

Distinguishing informational from value-related encoding of rewarding and punishing outcomes in the human brain

Ryan K. Jessup^{1,2,3} and John P. O'Doherty^{1,2,*}

¹Trinity College Institute of Neuroscience, Trinity College Dublin, Dublin, Ireland

²Division of Humanities and Social Sciences and Computation and Neural Systems Program, California Institute of Technology, Pasadena, CA, USA

³Department of Management Sciences, Abilene Christian University, Abilene, TX, USA

Keywords: decision making, insula cortex, orbitofrontal cortex, outcome coding, ventral striatum

Abstract

There is accumulating evidence implicating a set of key brain regions in encoding rewarding and punishing outcomes, including the orbitofrontal cortex, medial prefrontal cortex, ventral striatum, anterior insula, and anterior cingulate. However, it has proved challenging to reach consensus concerning the extent to which different brain areas are involved in differentially encoding rewarding and punishing outcomes. Here, we show that many of the brain areas involved in outcome processing represent multiple outcome components: encoding the value of outcomes (whether rewarding or punishing) and informational coding, i.e. signaling whether a given outcome is rewarding or punishing, ignoring magnitude or experienced utility. In particular, we report informational signals in the lateral orbitofrontal cortex and anterior insular cortex that respond to both rewarding and punishing feedback, even though value-related signals in these areas appear to be selectively driven by punishing feedback. These findings highlight the importance of taking into account features of outcomes other than value when characterising the contributions of different brain regions in outcome processing.

Introduction

A burgeoning literature aims to establish how the brain encodes rewarding and punishing outcomes. Human neuroimaging studies have identified a role for a number of brain areas in the encoding of such outcomes, including the orbitofrontal cortex, ventral striatum and insular cortex (Delgado *et al.*, 2000; Elliott *et al.*, 2000; Breiter *et al.*, 2001; Knutson *et al.*, 2001; O'Doherty *et al.*, 2001). One of the major questions is the extent to which rewarding and punishing outcomes are represented in overlapping areas vs. being encoded in different brain regions. Some studies have found evidence to support a degree of separation in the encoding of rewarding and punishing outcomes (Small *et al.*, 2001; Anderson *et al.*, 2003; O'Doherty *et al.*, 2003c). For example, on the one hand, an area of the medial orbitofrontal cortex has been consistently implicated in encoding rewarding outcomes, with this area often found to actively decrease in its responses to punishing outcomes (O'Doherty *et al.*, 2001; Smith *et al.*, 2010). On the other hand, the lateral orbitofrontal cortex has been found to encode aversive outcomes (O'Doherty *et al.*, 2001, 2003a; Small *et al.*, 2001; Ursu & Carter, 2005; Elliott *et al.*, 2010; Hampshire *et al.*, 2012), leading to the proposal of a medial

to lateral dissociation in orbitofrontal cortex function where rewarding outcomes are encoded in the medial and aversive outcomes are encoded in the lateral orbitofrontal cortex (Kringelbach & Rolls, 2004). Studies have also implicated the anterior insula in aversive responding, such as in encoding monetary losses, or aversive tastes or pain (Ploghaus *et al.*, 1999; O'Doherty *et al.*, 2003a; Nitschke *et al.*, 2006a,b), leading to the suggestion that the anterior insula may play a role in aversive processing.

However, the extent to which punishing outcomes are truly represented in brain areas different from those representing rewarding outcomes has been challenged by other findings. A single-unit neurophysiology recording study in non-human primates reported no evidence for a topography in the distribution of reward- and aversive-responding neurons within the central orbitofrontal cortex, although the full extent of the orbital surface was not sampled in that case, leaving the large-scale topography unaddressed (Morrison & Salzman, 2009). In addition, the lateral orbitofrontal cortex has also been found to respond during the receipt of rewarding outcomes in a number of human neuroimaging studies (O'Doherty *et al.*, 2003a; Noonan *et al.*, 2011). This has led to the proposal that the lateral orbitofrontal cortex may play a role in learning to assign 'credit' to stimuli on the basis of their predictive relationship to rewarding or punishing outcomes (Noonan *et al.*, 2012). Furthermore, the role of the insular cortex in aversive processing has been challenged by findings that this area can occasionally be activated by salient positive as well as negative outcomes, indicative of a more general role for this region in arousal (Elliott *et al.*, 2000;

Correspondence: John P. O'Doherty, present address below.
E-mail: johndoherty@caltech.edu

*Present address: MC 228-77, California Institute of Technology, 1200 E. California Boulevard, Pasadena, CA 91125, USA

Received 12 December 2013, revised 15 April 2014, accepted 16 April 2014

Bartra *et al.*, 2013). Finally, multi-variate statistical approaches to functional magnetic resonance imaging (fMRI) data analysis have revealed distributed yet overlapping signals within the orbitofrontal cortex and elsewhere in the brain, coding for rewarding and punishing outcomes (Kahnt *et al.*, 2010; Vickery *et al.*, 2011). However, a significant issue with much of the experimental literature to date on outcome coding is that outcomes can be represented in several different ways in the brain; these coding schemes include information signaling and salience as well as value signaling.

Informational signaling about an outcome in its simplest form distinguishes the valence of the outcome that is received, i.e. whether the outcome is a reward or a punishment, or is an affectively neutral outcome irrespective of its magnitude. This type of coding mechanism is independent of the subjective value assigned to that outcome. An example of such a signal would be a signal that responds differently to a gain or a loss outcome relative to a neutral outcome, or a signal that differentiates gains compared with losses independently of the magnitude (or subjective value) of those gains or losses. Such signals may be present in many areas of the brain that have previously been reported as being involved in encoding outcome value signals and, unless these signals are separately characterised and explicitly differentiated, it would be difficult to determine which type of signal is being found and reported in a given brain region in response to outcome reception.

Another type of outcome signal is a salience code in which neural responses increase as a function of an increasing magnitude of outcome received, but in which both rewarding and punishing outcomes result in a similar increase in activity (Zink *et al.*, 2004; Litt *et al.*, 2011; Bartra *et al.*, 2013). Thus, salience-related activity in response to an outcome does not distinguish the valence of that outcome, but rather encodes its intensity independently of its value *per se*.

Outcome value signals produce activity proportional to increasing magnitudes of outcomes as in salience codes, but in this case the activity profile distinguishes between the magnitude of rewarding and punishing outcomes. A reward-related outcome value response would yield increasing activity in response to increasing magnitudes of reward but not show an increasing response to an aversive outcome. However, an aversion-related outcome value response would manifest an increasing activity in response to increasing magnitudes of an aversive outcome but not a rewarding one.

The goal of the present study was to distinguish informational, salience and value-related accounts for neural responses to rewarding and punishing outcomes as measured with fMRI. We aimed to examine the distributions of these signals across the brain in order to ascertain whether informational, salience or value signals for rewarding and punishing outcomes are differentially represented across the brain. We used a task wherein participants played slot machines in which they could win or lose money drawn from one of two uniform distributions. One distribution had a positive mean payoff, whereas the other had a negative mean payoff. In some cases, participants had a choice of two slot machines, whereas on other trials participants saw only a single slot machine that they must select. By using varying outcome magnitudes, the task design enabled us to separate out responses to binary feedback from value codes, and by using both rewarding and punishing outcomes of varying magnitudes we could distinguish value codes from salience codes (see Fig. S1 for illustration of different outcome coding schemes). We hypothesised that informational signals for rewarding and punishing outcomes would be present in a number of brain areas hitherto implicated in encoding outcome value and salience signals, including the lateral orbitofrontal cortex and anterior insular cortex. Furthermore, we hypothesised that differential value signals

for rewarding and punishing outcomes would be present within a subset of these areas.

Materials and methods

Participants

A total of 35 participants (16 female and 19 male) completed four 13 min 50 s runs of a decision-making task while lying in a whole-body magnetic resonance imaging scanner (one additional participant began but did not complete the experiment). The fMRI data of five participants were excluded because of technical problems with the magnetic resonance imaging scanner resulting in usable data from 30 participants. The study was approved by the Trinity College Dublin School of Psychology research ethics committee, and all participants gave informed consent. The study conformed to the guidelines set out in the 2013 WMA Declaration of Helsinki.

Experimental task

The task consisted of four runs, each composed of 60 experimental and 30 neutral trials, all randomly intermixed. Each run began with 6000 ms of a fixation cross and ended with a final fixation period of variable length. On each trial participants were presented with stimuli resembling slot machines (Fig. 1). These machines were presented to the left and right of a central fixation cross. The experimental task contained two factors: set size and expected value (EV). The set size factor had two conditions (single and double) and the EV factor had two trial types (low and high); both factors were fully crossed yielding four different trial types: single low, single high, double low and double high. In the double condition, two slot machines were presented at the same time, whereas in the single condition only one slot machine was presented on either the left or the right. In both conditions, the side of presentation (left vs. right) of each specific slot machine was determined at random. Participants had 2000 ms in which to push a button to select either the left or the right slot machine (in the single condition, only the button corresponding to the side on which the machine was presented was valid). The button press resulted in an animation of the selected machine indicating that a response had been registered. At the end of the 2000 ms choice window on trials in which a valid button was pressed, an animation simulating the spinning wheels on slot machines was shown for 3000 ms. This was followed by a 1000 ms outcome presentation, revealing the amount of money won by the participant on that trial, shown in cents. The ensuing intertrial interval was drawn from a quasi-exponential distribution, ranging from 2000 to 7000 ms with a mean of 3000 ms, during which time a fixation cross was presented on-screen.

If no valid button was pressed within the 2000 ms choice time window, the text 'Too Late! Choose More Quickly' was presented on the screen for 4000 ms and then followed by the intertrial interval. The same trial was then repeated.

