

An additional difference between neuroimaging and neurophysiology studies is that, with human subjects, experiments use either abstract reward and punishment, typically monetary gain and loss (Liu et al., 2007; O'Doherty et al., 2001) or rewards and punishment drawn from the same sensory modality such as pleasant and unpleasant smells (Gottfried, O'Doherty, & Dolan, 2002) or tastes (Small, Zatorre, Dagher, Evans, & Jones-Gotman, 2001; Zald, Hagen, & Pardo, 2002). In contrast, neurophysiology studies in monkeys have used rewards and punishments drawn from different modalities. The reward is usually fruit juice, but the punishment can be air puffs to the face (Morrison & Salzman, 2009), time-outs (Roesch & Olson, 2004) or electric shocks (Hosokawa, Kato, Inoue, & Mikami, 2007). Differences in sensory modality may influence neural responses and obscure patterns of valence encoding. To control for this, we trained two monkeys to perform a visuomotor association task for secondary reinforcement. Subjects learned that the length of a reward bar shown on their task screen corresponded to the amount of juice they would receive after completing a block of six trials. Once the subjects learned this association, actions could be rewarded by increasing the length of the bar, or punished by decreasing it. By using secondary reinforcement we ensured that the sensory properties of the rewards and punishments were matched in a way that is not possible using primary reinforcers. In addition, it draws closer parallels to the majority of human paradigms that motivate subjects through monetary gains and losses.

Methods

Subjects and behavioral task

We trained two male rhesus monkeys (*Macaca mulatta*), aged 10 and 6 years and weighing approximately 11.0 and 14.5 kg at the time of recording, on a visuomotor association task (Figure 1B). Subjects sat in a primate chair and viewed a computer screen. Affixed to the front of the chair was a joystick that could be displaced to the right or left with minimal force. Stimuli were presented on the computer screen and behavioral contingencies were implemented with MonkeyLogic software (Asaad & Eskandar, 2008). Eye movements were tracked with an infrared system (ISCAN). All procedures were in accord with the National Institute of Health guidelines and recommendations of the University of California at Berkeley Animal Care and Use Committee.

To begin each trial, subjects maintained gaze within a 1.3° radius of a central fixation spot for 650-ms. After fixation, one of four familiar stimuli appeared. Stimuli were images of natural scenes, approximately $2^\circ \times 3^\circ$ in size. Subjects responded by moving the joystick to the right or to the left. Responses within 150-ms of stimulus presentation were punished with a 5-s time-out. This contingency was meant to allow the subjects to respond as quickly as they wanted, but to discourage arbitrary responding. After a response, the subject received feedback in the form of either an increase or decrease in the length of the reward bar. Correct responses to positive pictures were rewarded with an increase in reward bar length (85% probability), and incorrect responses resulted in no change to the reward bar (no feedback). Correct responses to negative pictures resulted in no feedback, and incorrect responses were punished with a decrease in reward bar length (85% probability, Figure 1B). Therefore, each picture was associated with only positive or negative outcomes (gains or losses respectively). No other penalties (e.g. time-out or repeat trials) were imposed for incorrect responses. There was then a 1-s inter-trial interval (ITI). Subjects completed blocks of six trials, receiving feedback only through the secondary reinforcer for each trial, and the size of the reward bar carried over from trial to trial. At the end of each block, the reward bar was exchanged for a proportional amount of juice and reset to an initial starting size. Juice volume was varied by altering flow rate over a fixed time interval. By using a fixed-interval exchange schedule, primary reward (juice) was equally likely to follow any trial type, and could follow a correct or incorrect response. Therefore, obtaining juice could not reliably reinforce any picture, response or visuomotor association, and the task could only be learned through secondary reinforcement. Our task design also ensured that learning from positive and negative reinforcement was separate, since the absence of one type of reinforcement did not indirectly provide the opposite type of reinforcement (Bischoff-Grethe, Hazeltine, Bergren, Ivry, & Grafton, 2009). For example, in a standard reinforcement learning paradigm, the absence of reward following a response is an unambiguous signal not to repeat the response. In contrast, in our paradigm the absence of reward could arise on both correct and incorrect trials, and sometimes is the optimal outcome.

There is an additional asymmetry that naturally arises in learning from positive and negative feedback. In learning from positive feedback, better performance results in more frequent

reinforcement, whereas the opposite is the case in learning from negative feedback. Thus, to ensure that we had sufficient trials where the animal received negative feedback, we delivered it with variable probability, so that at least 15% of all negative picture trials resulted in negative feedback. Consequently, if performance was poor, all negative feedback would follow incorrect trials, but if performance was high, some correct trials could result in negative feedback. This also draws closer parallels to the positive pictures, where positive feedback was omitted on 15% of correct trials, so there was always a small probability of receiving a sub-optimal outcome by chance.

Behavioral analysis

We compared reaction times (RT) and accuracy for each subject when they executed leftward or rightward responses to positive or negative pictures with valence $\times$ response ANOVAs. Trials with no fixation or no joystick response were excluded.

We examined the effects of two additional variables on behavioral performance: the size of the reward bar and the number of trials completed within a block. We assessed 15 different models to determine how these two factors affected behavioral performance. All models included picture valence and response, and models 2–15 included one or two additional measures shown in Table 1. Because relationships between behavioral measures and bar size or number of completed trials appeared to asymptote at higher values (Figure 1D) and because the relationship between a sensory stimulus and its subjective perception is frequently logarithmic (Krueger, 1989), we included both linear and logarithmic values in our set of candidate models.

Two candidate models hypothesized that the number of trials remaining in a block might discount the value of the reward bar. Temporal discounting is a phenomenon observed in both humans and animals in which future rewards are judged to be of lower value than immediately available rewards. By varying the time to reward and reward amount, one can fit a simple function to behavior that describes the value depreciation as a function of time. Two discount functions, one hyperbolic and one exponential, have been successful in describing temporal discounting behavior (Frederick, Loewenstein, & O'Donoghue, 2002). Here, we hypothesized that subjects may discount the value of the reward bar as a function of the number of trials remaining in a block (i.e. the number of trials until he obtains juice). We tested this using a hyperbolic discount function: (discounted bar value) = bar length / (1+(γ K)), and an exponential discount function: (discounted bar value) = bar length * exp(- γ K). In both equations, K is the number of trials remaining in a block (i.e. time to juice), and γ is a free parameter that varies between 0 and 1 and indicates how steeply subjects discount the reward. We conducted an exhaustive search of γ 's ranging from 0 to 1, in .01 increments, and accepted the γ that minimized the deviance of a multiple linear regression in the case of RTs, or a logistic regression in the case of accuracy.

For each model, RTs were fit by multiple linear regression, and accuracy by logistic regression. Model fits were compared using Akaike's Information Criterion (AIC):

$$AIC=2k - 2\ln L \quad (1)$$

where k is the number of model parameters and L is the maximized likelihood. Weights for each model i were then computed as follows:

$$\Delta AIC_i = AIC_i - AIC_{\min} \quad (2)$$

$$w_i = e^{\frac{1}{2} \times \Delta AIC_i} \quad (3)$$

$$weight_i = \frac{w_i}{\sum w} \quad (4)$$

These weights represent the relative probability of a model compared to the other candidate models (Anderson, 2008).

Neurophysiological recording

We used standard methods for acute neurophysiological recording that have been described in detail elsewhere (Lara, Kennerley, & Wallis, 2009). Briefly, each subject was surgically implanted with two recording chambers and a titanium head post to maintain head position during recording. Chamber positions were calculated based on images obtained from 1.5-T magnetic resonance imaging (MRI) scans of each subject's brain (Figure 1A). In each subject, one chamber was centered over lateral orbital regions, allowing access to IC and OFC. In the opposite cerebral hemisphere, the other chamber was centered medially, allowing access to mOFC and OFC. For the two subjects, chamber placement was counterbalanced across hemispheres.

