doi:10.1093/brain/awv409 BRAIN 2016: 139; 1295–1309 | 1295

BRAIN
A JOURNAL OF NEUROLOGY

The neural dynamics of reward value and risk coding in the human orbitofrontal cortex

Yansong Li,[1,2,*] Giovanna Vanni-Mercier,[1,2] Jean Isnard,[2,3] François Mauguière[2,3] and Jean-Claude Dreher[1,2]

See Kringelbach and Rapuano (doi:10.1093/brain/aww049) for a scientific commentary on this article.

The orbitofrontal cortex is known to carry information regarding expected reward, risk and experienced outcome. Yet, due to inherent limitations in lesion and neuroimaging methods, the neural dynamics of these computations has remained elusive in humans. Here, taking advantage of the high temporal definition of intracranial recordings, we characterize the neurophysiological signatures of the intact orbitofrontal cortex in processing information relevant for risky decisions. Local field potentials were recorded from the intact orbitofrontal cortex of patients suffering from drug-refractory partial epilepsy with implanted depth electrodes as they performed a probabilistic reward learning task that required them to associate visual cues with distinct reward probabilities. We observed three successive signals: (i) around 400 ms after cue presentation, the amplitudes of the local field potentials increased with reward probability; (ii) a risk signal emerged during the late phase of reward anticipation and during the outcome phase; and (iii) an experienced value signal appeared at the time of reward delivery. Both the medial and lateral orbitofrontal cortex encoded risk and reward probability while the lateral orbitofrontal cortex played a dominant role in coding experienced value. The present study provides the first evidence from intracranial recordings that the human orbitofrontal cortex codes reward risk both during late reward anticipation and during the outcome phase at a time scale of milliseconds. Our findings offer insights into the rapid mechanisms underlying the ability to learn structural relationships from the environment.

1 Neuroeconomics, Reward and Decision-making Team, Cognitive Neuroscience Centre, CNRS UMR 5229, Bron 69675, France
2 Université Claude Bernard Lyon 1, Lyon 69100, France
3 Neurological Hospital, Bron 69675, France

*Present address: Department of Psychology, School of Social and Behavioural Sciences, Nanjing University, Nanjing, China

Correspondence to: Jean-Claude Dreher, PhD,
Reward and Decision-making Team,
Cognitive Neuroscience Centre,
CNRS, UMR 5229,
67 Bd Pinel, 69675 Bron, France,
E-mail: dreher@isc.cnrs.fr

Correspondence may also be addressed to: Yansong Li,
Department of Psychology, School of Social and Behavioural Sciences, Nanjing University, Nanjing, China
E-mail: yansongli@nju.edu.cn

Keywords: OFC; iEEG; reward probability; experienced value; risk

Abbreviations: LFP = local field potential; OFC = orbitofrontal cortex; vmPFC = ventromedial prefrontal cortex

Received June 29, 2015. Revised November 3, 2015. Accepted November 25, 2015. Advance Access publication January 25, 2016

Introduction

Predicting the outcome of potentially rewarding events is a critical ability for adaptive behaviour. The orbitofrontal cortex (OFC) is known to code at least three different types of reward-related information: reward probability, risk and experienced value (or outcome value) (Doya, 2008; Rushworth and Behrens, 2008; Peters and Büchel, 2010; Padoa-Schioppa and Cai, 2011; Levy and Glimcher, 2012). The reward probability of a potential reward indicates the prospect of a reward that will occur within a specified time period. The risk of an upcoming outcome, defined as the outcome variance, measures the unpredictability of possible outcomes, and follows an inverted U-shaped relationship with reward probability (maximal for reward probability = 0.5). Experienced value (also called outcome value) reflects the value of consumption experienced at the time of reward delivery. In choice situations, expected utility theory (Von Neumann and Morgenstern, 1945) and prospect theory (Kahneman and Tversky, 1979) provide descriptions of subjective value and measure it with individuals' preferences for choice options. However, assessment of subjective value occurs not only in choice but also in no-choice 'imperative' situations (Tobler *et al.*, 2009). A number of functional MRI studies indicate that expected value is represented in a 'common currency' network encompassing the medial part of the OFC/ventromedial prefrontal cortex (vmPFC) and ventral striatum (O'Doherty *et al.*, 2002; Kim *et al.*, 2011; Sescousse *et al.*, 2015; Metereau and Dreher, 2015). The medial OFC has been shown to be a core component of a risk-sensitive processing network comprising the basal ganglia, amygdala, parietal cortex, anterior cingulate cortex and insular cortex (Fiorillo *et al.*, 2003; Hsu *et al.*, 2005; McCoy and Platt, 2005; Preuschoff *et al.*, 2006; Christopoulos *et al.*, 2009; Tobler *et al.*, 2009; O'Neill and Schultz, 2010). Furthermore, the lateral part of the OFC has been found to play a critical role in coding experienced value at the time of monetary reward delivery (Sescousse *et al.*, 2010; Li *et al.*, 2015). Lesion studies also emphasize the crucial role of the OFC in guiding adaptive behaviour on the basis of reward value, both in animals (Stopper *et al.*, 2014) and in humans (Hsu *et al.*, 2005; Clark *et al.*, 2008). However, OFC lesions in humans are often extended and are not restricted to the medial OFC or to the lateral OFC only. Moreover, simple extrapolation from monkey to human OFC is not straightforward because OFC homologies between species remain elusive (Öngür *et al.*, 2003; Petrides *et al.*, 2012).

In addition, the timing of neural computation of reward value and risk in the human OFC remains to be characterized. Scalp EEG and MEG studies have demonstrated that the anterior cingulate cortex can process reinforcement information as early as 200 ms (Thomas *et al.*, 2013), but neural activity in the OFC cannot be measured directly from the scalp, so the timing of neural activity in the human OFC is still unclear. Furthermore, although functional MRI can resolve brain activity changes in the order of seconds, intracranial EEG recordings can provide relatively more precise insights into the speed with which information is processed in the order of tens or hundreds of milliseconds. In particular, given the relatively poor temporal resolution of functional MRI, it has not been possible to specify whether the risk signal emerges only during reward anticipation or can also be found at the time of reward outcome. Finally, when considering the role of specific subdivisions of the OFC, an open question is whether risk is coded in the medial, lateral or in both parts of the OFC. Several functional MRI studies have reported an engagement of the lateral prefrontal or lateral OFC for risk coding (Hsu *et al.*, 2005; Tobler *et al.*, 2007, 2009). However, other functional MRI studies, together with lesion studies and monkey electrophysiological studies do not support such a clear-cut subdivision in the OFC. For example, human functional MRI studies indicate that the medial OFC responds to risk-related signal during anticipation of uncertain rewards (Abler *et al.*, 2009), but risk and reward value signals have been reported in the lateral part of the monkey OFC (O'Neill and Schultz, 2010).

Building on these considerations, we performed an intracranial EEG study to characterize the spatio-temporal dynamics of reward probability, risk and experienced value signals in the human OFC. We recorded local field potentials (LFPs) in epileptic patients with implanted depth electrodes in the OFC while they learned to associate cues of different slot machines with distinct reward probabilities. In this experiment, participants made no choice that was material to the reward outcome. Intracranial EEG provides a unique opportunity to examine the functioning of the human OFC, as it can circumvent some of the inherent limitations of other techniques (e.g. brain lesion and functional MRI), and combines the excellent temporal resolution (in milliseconds) of electrophysiological methods with high spatial resolution (Mukamel and Fried, 2012).

Materials and methods

Participants

Eight participants [four female; aged 19–61, average: 32, standard deviation (SD): ± 13.3 years] suffering from drug-refractory partial epilepsy took part in our study. Two of them were excluded due to very bad quality of the raw data. The remaining six participants (four female; aged 19–61, average: 34, SD: ± 15 years) had normal or corrected-to-normal vision. All of them were fully informed of the purpose of the study and provided their written informed consent. The study was approved by the ethics committee at the Epilepsy Department of the Neurological Hospital, where the recordings from all patients were collected. The patients were stereotaxically implanted with depth electrodes as part of a presurgical evaluation. No seizures occurred in any of the patients during the 12 h preceding the experiment. In all six

remaining patients, no seizure zones were found in the OFC, so artefact contamination due to epileptogenic focus was excluded. Specifically, Patient 1 suffered from right parietal epilepsy, with a focus on the external right parietal cortex; Patient 2 suffered from left parietal epilepsy, with a focus in the left superior parietal cortex; Patient 3 suffered from left frontal epilepsy, with a focus in the left cingulate gyrus; Patient 4 suffered from left frontal epilepsy, with a focus in F1 intern; Patient 5 suffered from right frontal epilepsy, with a focus in F1 (posterior lateral); and Patient 6 suffered from right temporal epilepsy with a focus in the right external temporal cortex. None of the patients had a visible lesion on MRI or X-ray, which might have a bearing on OFC function through closely connected regions. All of them were cured by corticectomy. The posology of the drugs were low because patients were in a weaning period of antiepileptic drugs at the time of testing, which was voluntarily prescribed at the time of intracranial EEG exploration to increase the chance of the emergence of epileptic seizures. Patients 1 to 6 were under the following antiepileptic therapies: Patient 1, lamotrigine: 300 mg/24 h, topiramate: 100 mg/24 h; Patient 2, levetiracetam 250 mg/24 h, topiramate: 200 mg/ 24 h, clobazam: 10 mg/24 h; Patient 3, levetiracetam: 2000 mg/24 h, gabapentine: 2400 mg/24 h and clobazam: 10 mg/24 h; Patient 4, carbamazepine: 1200 mg/24 h, clobazam: 40 mg/24 h, primidone: 500 mg/24 h; Patient 5, oxcarbazepine: 300 mg/24 h, levetiracetam: 1000 mg/24 h; and Patient 6, carbamazepine: 100 mg /24 h.