For two out of the four runs, the participants encountered the double condition, and for the other two runs the participants encountered the single condition. The order of presentation of these conditions was randomised, although the two runs of each type were grouped together (i.e. in some subjects the first two runs were from the double condition and the last two from the single condition, whereas in other subjects the reverse order was implemented). The double and single conditions were organised in this fashion as opposed to being intermixed within a run in order to (i) reduce the complexity and difficulty of the task by requiring participants to

keep track of fewer stimuli (as would have been required in a fully intermixed design) and (ii) avoid interactions between double and single trials that might have occurred if participants had to switch between these conditions on a trial-by-trial basis (e.g. by experiencing response conflict when transitioning from a double trial where a choice was available to a single trial where no choice was available). Keeping these conditions separate avoids such potentially confounding interactions.

In the double condition there were four experimental and two neutral slot machine stimuli, each differentiable by a randomly assigned color (red, yellow, blue, green, orange, and purple). Out of the four experimental machines, two machines were designated to yield a positive EV of 125 cents (high EV) by yielding outcomes from a uniform distribution ranging from -250 to $+500$ cents in 1 cent increments, whereas two of the machines were selected to have a negative EV of -120.5 cents (low EV) by yielding outcomes drawn from a uniform distribution from -496 to $+255$ cents in 1 cent increments. The neutral slots yielded no monetary outcomes but instead text reading 'No Change'. Participants were presented with pairs of machines of each type, i.e. a pair of high EV stimuli and a pair of low EV stimuli, as well as a pair of neutral stimuli. High and low EV stimuli were never intermixed with each other or with neutral trials. This was because we wanted the environment at the time of choice to be constant within trials of the same type, i.e. so that there was no overall difference in the EV between decision options. The result is that the value of different choice options in a choice set should remain equal except for local variations based on random changes in the outcome distributions. This allowed us to simplify the analysis so as to focus on (i) outcome signals, which are the main goal of the study, and (ii) the effects of having a choice *per se* on the outcomes without the confounding effects of a change in overall EV as a function of choice. Over the course of the two double condition runs, there were 60 trials for each pair of slots, randomly intermixed, totaling 180 double trials.

The single condition also had six different slots (in addition to being differentiated from one another by randomly assigned color, they also had a different shape than those in the double condition). Once again, two of the slots yielded high EV, and two of the slots yielded low EV (payoffs were drawn from the same distributions as for the double condition); the remaining two were neutral slots leading to no monetary outcomes. In this condition, each of the slots was presented alone on the left or right of a fixation cross and the participant was forced to select that slot. Over the course of the two single condition runs, each slot machine was presented 30 times, randomly intermixed, totaling 180 single trials. Trials on which the participant failed to select the slot were treated identically to missed trials in the double condition.

It should be noted that one other feature of the design is that decision conflict (which would be maximal when two options have the same value and minimal when there is a large value difference between the options) is controlled for in the double condition because the values of the options available on each trial are kept the same (except for small local variations). Thus, decision conflict is always constant within the double condition, and absent in the single condition as only one option is available.

Test trials

At the end of the second run of each condition we wanted to confirm whether or not participants had learned the differences in EVs assigned to the different slot machines from that condition. This was accomplished by presenting binary choices between pairs of slot

machines containing one slot with a positive and one with a negative EV, a pairing never encountered throughout either of the two conditions. Each of the two machines with high EV was paired with each of the two machines with low EV and these binary choice test trials were repeated twice. This resulted in eight trials per condition and 16 overall. Crucially, because we wanted to test for the effects of prior learning, the outcomes on these trials were registered but not revealed to the participants until the end of the experiment. These test trials were conducted after the scanner had stopped running and hence no corresponding imaging data were collected. These test trials were conducted at the end of each condition instead of at the end of the entire experiment because the stimuli used differed between conditions and the order in which participants observed the conditions differed between participants. Had we waited until the end to conduct all test trials the learning results would have been confounded by recency effects.

Instructions and payoff

At the beginning of the experiment, participants were informed that they should learn which machines paid better and choose accordingly because their payment depended on their choices during the experimental runs as well as the test trials. They further received detailed instructions on the task timing, the two different conditions, and appropriate buttons to press. They then completed four practice trials, two from each condition, in real time (no feedback regarding the outcomes of their choices was revealed during any of these practice trials). They were also informed that they would receive no less than €20 and up to €50. This was done to meet the minimum reimbursement requirements of the School of Psychology research ethics committee, which required a payment of €10 for each hour in the study. This range of possible outcomes was made explicit to the participants to clearly demonstrate the range and large variance of possible outcomes in order to encourage them to learn. Assuming that individuals chose optimally in the 16 test trials, the EV (SD) was €26 (€33.54). Conditioned on the above payoff parameters (i.e. €10/h plus earnings but not less than €20 or more than €50), participants were then paid exactly what they earned.

Functional magnetic resonance imaging acquisition and preprocessing

The task was conducted in a Philips Achieva 3T scanner using a phased-array eight-channel head coil. The imaging data were acquired with a 30° tilt from the anterior commissure–posterior commissure line in order to minimise signal loss in the orbitofrontal cortex (Deichmann *et al.*, 2003) using a gradient echo T2*-weighted echo planar imaging sequence (39 3.55 mm ascending slices were acquired with a 0.35 mm slice gap; TE (Echo Time), 28 ms; TR (Repetition Time), 2000 ms; voxel size, 3 mm, 3 mm, 3.55 mm; matrix, 80×80 voxels; field of view, 240×240 mm). A total of 415 images were collected in each of the four runs totaling 1660 functional scans. High-resolution T1 images were collected at the beginning of each participant's session.

SPM8 (Wellcome Department of Imaging Neuroscience, London, UK; www.fil.ion.ucl.ac.uk/spm) was used to preprocess and analyse the data. The functional data for each participant were spatially realigned using a six-parameter rigid-body spatial transformation and then slice time corrected to the middle (i.e. 20th) slice. The high-resolution structural image was then coregistered to the mean functional image generated by the realignment phase. The mean functional was spatially normalised to the echo planar imaging tem-

plate, and the resulting warping parameters were then applied to the realigned functionals and coregistered structural. The functionals were then spatially smoothed using an 8 mm Gaussian kernel.

Functional magnetic resonance imaging analysis

Changes in the blood oxygenation level-dependent response were examined using a general linear model (GLM) with random effects. We estimated multiple GLMs for each participant. A 128 s high-pass cutoff filter and an AR(1) correction for serial correlation were applied to each model. Task conditions were convolved with a canonical hemodynamic response function and then entered into the GLM matrix. Each scan of the GLM included a baseline regressor and six movement regressors, estimated via the spatial realignment preprocessing step.

Our GLM was designed to examine differences in responsiveness based on the EV and set size at the time of choice and set size and valence at the time of outcome. Separate regressors were used to model blood oxygenation level-dependent activity at both the choice and outcome phases. At the choice phase, regressors for experimental trials were separated according to set size (single or double) and EV (low or high); regressors for neutral trials were also separately modeled for double and single trials. A single regressor was also used for all trials on which no option was selected. This resulted in $2 \cdot 2 + 2 + 1 = 7$ choice regressors. At the outcome phase, regressors for experimental trials were separated according to both set size and valence (gain or loss) and each of these regressors was parametrically modulated by the actual outcome value. (For the experimental trials, any outcome > 0 was considered as a gain, otherwise it was coded as a loss; because the range for both high and low EV trials covered 0, both yielded gain and loss events.) Regressors for the neutral trials were also separately modeled as a function of set size. This resulted in $(2 \cdot 2) \cdot 2 + 2 = 10$ outcome regressors and, together with the choice phase regressors, a maximum of 17 total regressors beyond the baseline and movement regressors. Inclusion of the choice regressors in the model made it possible for us to exclude effects elicited at the time of cue presentation from the effects being tested at the time of outcome receipt, as both of these regressor types competed for variance and, due to the extra sum of squares principle, activity reported in response to outcome contrasts reflects effects of outcomes over and above any effects on the blood oxygenation level-dependent signal arising from the cue phase.

For the main statistical analysis of the fMRI data, we performed voxel-wise tests across the whole brain using an omnibus height threshold of $P < 0.005$ uncorrected and a cluster size $k \geq 5$. Following this, we then correcting for multiple comparisons at $P < 0.05$ family-wise error within small volumes defined on *a priori* regions. The coordinates for these regions were obtained from three meta-analyses (Kringelbach & Rolls, 2004; Hayes & Northoff, 2011; Diekhof *et al.*, 2012). The *a priori* regions were the lateral orbitofrontal cortex, insula, anterior cingulate, ventral striatum, and medial prefrontal cortex. For each of these we performed corrections for multiple comparisons within a 10 mm sphere. Both the extent threshold (P -value family-wise error) and cluster size k are reported for each comparison. The Montreal Neurological Institute (MNI) coordinates for the left and right lateral orbitofrontal cortex $[-33, 42, -5]$ and $[34, 41, -5]$, respectively, were found in the meta-analysis of Kringelbach & Rolls (2004), rounded to the nearest integer. From the meta-analysis of Hayes & Northoff (2011) we used the MNI coordinates for the insula $[-40, 16, 4]$ and anterior cingulate $[4, 24, 30]$. Contralateral coordinates for the insula were obtained by flipping the sign of the x coordinate. The MNI coordi-

nates for the left and right ventral striatum $[-10, 12, -6]$ and $[16, 12, -12]$, respectively, as well as the medial prefrontal cortex $[0, 46, -10]$ were obtained from the meta-analysis of Diekhof *et al.* (2012). Figure S2 shows these region of interest locations overlaid on a structural brain image.