We recorded simultaneously from 4 – 20 tungsten microelectrodes (FHC Instruments) distributed across brain areas. Each recording day, electrodes were lowered manually with custom-built microdrives to a target depth. Depths were calculated from MRI images and confirmed by mapping gray and white matter boundaries. Once electrodes entered the target area, fine adjustments were made to isolate waveforms from single neurons. We recorded all well-isolated neurons in a target area, resulting in a random sample of neurons. Waveforms were digitized by an acquisition system (Plexon Instruments) and saved for off-line analysis.

Neurophysiological analysis

For each neuron, we first calculated its firing rate during three trial epochs: fixation, sample and feedback. The fixation epoch was 650-ms at the beginning of each trial, when subjects were required to fixate a central point. If a subject broke then re-initiated fixation, only the final 650-ms were included. No stimuli were present on the screen except a fixation point and the reward bar. The sample epoch was 400-ms immediately preceding the subject's response. This epoch was time-locked to responses because the two subjects had different RTs, suggesting that they processed the stimuli at different rates. Initial assessment of neural activity suggested that encoding of stimulus information appeared later in the subject with longer RTs. Therefore, to best capture neural encoding related to the processing of the stimuli, sample epochs were time-locked to responses. The feedback epoch was 600-ms beginning 100-ms after the presentation of feedback.

We used regression models to determine which aspects of the task each neuron encoded. Because some potentially encoded variables were correlated with each other, we performed stepwise regression. Using a stepwise approach, the regressor that significantly predicted firing rate and explained the most variance was included in the model first. Additional regressors were included if they significantly improved the fit of the model. The statistical criterion for entering a predictor in a model was $p = 0.01$. Predictor variables included in each analysis are shown in Table 2. For fixation and sample epochs, error trials were excluded. Feedback epochs included all completed trials regardless of accuracy, so that different types of feedback could be analyzed. Mean firing rates were standardized to allow comparison of beta coefficients across neurons. Because behavioral data were better fit by logarithmically transforming the bar size and trial number (see Results), these two regressors were logarithmically transformed in the regressions.

We determined the proportion of neurons for which each predictor was included in the final model. Because fewer than 5% of neurons encoded information about the preceding trial, these variables were not considered further. Neurons encoding picture valence were subdivided into those responding to positive or negative pictures based on whether the beta coefficient for the picture valence regressor was positive (i.e. firing rates were higher when positive pictures were shown) or negative respectively. The same approach was used to divide neurons encoding feedback as responding to positive events (positive feedback, lack of negative feedback or both) versus negative events (negative feedback, lack of positive feedback or both). Chi-squared tests compared proportions of neurons encoding different variables.

To assess encoding strength, we calculated the coefficient of partial determination (CPD), which quantifies the amount of variance in a neuron's firing rate attributed to a specific predictor variable in a multiple regression model. During the fixation and sample epochs, we calculated CPDs from a multiple regression with the following predictors explaining firing rates: picture valence, response direction, bar size and trial number. A similar approach could not be taken for the feedback epoch because correlated predictors (e.g. picture valence and feedback valence) caused multicollinearity in a multiple regression.

Latency to encode picture valence or response direction was determined with sliding multiple regressions. We regressed valence and response on neural firing rates averaged over a 200-ms window. The window started 500-ms before stimulus onset, then was stepped forward by 10-ms. This process was repeated through 700-ms post-stimulus. Significant encoding was defined as $p = 0.002$ for four consecutive time windows, and latency as the first of these windows. To establish this criterion, we ran the same sliding regression on the fixation epoch, before the picture appeared on each trial and selected a criterion that resulted in false discovery rates $\leq 1\%$ for both picture valence and response direction.

We also wished to assess neural responses when the reward bar is exchanged for juice. At this time, the bar is the only stimulus on the screen, however its length directly corresponded to the amount of juice delivered. Despite this confound, we reasoned that neurons specifically responding to juice reward would be selective at the time of juice delivery, but not during the immediately preceding feedback epoch, when only the reward bar is present.

To make this comparison, we analyzed only the 6th trial of each block in two 600ms epochs: feedback and juice delivery. Each epoch began 100ms after the relevant event, and mean firing rates were regressed against reward bar size, which was equivalent to juice volume. Beta coefficients with $p \leq .01$ were considered significant.

Anatomical demarcation of orbital regions and construction of flattened cortical maps

We grossly divided the orbital surface into three regions, mOFC, OFC and IC, which correspond approximately to area 14 (mOFC), areas 11 and 13 (OFC) and area 47/12 (IC) (Figure 1A). To ensure that our division of the region did not obscure true patterns of neural coding by misplacing a boundary, we also analyzed neuronal data with respect to the neurons' medial-lateral position on the orbital surface regardless of region, and present flattened maps for visualization. To construct these maps, the following landmarks were identified and measured on coronal MRI images at 1-mm AP intervals: the medial and lateral convexities (where the orbital surface bends around onto the medial and lateral frontal lobe surfaces respectively) and the medial and lateral orbital sulci (MOS and LOS respectively). All measurements were taken relative to the LOS, and the flatmap location of each orbital neuron was measured relative to the LOS.

Results

Behavior

The two subjects completed a mean ($\pm$ SEM) of 567 ($\pm$ 17) and 616 ($\pm$ 16) trials per session. For both subjects, accuracy was higher when responding to positive pictures compared to negative pictures (Figure 1C, Subject C: $F_{(1,140)} = 130$, $p < 5 \times 10^{-22}$, Subject M: $F_{(1,124)} = 59$, $p < 5 \times 10^{-12}$), and reaction times (RT) were faster for positive compared to negative pictures (Subject C: $F_{(1,140)} = 120$, $p < 1 \times 10^{-20}$, Subject M: $F_{(1,124)} = 34$, $p < 5 \times 10^{-8}$).

Behavior was also influenced by two other sources of information. Performance became faster and more accurate as the size of the reward bar increased, indicating that a greater amount of juice had been earned. Performance also improved as the number of trials completed within a block increased, indicating that the subject was closer to exchanging the reward bar for juice (Figure 1D). Although bar size and trial number were weakly correlated, they were sufficiently uncorrelated that we could determine their independent effects on behavior (variance inflation factor = 1.33 and 1.39 for subjects C and M).

To determine precisely how these factors affected behavior, we assessed the ability of different models to predict either RT or accuracy, and compared model fits using Akaike's Information Criterion (AIC). Two models included temporally discounted values of the reward bar, which were derived by optimizing the fit of hyperbolic or exponential discount functions. Fits were separately optimized for RT and accuracy for each subject, yielding 8 different values of γ (2 subjects $\times$ 2 discount functions $\times$ 2 behavioral measures). If $\gamma = 0$, there is no discounting and no change in the bar value over time. As γ increases, the discount curve becomes increasingly steep (i.e. there is more value depreciation with time). For subject M, all models failed to optimize within the range 0 – 1, and the best fit was found when $\gamma = 0$, meaning that there was no behavioral evidence that M discounted the value of

the reward bar over time. In the case of subject C, the functions did optimize at values close to 0, suggesting that C may have discounted the bar value slightly (γ values: hyperbolic RT 0.13, accuracy 0.09; exponential RT 0.12, accuracy 0.09). However, compared to other simpler models of behavior, the temporal discounting models predicted behavior poorly (Figure 1E). This is perhaps not surprising given that the optimized γ was close to or equal to 0, indicating that very little temporal discounting was occurring. This result indicated that the reward bar did, indeed, act as a secondary reinforcer. If the bar simply represented an amount of juice to be received some time in the future, its value should increase as the time to exchange decreased. However, the bar value stayed constant across trials, suggesting that subjects learned to value the bar itself independent of the juice it predicted.