Stereotaxic implantation and electrode location

Depth electrodes used to record EEG activity were 0.8 mm multi-contact cylinders (DIXI Medical). They were implanted into several brain areas, perpendicular to the midsagittal plane according to the Talairach and Bancaud's (1973) stereotaxic technique, described in earlier studies (Krolak-Salmon *et al.*, 2004). Contacts (5–15 per electrode) were 2 mm long and spaced every 1.5 mm. For four patients having MRI-compatible implanted electrodes, electrode locations were directly identified with the post-implantation structural MRI images containing the traces of the electrodes using the Mango software (http://ric.uthscsa.edu/mango/index.html). For the two others participants implanted with non-compatible MRI electrodes, electrode locations were reconstructed onto the subject's individual MRI through the superimposition of the frontal skull X-ray images with the electrodes in place on the patient's structural frontal MRI slices, corresponding to each set of electrode coordinates, using in-house software ('Activis' software, Lyon, France). We used the Chiavaras atlas of the orbitofrontal cortex, defined in a normalized Talairach space, to identify the exact locations of contacts involved in reward value and risk information within the OFC (Chiavaras *et al.*, 2001). Additionally, each participant's contacts in the OFC with reward value (expected and experienced value) and risk signals were shown in the normalized MNI (Montreal Neurological Institute) brain space, respectively to help compare with other brain imaging studies (Ossandon *et al.*, 2012). MNI and Talairach coordinates were computed using the SPM (http://www.fil.ion.ucl.ac.uk/spm/) toolbox.

Experimental design

Our current experimental protocol was the same as the one used in our previous intracranial EEG study (Vanni-Mercier *et al.*, 2009). The experimental paradigm was implemented with the software Presentation (version 9, Neurobehavioral Systems). The participants performed the experiment in a noise-shielded room in the hospital. Before starting the experiment, the experimenter explained the procedures of the task to each participant. The experiment was composed of two sessions: a practice session and an experimental session containing eight runs, each of which was comprised of five blocks, corresponding to the five different types of slot machines. Each of them was associated with one of five reward probabilities [P0 (having no rewards), P0.25, P0.5, P0.75, and P1 (always having rewards)]. Thus, there were 40 different slot machines in eight runs in total. The participants were presented with five different kinds of slot machines randomly in each run. Each block had the same structure, which contained 20 consecutive trials. In each block, rewarded and unrewarded trials were pseudorandomized for each participant. Each trial in a block was composed of four phases as follows.

(i) Presentation of the slot machine phase. In the first phase, a picture comprising a single slot machine image and a fractal image on top of the slot machine image was presented to the participants at the centre of the screen on the black ground. Each slot machine included three spinners. At the beginning, the slot machine showed the symbols '7 – 7' on each spinner separately from left to right. The picture would be erased when the participants made responses. The patients' responses were self-paced.

(ii) Delay period phase. After the participants' responses, the three spinners in the slot machine started to roll from left to right successively. When the first spinner stopped, the second would subsequently start. Each of them stopped at an interval of 500 ms successively. So, the delay period from responses to the stopping of the third spinner was 1500 ms.

(iii) Rolling spinners' outcome phase. In the third phase, the participants would know whether they had gotten reward or not according to the information on the third spinner. There were two types of spinners' results: BAR BAR SEVEN (- - 7) and BAR BAR BAR (- - -). The former indicated no subsequent reward delivery and the latter depicted reward delivery subsequently. In other words, the participants were fully informed of subsequent reward or no reward delivery according to information shown on the third spinner. When the third spinner stopped, it was still on the screen for another 500 ms, which was followed by the reward or no reward delivery.

(iv) Reward or no reward delivery phase. In the last phase, either reward (a picture of a 20€ bill) or no reward (rectangle with '0€' written inside which is the same size as the reward) was shown at the centre of the screen for 1000 ms. The intertrial interval was 1.5 s plus ± 0.5 s (Fig. 1).

In the experiment, the participants were instructed to make an estimation of the reward probability of each slot machine at each trial on the basis of all the outcomes of the slot machines that happened previously until the current trial (i.e. estimate of cumulative probability since the first trial). Participants were also informed that their current responses had no effect on subsequent occurrence of reward. During the experiment, no feedback relating to whether their estimation about the

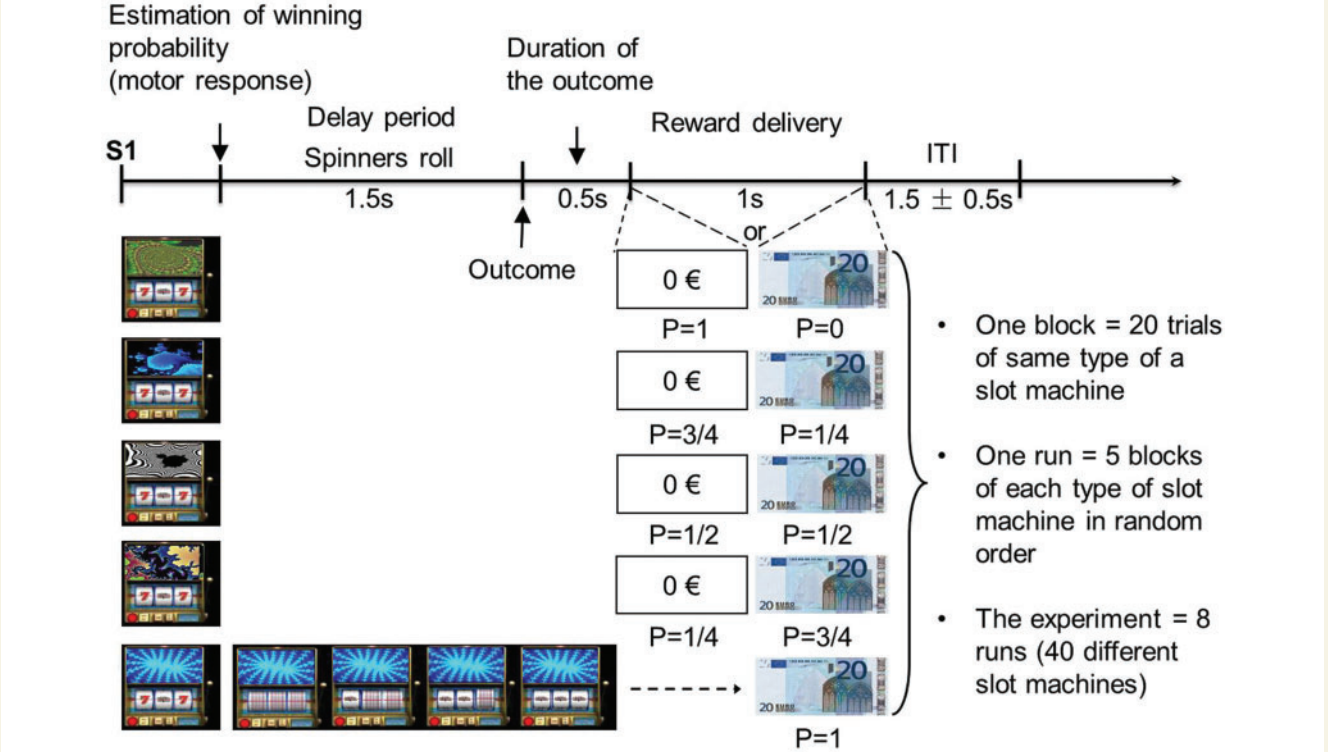

Figure 1 Experimental paradigm. Each trial (self-paced) can be decomposed in four different phases: (i) Presentation of slot machines phase (S1): participants were asked to estimate whether a given slot machine was frequently associated with 20€ delivery or not by pressing one of two keys. There were five types of slot machines, distinguishable by different fractals on their top, each one associated with one of five reward probabilities (P0, P0.25, P0.5, P0.75 and P1), unbeknownst to the participants; (ii) Delay period phase (1.5 s): participants' responses made three spinners begin to roll around and successively stop every 0.5 s during 0.5 s; (iii) Rolling spinners' outcome phase (0.5 s): the stopping of the third spinner revealed the trial outcome (i.e. informing participants of subsequent reward or no reward delivery), which was indicated by two configurations of the three spinners: 'BAR, BAR, 7' (no reward) or 'BAR, BAR, BAR' (rewarded); (iv) Reward/no reward delivery phase (1 s): a 20 € bill picture or a rectangle of the same size with '0 €' written inside was shown to the participants. The intertrial interval (ITI) was 1.5 ± 0.5 s.

winning probability of the slot machine was correct or not were shown to the participants. In addition, the task was not concerned with the judgement about the predication of the slot machine at the current trial. To perform the task, participants were asked to make one of two button presses: one button referring to having a high winning probability of a slot machine and the other one indicating that, overall, the slot machine had a low winning probability. Finally, at the end of each block, the participants were asked to rate this slot machine on a scale from 0 to 4 (0 indicating no wining probability and 4 meaning definitive 100% winning probability) according to their global estimation of reward delivery.