Parameter estimate plots were obtained by extracting the signal from the 10 mm sphere centered on the *a priori* coordinates derived from the meta-analyses described above, producing images independent of the statistical comparison that inspired them. When presenting the results in Figs 2–6, we have followed these guidelines: left-sided parameter estimates are presented on the left and right-sided parameter estimates are presented on the right (when appropriate, parameter estimates for structures located along the midline are presented in the center). The color for the parameter estimates from a particular *a priori* region matches the color of that highlighted region presented in the cluster image, which is overlaid onto the mean (over all participants) high-resolution T1 image. The anterior cingulate region is always presented in blue, the medial prefrontal cortex is always presented in green, and due to the fact that the lateral orbitofrontal cortex, insula, and ventral striatum have activation loci that extend into one another in these results, these structures are presented in red. Moreover, because some of the regions surpassed our significance threshold in multiple contrasts, we chose to present each region in a maximum of two contrast figures: once for the informational signal figures and once for the value-related signals. The figure that we displayed the region in was the contrast in which it had the largest number of significant contiguous voxels (k , for all $k \geq 5$). The set size factor was included to allow us to determine the extent to which outcome signals are modulated by the presence or absence of a choice. Because we did not find a significant effect of set size (single vs. double) at the time of outcome for gains or losses in any of our predefined regions at our significance threshold, nor any significant interaction between outcome valence and set size for either the information or value signal regressors, all

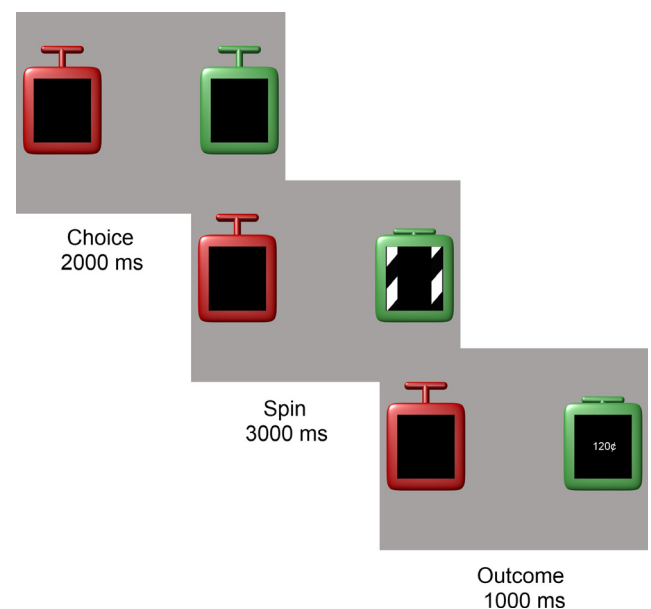


FIG. 1. Illustration of trial structure for double condition. Here, a 2000 ms choice phase, in which the option on the right is selected, is followed by a 3000 ms spin phase that is designed to simulate the spinning of a slot machine. The 1000 ms outcome phase displays the gain or loss amount for the current trial, here, a win of 120 cents.

of the results shown collapse across the set size conditions. Inspection of the parameter plots does reveal apparent effects of set size; however, we did not interpret these effects further owing to the absence of a significant effect for this factor at the voxel-wise level.

Results

Behavioral evidence of preference learning

We tested whether participants had learned to prefer the stimuli presented in the high EV conditions compared with those presented in the low EV conditions in the binary 'test trials' choice test conducted after the fMRI sessions. The mean ($\pm$ SEM) probability across participants of choosing the higher EV option over the low EV option was 0.67 ± 0.04 . A one-sample *t*-test on the choice proportions indicated that the high EV option was significantly favored over the low EV option [$t(34) = 4.87$, $P < 1.28 \times 10^{-5}$, one-tailed]. These results demonstrate that the participants acquired preferences

for the stimuli, indicating that they had learned the EVs for the different stimuli during the task performance.

Imaging data

Figures 2–6 present the results of our imaging analysis. As described in the Materials and methods, all of the results described below are significant at $P < 0.05$ family-wise error using a small volume correction on coordinates derived from published meta-analyses. Coordinates and related statistics for each of the activated areas are given in Table 1. If multiple peak voxels were found within a region for a specific contrast, only the highest peak is presented.

Informational outcome signals: binary (non-value scaling) outcome responses

We first tested for blood oxygenation level-dependent signals related to the receipt of a gain or loss outcome irrespective of the

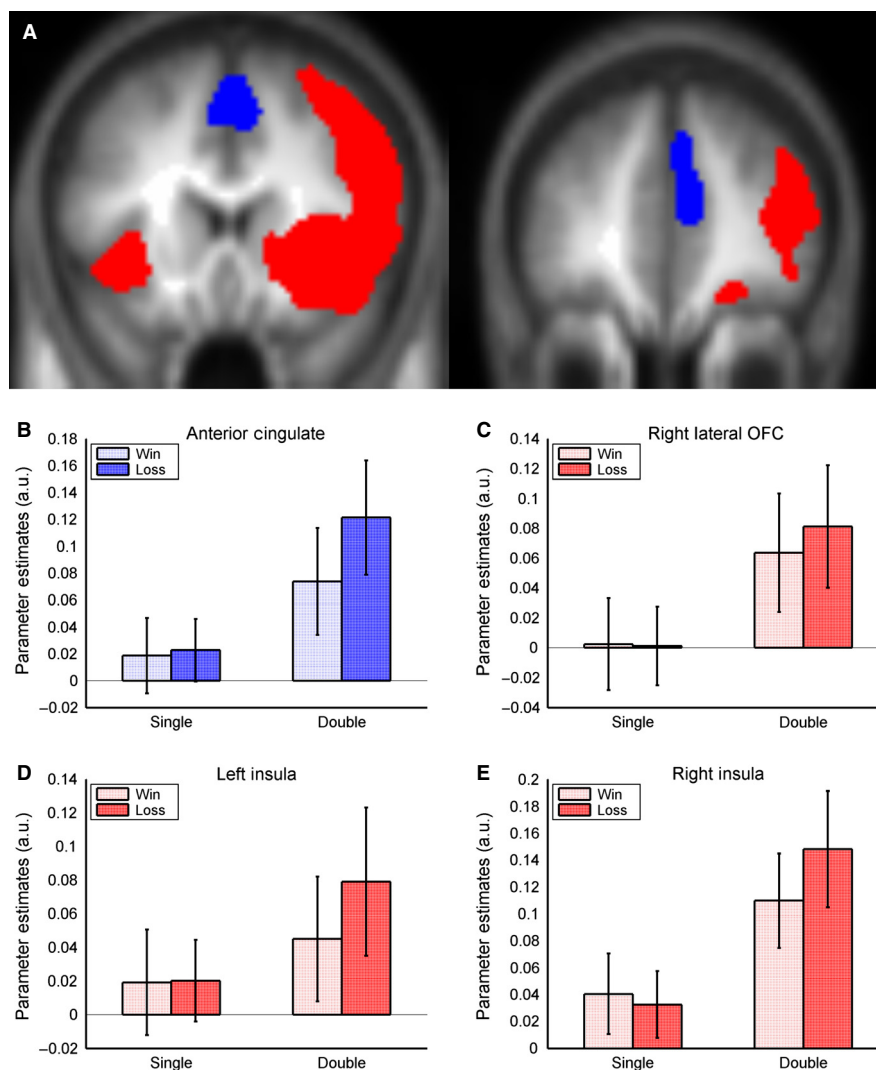


FIG. 2. Informational (non-value) signals for gain and loss outcomes relative to a neutral outcome at the time of outcome. Using a small volume correction on the respective *a priori* coordinates the activities for each of the regions shown were significant at $P < 0.05$ family-wise error. Each of the displayed parameter estimates was obtained by running an independent analysis on the 10 mm sphere at the *a priori* coordinates. (A) Two coronal slices at $Y = 16$ (left) and $Y = 41$ (right) displaying significant clusters in the anterior cingulate cortex (ACC) (blue) and the insula and lateral orbitofrontal cortex (OFC) (red). The cluster images are thresholded at Uncorrected probability (p_{Unc}) $p_{Unc} < 0.005$, $k \geq 5$ and overlaid onto the mean high-resolution T1 image. (B) Parameter estimates for the ACC. (C) Parameter estimates for the right lateral OFC. (D) Parameter estimates for the left insula. (E) Parameter estimates for the right insula. a.u., arbitrary units.

magnitude of the outcome received [contrast: (outcome single gain, outcome double gain, outcome single loss, outcome double loss) – (outcome single neutral, outcome double neutral)]. Hence, these areas showed more activity for either a gain or a loss outcome compared with a neutral outcome. We found activity corresponding with this informational outcome signal in a region of the anterior cingulate cortex (Fig. 2B), right lateral orbitofrontal cortex (Fig. 2C), and bilateral insula (Fig. 2D and E). Inspection of the parameter estimates indicated that these areas were all showing responses to both losses and gains, particularly in the double condition (to improve clarity, experimental trial parameter estimates are plotted relative to their respective neutral outcome parameter estimates).

In addition, we also tested for areas responding significantly more for gains than losses, irrespective of magnitude [contrast: (outcome single gain, outcome double gain) – (outcome single loss, outcome double loss)]. For this contrast, we found two areas: a region of the medial prefrontal cortex, and a region of the bilateral ventral striatum (Fig. 3; for contiguity with Fig. 2, the parameter estimates for the experimental trials are shown relative to their respective neutral trial parameter estimates). No area survived the reverse contrast testing for significantly greater binary responses to losses relative to gains.