The best-fitting model for both accuracy and RT for both subjects included the logarithm of the reward bar size and the logarithm of the number of trials as independent factors (model 12). To quantify the weight of evidence in favor of the best-fit model, we calculated AIC weights, which give the probability that the model is the best fit among the set of candidate models (Figure 1E). For subject C, the evidence overwhelmingly favored model 12 (AIC weight > 0.95 for both accuracy and RT). For subject M, model 12 was also the best fit (AIC weight: accuracy = 0.58, RT = 0.57), but there were reasonable fits from models that either included the number of trials completed as a linear relationship (model 8: accuracy = 0.20, RT = 0.23) or omitted it entirely (model 3: accuracy = 0.22, RT = 0.20). This slight variability suggests that while subject M's behavior was affected by the number of trials completed, the effect was smaller than it was in subject C, and smaller than the effect of bar size.

Overall, behavioral analysis showed that subjects responded to two sources of value information beyond the valence of the picture presented – the size of the reward bar and the number of trials completed within a block – and that these two values were tracked independent of one another.

Neurophysiology

We recorded from a total of 648 neurons throughout orbital cortex (316 subject C, 332 subject M), and divided them into three regions based on their anatomical location (Figure 1A)(Petrides & Pandya, 1994). Neurons recorded medial to the depth of the medial orbital sulcus (MOS) were grouped as mOFC. This included neurons located primarily in area 14 and the ventral part of the medial wall. Neurons located between the depth of the medial and lateral orbital sulcus were grouped as OFC, and comprised neurons in area 11 and 13. Neurons lateral to the depth of the lateral orbital sulcus (LOS) were grouped as those in the inferior convexity (IC). The majority were in area 47/12, but some were also in area 45. We excluded neurons with a mean firing rate <1-Hz, since a low number of action potentials prevent us from statistically characterizing neural encoding. The final sample size for each area was as follows: mOFC 129, OFC 192, IC 226.

Encoding of picture valence and response direction

To quantify how neurons encoded task-related information, we first analyzed neural activity during the sample presentation. At this time, the two most commonly encoded variables

were the picture valence and corresponding motor response (Figure 2A–D). Within each brain area, similar proportions of neurons encoded valence and response direction (Figure 2E *top panel*, all $\chi^2 < 2.6$, $p > 0.1$). Among OFC neurons, picture valence accounted for more variance as measured by CPDs (Figure 2E *bottom panel*, Wilcoxon rank-sum test, $p < .05$). In other areas there were no differences between encoding strength for picture valence compared to response direction (all comparisons $p > 0.1$).

In the neurons in Figure 2A–D, valence selectivity appeared earlier than response selectivity, even though both types of information were conveyed by a single stimulus. Therefore, we compared latencies to encode different types of information with 3-way ANOVA of variable encoded $\times$ brain area $\times$ subject. Overall, response encoding appeared at longer latencies relative to picture valence encoding (median latency 201-ms valence; 321-ms response; main effect of valence vs. response $F_{(1,406)} = 17$, $p < 0.001$). Latencies also differed among brain areas ($F_{(2,406)} = 4.5$, $p = .01$), with IC encoding information significantly earlier than OFC (Figure 2E, pairwise comparisons $p < .01$, Bonferroni corrected). mOFC latencies did not significantly differ from OFC or IC ($p > .05$), though the small number of mOFC neurons encoding either variable resulted in relatively low power to detect latency differences.

Valence coding in orbital cortex

Though many neurons encoded the valence of a picture, no orbital area responded predominantly to positive or negative pictures during the sample epoch (Figure 3A *top panel*, all $\chi^2 < 3.5$, $p > 0.05$). It is possible for different brain areas to have similar numbers of selective neurons, but differ in the strength of their encoding. If a neuron strongly encodes a particular experimental parameter, then that parameter should account for a large amount of variance in the neuron's firing rate. Therefore, we sorted neurons according to whether they had higher firing rates for positive or negative pictures and quantified how much variance picture valence accounted for in each group by calculating CPDs. Using this approach, no area sampled showed significant differences between encoding strength of positive and negative pictures (Figure 3A *bottom panel*, Wilcoxon rank-sum tests, all $p > 0.1$). We also looked for differences in latency to encode picture valence among neurons that responded to positive versus negative pictures. A 3-way ANOVA of valence $\times$ brain area $\times$ subject revealed a significant effect of valence (Figure 3B, $F_{(1,128)} = 6.6$, $p = .01$), with shorter latencies to respond to positive pictures (median 171ms) versus negative pictures (median 241ms). However this effect did not differ across brain areas. All other main effects and interactions were not significant (all $p > 0.05$).

To ensure that our division of anatomical areas did not misplace a functional boundary, we plotted the anatomical location of each neuron encoding either positive or negative pictures. Here, we observed more valence selective neurons lateral to the MOS. However, neurons responding to positive pictures appeared to be randomly intermingled with those responding to negative pictures (Figure 3C). There was no significant difference in the median medial-lateral position of neurons coding positive and negative pictures (Figure 3C *inset*, Wilcoxon rank-sum test, $p > 0.1$).

Feedback encoding in orbital cortex

After feedback, there are several potential ways valence could be encoded. For example, the neuron in Figure 4A showed a slight anticipatory effect prior to feedback onset, and subsequently responded only when positive feedback was obtained. In contrast, the neuron in Figure 4B responded only to negative feedback. The neuron in Figure 3C encoded the entire value scale, showing its maximal response to positive feedback, a smaller response to neutral outcomes and minimal response to a negative outcome. Other neurons encoded the outcomes relative to the subjects' expectations. For example, the neuron in Figure 4D increased its firing rate when the subject saw a positive picture but received no feedback. Since this neuron did not respond to no feedback that followed a negative picture, the response must be coding the absence of an expected reward. While this pattern is consistent with prediction error coding, the task did not systematically vary magnitudes of outcomes and expectations in order to definitely identify prediction errors. Of note, no orbital neurons were found that responded to omitted reward and punishment in opposite directions, as expected from a fully signed prediction error signal. Therefore, we refer to responses such as 4D as coding omitted reward. Similarly, the neuron in Figure 4E encoded omitted punishment. Finally, Figure 4F shows a neuron encoding a relative value signal. It responded to whichever was the worse potential outcome (a loss in the case of negative pictures and a neutral outcome in the case of positive pictures). Table 2 summarizes the complete list of parameters tested.

Across orbital cortex, over 30% of neurons responded to feedback, but there were no clear differences in the predominant type of coding between areas (Figure 4G). Therefore, in order to evaluate valence coding, we pooled feedback responses, distinguishing neurons that increased activity when outcomes were positive or negative. For example, neurons that responded to positive feedback alone or on a continuous value scale, or responded to the better relative value of feedback, were grouped as responding to positive outcomes. In all areas, there were approximately equal proportions of neurons responding to positive and negative outcomes (Figure 4H) (IC: $\chi^2 < 2$, $p > 0.1$, OFC: $\chi^2 < 1$, $p > 0.5$, mOFC: $\chi^2 < 2$, $p > 1$).

In addition to valence, we tested whether neurons were more or less likely to encode unexpected outcomes, in other words, those that were experienced less often. Since most trials were performed correctly, subjects typically received positive feedback for responding to positive pictures and no feedback for negative pictures. Therefore these results were grouped as expected outcomes, and unexpected outcomes included no feedback for positive pictures and negative feedback for negative pictures. Note that expected outcomes are also preferred relative to unexpected outcomes, but this analysis differed from the previous analysis, which tested whether neurons were activated by a particular valence. Here, we tested whether neurons were selective for a given outcome, either by activation or suppression. While IC neurons encoded expected and unexpected outcomes equally ($\chi^2 < 1$, $p > .5$), OFC and mOFC neurons tended to encode expected outcomes more than those that were unexpected (Figure 4I). This difference reached significance in OFC ($\chi^2 = 5.9$, $p = .015$) but not mOFC ($\chi^2 = 3.1$, $p = .079$).

In summary, we looked for differences in valence coding across the entire orbital surface using a variety of neural measures during both the sample and feedback epochs, and found little evidence for cortical organization based on valence. Instead, we found a tendency for OFC neurons to encode expected over unexpected outcomes.