Electrophysiological data recording

We started our experiment 8 days after the electrode implantation. During this period, anticonvulsive drug treatment had been drastically reduced for at least 1 week to record spontaneous epileptic seizures during continuous video-scalp EEG recordings performed in specially equipped rooms. Patients with depth recording electrodes seated in front of a computer screen. Continuous-LFP recordings were collected using a 128-channel device (Brain Quick System Plus; Micromed) at a sampling rate of 512 Hz, amplified and filtered (0.1–200 Hz bandwidth). The intracranial EEG was referenced to another electrode contact located outside the brain, near the skull. Those continuous EEG recordings were stored with the digital event markers indicating the different events of the experiment. Those event markers were composed of three categories: five cue markers reflecting appearance of the slot machine (S1), two response markers depicting the patients' button responses (R) and eight outcome markers [when the third spinner stopped spinning (S2)]. Those five cue markers corresponded to each of five reward winning probabilities of the slot machines (P0, P0.25, P0.5, P0.75, and P1). The two response markers referred to the patients' high or low winning probability estimation. And eight outcome markers were used to differentiate all possible reward/no reward delivery corresponding to five reward probabilities of the slot machines [three slot machines associated with P0.25, P0.5 and P0.75 containing either rewarded or unrewarded trials, one (P1) with only rewarded trials, and one (P0) with only unrewarded trials].

Electrophysiological data analysis

All EEG data analysis was performed with EEGLAB 9.04 (http://sccn.ucsd.edu/eeglab/) (Delorme and Makeig, 2004)

which runs on Matlab. In each participant, the raw EEG recordings were first notch-filtered at a frequency of 50 Hz based on the distribution of power spectrum. The resulting EEG data were low-pass filtered (30 Hz), which was followed by the exclusion of visual inspection of artefacts showing epileptic spikes and other artefacts. Then the data were segmented into three epochs: (i) cue-locked epochs lasting 1000 ms, which started 200 ms prior to the presentation of the cues and ended 800 ms after the cue presentation; (ii) response-locked epochs lasting 3000 ms that started 500 ms prior to the response and ended 2500 ms after the response; and (iii) 1200 ms reward/non-reward delivery-locked epochs lasting 1200 ms, which started 200 ms prior to the reward/non-reward delivery and ended 1000 ms after the delivery. Subsequently, data artefacts were further removed. Specifically, to detect EEG segments containing 'improbable data', we excluded the epochs having 5 SD from the epochs mean probability distribution for the subsequent analysis. The 200–0 ms pre-cue, 500–200 ms pre-response and 200–0 ms pre-delivery time window were used to perform baseline correction. Before averaging, we further excluded the epochs having the voltage above +200 μV and below −200 μV. Afterwards, these artefact-removed data were submitted to averaging. First, the averaging of cue-locked EEG signals was performed for each type of reward probability (P0.25, P0.5, P0.75, and P1) and reward/non-reward delivery-locked EEG signals were averaged in each participant. Then, based on previous studies bearing functional similarities (Axmacher *et al.*, 2010), the grand-averaged LFPs analysis across all participants was derived from the contacts in the OFC with maximal cue-locked LFP signals and reward/non-reward delivery-locked LFP signal, respectively. After the signal averaging step, we analysed the mean amplitude of LFPs during the interval 400–600 ms after the cue presentation and during the interval 0–800 ms after the reward/non-reward delivery. Second, we performed signal averaging of EEG recordings for each level of either unrewarded or rewarded trials (unrewarded trials: P0, P0.25, P0.5, P0.75; rewarded trials: P0.25, P0.5, P0.75, P1) in each participant to probe risk signals in the OFC. The grand-averaged LFP analysis across all participants was derived from the contacts in the OFC with maximal risk signals from each participant. After the signal averaging step, we analysed the peak amplitude of LFPs during the interval 1000–2000 ms because risk signals peaked in this time window.

For the group statistical analysis, regarding the reward probability and experienced value signals, we performed two separate one-way repeated-measures ANOVA on the amplitudes. Tukey's HSD *post hoc* comparisons were then carried out to clarify the significant difference between cue-induced LFP amplitudes as a function of probability when the main effect was significant. Regarding risk signals, under unrewarded and rewarded trials, we performed two-way repeated-measures ANOVA on the peak amplitudes with reward probability and outcome (reward/unreward) as independent factors. Tukey's HSD *post hoc* comparisons were then carried out to clarify the significant difference between risk-induced LFP amplitudes as a function of probability and outcome.

Behavioural data analysis

Response times were analysed as a function of the reward probabilities of the slot machines and the trial rank.

The percentages of correct estimations of the high/low probability of winning for each slot machine were analysed as a function of trial rank (1–20) averaged across participants and runs. The estimations were defined as correct for the slot machines with low reward probabilities (P0 and P0.25) if participants identified them as 'low winning' and were defined as correct for the slot machines with high reward probabilities (P0.75 and P1) if participants identified them as 'high winning.' The slot machine with a reward probability of P0.5 had neither 'low' nor 'high' winning probability. The choice being binary, the percentage of 50% estimates of 'high,' or symmetrically, of 'low' winning probability corresponded to the correct estimate of winning probability for this slot machine.

For the probabilities P0, P0.25, P0.75, and P1, the trial rank when learning occurred was defined as the trial rank with at least 80% correct responses and for which the percentage of correct estimation did not decrease below this limit for the remaining trials. For the probability P0.5, the trial rank when learning occurred was defined as the trial rank with 50% of the responses being either 'high' or 'low' winning probability, with responses then oscillating around this value for the remaining trials. Moreover, results from participants' classifications of the slot machines at each of the 20 successive presentations of a single type of slot machine within runs were compared with their estimations made at the end of each block.

Results

Behavioural results

Response times

A two-way ANOVA with reward probability (P) of the slot machines and trial rank (R) as repeated-measures was performed on the response times. The results revealed that reward probability had a significant influence on the participants' response times [$F(4,20) = 8.15$, $P < 0.001$]. A Tukey's HSD *post hoc* test on reward probability showed that the mean response times for P0.5 (maximal risk) was significantly slower than for all other lower levels of risk (P0, P0. 25, P0.75 and P1), indicating that the participants' response times were modulated by the levels of risk (Fig. 2A). The main effect of trial rank on response times also reached statistical significance [$F(19,95) = 11.68$; $P < 0.001$], but the reward probability × trial rank interaction effect was not significant [$F(76,380) = 1.64$; $P = 0.23$]. Note that due to the task being self-paced and that we did not set an explicit incentive in the task, the sensitivity of response times to the cue may potentially be low. Despite this, we still observed a modulation of response times by reward probability in the absence of explicit incentive in the task. This reflects that participants were slower for the slot machine with $P = 0.5$, being more uncertain regarding the outcome of this slot machine.

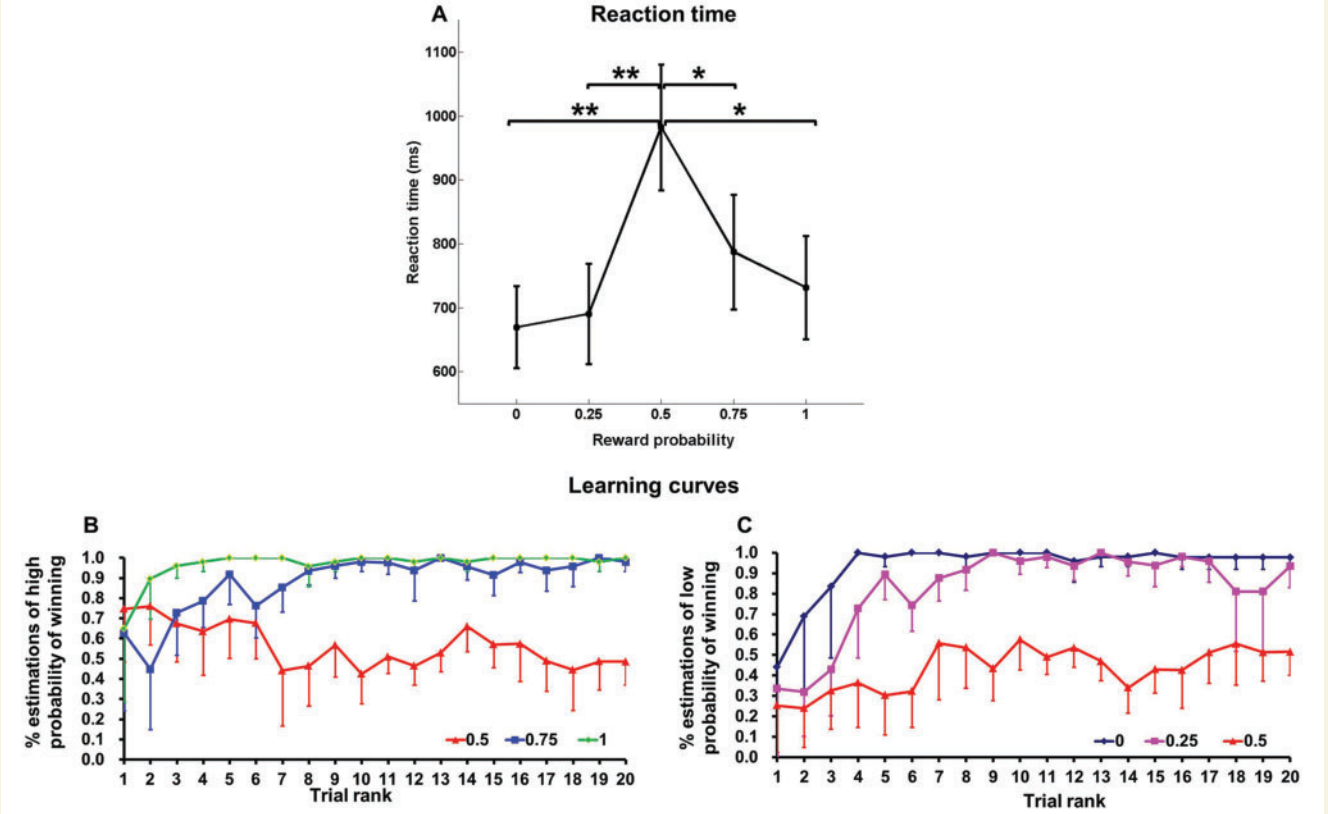

Figure 2 Behavioural performance. (**A**) Mean reaction times as a function of reward probability. (**B**) Mean learning curves averaged across participants, expressed as the mean percentage of 'high winning probability' (**C**) and 'low winning probability'. $^{**}P < 0.01$, $^{*}P < 0.05$. Error bars indicate SEM. Note that participants' task was simply to estimate at each trial the reward probability of each slot machine at the time of its presentation, based upon the previous outcomes of the slot machine until this trial. To do so, participants had to press one of two response buttons: 'high winning probability' and 'low winning probability.' In particular, the estimation of the slot machine with $P = 0.5$ of winning reached the learning criterion (i.e. >80% correct estimations) after the seventh trial (estimations oscillating around 50% as 'high' or 'low' probability of winning).