Bivalent value-related outcome signals: areas showing increasing activity to both increasing magnitude of gains and increasing magnitude of losses

Next we tested for areas responding to both gains and losses in a parametric fashion, i.e. showing increasing activity to increasing magnitudes of both losses and gain outcomes [contrast: (parametric modulator (pmod) outcome single gain, pmod outcome double gain, pmod outcome single loss, pmod outcome double loss) > 0]. Our contrast revealed activity in a region of the bilateral ventral striatum as well as in the bilateral insular cortex and anterior cingulate cortex (Fig. 4). However, inspection of the parameter estimates in each of these regions revealed that the only region showing a truly bivalent response to magnitudes of gains and losses was the ventral striatum bilaterally. Plots of parameter estimates revealed that parametric activity in the other regions (insular cortex and anterior cingulate cortex) seemed predominantly driven by loss magnitude as opposed to gain magnitude and hence these regions are highlighted below. Plots of the parametric effects for gains and losses for voxels in the ventral striatum are shown in Fig. S3.

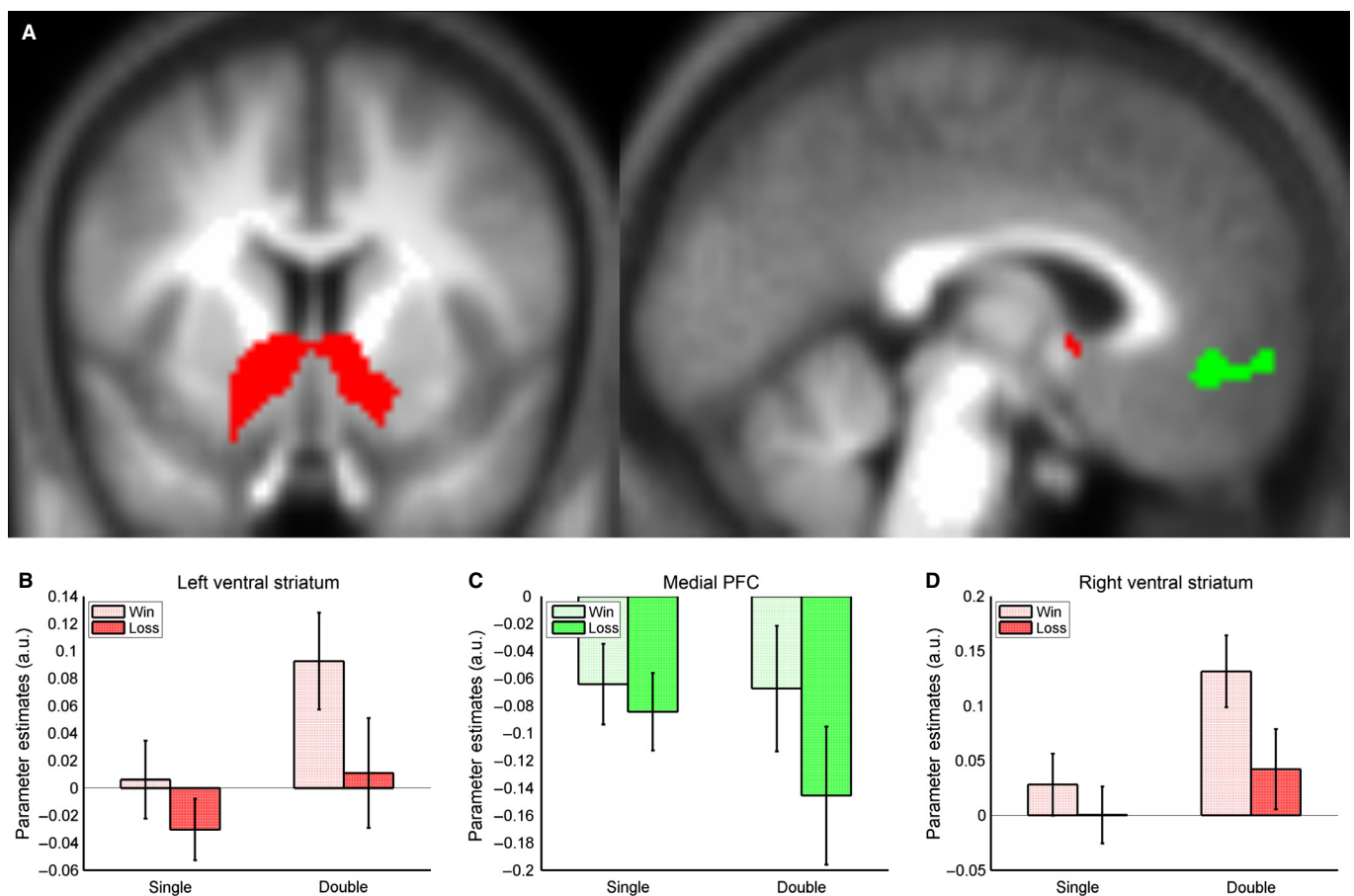


FIG. 3. Regions showing significantly greater informational (non-value) activity for gains compared with losses at the time of outcome. Using a small volume correction on the respective *a priori* coordinates the activities for each of the regions shown were significant at $P < 0.05$ family-wise error. Each of the displayed parameter estimates was obtained by running an independent analysis on the 10 mm sphere at the *a priori* coordinates. For contiguity with Fig. 2, activity on neutral trials is here removed from the parameter estimates. (A) Coronal (left) and sagittal (right) slices at $X = -2$ and $Y = 6$ displaying significant clusters in the ventral striatum (red) and medial prefrontal cortex (PFC) (green). The cluster images are thresholded at $p_{\text{Unc}} < 0.005$, $k \geq 5$ and overlaid onto the mean high-resolution T1 image. (B) Parameter estimates for the left ventral striatum. (C) Parameter estimates for the medial PFC. (D) Parameter estimates for the right ventral striatum. a.u., arbitrary units.

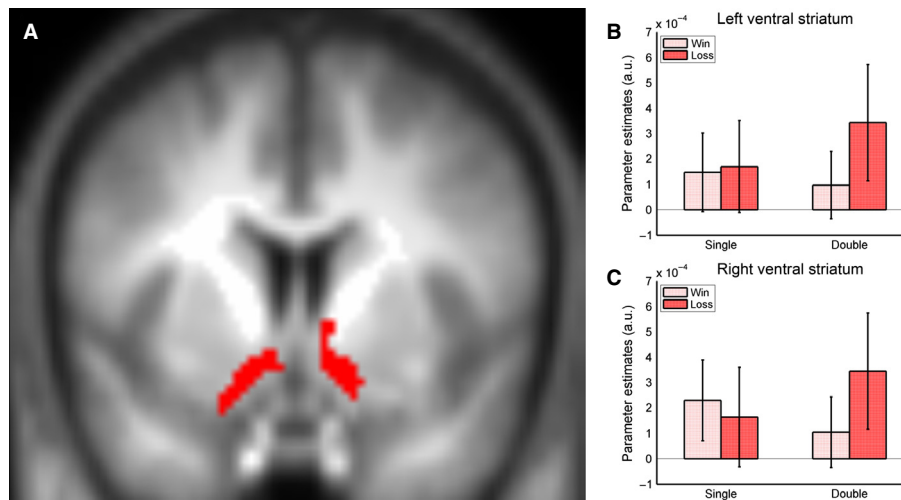


FIG. 4. Regions showing significant value-related signals to gains and losses at the time of outcome. Activity in this area increases as both gains and losses become more extreme. Using a small volume correction on the respective *a priori* coordinates the activities for each of the regions shown were significant at $P < 0.05$ family-wise error. Each of the displayed parameter estimates was obtained by running an independent analysis on the 10 mm sphere at the *a priori* coordinates. (A) Coronal slice at $Y = 5$ displaying significant clusters in the ventral striatum. The cluster image is thresholded at $p_{Unc} < 0.005$, $k \geq 5$ and overlaid onto the mean high-resolution T1 image. (B) Parameter estimates for the left ventral striatum. (C) Parameter estimates for the right ventral striatum. a.u., arbitrary units.

Value-related loss signals: areas showing increasing activity to increasing magnitude of losses

We then tested for areas showing increasing activity to increasing magnitude of losses [contrast: (pmod outcome single loss, pmod outcome double loss) > 0]. This contrast revealed activity in three areas only: the lateral orbitofrontal cortex bilaterally, insular cortex bilaterally and anterior cingulate cortex (Fig. 5). Plots of parameter estimates in all three regions showed that value-related activity in these areas was driven predominantly by losses and not gains. Using a more strict comparison, all three of these areas also showed significantly greater responsiveness to changes in loss magnitude and not changes in gain magnitude (this comparison was conducted by exclusively masking the parametrically modulated signal for losses with the parametrically modulated signal for gains at a $P < 0.10$ threshold, and then using the same statistical thresholds as all previous comparisons; see Table 1 for relevant statistics), further confirming that these regions have selective parametric magnitude signals to loss outcomes. Plots of the parametric effects for gains and losses for voxels identified in this contrast are shown in Fig. S4. Using the strictest comparison, we tested for areas showing a significantly greater increase to increasing magnitude of losses when compared with increasing activity for increasing magnitude of gains [contrast: (pmod outcome single loss, pmod outcome double loss) – (pmod outcome single gain, pmod outcome double gain)]. This contrast revealed activity in the same three regions: significant effects in both the right lateral orbitofrontal cortex and anterior cingulate cortex, and at marginal significance ($P < 0.06$) activity in the left insula. Interestingly, no voxels survived our whole brain threshold when testing for areas that showed increasing activity to decreasing magnitude of losses.

Signals related to differences in expected value at the time of choice

Although not a main focus of our study, we also tested whether activity at the time of choice was different in the medial prefrontal cortex as a function of whether participants were choosing from high compared with low EV trials [contrast: (choice single high EV,

choice double high EV) – (choice single low EV, choice double low EV)]. We therefore compared average activity at the time of choice in high vs. low EV trials and found significant activity in a region of the anterior medial prefrontal cortex, a region overlapping those found in numerous previous studies examining neural responses to EV during choice (Fig. 6). This finding validated our paradigm, indicating that participants were encoding differences in EV during the decision-making phase of the task. There were no areas reaching significance for the reverse contrast at the time of choice (e.g. low EV $>$ high EV).