Consistent valence selectivity for pictures and feedback

Theories of OFC function argue that neurons encode outcomes predicted by sensory stimuli (Schoenbaum & Roesch, 2005; Takahashi et al., 2011). If this is the case, we might expect neurons selective for picture valence to also encode feedback of the same valence. In other words, if a neuron's response to a picture reflects its prediction of the likely outcome, we would expect its response to the picture to have the same valence as the response to the outcome. To assess this, we again considered all types of valence-related feedback encoding together. If neurons coded picture and feedback valence independently, by chance a certain percentage of neurons would be observed with the same valence selectivity during both epochs. The proportion expected by chance is the proportion of neurons selective for pictures multiplied by the proportion selective for feedback. The number of neurons expected to show temporal consistency (i.e. the same valence selectivity during both the sample and feedback epochs) was compared to the actual number of consistent neurons using binomial tests. Similarly, we compared the number of neurons observed and expected by chance to have inconsistent valence coding (e.g. responding to positive pictures and negative outcomes).

Across all areas, 28% of neurons encoding picture valence maintained consistent valence encoding at the time of feedback, and this was more than expected by chance (binomial test, $p < 3 \times 10^{-5}$). When examined on an area-by-area basis, this effect was driven primarily by OFC, which showed more consistency than expected by chance (Figure 5A, binomial test, consistent encoding of positive valences $p = 0.0005$, negative $p < 0.002$). In all other areas, the prevalence of consistent valence coding was not significantly different from chance (all $p > 0.07$). In contrast, only 6.5% of neurons had inconsistent valence coding, not significantly different from chance ($p > 0.05$), and no individual area differed from chance levels of inconsistent valence encoding (Figure 5B, all $p > 0.1$). Thus, only in OFC did the valence of picture encoding match the valence of outcome encoding, consistent with the notion that this area is responsible for encoding the outcome predicted by a stimulus.

Orbital neurons code additional sources of value information

Our behavior analysis showed that subjects were affected by both the size of the reward bar and the number of trials until exchange of the secondary reinforcer, and that these variables had independent effects. Therefore, we hypothesized that separate populations of neurons would respond to each variable. We focused on the fixation epoch, since it provides a relatively clean measure of these variables uncontaminated by sensory stimuli or motor responses. Figure 6A illustrates a neuron that encoded the size of the reward bar but did not differentiate the number of trials until the bar could be exchanged for juice. In contrast, the neuron in Figure 6B showed the opposite response. It did not encode the size of the reward bar, but it tracked the number of trials completed, with its activity becoming progressively higher as the subject got closer to obtaining juice.

Consistent with the independent effects on behavior, most neurons encoded either reward bar size or trial number: very few encoded both (2.7% during fixation). Consequently, in a multiple regression, neurons either had high beta coefficients for bar size, but low betas for trial number, or vice versa, but few neurons lay along the diagonal (Figure 6C). However, not every area represented these values equally. Significantly more OFC neurons encoded the number of trials completed relative to the value of the reward bar (Figure 6D, OFC: $\chi^2 = 12$, $p = 0.0005$). In contrast, more IC neurons encoded bar size than trial number ($\chi^2 = 3.9$, $p < 0.05$). There were no differences between the relative proportions of neurons coding the reward bar and trial number in mOFC (all $\chi^2 < 2.8$, $p > 0.05$). We found neurons whose activity increased with increasing bar size or trial number, and others whose activity decreased. There was a trend among IC neurons toward encoding trial number inversely (i.e. with a negative beta coefficient) ($\chi^2 = 4.2$, $p = 0.04$), but all other comparisons were not significant (all $\chi^2 < 2$, $p > 0.2$). Across all neurons, OFC also had stronger encoding of trial number compared to bar size, as measured by CPDs (Figure 6D *insets*, Wilcoxon rank-sum tests: $p < 0.005$, all other areas $p > 0.1$), and encoded trial number more than bar size during the sample epoch ($\chi^2 = 8.3$, $p < 0.005$). There were no differences between bar size and trial number in other areas (all $\chi^2 < 1$, $p > 0.1$).

More fine-grained anatomical analysis showed that a clear lateral-medial organization was apparent among orbital neurons encoding bar size or trial number (Figure 6E). Neurons lateral to the LOS encoded value information associated with the reward bar as well as trial number, but encoding of bar size decreased significantly among neurons medial to the LOS. We confirmed this by comparing the lateral-medial position of the neurons encoding the bar size or the number of completed trials. Neurons encoding the bar size were located significantly more laterally than neurons encoding the number of completed trials (Wilcoxon rank-sum test, $p = .004$, Figure 6E *inset*).

Because our behavior analysis suggested that subject M's behavior was somewhat less influenced by trial number than subject C's, we assessed whether trial number encoding was less frequent in subject M. However, the opposite was true. Encoding of trial number was slightly more common in subject M (20% of neurons) than subject C (18% of neurons).

These variables, bar size and trial number, represent two potential sources of value information. The reward bar's value was indicated by a specific visual feature, its length, whereas trial number was tracked internally and was not indicated by any external cues. To ensure that IC encoding of bar size did not constitute simple visual responses, we analyzed neural activity during the ITI that immediately preceded the fixation epoch. During this time, the reward bar was visible and subjects were allowed to freely view the screen, however they were not engaged in the task. During the ITI, fewer IC neurons encoded bar size (18% ITI vs. 28% fixation), and only 4.4% (10 neurons) encoded bar size during both epochs, even though the bar was present continuously. Thus, IC neuronal responses to the bar were not simply driven by the visual sensory input.

Neural responses during secondary reinforcer exchange

In order to determine whether there were populations of neurons activated in response to juice delivery, we compared neural coding of reward bar size, which directly correlated with

juice amounts, between epochs of juice delivery and the feedback epoch immediately preceding juice delivery. If a neuron encoded juice and not bar value, we would expect it to show selectivity during juice delivery but not feedback, when only the reward bar is present. Instead, we found evidence for the opposite pattern. Among IC neurons, the strength of bar size encoding decreased significantly when juice was delivered, compared to the preceding feedback epoch (Figure 7, Wilcoxon rank-sum test of beta coefficients, $p < 8 \times 10^{-5}$). The proportion of neurons with significant encoding of bar size also declined in IC (Figure 7 insets, $\chi^2 = 5.8$, $p = .016$). Although mOFC neurons appeared to show the opposite pattern, the coding differences did not reach significance (rank-sum test $p = .45$, $\chi^2 = 2.7$, $p = .099$). OFC neurons also showed no difference (rank-sum test $p = .18$, $\chi^2 = 0.4$, $p = .55$).

Discussion

Across species, the orbitofrontal cortex plays a crucial role in evaluating relationships between stimuli and predicted outcomes (Murray, O'Doherty, & Schoenbaum, 2007), but its functional organization remains unclear. Here, we investigated a theory of valence selective processing (Kringelbach & Rolls, 2004) by directly recording neural activity, and found no consistent selectivity for either valence across the orbital cortex. Many neurons encoded valence, but these were intermingled anatomically. Instead, we found distinctions among orbital areas in coding stimuli and feedback of the same valence, and coding the reward bar or trial number within a block. These results suggest that orbital cortex is not organized according to the valence of information being processed, but rather by different types of value computations being performed by the different subregions.

Positive and negative valence processing

Positive and negative outcomes can be operationally identified as those that an animal will work to obtain or avoid (Seymour, Singer, & Dolan, 2007). Here, subjects learned arbitrary stimulus-response mappings in order to obtain or avoid losing secondary reinforcers. In order to learn a correct response to a positive picture, subjects must have been motivated only to obtain secondary reinforcement, since there was never a potential for loss on these trials. Likewise, on negative picture trials subjects were motivated to avoid a loss, since these pictures were never associated with gains. While it is true that completing a negative picture trial advanced the subject within a block, the same result could be obtained without learning the stimulus-response mapping and executing an arbitrary response, since no penalties other than loss of a secondary reinforcer were imposed. Therefore, loss of a secondary reinforcer was an aversive outcome that they learned to avoid. Similar results have been obtained using paradigms studying competitive games. Gain and loss of secondary reinforcers had approximately equal and opposite effects on monkeys' choices in a mixed-strategy game, such that gains were rewarding and losses were aversive (Seo & Lee, 2009).