Estimation of reward probability

We performed a two-way repeated measures ANOVA on the percentage of correct estimates of the probability of winning, including reward probability (P) and trial rank (R) as factors. The learning curves corresponding to the correct estimates for high (slot machines P0.75 and P1) and low (slot machines P0 and P0.25) probability of winning are illustrated in Fig. 2B and C. The results revealed that reward probability and trial rank influenced the correct estimation of slot machines, respectively [$F(4,20) = 69.18$, $P < 0.001$; $F(19,95) = 28.21$; $P < 0.001$]. Moreover, there was an interaction reward probability × trial rank [$F(76,380) = 1.94$; $P < 0.001$], indicating that reaching the learning criterion (>80% correct estimations) depended on reward probability. The estimates of the slot machines with reward probabilities P0 and P1 reached the learning criterion after the second and first trial, respectively (>80% correct estimation). In contrast, the estimates of the slot machines with reward probabilities P0.25 and P0.75 reached the criteria for learning after the fourth trial (>80% correct estimations). The estimation of the reward winning probability P0.5 oscillated around 50% as 'high' or 'low' winning probability. Furthermore, the classification of the slot machines based on the scores (scale range: 0–4) confirmed that participants learned the actual reward probability (correct estimation: 98% for P0, 100% for P1, 83% for P0.25, 88% for P0.75, and 90% for P0.5).

Electrophysiological results

Reward probability signal

As indicated in Fig. 3A, positive or negative LFPs emerged in the OFC following the cue presentation. These signals began ~400 ms after onset of the cue and continued until ~600 ms. We performed a one-way repeated-measures ANOVA on the mean amplitude in this time window with reward probability as an independent factor. Our analysis revealed a significant main effect of reward probability [$F(4,20) = 5.45$; $P < 0.005$]. Furthermore, Tukey's HSD

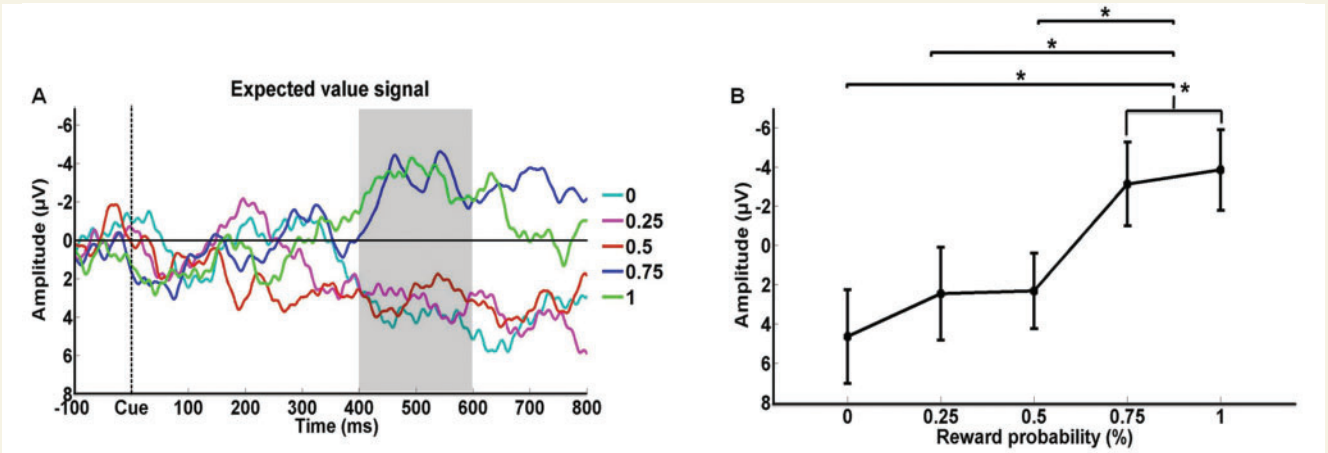

Figure 3 Reward probability coding in the human OFC. (**A**) The reward probability-related orbitofrontal LFPs signals occurred after the presentation of the cues, which were obtained by averaging the contacts with maximal reward probability-like potentials across participants. (**B**) The monotonic increase of LFP amplitude with reward probability. *$P < 0.05$. Error bars indicate SEM.

post hoc tests revealed larger amplitudes for P1 than for P0.5, P0.25 and P0 in this time window (Fig. 3B). The amplitude of these LFPs increased monotonically with reward probability, consistent with the characteristics of an expected value signal. Note that the LFP signals observed at the cue are unlikely to represent the neural response to low level stimulus attributes of the fractal displayed on the slot machine because each of these fractals was associated with a distinct reward probability in each block, and the LFPs associated with each probability represents the mean response to all the different fractals having the same reward probability averaged over the eight runs.

Risk signal

Under both rewarded and unrewarded conditions, robust risk signals were observed, which started from the late reward anticipation phase (1000–1500 ms), i.e. after the second spinner stopped, reaching a maximum during the rolling spinners' outcome phase (1500–2000 ms) (Fig. 4A and B, shaded areas). These risk-related LFP signals peaked at 25.47 ± 40.85 ms (unrewarded trials) and 89.11 ± 48.93 ms (rewarded trials), respectively, after the third spinner stopped (i.e. at the rolling spinners' outcome phase). Then, the signals began to gradually decrease during the reward/no reward delivery phase (2000–2500 ms). A two-way ANOVA was performed on the peak amplitudes during the rolling spinners' outcome phase with reward probability and outcome as independent factors, both for the rewarded and unrewarded conditions. The results revealed a main effect of reward probability [$F(3,15) = 10.28$; $P < 0.005$], but no main effect of outcome [$F(1,5) = 0.54$; $P > 0.1$] and no reward probability × outcome interaction [$F(3,15) = 1.20$; $P > 0.1$) in the outcome phase time window. More importantly, Tukey's HSD *post hoc* tests revealed significantly larger LFP amplitude elicited by P0.5 as compared with other reward probabilities in this same time window (Fig. 4C). The peak LFP amplitudes of rewarded and unrewarded trials followed an inverted U-curve relationship with reward probability, varying non-linearly with reward probability, being maximal when risk is highest ($P = 0.5$), and minimal when risk is lowest ($P = 0$ and $P = 1$), during both the late phase of reward anticipation and during the rolling spinners' outcome phase. Moreover, it could be argued that the risk signal is correlated with ease of learning in our study. To rule out this hypothesis, we ran an additional analysis on the event related potentials for the last 10 trials of each run, after learning of stimuli-outcomes associations was established. We observed that the amplitudes of risk signals in the OFC follow a similar inverted U-shaped relationship as a function of reward probability. Specifically, we found a main effect of reward probability [$F(3,15) = 8.62$; $P < 0.01$], but no main effect of outcome type [$F(1,5) = 0.41$; $P > 0.1$] and no reward probability × outcome interaction [$F(3,15) = 0.86$; $P > 0.1$]. More importantly, Tukey's HSD *post hoc* tests revealed significantly larger LFPs amplitude elicited by P0.5 as compared with other reward probabilities (Supplementary Fig. 1).