Overlap between different types of outcome coding

In an effort to better visualise areas that have exclusive responses vs. those that respond to multiple components of outcomes, Fig. 7 presents areas showing binary responses to outcomes alongside areas showing value-related responses in order to demonstrate the degree of overlap between the representations. As can be seen, a number of regions are involved in encoding informational signals about outcomes as well as parametric signals. In particular, both the insular cortex and lateral orbitofrontal cortex are involved in encoding binary outcome signals (for both gains and losses) as well as encoding parametric loss signals. Furthermore, the ventral striatum is involved in uniquely encoding both binary gain signals and parametric gain and loss signals.

Discussion

Neuroimaging studies investigating neural responses to rewarding and punishing feedback have been ongoing for more than 15 years (Bartra *et al.*, 2013). In that time, a number of brain structures have been implicated as encoding rewarding and punishing outcomes including the orbitofrontal and medial prefrontal cortices, ventral striatum and insular cortex. In spite of the large number of studies, there is still considerable disagreement in the literature concerning precisely how these outcomes are encoded and, in particular, whether it is the case that different regions or subregions are

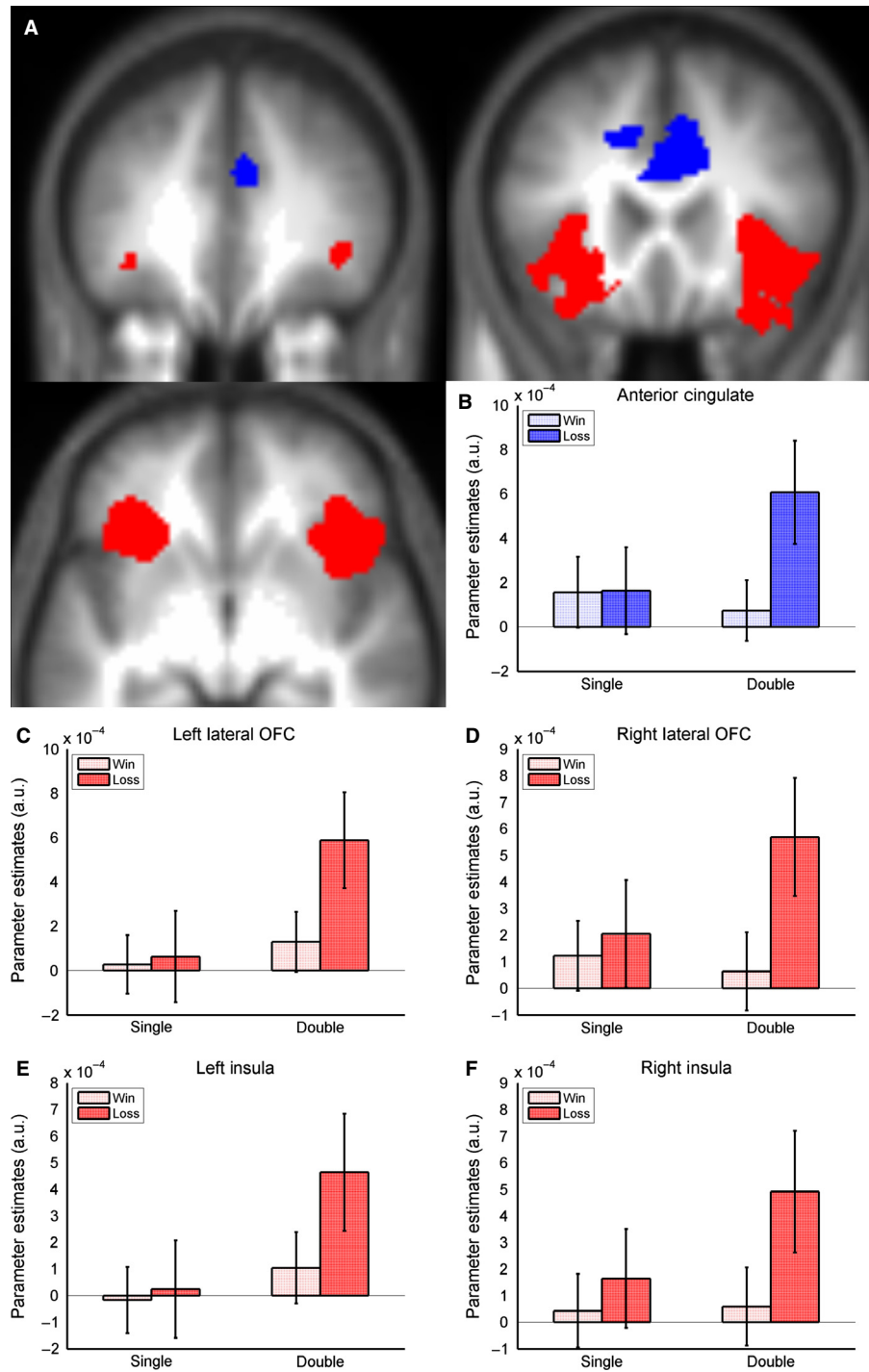


FIG. 5. Regions showing significant value-related signals to losses at the time of outcome. Activity in these areas increases as losses become more extreme. Using a small volume correction on the respective *a priori* coordinates the activities for each of the regions shown were significant at $P < 0.05$ family-wise error. Each of the displayed parameter estimates was obtained by running an independent analysis on the 10 mm sphere at the *a priori* coordinates. (A) Two coronal slices at $Y = 35$ and $Y = 20$, respectively, and an axial slice ($Z = -9$) displaying significant clusters in both the anterior cingulate cortex (blue) and insula and lateral OFC (red). The cluster images are thresholded at $p_{unc} < 0.005$, $k \geq 5$ and overlaid onto the mean high-resolution T1 image. (B) Parameter estimates for the ACC. (C) Parameter estimates for the left lateral orbitofrontal cortex (OFC). (D) Parameter estimates for the right lateral OFC. (E) Parameter estimates for the left insula. (F) Parameter estimates for the right insula. a.u., arbitrary units.

uniquely involved in encoding rewarding and punishing outcomes. The present study makes clear the importance of distinguishing between the informational coding of outcomes and the separable value-related signals that encode the magnitude of a particular outcome.

We found evidence for the involvement of a large corpus of brain areas in encoding informational signals that distinguish rewarding or punishing outcomes (gains and losses) from a status quo outcome (in which neither a monetary gain or a loss is obtained). Most of these areas responded equally to rewarding and punishing events.

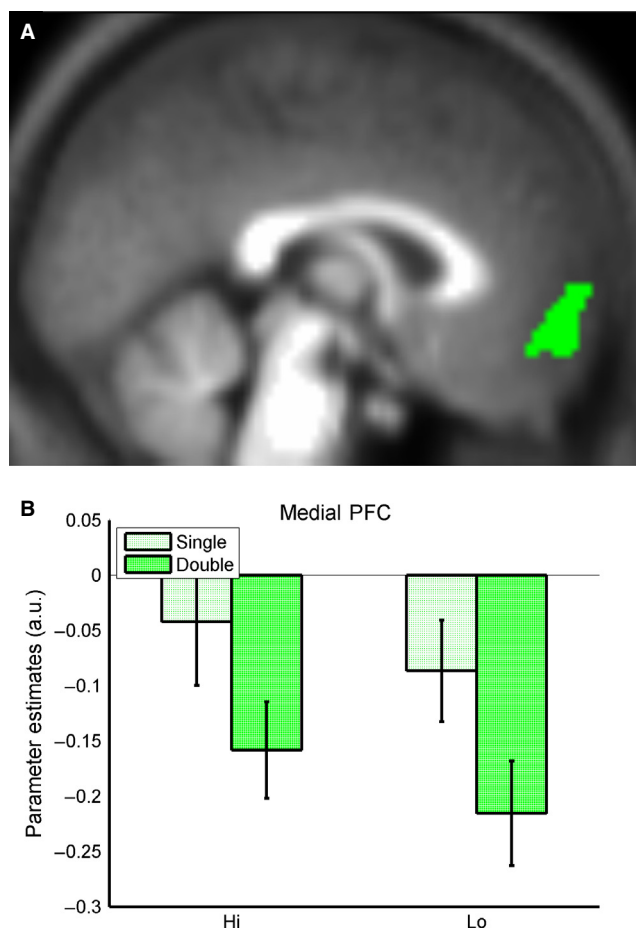


FIG. 6. Responses to EV. Region showing significantly greater activity at the time of choice when choosing from high EV compared with low EV trials. Using a small volume correction on the respective *a priori* coordinates the activity for the region shown was significant at $P < 0.05$ family-wise error. Each of the displayed parameter estimates was obtained by running an independent analysis on the 10 mm sphere at the *a priori* coordinates. (A) Sagittal slice at $X = 0$ displaying a significant cluster in the medial prefrontal cortex (PFC). The cluster image is thresholded at $p_{\text{Unc}} < 0.005$, $k \geq 5$ and overlaid onto the mean high-resolution T1 image. (B) Parameter estimates for the medial PFC. a.u., arbitrary units.

This included the lateral orbitofrontal cortex, anterior insular cortex, and anterior cingulate cortex, each showing an increase in activity to both rewarding and punishing events irrespective of the magnitude of those outcomes. A different set of regions also responded to informational signals about outcomes in a manner that differentiated rewarding from punishing outcomes: a region of the medial prefrontal cortex as well as part of the ventral striatum responded selectively to the receipt of rewarding but not punishing outcomes. These findings are broadly consistent with a report by Vickery *et al.* (2011) who used multi-variate analysis techniques to decode rewarding and punishing outcome signals. They reported outcome signals to be present in a wide variety of brain areas that overlapped with some of the areas listed here. However, because we used outcomes with differing magnitudes as opposed to binary outcomes we were further able to distinguish informational outcome signals from value signals, whereas in the previous study those signals were probably aliased together.