Despite this, we found no evidence that mOFC preferentially responded to rewards or that IC responded to punishment. This discrepancy between our results and the theory of valence selectivity in orbital cortex could have a number of causes. The theory is based on imaging studies in humans (primarily fMRI), while our data consist of single neuron activity recorded in monkeys. It is impossible to rule out species differences in function or anatomy.

However, current evidence suggests remarkable homology between human and monkey OFC (Jbabdi, Lehman, Haber, & Behrens, 2013; Mackey & Petrides, 2010; Wallis, 2012). In contrast, the relationship between the fMRI BOLD signal and spiking activity is less straightforward (Goense & Logothetis, 2008; Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001; Sirotin & Das, 2009). In addition, recent neuroimaging studies have cast doubt on a valence-based organization (Elliott et al., 2010). For example, mOFC BOLD correlates with both appetitive and aversive food values (Plassmann, O'Doherty, & Rangel, 2010) or monetary wins and losses (Tom, Fox, Trepel, & Poldrack, 2007). Finally, one neurophysiology study did find separate groups of neurons responding to appetitive and aversive stimuli, but they were in close proximity to one another within the most posterior regions of mOFC and ventromedial PFC (Monosov & Hikosaka, 2012). We did not find similar distinctions, however our recordings were anterior to these sites. Overall, our results confirm emerging data suggesting that orbital subregions are not organized according to the valence of information they process.

Nonetheless, there is reason to believe that different neural circuits should be involved in evaluating appetitive and aversive stimuli. In particular, these stimuli have opposite effects on behavior. Rewarding stimuli promote approach responses; punishing stimuli encourage avoidance (Huys et al., 2011). This is illustrated by classic studies of rats, who easily learn to press a lever for reward, but have difficulty learning to lever press to avoid a footshock (Bolles, 1970). In contrast, they readily learn to run or jump to a safe location to escape the same shock. Thus, learning based on negative stimuli is hampered when the necessary response is approach (lever press), and facilitated when the necessary response is avoidance (run away, hide, withhold responding), suggesting that dissociable circuits in the brain are prepared to learn certain responses to oppositely valenced stimuli (also see Hershberger, 1986)). It stands to reason that this valence separation could occur in higher brain areas. While this was not the case in orbital cortex, neurons in a region of subgenual anterior cingulate cortex appear to respond selectively to negative values (Amemori & Graybiel, 2012). Such neurons may play a role in enabling different behavioral “modules” (Amemori, Gibb, & Graybiel, 2011), such as avoidance responses, through connections with the striatum (Eblen & Graybiel, 1995). In contrast, orbital areas project to the ventral striatum (Haber, Kunishio, Mizobuchi, & Lynd-Balta, 1995; Selemon & Goldman-Rakic, 1985), important in value-based learning and decision-making. In this circuit, hard-wired valence-specific responses may be undesirable. While valence-specific motor responses can safeguard an organism from approaching potential harm, decision-making is more nuanced and often involves accepting an undesirable cost to obtain a desired benefit.

Functional differences between orbitofrontal areas

Though our results did not reveal valence-specific processing in orbital regions, we did observe interesting area differences. Few mOFC neurons encoded any variables assessed, suggesting that it was relatively less engaged by the task. The precise functions of mOFC are unclear, but recent data point to a role in comparing option values (Noonan et al., 2010), or coding internal motivational values, particularly in the absence of external prompts (Bouret & Richmond, 2010). In either case, mOFC may not have been engaged by the

present task, since subjects were not presented with stimulus choices and all responses were cued.

Neurons in both OFC and IC were active during the task, and showed interesting distinctions. First, OFC tracked progress through trial blocks but did not encode the size of the reward bar; only IC showed appreciable encoding of bar size. Second, single OFC neurons encoded pictures and feedback with the same valence, and at the time of feedback they lost selectivity when an unexpected outcome occurred. In the present task, whether an outcome is unexpected is confounded with whether it is the less preferred outcome, and further studies are needed to resolve this issue. However, a previous study that found that rat OFC neurons show little or no response to unexpected rewards as well as unexpected omission of reward (Takahashi et al., 2009), supporting the view that it is the degree to which an outcome is expected or not that accounts for our results.

We interpret the overall pattern of encoding in OFC as follows. Although OFC is associated with reward processing, reward-related responses can be heavily dependent on the task in which the subject is engaged (Luk & Wallis, 2013). Such observations support the idea that OFC uses knowledge about the task structure and environment to make outcome predictions (Jones et al., 2012; Takahashi et al., 2011). From this view, OFC neurons with consistent valence encoding may play a role in predicting feedback based on sample pictures. Selectivity at feedback time represents the realization of these predictions, and when they are not met selective coding diminishes. Finally, OFC tracking of trial number may reflect knowledge of the task structure or predictions about when primary reward will be obtained.

However, this raises the question as to why OFC neurons showed only weak encoding of the reward bar, since it predicts the amount of juice to be delivered. One explanation compatible with this account is that the bar is interpreted as an outcome, rather than as an outcome-predictive cue. Supporting this, our behavioral analysis showed that the effect of the bar size on subjects' behavior was independent of how close they were to exchanging it for juice. If the bar were treated as a prediction of juice, one would expect that it should be temporally discounted like other reward-predicting stimuli. A related interpretation is that, although the bar reinforced subjects' behavior, its value (how much juice it predicted) was independent of the task at hand. Supporting this, the value of the bar was tracked preferentially by IC neurons, and it was precisely in this population that we observed a significant loss of value-related encoding during juice delivery. We believe that this happened because the amount of juice was fully predicted by the reward bar, and it was the bar, not the delivery of juice that reinforced specific behaviors in the task.

In contrast to OFC, IC neurons strongly encoded the reward bar and showed no evidence of consistent coding between sample and feedback epochs. These observations suggest that IC neurons do not perform the same functions as OFC. IC receives highly processed sensory information from temporal and parietal cortex (Carmichael & Price, 1995; Cavada & Goldman-Rakic, 1989; Petrides & Pandya, 2002), and lesions impair the use of visual information to guide motor responses and behavioral strategies (Baxter, Gaffan, Kyriazis, & Mitchell, 2009; Bussey, Wise, & Murray, 2001, 2002). IC may play a role in attention processes (Kennerley & Wallis, 2009) or determining the behavioral significance of stimuli

(Rushworth et al., 2005). As such, IC neurons may assign meaning to stimuli such as the reward bar.

Conclusion

In contrast to a commonly held notion that medial and lateral orbital areas process positive and negative stimuli, at the single neuron level, we found no evidence that orbital processing is organized with respect to valence. Instead, we report differences between distinct orbital regions in how they represent aspects of the task that support the view that different areas use value information in markedly different ways.