Finally, we performed a two-way ANOVA on the peak LFP amplitudes with reward probability and outcome as independent factors and with response time as a covariate of no interest to control for the possibility that LFPs track response times rather than risk. The results revealed a main effect of reward probability [$F(3,12) = 6.43$; $P < 0.05$], but no main effect of outcome [$F(1,4) = 0.15$; $P > 0.1$] and no reward probability × outcome interaction [$F(3,12) = 0.47$; $P > 0.1$] in the outcome phase time window. More importantly, Tukey's HSD *post hoc* tests revealed significantly larger LFP amplitudes elicited by P0.5 as compared with other reward probabilities in this same time window. The peak LFP amplitudes of rewarded and unrewarded trials followed an inverted U-curve relationship with reward probability, varying non-linearly with reward probability, being maximal when risk is highest ($P = 0.5$), and minimal when risk is lowest ($P = 0$ and $P = 1$), during both the late

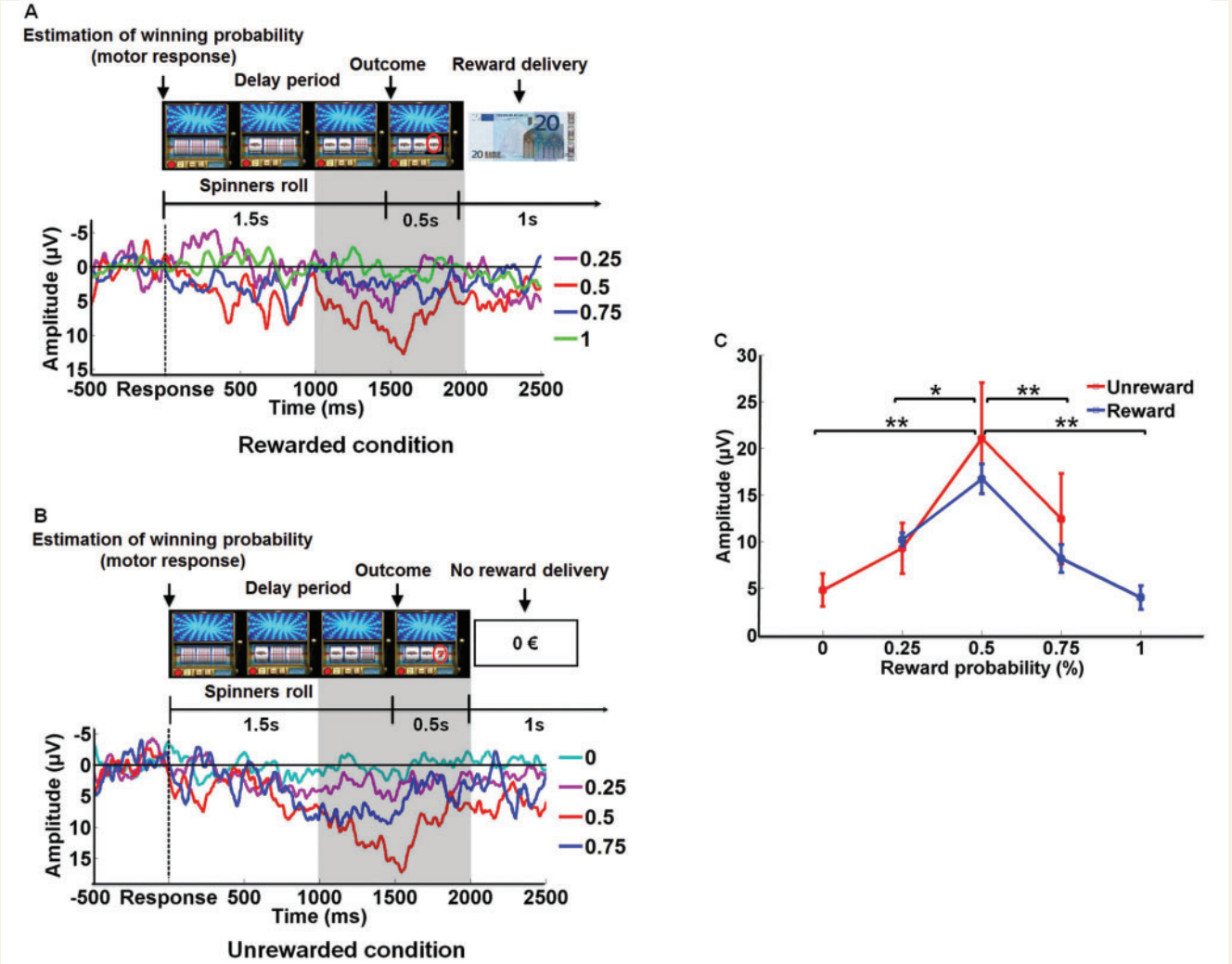

Figure 4 Risk coding in the human OFC. (**A** and **B**) The risk-related orbitofrontal LFP signals occurred during the late phase of reward anticipation and during the rolling spinners' outcome phase for each type of five slot machines under rewarded condition (**A**) and unrewarded condition (**B**). The signals were obtained by averaging the contacts with maximal risk-like potentials across subjects. (**C**) Inverted U shape relationship of LFP amplitude with reward probability. Mean amplitudes of LFPs during the rolling spinners' outcome phase, as a function of reward probability, varied as an inverted U-shaped curve, both for rewarded and unrewarded conditions. $^{**}P < 0.01$, $^{*}P < 0.05$. Error bars indicate SEM.

phase of reward anticipation and during the rolling spinners' outcome phase. This result clearly demonstrates that the OFC tracks risk information rather than response times, as the OFC still tracks risk when regressing out response times.

Experienced value signal

We observed a robust experienced value-related signal in the OFC. As shown in Fig. 5, a difference between reward and non-reward delivery LFPs emerged rapidly in the OFC after the presentation of the bill or after 0 €. This signal started immediately at the time of reward/non-reward delivery and continued until ~800 ms. A one-way repeated-measures ANOVA in this time window with reward/non-reward delivery as an independent factor revealed a main effect of rewarded outcome [$F(1,5) = 19.5$; $P < 0.01$]. It could be argued that the higher LFP amplitudes observed at the time of reward delivery relative to the no-reward delivery could be confounded with feedback updating of one's predictions. However, if this was the case, one would expect the OFC to encode a reward prediction error, known to show decreasing activity with increasing reward probability at the time of outcome (Fiorillo *et al.*, 2003). When performing such analysis, we found no contact responding as a prediction error in the OFC. It should be noted that the fourth phase does not uniquely encode reward/no reward delivery, because it differs from the third phase in containing an image of money. Any difference between these two phases is therefore better explained by associations with the image of money than with perceived reward outcome, which is signalled by the two phases equally.

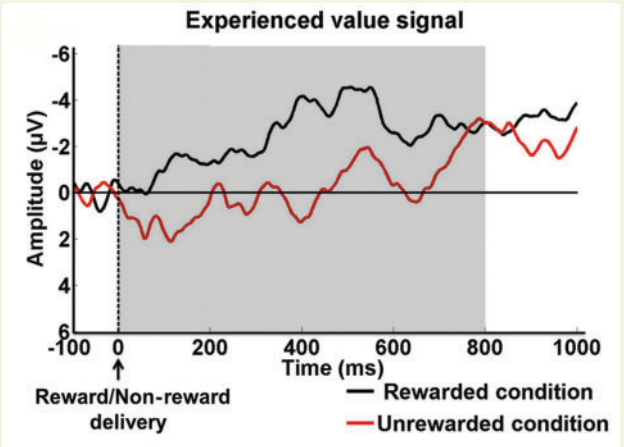

Figure 5 Experienced value coding in the human OFC. The experienced value-related orbitofrontal LFP signals occurred immediately at the time of reward/non-reward delivery. The signals were obtained by averaging the contacts with maximal experienced value-like potentials across subjects. The amplitude elicited by reward delivery was larger than that elicited by non-reward delivery.

Proportion of contacts responding to different signals

Among our participants, four had unilateral implantation in the left OFC and two had unilateral implantation in the right OFC. We recorded from a total of 83 contacts of six depth electrodes covering the OFC from the most medial to the lateral part. The maximal reward probability-like signal, maximal risk-like signal and maximal experienced value-like signal were all defined as the maximal positive or negative peak event-related potentials in their respective time windows. That is, the contact with the maximal peak reward probability signal was in a window 400–600 ms after the onset of the cue; the contact with the maximal peak risk signal was in a window of 1000–2000 ms after the motor response; and the contact with the maximal peak experienced value signal was in a window of 0–800 ms after the onset of reward delivery. The contacts with maximal reward probability signals (Fig. 6A), maximal risk signals (Fig. 6C) and maximal experienced value signals (Fig. 6E) are shown in red on each participant's anatomical images in Talairach space. Coordinates of the corresponding contacts showing maximal reward probability, risk and experienced value signals are listed in Supplementary Tables 1–3, respectively, for each participant. To specify the exact locations of the contacts responding to the expected value, risk and experienced value signals in all participants, we converted the Talairach anatomical locations of the contacts responding to these three signals to the normalized MNI (Montreal Neurological Institute) space. The corresponding converted contacts are shown on a human OFC MNI template for each type of these three signals (Fig. 6B, D and F).

The definition of the medial and lateral OFC was based on previous studies (Zald *et al.*, 2014). Specifically, the medial orbital sulcus was used as the primary division between the medial OFC and lateral OFC. We found 29 contacts (35% of total number of contacts) coding reward probability. Among them, 11 contacts were distributed in the medial OFC and the remaining 18 contacts were in the lateral OFC (Fig. 6B). There was no significant difference in the distribution of reward probability signals between the medial and lateral OFC (chi-square test, $P = 0.06$), suggesting that the medial and lateral OFC played a similar role in coding reward probability information. With regard to the location of contacts responding to risk, we found 22 contacts (27% of total number of contacts) coding risk information. Among them, eight contacts were distributed in the medial OFC and the remaining 14 contacts were located in the lateral OFC (Fig. 6D). Again, a chi-square test did not reveal a statistically significant difference in the distribution of risk signals between the medial and lateral OFC ($P = 0.21$), indicating that both parts of the OFC played a similar role in coding risk signals. Finally, regarding the location of experienced value, we found 45 contacts (54% of total number of contacts) responding to this signal. Among them, 18 contacts were distributed in the medial OFC and the remaining 27 contacts were in the lateral OFC (Fig. 6F). We observed a significant difference in the distribution of experienced value signal between the medial and lateral parts of the OFC (chi-square test, $P < 0.001$) (Fig. 6G), indicating that the experienced value signal was more predominantly localized in the lateral OFC.