Another previous study by Elliott *et al.* (2010) aimed to distinguish a different kind of informational signal from value signals. In that particular study, the informational signal in question corre-

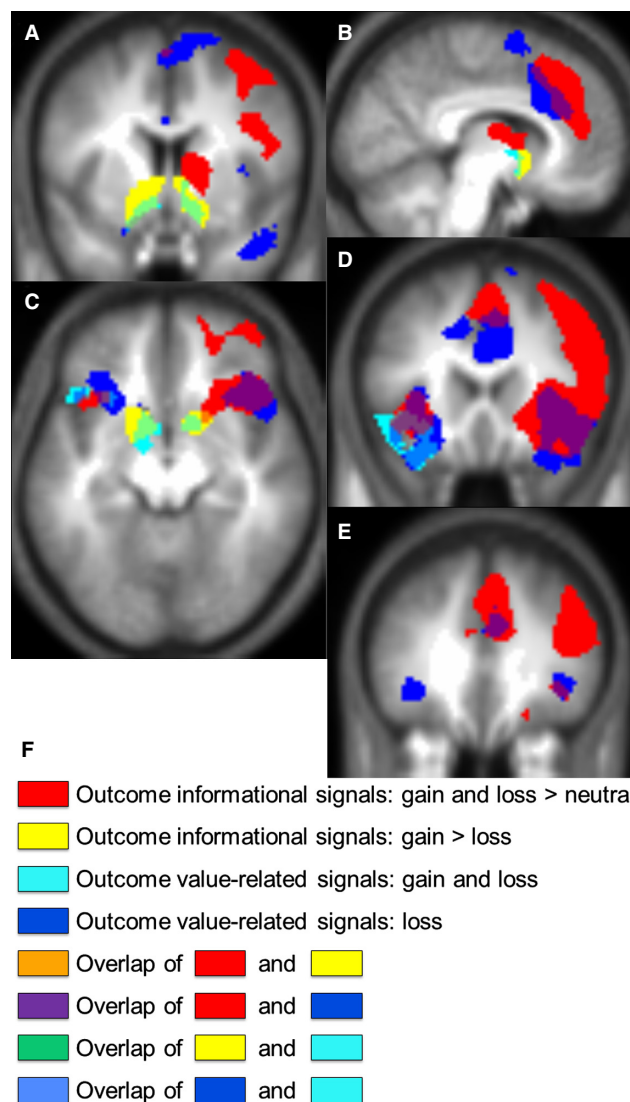


FIG. 7. Overlap of informational and valued-related outcome signals in the brain. (A) Coronal slice at $Y = 5$, highlighting ventral striatum activity. (B) Sagittal slice at $X = 9$, highlighting anterior cingulate and ventral striatum activity. (C) Axial slice at $Z = -14$, highlighting ventral striatum, insula, and lateral orbitofrontal cortex activity. (D) Coronal slice at $Y = 18$, highlighting insula activity. (E) Coronal slice at $Y = 32$, highlighting lateral orbitofrontal cortex activity. (F) Legend displaying colors representing each contrast and overlap between contrasts. Note that, whereas the left insular cortex is shown for the contrast of outcome value signals to gains and losses (teal) as well as in the contrast of outcome value signals to losses only (blue), inspection of parameter estimates show that this area responds especially to losses and is not at all responsive to gains; indeed in a direct test this area responds significantly to losses and not to gains (see Table 1). Thus, the only area in our study found to actually respond in a value-related manner to both increasing losses and increasing gains is the ventral striatum.

sponded to whether or not the participant made a strategic decision to stay or switch their choice of behavior following detection of a reversal of contingencies in the context of a reversal learning task. The informational signal in the present study concerns whether or not a particular outcome is rewarding or punishing feedback, as opposed to whether or not an individual will maintain or switch current behavior. When taken together, these findings support the notion that the lateral orbitofrontal cortex represents multiple types of task-relevant signals, suggesting that this area may play a very general role in encoding task-relevant information.

TABLE 1. Statistical results for significant clusters

Cluster level		Voxel level		x;y;z	L/R/C	Structure
<i>pFWE</i>	<i>k</i>	<i>pFWE</i>	<i>Z</i>			
Outcome informational signals: gain and loss > neutral						
0.016	203	0.0001	4.51	[4;30;38]	R	ACC
0.030	53	0.0009	3.96	[42;44;0]	R	IOFC
0.009	368	0.0000	5.27	[36;20;−4]	R	Insula
0.025	91	0.0028	3.62	[−36;20;−4]	L	Insula
Outcome informational signals: gain > loss						
0.029	63	0.0161	3.02	[−2;42;−8]	L	mPFC
0.021	136	0.0007	4.01	[14;4;−16]	R	Ventral striatum
0.016	205	0.0000	4.79	[−12;6;−12]	L	Ventral striatum
Outcome value-related signals: gain and loss						
0.031	40	0.0132	3.04	[14;6;−16]	R	Ventral striatum
0.034	21	0.0149	3.00	[−10;6;−14]	L	Ventral striatum
Outcome value-related signals: loss						
0.008	402	0.0065	3.35	[10;22;28]	R	ACC
0.042	6	0.0324	2.74	[−36;34;−10]	L	IOFC
0.038	15	0.0137	3.08	[38;32;−4]	R	IOFC
0.013	262	0.0102	3.19	[34;18;2]	R	Insula
0.023	108	0.0084	3.26	[−32;20;4]	L	Insula
Outcome value-related signals: loss and not gain						
0.010	343	0.0063	3.35	[10;22;28]	R	ACC
0.041	6	0.0313	2.74	[−36;34;−10]	L	IOFC
0.037	15	0.0132	3.08	[38;32;−4]	R	IOFC
0.014	251	0.0099	3.19	[34;18;2]	R	Insula
0.024	99	0.0081	3.26	[−32;20;4]	L	Insula
Outcome value-related signals: loss > gain						
0.0489	8	0.0452	2.68	[12;28;26]	C	ACC
0.0494	5	0.0440	2.68	[38;32;−6]	R	IOFC
0.0528	3	0.0505	2.63	[−34;12;−2]	L	Insula
Choice EV: high > low						
0.022	113	0.0020	3.65	[0;54;−12]	C	mPFC

All clusters listed were significant at the $P < 0.05$ corrected level. Only the peak voxel statistics for each region and contrast combination are reported. *k*, number of contiguous voxels; *pFWE*, family-wise error *P*-value; *pFDR*, false discovery rate *P*-value; L, left; R, right; C, center; ACC, anterior cingulate cortex; IOFC, lateral orbitofrontal cortex; mPFC, medial prefrontal cortex.

In addition to informational encoding of outcomes, we also observed value-related outcome signals in a circumscribed set of brain regions. In particular, we found aversive-going value signals (activity that increased as a function of increasing magnitude of losses) in only three brain regions: the bilateral lateral orbitofrontal cortex, anterior insular cortex, and anterior cingulate cortex. Regarding the lateral orbitofrontal cortex, a number of previous studies have revealed activity in this area correlating with receipt of aversive outcomes (O'Doherty *et al.*, 2001; Small *et al.*, 2001; Ursu & Carter, 2005), findings that led to a proposal that the lateral orbitofrontal cortex is specialised for encoding aversive outcomes (Kringelbach & Rolls, 2004). The present findings do confirm a role for the lateral orbitofrontal cortex in value-related encoding of aversive outcomes. However, our findings also show that this area is involved more generally in encoding informational properties of outcomes, irrespective of whether those outcomes are rewarding or punishing. Thus, in order to understand the functions of the lateral orbitofrontal cortex overall, it is necessary to take into account the fact that this region may play multiple roles in encoding different features of outcomes. A recent study by Noonan *et al.* (2011) reported activity in the lateral orbitofrontal cortex in response to 'errors', i.e. negative feedback after participants made an incorrect choice during performance of an associative learning task. They also reported activity in response to correct feedback in the same region,

interpreting it as evidence against the notion that the lateral orbitofrontal cortex is involved specifically in encoding aversive outcomes. The present findings do corroborate those results in that we do find lateral orbitofrontal cortex engagement in response to both rewarding and punishing feedback. However, our findings also reconcile those results with past findings on aversive-related encoding in the lateral orbitofrontal cortex by showing that this region plays two different roles in outcome processing: one in encoding informational properties of outcomes and another in encoding loss magnitude.

The anterior insular cortex was also found to encode aversive values in the present study. As with the lateral orbitofrontal cortex, this area was found to encode informational outcome signals for both rewarding and punishing outcomes. These findings also potentially help to resolve ambiguity in the literature on the functions of the anterior insula. Although some have suggested that this region is involved in encoding 'arousal' or 'salience' such that this area responds to sufficiently arousing rewarding and punishing events irrespective of the valence (Bartra *et al.*, 2013), others have emphasised a role for the anterior insula in aversive encoding in particular (O'Doherty *et al.*, 2003a; Seymour *et al.*, 2004; Palminteri *et al.*, 2012). The present study suggests that the insula may indeed respond to both rewarding and punishing outcomes; however, such responses do not reflect 'salience' *per se*, because they do not scale

with the absolute magnitude of the outcome but rather conform to informational representations about outcome identity, in that they simply encode a binary response (showing an increase for both rewarding and punishing outcomes irrespective of magnitude). However, value-related outcome signals (that do scale with magnitude) in the insula appear to be predominantly aversive-related and not reward-related. Previous studies reporting 'salience' type codes may have potentially conflated informational signals with value signals.