Acknowledgments

The project was funded by NIDA grant R01 DA19028 and NINDS grant P01 NS040813 to J.D.W. and a grant from the Hilda and Preston Davis Foundation to E.L.R. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

- Amemori K, Gibb LG, Graybiel AM. Shifting responsibly: the importance of striatal modularity to reinforcement learning in uncertain environments. *Front Hum Neurosci.* 2011; 5:47. [PubMed: 21660099]
- Amemori K, Graybiel AM. Localized microstimulation of primate pregenual cingulate cortex induces negative decision-making. *Nat Neurosci.* 2012; 15(5):776–785. [PubMed: 22484571]
- Anderson, DR. *Model Based Inference in the Life Sciences: A Primer on Evidence.* New York, New York: Springer Science+Business Media LLC; 2008.
- Asaad WF, Eskandar EN. A flexible software tool for temporally-precise behavioral control in Matlab. *Journal of neuroscience methods.* 2008; 174(2):245–258. [PubMed: 18706928]
- Baxter MG, Gaffan D, Kyriazis DA, Mitchell AS. Ventrolateral prefrontal cortex is required for performance of a strategy implementation task but not reinforcer devaluation effects in rhesus monkeys. *European Journal of Neuroscience.* 2009; 29(10):2049–2059. [PubMed: 19453635]
- Bischoff-Grethe A, Hazeltine E, Berggren L, Ivry RB, Grafton ST. The influence of feedback valence in associative learning. *NeuroImage.* 2009; 44(1):243–251. [PubMed: 18834944]
- Bolles RC. Species-specific defense reactions and avoidance learning. *Psychological Review.* 1970; 77(1):32–48.
- Bouret S, Richmond BJ. Ventromedial and orbital prefrontal neurons differentially encode internally and externally driven motivational values in monkeys. *J Neurosci.* 2010; 30(25):8591–8601. [PubMed: 20573905]
- Bussey TJ, Wise SP, Murray EA. The role of ventral and orbital prefrontal cortex in conditional visuomotor learning and strategy use in rhesus monkeys (*Macaca mulatta*). *Behav Neurosci.* 2001; 115(5):971–982. [PubMed: 11584930]
- Bussey TJ, Wise SP, Murray EA. Interaction of ventral and orbital prefrontal cortex with inferotemporal cortex in conditional visuomotor learning. *Behav Neurosci.* 2002; 116(4):703–715. [PubMed: 12148937]
- Carmichael ST, Price JL. Architectonic subdivision of the orbital and medial prefrontal cortex in the macaque monkey. *J Comp Neurol.* 1994; 346(3):366–402. [PubMed: 7527805]
- Carmichael ST, Price JL. Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. *Journal of Comparative Neurology.* 1995; 363:642–664. [PubMed: 8847422]
- Cavada C, Goldman-Rakic PS. Posterior parietal cortex in rhesus monkey: II. Evidence for segregated corticocortical networks linking sensory and limbic areas with the frontal lobe. *J Comp Neurol.* 1989; 287(4):422–445. [PubMed: 2477406]

- Chib VS, Rangel A, Shimojo S, O'Doherty JP. Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *J Neurosci*. 2009; 29(39): 12315–12320. [PubMed: 19793990]
- Crunelle CL, Veltman DJ, Booij J, Emmerik-van Oortmerssen K, van den Brink W. Substrates of neuropsychological functioning in stimulant dependence: a review of functional neuroimaging research. *Brain Behav*. 2012; 2(4):499–523. [PubMed: 22950052]
- Eblen F, Graybiel AM. Highly restricted origin of prefrontal cortical inputs to striosomes in the macaque monkey. *J Neurosci*. 1995; 15(9):5999–6013. [PubMed: 7666184]
- Elliott R, Agnew Z, Deakin JF. Hedonic and informational functions of the human orbitofrontal cortex. *Cereb Cortex*. 2010; 20(1):198–204. [PubMed: 19435707]
- Frederick S, Loewenstein G, O'Donoghue T. Time discounting and time preference: a critical review. *J Econ Lit*. 2002; 40:50.
- Fujiwara J, Tobler PN, Taira M, Iijima T, Tsutsui K. Personality-dependent dissociation of absolute and relative loss processing in orbitofrontal cortex. *Eur J Neurosci*. 2008; 27(6):1547–1552. [PubMed: 18336572]
- Gläscher J, Adolphs R, Damasio H, Bechara A, Rudrauf D, Calamia M, et al. Lesion mapping of cognitive control and value-based decision making in the prefrontal cortex. *Proc Natl Acad Sci U S A*. 2012; 109(36):14681–14686. [PubMed: 22908286]
- Goense JB, Logothetis NK. Neurophysiology of the BOLD fMRI signal in awake monkeys. *Current biology : CB*. 2008; 18(9):631–640. [PubMed: 18439825]
- Gottfried JA, O'Doherty J, Dolan RJ. Appetitive and aversive olfactory learning in humans studied using event-related functional magnetic resonance imaging. *J Neurosci*. 2002; 22(24):10829–10837. [PubMed: 12486176]
- Haber SN, Kunishio K, Mizobuchi M, Lynd-Balta E. The orbital and medial prefrontal circuit through the primate basal ganglia. *J Neurosci*. 1995; 15(7 Pt 1):4851–4867. [PubMed: 7623116]
- Hayes DJ, Northoff G. Common brain activations for painful and non-painful aversive stimuli. *BMC Neurosci*. 2012; 13:60. [PubMed: 22676259]
- Hershberger W. An approach through the looking-glass. *Animal Learning & Behavior*. 1986; 14(4): 443–451.
- Hosokawa T, Kato K, Inoue M, Mikami A. Neurons in the macaque orbitofrontal cortex code relative preference of both rewarding and aversive outcomes. *Neuroscience research*. 2007; 57(3):434–445. [PubMed: 17239463]
- Huys QJ, Cools R, Gölzer M, Friedel E, Heinz A, Dolan RJ, et al. Disentangling the roles of approach, activation and valence in instrumental and pavlovian responding. *PLoS Comput Biol*. 2011; 7(4):e1002028. [PubMed: 21556131]
- Jbabdi S, Lehman JF, Haber SN, Behrens TE. Human and Monkey Ventral Prefrontal Fibers Use the Same Organizational Principles to Reach Their Targets: Tracing versus Tractography. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 2013; 33(7):3190–3201. [PubMed: 23407972]
- Jones JL, Esber GR, McDannald MA, Gruber AJ, Hernandez A, Mirenzi A, et al. Orbitofrontal cortex supports behavior and learning using inferred but not cached values. *Science*. 2012; 338(6109): 953–956. [PubMed: 23162000]
- Kennerley SW, Wallis JD. Reward-Dependent Modulation of Working Memory in Lateral Prefrontal Cortex. *Journal of Neuroscience*. 2009; 29(10):3259–3270. [PubMed: 19279263]
- Kim H, Shimojo S, O'Doherty JP. Overlapping responses for the expectation of juice and money rewards in human ventromedial prefrontal cortex. *Cereb Cortex*. 2011; 21(4):769–776. [PubMed: 20732900]
- Kringelbach ML, Rolls ET. The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. *Prog Neurobiol*. 2004; 72(5):341–372. [PubMed: 15157726]
- Krueger LE. Reconciling Fechner and Stevens: Toward a unified psychophysical law. *Behav Brain Sci*. 1989; 12:251–320.
- Lara AH, Kennerley SW, Wallis JD. Encoding of gustatory working memory by orbitofrontal neurons. *J Neurosci*. 2009; 29(3):765–774. [PubMed: 19158302]