Finally, the number of contacts coding reward value (expected and experienced value), risk information or both are illustrated in Fig. 6H. As shown in this figure, there were 74 contacts coding reward value and 22 contacts coding risk information. Among them, nine contacts were involved in coding both reward value and risk information.

Discussion

To the best of our knowledge, the present study provides the first intracranial EEG evidence characterizing the neural dynamics of expected value, risk and experienced value signals in the humans OFC during a probabilistic reward learning task. Several important results emerge from the present study: (i) different anatomical sub-regions of the OFC are predominantly involved in coding these reward information signals; (ii) the reward probability signal emerges ~400 ms after cue presentation; (iii) the risk signal is reflected in slowly growing LFPs during the late phase of reward anticipation after cue presentation and during the rolling spinners' outcome phase; and (iv) the experienced value is coded immediately at the time of reward delivery. Together, these results shed new light on the spatio-temporal dynamics of reward probability, risk and experienced value representations in the human OFC.

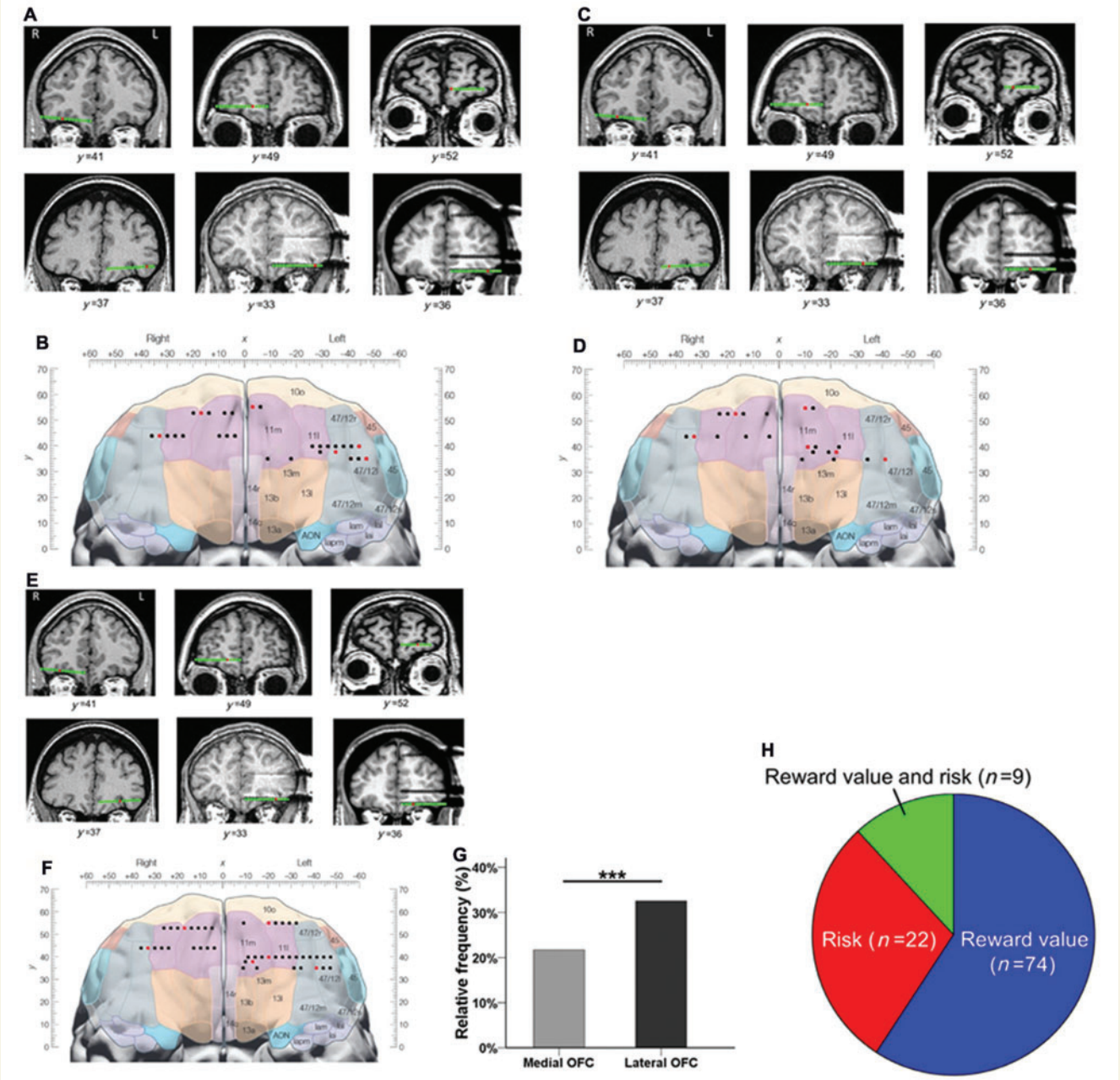

Figure 6 Location of signals increasing with reward probability, risk and experienced value in the human OFC. (**A**) Coronal MRI slices from the six participants showing locations of contacts (in red) yielding a maximal reward probability signal during the time window between 400-600 ms after the onset of the cue. Intracranial electrodes in the OFC are shown in the Talairach brain space. (**B**) The recording contacts across all participants with LFP signals elicited by reward probability are shown in a normalized MNI brain space. For each patient, the contacts exhibiting the maximal LFPs amplitudes with increasing reward probability during the time window between 400–600 ms after the onset of the cue are shown as red dots. The black dots denote the contacts exhibiting a significant increase of the reward probability signal in this same time window. (**C**) Coronal MRI slices from the six participants showing locations of contacts (in red) yielding a maximal risk signal between 1000–2000 ms after the motor response. (**D**) The recording contacts across all participants with risk LFP signals are shown in a normalized MNI brain space. For each patient, the contacts exhibiting the maximal risk LFPs amplitudes during the time window between 1000–2000 ms after the motor response are shown as red dots. The black dots depict the contacts exhibiting a significant risk signal in this same time window. (**E**) Coronal MRI slices from the six participants showing locations of contacts (in red) yielding a maximal experienced value signal between 0–800 ms after the onset of reward delivery. (**F**) The recording contacts across all participants with LFP signals elicited by experienced value are shown in a normalized MNI brain space. For each patient, the contacts exhibiting the maximal experienced value LFP amplitudes during the time window between 0–800 ms after the onset of reward delivery are shown as red dots while the black dots denote the contacts exhibiting a significant experienced value signal in this same time window. (**G**) Relative frequency of contacts with experienced value signals in the medial and lateral parts of the OFC. ***$P < 0.001$. (**H**) Pie chart of the percentage of contacts coding reward value and risk information in the human OFC.

Reward probability coding in the orbitofrontal cortex

Electrophysiological recordings in animals indicate that value is coded in the firing rates of orbitofrontal neurons (Padoa-Schioppa and Assad, 2006; Wallis, 2012). A large body of evidence from human functional MRI studies also confirms that reward probability is coded in this region (Rolls *et al.*, 2008; Kahnt *et al.*, 2010). By overcoming some of the shortcomings associated with functional MRI, our results extend these prior findings in several ways. First, we found that maximum amplitude of this signal increased monotonically with reward probability, demonstrating directly that the OFC indeed codes the reward probability after the presentation of the slot machine. Second, the precise time resolution of our intracranial EEG recordings reveal that such reward probability signal starts to rise around 400 ms after cue presentation in the medial and lateral OFC for higher reward probabilities ($P = 0.75$ and $P = 1$). This latency is similar to the peak latency of the firing rates of neurons coding reward probability previously observed in monkey OFC neurons (O'Neill and Schultz, 2010). This observation indicates that the OFC codes the reward probability relatively late after the presentation of the slot machine. In animals, the latency of the reward probability signal in the OFC is also slower than the reward probability response from the midbrain dopaminergic neurons (Fiorillo *et al.*, 2003) and neurons in the visual cortex (Shuler and Bear, 2006). Our finding of slowly rising reward probability responses in the human OFC is also consistent with current conceptualizations of OFC-dopamine neuron relationships supporting that the role of the OFC is to represent states in partially observable scenarios (Takahashi *et al.*, 2011; Metereau and Dreher, 2015).

Risk coding in the orbitofrontal cortex

The risk signal appears both before and after risk is resolved. This finding is consistent with results from a recent single-unit electrophysiological study in monkey OFC (O'Neill and Schultz, 2010). During the rolling spinners' outcome phase, the peak LFP amplitude followed an inverted U-curve relationship with reward probability, both for rewarded and unrewarded conditions. Data from neurophysiological studies in non-human primates and human functional MRI studies have implicated midbrain dopaminergic neurons and their cortical and subcortical projections in coding risk information (McCoy and Platt, 2005; Dreher *et al.*, 2006; Preuschoff *et al.*, 2006; Tobler *et al.*, 2009; Vanni-Mercier *et al.*, 2009; Mohr *et al.*, 2010; Rudorf *et al.*, 2012; Sugam *et al.*, 2012; Wright *et al.*, 2012; Monosov and Hikosaka, 2013). Within this risk-sensitive circuit, the literature in non-human primates suggests that risk-sensitive neurons in the OFC may transmit an early signal to other brain structures responding to risk information, such as dopamine neurons, anterior insula and anterior cingulate cortex, for further processing. This argument is supported by the fact that risk response latency in the monkey OFC (O'Neill and Schultz, 2010) appears to be shorter than the risk-related responses in dopamine neurons (Fiorillo *et al.*, 2003) and cingulate neurons (McCoy and Platt, 2005). Moreover, anticipation-related firing rates in dopamine neurons depend on the OFC inputs in rodents (Lodge, 2011; Takahashi *et al.*, 2011). Thus, the early latency of the risk response in the OFC may allow downstream neurons to participate in detecting risk information in decision situations. It is likely that during phylogeny, the circuit involved in coding risk information has been well preserved across species. Confirming this hypothesis, the risk-elicited LFPs observed in the human OFC occurred early after the second spinner of the slot machine stopped.