The only region where we did find clear evidence of bivalent encoding of outcome values was the ventral striatum: this area encoded both reward-related and aversive-related outcome value signals. Intriguingly, this region also encoded informational outcome signals but only for rewarding and not punishing outcomes. The ventral striatum has been consistently implicated in reward-related processing, particularly anticipation of rewarding outcomes, as well as reward-prediction errors (Knutson *et al.*, 2001). However, as with the other brain regions discussed above, the literature has been somewhat inconsistent with regard to the specific role of the ventral striatum in aversive coding. Although some studies have exclusively reported activity in the ventral striatum in response to reward-related processes (Delgado *et al.*, 2000; Knutson *et al.*, 2001; Kim *et al.*, 2006), others have found increased activity in the striatum in response to aversive events such as aversive outcomes, aversive predictions and aversive prediction errors (Becerra *et al.*, 2001; Jensen *et al.*, 2003; Seymour *et al.*, 2004, 2005). The present findings indicate that outcomes might be represented in two different ways in this region: whereas informational signals might be predominantly reward-driven in the ventral striatum, both gains and losses elicit value-related outcome responses in this area, a response profile that could be interpreted as being more consistent with a 'salience'-type response (Zink *et al.*, 2004). Unlike many previous studies from ourselves and others (McClure *et al.*, 2003; O'Doherty *et al.*, 2003b, 2004; Abler *et al.*, 2006; Kim *et al.*, 2006), we did not explicitly design our current experiment to maximise sensitivity to detecting prediction error signals. This does not preclude the presence of prediction errors in the striatum during this task, and it is indeed possible that aspects of the signals that we report as responding to outcomes reflect elements of prediction error responses. However, given our task design, the 'salience' type of value response that we found here does not correspond to the type of signal that would be expected in a prediction error code, as a reward-related (or punishment-related) prediction error should differentially respond to prediction errors generated by unexpected gains and losses (e.g. a signed reward prediction error should increase for unexpected gains, and decrease for unexpected losses, which is not what we found in the ventral striatum for the value-related signals). Our preferred interpretation is that it is likely that the ventral striatum is involved in multiple functions, i.e. not only are prediction errors represented in this region (as shown in many previous studies) but also outcome as well as anticipatory value signals. The present study, by virtue of its design, is most sensitive to detecting outcome signals as opposed to prediction errors or anticipatory value signals.

Another feature of the experimental design is that we included two different trial types: double trials, in which participants could make choices between two different slot machines, and single trials, in which only a single slot machine was presented and, thus, participants had no choice about which slot machine to select. We included this feature to enable us to determine the extent to which outcome signals are modulated by the presence or absence of a choice. We did not find any significant main effect of choice (i.e. set size) on activity at the time of outcome in any of the areas in which we found outcome responses, and nor did we find any

significant outcome valence $\times$ set size interactions. As a consequence, we collapsed across the set size conditions in the analyses. Nevertheless, in the plots of parameter estimates, an effect of set size is apparent throughout the different brain regions assayed, such that activity appears to be elevated in response to outcomes in the double condition compared with the single (no choice) condition. However, given that these effects did not reach significance in our statistical tests, we refrain from drawing strong conclusions about them other than acknowledging that there appears to be a trend in this direction throughout the areas in which outcome signals were detected.

In the present study we focused on responses to outcomes, differentiating informational encoding of outcomes from value-related encoding. Although we did replicate previous findings of a role for the medial prefrontal cortex in encoding expectations of future reward (Tanaka *et al.*, 2004; Daw *et al.*, 2006; Hampton *et al.*, 2006), the study was not designed to focus on encoding of predictions about future outcomes, nor on encoding of decision utilities. The coding mechanisms used for encoding anticipatory value signals and for encoding decision utilities may be different to that used for coding outcome values. Indeed, previous studies have found that parametric signals throughout the ventromedial prefrontal cortex and striatum correlated with decision values that increase for increasing prospective gains or appetitive food items, and decrease for an increasing magnitude of prospective losses or aversive food items (Tom *et al.*, 2007; Plassmann *et al.*, 2010). However, in those studies no region was found to show increasing activity to an increasing magnitude of prospective monetary losses or to the prospect of receiving increasingly aversive food outcomes. When taken together with the present results, these findings suggest that outcome value and decision values may deploy very different coding strategies in the brain. Whereas decision values show an increasing response to increasing predictions of gain and a decreasing response to increasing predictions of loss, outcome values may be coded in a more heterogeneous manner: some outcome codes show increasing activity to both increasing gains and losses [such as in the ventral striatum in the present study; see also Weis *et al.* (2013)], other outcome codes also show increasing activity to increasing losses, such as the lateral orbitofrontal cortex in the present study. It is also commonly reported that activity in the medial orbitofrontal and prefrontal cortices correlates positively with increasing outcome value (Knutson *et al.*, 2001; O'Doherty *et al.*, 2001; Rohe *et al.*, 2012). Such a difference in coding strategy for decision utilities vs. outcomes might reflect differences in the type of computations that animals need to perform when making a decision vs. receiving and interpreting an outcome. For example, at the time of decision making, an animal needs to integrate costs and benefits into a unitary common currency for the purposes of making a choice (Shizgal, 1997; Montague & Berns, 2002; Chib *et al.*, 2009), and this would benefit by utilising unitary neural coding. Continuing, at the time of outcome reception, the animal needs to organise its behavioral responses to that outcome, a response that may differ dramatically depending on whether (i) the outcome is appetitive or aversive and (ii) the magnitude of the outcome, irrespective of valence. Thus, it is often under-appreciated in the literature that outcome value codes might need to be represented in a very different way in the brain than decision or anticipatory outcome signals, a proposition supported by our present results.

Nonetheless, distinguishing informational signals from value-related signals is also pertinent when examining responses during anticipation of or when making decisions about outcomes because such cues or actions associated with particular outcomes can also retrieve representations of outcomes related to identity or other

non-value-related features beyond merely retrieving outcome values. Thus, it would not be surprising if some of the areas identified in the present study found to encode non-value-related properties of an outcome are also involved in responding to cues that predict outcomes in a manner independent of value. Precisely such cue-elicited stimulus identity signals have been reported during a simple associative learning task (Klein-Flugge *et al.*, 2013), and a value-based decision-making task (McNamee *et al.*, 2013) in the lateral orbitofrontal cortex, as well in parts of the medial orbitofrontal cortex.

To conclude, the present findings add an important dimension to our understanding of how outcome signals are encoded in the brain. We find informational signals that respond differentially to rewarding outcomes and punishing outcomes compared with neutral outcomes in a number of brain areas – these signals overlap substantially with signals encoding value for punishing and/or rewarding feedback. More specifically, our findings help to clarify the functions of the lateral orbitofrontal cortex and insular cortex in outcome coding by showing that, whereas both rewarding and punishing outcomes are represented in these areas in a non-value-related manner, these areas also appear to be especially involved in encoding aversive outcome value signals. Furthermore, the implications of the present study go beyond neuroimaging studies to other approaches used to study value-related coding in the brain such as in single- or multi-unit neurophysiological recordings. In order to identify the actual functions of a particular neuron in responding to an outcome (or to predictors of outcomes), it is important to appreciate that such neural responses may correspond not only to value-related responses but also to other outcome features such as stimulus properties of the outcome that might differ depending on the type of outcome involved (e.g. sensory differences between a sweet and bitter taste). Therefore, even in neurophysiological studies, reports of differential neuronal responses to outcomes may sometimes conflate apparent differences in value coding with actual differences in stimulus feature coding unless careful steps are taken to dissociate value from other stimulus properties [see Padoa-Schioppa & Assad (2006) for an example of successfully differentiating expected utility and stimulus properties in central parts of the monkey orbitofrontal cortex]. We argue that only by considering multiple features of outcome representations – including but not limited to both informational signals and value-related signals – will it be possible to gain a more complete understanding of the contributions made by multiple neural systems and their constituent neurons when encoding outcomes.

Supporting Information

Additional supporting information can be found in the online version of this article:

Fig. S1. Schematic illustrating informational signals, salience and value-signals in the context of the present task.

Fig. S2. Schematic indicating the location of each ROI used to correct for small volume in each of our *a priori* areas of interest, shown overlaid on the subject averaged structural scan in MNI space.

Fig. S3. Plot of activity in the ventral striatum showing significant value-related signals to gains and losses at the time of outcome, indicative of a 'salience' type response.

Fig. S4. Parametric modulations from regions showing significant value-related signals to losses at the time of outcome.

Acknowledgements

This work was funded by the Science Foundation Ireland (grant 08/IN.1/B1844) and a grant from the Gordon and Betty Moore Foundation to J.P.O'D. We thank Teresa Furey, Simon Dunne, and Sojo Josephs for their help with participant recruitment and data collection.

Abbreviations

EV, expected value; fMRI, functional magnetic resonance imaging; GLM, general linear model; pmod, parametric modulator.