- Liu X, Hairston J, Schrier M, Fan J. Common and distinct networks underlying reward valence and processing stages: a meta-analysis of functional neuroimaging studies. *Neurosci Biobehav Rev*. 2011; 35(5):1219–1236. [PubMed: 21185861]
- Liu X, Powell DK, Wang H, Gold BT, Corbly CR, Joseph JE. Functional dissociation in frontal and striatal areas for processing of positive and negative reward information. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 2007; 27(17):4587–4597. [PubMed: 17460071]
- Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A. Neurophysiological investigation of the basis of the fMRI signal. *Nature*. 2001; 412(6843):150–157. [PubMed: 11449264]
- Luk CH, Wallis JD. Choice coding in frontal cortex during stimulus-guided or action-guided decision-making. *J Neurosci*. 2013; 33(5):1864–1871. [PubMed: 23365226]
- Ma N, Liu Y, Li N, Wang CX, Zhang H, Jiang XF, et al. Addiction related alteration in resting-state brain connectivity. *Neuroimage*. 2010; 49(1):738–744. [PubMed: 19703568]
- Mackey S, Petrides M. Quantitative demonstration of comparable architectonic areas within the ventromedial and lateral orbital frontal cortex in the human and the macaque monkey brains. *The European journal of neuroscience*. 2010; 32(11):1940–1950. [PubMed: 21050280]
- McCabe C, Cowen PJ, Harmer CJ. Neural representation of reward in recovered depressed patients. *Psychopharmacology (Berl)*. 2009; 205(4):667–677. [PubMed: 19529923]
- McCabe C, Mishor Z, Cowen PJ, Harmer CJ. Diminished neural processing of aversive and rewarding stimuli during selective serotonin reuptake inhibitor treatment. *Biol Psychiatry*. 2010; 67(5):439–445. [PubMed: 20034615]
- Milad MR, Rauch SL. The role of the orbitofrontal cortex in anxiety disorders. *Ann N Y Acad Sci*. 2007; 1121:546–561. [PubMed: 17698998]
- Monosov IE, Hikosaka O. Regionally distinct processing of rewards and punishments by the primate ventromedial prefrontal cortex. *J Neurosci*. 2012; 32(30):10318–10330. [PubMed: 22836265]
- Morrison SE, Salzman CD. The convergence of information about rewarding and aversive stimuli in single neurons. *J Neurosci*. 2009; 29(37):11471–11483. [PubMed: 19759296]
- 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(31):8166–8169. [PubMed: 17670960]
- Noonan MP, Walton ME, Behrens TE, Sallet J, Buckley MJ, Rushworth MF. Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. *Proc Natl Acad Sci U S A*. 2010; 107(47):20547–20552. [PubMed: 21059901]
- 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(1):95–102. [PubMed: 11135651]
- Padoa-Schioppa C. Neurobiology of economic choice: a good-based model. *Annual review of neuroscience*. 2011; 34:333–359.
- Petrides, M.; Pandya, DN. Comparative architectonic analysis of the human and macaque frontal cortex. In: Boller, F.; Grafman, J., editors. *Handbook of Neuropsychology*. Vol. Vol. 9. New York: Elsevier; 1994. p. 17-57.
- Petrides M, Pandya DN. Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. *Eur J Neurosci*. 2002; 16(2):291–310. [PubMed: 12169111]
- Plassmann H, O'Doherty JP, Rangel A. Appetitive and Aversive Goal Values Are Encoded in the Medial Orbitofrontal Cortex at the Time of Decision Making. *Journal of Neuroscience*. 2010; 30(32):10799–10808. [PubMed: 20702709]
- Rangel A, Hare T. Neural computations associated with goal-directed choice. *Current opinion in neurobiology*. 2010; 20(2):262–270. [PubMed: 20338744]
- Roesch MR, Olson CR. Neuronal activity related to reward value and motivation in primate frontal cortex. *Science*. 2004; 304(5668):307–310. [PubMed: 15073380]
- Rushworth MF, Buckley MJ, Gough PM, Alexander IH, Kyriazis D, McDonald KR, et al. Attentional selection and action selection in the ventral and orbital prefrontal cortex. *J Neurosci*. 2005; 25(50):11628–11636. [PubMed: 16354921]

- Rushworth MF, Noonan MP, Boorman ED, Walton ME, Behrens TE. Frontal cortex and reward-guided learning and decision-making. *Neuron*. 2011; 70(6):1054–1069. [PubMed: 21689594]
- Schoenbaum G, Roesch M. Orbitofrontal cortex, associative learning, and expectancies. *Neuron*. 2005; 47(5):633–636. [PubMed: 16129393]
- Schoenbaum G, Takahashi Y, Liu TL, McDannald MA. Does the orbitofrontal cortex signal value? *Annals of the New York Academy of Sciences*. 2011; 1239:87–99. [PubMed: 22145878]
- Selemon LD, Goldman-Rakic PS. Longitudinal topography and interdigitation of corticostriatal projections in the rhesus monkey. *J Neurosci*. 1985; 5(3):776–794. [PubMed: 2983048]
- Seo H, Lee D. Behavioral and Neural Changes after Gains and Losses of Conditioned Reinforcers. *Journal of Neuroscience*. 2009; 29(11):3627–3641. [PubMed: 19295166]
- Seymour B, O'Doherty JP, Koltzenburg M, Wiech K, Frackowiak R, Friston K, et al. Opponent appetitive-aversive neural processes underlie predictive learning of pain relief. *Nat Neurosci*. 2005; 8(9):1234–1240. [PubMed: 16116445]
- Seymour B, Singer T, Dolan R. The neurobiology of punishment. *Nat Rev Neurosci*. 2007; 8(4):300–311. [PubMed: 17375042]
- Sirotin YB, Das A. Anticipatory haemodynamic signals in sensory cortex not predicted by local neuronal activity. *Nature*. 2009; 457(7228):475–479. [PubMed: 19158795]
- Small DM, Zatorre RJ, Dagher A, Evans AC, Jones-Gotman M. Changes in brain activity related to eating chocolate: from pleasure to aversion. *Brain*. 2001; 124(Pt 9):1720–1733. [PubMed: 11522575]
- Takahashi YK, Roesch MR, Stalnaker TA, Haney RZ, Calu DJ, Taylor AR, et al. The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. *Neuron*. 2009; 62(2):269–280. [PubMed: 19409271]
- Takahashi YK, Roesch MR, Wilson RC, Toreson K, O'Donnell P, Niv Y, et al. Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. *Nature neuroscience*. 2011; 14(12):1590–1597.
- Tom SM, Fox CR, Trepel C, Poldrack RA. The neural basis of loss aversion in decision-making under risk. *Science*. 2007; 315(5811):515–518. [PubMed: 17255512]
- Wallis JD. Cross-species studies of orbitofrontal cortex and value-based decision-making. *Nature neuroscience*. 2012; 15(1):13–19.
- Wallis JD, Kennerley SW. Heterogeneous reward signals in prefrontal cortex. *Curr Opin Neurobiol*. 2010; 20(2):191–198. [PubMed: 20303739]
- Zald DH, Hagen MC, Pardo JV. Neural correlates of tasting concentrated quinine and sugar solutions. *J Neurophysiol*. 2002; 87(2):1068–1075. [PubMed: 11826070]

Figure 1. Recording locations and behavioral task

A) Top panel: Schematic of the ventral view of the macaque brain grossly distinguishing mOFC (green), OFC (cyan) and IC (dark blue). **Bottom panel:** Coronal MRI image corresponding to the level of the gray line in the top panel. Yellow lines depict electrode tracks and shaded regions are areas from which we recorded. **B) Task sequence** for a positive (left) and negative (right) picture trial. The reward bar is in blue at the bottom of the screen and remains visible throughout the trial. On each trial, the length of the reward bar could increase (left), decrease (right) or remain the same size (not shown). The size of the bar carries over until the next trial within every block of six trials. **C) Behavioral effects** of positive and negative stimuli for each subject. Plots show the mean percentage of trials in which a correct joystick response is executed (top) or the median RT to make a response (bottom), separately for positive and negative pictures, and right and left joystick movements. **D) Mean percent correct** ($\pm$ SE, top) and median RT ($\pm$ SE, bottom) for 6 different levels of reward bar size or number of trials completed within each block, shown separately for positive and negative pictures. **E) AIC weights** averaged across both subjects for each of 15 models of behavior. Model numbers refer to Table 1. O = weights for Subject C. X = weights for Subject M. Model 12 was the best fitting model for both subjects.

Figure 2. Neuronal responses during the sample period

A–D) Spike density histograms for two neurons encoding picture valence (A and B) and two neurons encoding response direction (C and D) during the sample epoch. Firing rates are aligned to stimulus onset (gray line). Neurons A and C were recorded from IC, B and D from OFC. E) The top panel shows the percentage of neurons in each region encoding the picture valence (gray) or response direction (black) during the sample epoch. The bottom panel shows mean CPDs ($\pm$ SE) for valence and response during the sample epoch. Mean CPDs are low, as they are population averages that included non-selective neurons (* = $p < 0.05$, Wilcoxon rank sum test). F) Plots show the cumulative number of neurons significantly encoding picture valence or response in mOFC (blue), OFC (red) and IC (black) across time.