Although human intracranial EEG recordings of risk coding are scarce, we recently observed a risk-related LFP signal in the human hippocampus during the rolling spinners' outcome phase, peaking around 410 ms after the third spinner stopped, using the same task as the one described in the present study (Vanni-Mercier *et al.*, 2009). Thus, the latency of risk-elicited LFPs in the human hippocampus is slower than that of the risk signal observed in the human OFC. Furthermore, in another human intracranial EEG study, unexpected events enhanced early (187 ms) and late (482 ms) hippocampal potentials as well as a late (475 ms) nucleus accumbens potential (Axmacher *et al.*, 2010). Moreover, during a reward learning task, LFPs recorded in the nucleus accumbens were higher for risky compared to safe stimulus ~400–600 ms after the cue presentation as well as experienced value at the time of feedback (Cohen *et al.*, 2008). In addition, a recent human intracranial EEG study recording dopamine neurons with microelectrodes reported relatively late latency (~200 ms after the onset of feedback) of firing rates after unexpected gains (Zaghloul *et al.*, 2009). Together, these findings in humans indicate that the risk signal in the OFC may constitute an early component of the risk-related system in humans, which transfers risk information to other components of the system such as the midbrain, ventral striatum and hippocampus. Although the time resolution of the blood oxygen level-dependant (BOLD) signal does not allow us to make precise inferences about timing, it should be noted that the same network, including the OFC, together with the ventral striatum and hippocampus, has been shown to code risk when varying reward probabilities (Abler *et al.*, 2009). It has also been reported that a risk BOLD signal is observed relatively early in the medial OFC, while expected value is coded in the hippocampus by a relatively later BOLD response (Abler *et al.*, 2009).

In addition, the observation that most of the risk-coding contacts in the human OFC differ from the contacts coding reward value (expected and experienced value) extends to the neuronal population level (LFPs domain) from previous single-unit recordings in monkeys showing that most orbitofrontal risk responses are distinct from

value responses (O'Neill and Schultz, 2010). Consistent with these findings, other studies have also reported that the population of neurons in the OFC, which are sensitive to probabilistic, costly or delayed rewards are insensitive to absolute reward values (Roesch *et al.*, 2006; Kennerley *et al.*, 2009).

Finally, our results offer an interpretation of the classical dysfunctional decision-making under risky conditions following damage of the OFC in humans (Hsu *et al.*, 2005; Clark *et al.*, 2008). Changes in decision-making due to the OFC lesions may result from an inability to accurately process risk information because OFC patients do not have the risk signal propagated by the OFC neurons. Such impaired risk processing would constitute a parsimonious explanation for the deficits in risky decision-making following damage to the OFC and would provide a potential pathophysiological account for these striking and severely incapacitating behavioural deficits.

Experienced value and comparison with the dopaminergic system

The increased LFP amplitudes observed for reward compared to no-reward delivery confirms that the OFC codes an experienced value signal. Previous human functional MRI and monkey electrophysiological findings have shown that experienced value is coded by the OFC as well as the amygdala and ventral striatum (Hikosaka *et al.*, 2008; Sescousse *et al.*, 2010). More recent functional MRI findings have reported functional dissociations in the OFC according to rewards types: the anterior OFC is more specifically engaged by secondary rewards than primary rewards, while the posterior OFC is more engaged by primary than secondary rewards (Sescousse *et al.*, 2013; Li *et al.*, 2015). Because the present study did not vary reward types, we cannot ascertain whether these functional divisions can be confirmed using intracranial EEG. However, we did observe an experienced value signal in the anterior OFC for monetary rewards.

Note that we did not observe a linear decrease in LFP amplitudes with increasing reward probability at the time of reward outcome, which would have been expected if OFC coded a reward prediction error. Thus, the experienced value signal observed in the OFC is not concomitant with an OFC reward prediction error, as coded by midbrain dopaminergic neurons showing a monotonic decrease in neuronal activities with increasing reward probability (Fiorillo *et al.*, 2003). Confirming our findings, at the single cell level, a monkey electrophysiological study showed that many vmPFC/OFC neurons do not code reward prediction error (Monosov and Hikosaka, 2012). Rather, it seems that vmPFC/OFC neurons are linked to the processing of the reception of their preferred outcomes (Bouret and Richmond, 2010; Noonan *et al.*, 2010; Rudebeck and Murray, 2011*b*; Rushworth *et al.*, 2011).

Functional interpretation

The OFC is crucial for changing established behaviour in the face of unexpected outcomes. Historically, this function has been attributed either to the role of the OFC in response inhibition or to the fact that the OFC is a rapidly flexible associative-learning area (Schoenbaum *et al.*, 2009; Rudebeck and Murray, 2011*a*, *b*). However, recent data indicate instead that the OFC is not crucial for response inhibition. Rather, it is key to signal outcome expectancies, as demonstrated by excitotoxic, fibre-sparing lesions confined to OFC, which do not alter behavioural flexibility (Rudebeck and Murray, 2011*b*). Thus, the function of the OFC in signalling expected outcomes can also explain its crucial role in changing behaviour in the face of unexpected outcomes (Rudebeck and Murray, 2011*a*). Our intracranial EEG findings in humans offer an electrophysiological basis to understand that this general function is based on different responses from neuronal populations as observed with LFPs, including expected value, risk and experienced/outcome value coding.

Based on brain imaging, lesion and anatomical studies, the lateral OFC has been proposed to be important for stimulus–value learning that is critical for motivating actions towards rewards (Noonan *et al.*, 2010; Rudebeck and Murray, 2011*a*), whereas ventral vmPFC may be concerned with evaluation, value-guided decision-making and maintenance of choices over successive decisions (Noonan *et al.*, 2010; Rudebeck and Murray, 2011*b*). A study recording neurons from the vmPFC and OFC in a task in which neuronal activity could be linked to external (stimulus value) or internal motivational factors (e.g. satiety) found that the vmPFC coded internal motivational processes, whereas the OFC coded external environment-centred value information (Bouret and Richmond, 2010). These distinct functions may be based on specific connectivity of OFC subdivisions. The OFC, situated laterally to vmPFC, receives inputs from sensory systems (Cavada *et al.*, 2000), whereas the vmPFC does not (Barbas *et al.*, 1999). Moreover, the OFC and vmPFC project to different regions of the striatum: the vmPFC has dense projections to the nucleus accumbens (Öngür and Price, 2000), whereas the OFC does not (Haber *et al.*, 1995). Both the OFC (Brodmann areas 13 and 11) and vmPFC have extensive projections to some limbic structures, such as the amygdala (Öngür and Price, 2000), and the vmPFC has particularly strong projections to the hypothalamus (Öngür and Price, 2000).

In the present study, we observed that the experienced value signal was predominantly localized in the lateral OFC, although this signal was coded throughout OFC. In contrast, the medial and lateral OFC played similar roles in coding risk and reward probability signals. These findings are mostly consistent with functional MRI studies reporting both the medial and lateral OFC activity for risk (Tobler *et al.*, 2007; Abler *et al.*, 2009) and for expected value (Rolls *et al.*, 2008; Sescousse *et al.*, 2015) while lateral

OFC activity has been mostly observed for experienced value (Sescousse *et al.*, 2013, 2015; Li *et al.*, 2015).

An important functional dissociation within the OFC is that the medial OFC is involved in positive reinforcers, whereas the lateral OFC would be concerned with the evaluation of punishments (Kringelbach and Rolls, 2004; Kringelbach, 2005). Because the present study only tested rewards and not punishments, we cannot judge whether this medial-lateral functional distinction in the OFC is valid. However, recent Pavlovian conditioning studies associating abstract cues to different types of rewards and punishments did not report such dissociation but found a common engagement of the medial OFC in expectation of rewards and punishments (Plassmann *et al.*, 2010; Metereau and Dreher, 2015). This is further corroborated by a recent single-unit recording study in monkeys reporting no convincing evidence for valence selectivity in the OFC (Rich and Wallis, 2014) and a recent meta-analytical connectivity modelling study reporting convergent co-activations with other areas in the medial and lateral OFC during reward tasks (Zald *et al.*, 2014).

Conclusion

Our intracranial EEG results bridge the gap between neuroimaging studies in healthy humans (Tobler *et al.*, 2007, 2009) and electrophysiological recordings in animals (O'Neill and Schultz, 2010). These findings shed new light on the spatio-temporal dynamics underlying reward value and risk coding in the human OFC. We provided evidence that the brain computes separate reward-related information signals when expecting and experiencing rewarding events, at a sub-second time scale. Although our study focused on a no-choice situation, we believe that it has important implications for disorders of decision-making under risk situations. Indeed, as proposed recently in an electrophysiological recording monkey study investigating risk signal in a no-choice situation, changes in decision-making under uncertainty due to the orbitofrontal damage may result from an inability to accurately process risk information because these patients do not have the risk signal propagated by OFC neurons (O'Neill and Schultz, 2010).

Acknowledgements

We thank the patients for their collaboration.