References

- Abler, B., Walter, H., Erk, S., Kammerer, H. & Spitzer, M. (2006) Prediction error as a linear function of reward probability is coded in human nucleus accumbens. *NeuroImage*, **31**, 790–795.
- Anderson, A.K., Christoff, K., Stappen, I., Panitz, D., Ghahremani, D.G., Glover, G., Gabrieli, J.D. & Sobel, N. (2003) Dissociated neural representations of intensity and valence in human olfaction. *Nat. Neurosci.*, **6**, 196–202.
- Bartra, O., McGuire, J.T. & Kable, J.W. (2013) The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage*, **76**, 412–427.
- Becerra, L., Breiter, H.C., Wise, R., Gonzalez, R.G. & Borsook, D. (2001) Reward circuitry activation by noxious thermal stimuli. *Neuron*, **32**, 927–946.
- Breiter, H.C., Aharon, I., Kahneman, D., Dale, A. & Shizgal, P. (2001) Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron*, **30**, 619–639.
- Chib, V.S., Rangel, A., Shimojo, S. & O'Doherty, J.P. (2009) Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *J. Neurosci.*, **29**, 12315–12320.
- Daw, N.D., O'Doherty, J.P., Dayan, P., Seymour, B. & Dolan, R.J. (2006) Cortical substrates for exploratory decisions in humans. *Nature*, **441**, 876–879.
- Deichmann, R., Gottfried, J.A., Hutton, C. & Turner, R. (2003) Optimized EPI for fMRI studies of the orbitofrontal cortex. *NeuroImage*, **19**, 430–441.
- Delgado, M.R., Nystrom, L.E., Fissell, C., Noll, D.C. & Fiez, J.A. (2000) Tracking the hemodynamic responses to reward and punishment in the striatum. *J. Neurophysiol.*, **84**, 3072–3077.
- Diekhof, E.K., Kaps, L., Falkai, P. & Gruber, O. (2012) The role of the human ventral striatum and the medial orbitofrontal cortex in the representation of reward magnitude – an activation likelihood estimation meta-analysis of neuroimaging studies of passive reward expectancy and outcome processing. *Neuropsychologia*, **50**, 1252–1266.
- Elliott, R., Friston, K.J. & Dolan, R.J. (2000) Dissociable neural responses in human reward systems. *J. Neurosci.*, **20**, 6159–6165.
- Elliott, R., Agnew, Z. & Deakin, J.F. (2010) Hedonic and informational functions of the human orbitofrontal cortex. *Cereb. Cortex*, **20**, 198–204.
- Hampshire, A., Chaudhry, A.M., Owen, A.M. & Roberts, A.C. (2012) Dissociable roles for lateral orbitofrontal cortex and lateral prefrontal cortex during preference driven reversal learning. *NeuroImage*, **59**, 4102–4112.
- Hampton, A.N., Bossaerts, P. & O'Doherty, J.P. (2006) The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *J. Neurosci.*, **26**, 8360–8367.
- Hayes, D.J. & Northoff, G. (2011) Identifying a network of brain regions involved in aversion-related processing: a cross-species translational investigation. *Front. Integr. Neurosci.*, **5**, 49.
- Jensen, J., McIntosh, A.R., Crawley, A.P., Mikulis, D.J., Remington, G. & Kapur, S. (2003) Direct activation of the ventral striatum in anticipation of aversive stimuli. *Neuron*, **40**, 1251–1257.
- Kahnt, T., Heinzle, J., Park, S.Q. & Haynes, J.D. (2010) The neural code of reward anticipation in human orbitofrontal cortex. *Proc. Natl. Acad. Sci. USA*, **107**, 6010–6015.
- Kim, H., Shimojo, S. & O'Doherty, J.P. (2006) Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. *PLoS Biol.*, **4**, e233.
- Klein-Flugge, M.C., Barron, H.C., Brodersen, K.H., Dolan, R.J. & Behrens, T.E. (2013) Segregated encoding of reward-identity and stimulus-reward associations in human orbitofrontal cortex. *J. Neurosci.*, **33**, 3202–3211.
- Knutson, B., Fong, G.W., Adams, C.M., Varner, J.L. & Hommer, D. (2001) Dissociation of reward anticipation and outcome with event-related fMRI. *NeuroReport*, **12**, 3683–3687.
- Kringelbach, M.L. & Rolls, E.T. (2004) The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. *Prog. Neurobiol.*, **72**, 341–372.
- Litt, A., Plassmann, H., Shiv, B. & Rangel, A. (2011) Dissociating valuation and saliency signals during decision-making. *Cereb. Cortex*, **21**, 95–102.

- McClure, S.M., Berns, G.S. & Montague, P.R. (2003) Temporal prediction errors in a passive learning task activate human striatum. *Neuron*, **38**, 339–346.
- McNamee, D., Rangel, A. & O'Doherty, J.P. (2013) Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. *Nat. Neurosci.*, **16**, 479–485.
- Montague, P.R. & Berns, G.S. (2002) Neural economics and the biological substrates of valuation. *Neuron*, **36**, 265–284.
- Morrison, S.E. & Salzman, C.D. (2009) The convergence of information about rewarding and aversive stimuli in single neurons. *J. Neurosci.*, **29**, 11471–11483.
- Nitschke, J.B., Dixon, G.E., Sarinopoulos, I., Short, S.J., Cohen, J.D., Smith, E.E., Kosslyn, S.M., Rose, R.M. & Davidson, R.J. (2006a) Altering expectancy dampens neural response to aversive taste in primary taste cortex. *Nat. Neurosci.*, **9**, 435–442.
- Nitschke, J.B., Sarinopoulos, I., Mackiewicz, K.L., Schaefer, H.S. & Davidson, R.J. (2006b) Functional neuroanatomy of aversion and its anticipation. *NeuroImage*, **29**, 106–116.
- Noonan, M.P., Mars, R.B. & Rushworth, M.F. (2011) Distinct roles of three frontal cortical areas in reward-guided behavior. *J. Neurosci.*, **31**, 14399–14412.
- Noonan, M.P., Kolling, N., Walton, M.E. & Rushworth, M.F. (2012) Re-evaluating the role of the orbitofrontal cortex in reward and reinforcement. *Eur. J. Neurosci.*, **35**, 997–1010.
- O'Doherty, J., Kringelbach, M.L., Rolls, E.T., Hornak, J. & Andrews, C. (2001) Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat. Neurosci.*, **4**, 95–102.
- O'Doherty, J., Critchley, H., Deichmann, R. & Dolan, R.J. (2003a) Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. *J. Neurosci.*, **23**, 7931–7939.
- O'Doherty, J., Dayan, P., Friston, K., Critchley, H. & Dolan, R.J. (2003b) Temporal difference models and reward-related learning in the human brain. *Neuron*, **38**, 329–337.
- O'Doherty, J., Winston, J., Critchley, H., Perrett, D., Burt, D.M. & Dolan, R.J. (2003c) Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia*, **41**, 147–155.
- O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K. & Dolan, R.J. (2004) Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, **304**, 452–454.
- Padoa-Schioppa, C. & Assad, J.A. (2006) Neurons in the orbitofrontal cortex encode economic value. *Nature*, **441**, 223–226.
- Palmeri, S., Justo, D., Jauffret, C., Pavlicek, B., Dauta, A., Delmaire, C., Czernecki, V., Karachi, C., Capelle, L., Durr, A. & Pessiglione, M. (2012) Critical roles for anterior insula and dorsal striatum in punishment-based avoidance learning. *Neuron*, **76**, 998–1009.
- Plassmann, H., O'Doherty, J.P. & Rangel, A. (2010) Appetitive and aversive goal values are encoded in the medial orbitofrontal cortex at the time of decision making. *J. Neurosci.*, **30**, 10799–10808.
- Ploghaus, A., Tracey, I., Gati, J.S., Clare, S., Menon, R.S., Matthews, P.M. & Rawlins, J.N. (1999) Dissociating pain from its anticipation in the human brain. *Science*, **284**, 1979–1981.
- Rohe, T., Weber, B. & Fließbach, K. (2012) Dissociation of BOLD responses to reward prediction errors and reward receipt by a model comparison. *Eur. J. Neurosci.*, **36**, 2376–2382.
- Seymour, B., O'Doherty, J.P., Dayan, P., Koltzenburg, M., Jones, A.K., Dolan, R.J., Friston, K.J. & Frackowiak, R.S. (2004) Temporal difference models describe higher-order learning in humans. *Nature*, **429**, 664–667.
- Seymour, B., O'Doherty, J.P., Koltzenburg, M., Wiech, K., Frackowiak, R., Friston, K. & Dolan, R. (2005) Opponent appetitive-aversive neural processes underlie predictive learning of pain relief. *Nat. Neurosci.*, **8**, 1234–1240.
- Shizgal, P. (1997) Neural basis of utility estimation. *Curr. Opin. Neurobiol.*, **7**, 198–208.
- Small, D.M., Zatorre, R.J., Dagher, A., Evans, A.C. & Jones-Gotman, M. (2001) Changes in brain activity related to eating chocolate: from pleasure to aversion. *Brain*, **124**, 1720–1733.
- Smith, D.V., Hayden, B.Y., Truong, T.K., Song, A.W., Platt, M.L. & Huetzel, S.A. (2010) Distinct value signals in anterior and posterior ventromedial prefrontal cortex. *J. Neurosci.*, **30**, 2490–2495.
- Tanaka, S.C., Doya, K., Okada, G., Ueda, K., Okamoto, Y. & Yamawaki, S. (2004) Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nat. Neurosci.*, **7**, 887–893.
- Tom, S.M., Fox, C.R., Trepel, C. & Poldrack, R.A. (2007) The neural basis of loss aversion in decision-making under risk. *Science*, **315**, 515–518.
- Ursu, S. & Carter, C.S. (2005) Outcome representations, counterfactual comparisons and the human orbitofrontal cortex: implications for neuroimaging studies of decision-making. *Brain Res.*, **23**, 51–60.
- Vickery, T.J., Chun, M.M. & Lee, D. (2011) Ubiquity and specificity of reinforcement signals throughout the human brain. *Neuron*, **72**, 166–177.
- Weis, T., Puschmann, S., Brechmann, A. & Thiel, C.M. (2013) Positive and negative reinforcement activate human auditory cortex. *Front. Hum. Neurosci.*, **7**, 842.
- Zink, C.F., Pagnoni, G., Martin-Skurski, M.E., Chappelow, J.C. & Berns, G.S. (2004) Human striatal responses to monetary reward depend on saliency. *Neuron*, **42**, 509–517.