Figure 3. Neural encoding of valence during sample epochs

A) The top panel shows the percentage of neurons encoding positive or negative valences, and the bottom panel shows mean CPDs ($\pm$ SE) for picture valence during the sample epoch. CPDs were averaged over all neurons with positive or negative beta coefficients in a multiple regression. B) Heat plots show significant beta coefficients among neurons responding to positive or negative pictures. Each horizontal line corresponds to the data from an individual neuron, and the color indicates the absolute value of the beta coefficient at that time point. Neurons were sorted by latency to encode each variable. Yellow lines indicate picture onset, and gray lines show the median encoding latency for each group. C) Flattened maps of the orbital surface outlining the LOS and MOS (shaded gray), averaged across subjects, and locations of neurons responding to positive (blue) and negative (red) pictures during the sample epoch. Gray lines indicate the lateral (top) and medial (bottom) convexities, where the orbital surface terminates and curves around onto the lateral and medial surfaces of the frontal lobe. Anterior-posterior (AP) positions are relative to the interaural line, and medial-lateral (ML) positions are relative to the LOS. For display, AP

positions were jittered by ± 0.2 -mm and offset 0.2-mm. Circle diameters represent strength of encoding, and are proportional to the absolute value of the beta coefficient for picture valence. Inset boxplots show the median (central line), 25th and 75th percentile (box top and bottom) ML position of neurons responding to positive (blue) and negative (red) pictures during the sample epoch (Wilcoxon rank-sum test, $p > 0.1$). Whiskers show the data spread, and '+' points identify outliers. The flatmap on the right shows labeled anatomical landmarks. Each 'x' indicates the location of a recorded neuron. Sulci are shaded gray; convexities are marked by a gray line. Inset is a schematic of the orbital region of a single coronal slice, demarcating landmarks shown on the flatmap.

Figure 4. Neural encoding of feedback valence

A–F) Spike density histograms show four neurons encoding feedback (FB) in different ways. Firing rates are aligned to feedback onset (gray line). Blue = positive feedback trials, cyan = no feedback following a positive picture, orange = no feedback following a negative picture, red = negative feedback trials. Neurons A, C, E and F were recorded from IC, B and D from OFC. G) Scatter plots show the distribution of feedback-related responses. Each point represents a selective neuron, and its position on the y-axis is determined by the beta coefficient. Positive and negative betas indicate that the neuron's response consisted of an increase or decrease in firing rate respectively. H) Percent of neurons with higher firing rates for positive or negative outcomes. I) Percent of neurons that encode expected or unexpected outcomes. Expected outcomes include positive feedback following a positive picture and no feedback following a negative picture (e.g. A and E). Unexpected outcomes include negative feedback or no feedback following a positive picture (e.g. B and D). * = $p < 0.05$, χ^2 test

Figure 5. Consistent and inconsistent valence coding across epochs

Percentage of neurons showing A) the same valence encoding for sample pictures and feedback or B) opposite valence encoding between these two epochs. The x-axis the percentage of neurons in each region expected to respond to sample pictures and feedback of the same or different valence if coding in each epoch were independent, and the y-axis is the percentage of neurons observed. Shapes indicate data from different brain areas, and color indicates the valence of encoding during the sample presentation (blue = positive, red = negative). * = $p < 0.05$, binomial test.

Figure 6. Neurons encoding other sources of value in the task

Spike density histograms for single neurons whose firing rate during the fixation epoch correlated with either A) the size of the reward bar in IC or B) the number of trials completed in OFC. The same neuron's activity was sorted by reward bar size (left panels), or number of trials completed (right panels), with colors indicating size or number respectively. Bar sizes of 0 and ≥ 6 consisted of very few trials, and were therefore excluded from plots. C) Scatterplots of beta coefficients for bar size (x-axes) versus trial number (y-axes) from a multiple regression that included both parameters. Each point represents a neuron selective for either bar size (purple) or trial number (green). Positive beta coefficients indicate that the neurons activity was positively correlated with value, while negative values indicated an anti-correlation. Data were taken from the fixation epoch. D) Percent of neurons encoding either reward bar size or trial number during fixation (Fix) and sample (S) epochs. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, χ^2 test. E) Flattened orbital map outlining the lateral and medial orbital sulci (shaded gray) as in Figure 3. Circles

show the location of neurons encoding bar size (purple) and trial number (green) during the fixation epoch. AP positions were jittered within ± 0.2 -mm and offset 0.2-mm for display. Circle diameters represent strength of encoding and are proportional to the absolute value of the beta coefficient for bar size or trial number respectively. Inset boxplots show ML distributions of neurons in orbital areas encoding bar size (purple) or trials completed (green). The central line shows the median, box top and bottom show the 25th and 75th percentile, whiskers show the data spread, and '+' points identify outliers. **= $p<0.01$, Wilcoxon rank-sum test.

Figure 7. Comparison of neurons encoding value during feedback and juice delivery

Scatterplot shows the absolute values of beta coefficients from a linear regression of reward bar size (which corresponded to juice volume) and neuron firing rates during feedback (x-axis) and juice delivery (y-axis). Each point represents a neuron that was selective during at least one epoch. Red=IC, Blue=OFC, Green=mOFC. Inset bar plots show the overall percent of neurons that were selective during each epoch (FB=feedback, juice=juice delivery). * = $p < .05$ χ^2 test.

Table 1

Behavioral models tested.

Model number	Parameters
1	No additional parameters
2	Bar
3	log(Bar)
4	Trials
5	log(Trials)
6	Bar, Trials
7	Bar $\times$ Trials
8	log(Bar), Trials
9	log(Bar) $\times$ Trials
10	Bar, log(Trials)
11	Bar $\times$ log(Trials)
12*	log(Bar), log(Trials)
13	log(Bar) $\times$ log(Trials)
14	Hyperbolically discounted Bar
15	Exponentially discounted Bar

* = best-fit model for both subjects

Model 1 is the reduced model and includes only picture valence and response direction. Models 2 through 15 include picture valence and response direction plus the additional parameters indicated in the table. *Bar* is the length of the reward bar. *Trials* are the number of trials completed within each block of six trials. Commas separate two independent variables in a model; x indicates a multiplicative interaction term.

Table 2

Variables explaining neural activity.

Variable	Definition	Percentage of neurons		
		Fixation	Sample	Feedback
Picture Valence	Positive or negative	<i>1</i>	28	9
Response	Right or left joystick movement	<i>1</i>	24	16
Bar	Reward bar length	17	10	21
Trials	Number of trials completed within a block	19	14	30
Feedback previous trial	Positive, negative or zero	<i>3</i>	2	2
Picture valence previous trial	Positive or negative	<i>2</i>	2	<i>1</i>
Response previous trial	Right or left joystick movement	<i>1</i>	2	<i>1</i>
Feedback value	+1, -1 or 0	n/a	n/a	6
Feedback relative value	Best or worst outcome, given the picture valence (e.g. positive feedback for a positive picture and no feedback for a negative picture)	n/a	n/a	<i>4</i>
Outcome salience	Positive or negative feedback = 1; no feedback = 0	n/a	n/a	<i>3</i>
Positive feedback	Positive feedback = 1; negative or no feedback = 0	n/a	n/a	<i>6</i>
Negative feedback	Negative feedback = 1; positive or no feedback = 0	n/a	n/a	<i>3</i>
Omitted reward	No feedback for positive pictures = 1; all other trials = 0	n/a	n/a	<i>5</i>
Omitted punishment	No feedback for negative pictures = 1; all other trials = 0	n/a	n/a	<i>7</i>
Correct/error	Correct or incorrect joystick response, regardless of feedback	n/a	n/a	<i>3</i>
		Total feedback encoding		31

Variables included in the analysis of neural activity and the overall percentage of neurons encoding each variable. Note that some variables are correlated with each other. Because of these correlations, we employed stepwise regressions to identify the variable(s) that best described neuron activity. Percentages are totals across all areas, during three trial epochs (fixation, sample and feedback) plus an epoch at the end of each trial block when the reward bar is cashed in for juice reward. Italicized percentages were not significantly above chance (5%) occurrence as determined by binomial tests ($p > 0.05$). Analysis of all epochs included the first 7 variables; the last 6 were only included in analysis of the feedback epoch.