Funding

This work was funded by ANR-14-CE13-0006-01 to JC D. It was performed within the framework of the LABEX ANR-11-LABEX-0042 of Université de Lyon, within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR). Y.L. was supported by a PhD fellowship obtained by JC D from Pari Mutuel Urbain (PMU). JCD was also funded by the EURIAS Fellowship Programme, the European Commission (Marie-Sklodowska-Curie Actions - COFUND Programme - FP7) and the Institute for Advanced Study 'Hanse-Wissenschaftskolleg'. We thank the staff of the epilepsy department (Neurological hospital, Lyon) for helpful assistance with data collection.

Supplementary material

Supplementary material is available at *Brain* online.

References

Abler B, Herrnberger B, Gron G, Spitzer M. From uncertainty to reward: BOLD characteristics differentiate signaling pathways. BMC Neurosci 2009; 10: 154.

Axmacher N, Cohen MX, Fell J, Haupt S, Dümpelmann M, Elger CE, et al. Intracranial EEG correlates of expectancy and memory formation in the human hippocampus and nucleus accumbens. Neuron 2010; 65: 541–9.

Barbas H, Ghashghaei H, Dombrowski S, Rempel-Clower N. Medial prefrontal cortices are unified by common connections with superior temporal cortices and distinguished by input from memory-related areas in the rhesus monkey. J Comp Neurol 1999; 410: 343–67.

Bouret S, Richmond BJ. Ventromedial and orbital prefrontal neurons differentially encode internally and externally driven motivational values in monkeys. J Neurosci 2010; 30: 8591–601.

Cavada C, Tejedor J, Cruz-Rizzolo RJ, Reinoso-Suárez F. The anatomical connections of the macaque monkey orbitofrontal cortex: a review. Cereb Cortex 2000; 10: 220–42.

Chiavaras MM, LeGoualher G, Evans A, Petrides M. Three-dimensional probabilistic atlas of the human orbitofrontal sulci in standardized stereotaxic space. NeuroImage 2001; 13: 479–96.

Christopoulos GI, Tobler PN, Bossaerts P, Dolan RJ, Schultz W. Neural correlates of value, risk, and risk aversion contributing to decision making under risk. J Neurosci 2009; 29: 12574–83.

Clark L, Bechara A, Damasio H, Aitken MR, Sahakian BJ, Robbins TW. Differential effects of insular and ventromedial prefrontal cortex lesions on risky decision-making. Brain 2008; 131 (Pt 5): 1311–22.

Cohen MX, Axmacher N, Lenartz D, Elger CE, Sturm V, Schlaepfer TE. Neuroelectric signatures of reward learning and decision-making in the human nucleus accumbens. Neuropsychopharmacology 2008; 34: 1649–58.

Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 2004; 134: 9–21.

Doya K. Modulators of decision making. Nat Neurosci 2008; 11: 410–16.

Dreher JC, Kohn P, Berman KF. Neural coding of distinct statistical properties of reward information in humans. Cereb Cortex 2006; 16: 561–73.

Fiorillo CD, Tobler PN, Schultz W. Discrete coding of reward probability and uncertainty by dopamine neurons. Science 2003; 299: 1898–902.

Haber S, Kunishio K, Mizobuchi M, Lynd-Balta E. The orbital and medial prefrontal circuit through the primate basal ganglia. J Neurosci 1995; 15: 4851–67.

Hikosaka O, Sesack SR, Lecourtier L, Shepard PD. Habenula: crossroad between the basal ganglia and the limbic system. J Neurosci 2008; 28: 11825–9.

Hsu M, Bhatt M, Adolphs R, Tranel D, Camerer CF. Neural systems responding to degrees of uncertainty in human decision-making. Science 2005; 310: 1680–3.
Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. Econometrica 1979; 47: 263–91.
Kahnt T, Heinzle J, Park SQ, Haynes JD. The neural code of reward anticipation in human orbitofrontal cortex. Proc Natl Acad Sci USA 2010; 107: 6010–15.
Kennerley SW, Dahmubed AF, Lara AH, Wallis JD. Neurons in the frontal lobe encode the value of multiple decision variables. J Cogn Neurosci 2009; 21: 1162–78.
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: 769–76.
Kringelbach ML. The human orbitofrontal cortex: linking reward to hedonic experience. Nat Rev Neurosci 2005; 6: 691–702.
Kringelbach ML, Rolls ET. The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. Prog Neurobiol 2004; 72: 341–72.
Krolak-Salmon P, Henaff MA, Vighetto A, Bertrand O, Mauguiere F. Early amygdala reaction to fear spreading in occipital, temporal, and frontal cortex: a depth electrode ERP study in human. Neuron 2004; 42: 665–76.
Levy DJ, Glimcher PW. The root of all value: a neural common currency for choice. Curr Opin Neurobiol 2012; 22: 1027–38.
Li Y, Sescousse G, Amiez C, Dreher J-C. Local morphology predicts functional organization of experienced value signals in the human orbitofrontal cortex. J Neurosci 2015; 35: 1648–58.
Lodge DJ. The medial prefrontal and orbitofrontal cortices differentially regulate dopamine system function. Neuropsychopharmacology 2011; 36: 1227–36.
McCoy AN, Platt ML. Risk-sensitive neurons in macaque posterior cingulate cortex. Nat Neurosci 2005; 8: 1220–7.
Metereau E, Dreher J-C. The medial orbitofrontal cortex encodes a general unsigned value signal during anticipation of both appetitive and aversive events. Cortex 2015; 63: 42–54.
Mohr PN, Biele G, Heekeren HR. Neural processing of risk. J Neurosci 2010; 30: 6613–19.
Monosov IE, Hikosaka O. Regionally distinct processing of rewards and punishments by the primate ventromedial prefrontal cortex. J Neurosci 2012; 32: 10318–30.
Monosov IE, Hikosaka O. Selective and graded coding of reward uncertainty by neurons in the primate anterodorsal septal region. Nat Neurosci 2013; 16: 756–62.
Mukamel R, Fried I. Human intracranial recordings and cognitive neuroscience. Annu Rev Psychol 2012; 63: 511–37.
Noonan M, Walton M, Behrens T, Sallet J, Buckley M, Rushworth M. Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. Proc Natl Acad Sci USA 2010; 107: 20547–52.
O'Doherty JP, Deichmann R, Critchley HD, Dolan RJ. Neural responses during anticipation of a primary taste reward. Neuron 2002; 33: 815–26.
O'Neill M, Schultz W. Coding of reward risk by orbitofrontal neurons is mostly distinct from coding of reward value. Neuron 2010; 68: 789–800.
Öngür D, Ferry AT, Price JL. Architectonic subdivision of the human orbital and medial prefrontal cortex. J Comp Neurol 2003; 460: 425–49.
Öngür D, Price J. The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. Cereb Cortex 2000; 10: 206–19.
Ossandon T, Vidal JR, Ciumas C, Jerbi K, Hamame CM, Dalal SS, et al. Efficient "pop-out" visual search elicits sustained broadband gamma activity in the dorsal attention network. J Neurosci 2012; 32: 3414–21.
Padoa-Schioppa C, Assad JA. Neurons in the orbitofrontal cortex encode economic value. Nature 2006; 441: 223–6.
Padoa-Schioppa C, Cai X. The orbitofrontal cortex and the computation of subjective value: consolidated concepts and new perspectives. Annals of the New York Academy of Sciences 2011; 1239: 130–7.
Peters J, Büchel C. Neural representations of subjective reward value. Behav Brain Res 2010; 213: 135–41.
Petrides M, Tomaiuolo F, Yeterian EH, Pandya DN. The prefrontal cortex: comparative architectonic organization in the human and the macaque monkey brains. Cortex 2012; 48: 46–57.
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. J Neurosci 2010; 30: 10799–808.
Preuschoff K, Bossaerts P, Quartz SR. Neural differentiation of expected reward and risk in human subcortical structures. Neuron 2006; 51: 381–90.
Rich E, Wallis J. Medial-lateral Organization of the Orbitofrontal Cortex. J Cogn Neurosci 2014; 26: 1347–62.
Roesch MR, Taylor AR, Schoenbaum G. Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. Neuron 2006; 51: 509–20.
Rolls ET, McCabe C, Redoute J. Expected value, reward outcome, and temporal difference error representations in a probabilistic decision task. Cereb Cortex 2008; 18: 652–63.
Rudebeck PH, Murray EA. Balkanizing the primate orbitofrontal cortex: distinct subregions for comparing and contrasting values. Ann NY Acad Sci 2011a; 1239: 1–13.
Rudebeck PH, Murray EA. Dissociable effects of subtotal lesions within the macaque orbital prefrontal cortex on reward-guided behavior. J Neurosci 2011b; 31: 10569–78.
Rudorf S, Preuschoff K, Weber B. Neural correlates of anticipation risk reflect risk preferences. J Neurosci 2012; 32: 16683–92.
Rushworth MF, Behrens TE. Choice, uncertainty and value in prefrontal and cingulate cortex. Nat Neurosci 2008; 11: 389–97.
Rushworth MF, Noonan MP, Boorman ED, Walton ME, Behrens TE. Frontal cortex and reward-guided learning and decision-making. Neuron 2011; 70: 1054–69.
Schoenbaum G, Roesch MR, Stalnaker TA, Takahashi YK. A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. Nat Rev Neurosci 2009; 10: 885–92.
Sescousse G, Caldú X, Segura B, Dreher J-C. Processing of primary and secondary rewards: a quantitative meta-analysis and review of human functional neuroimaging studies. Neurosci Biobehav Rev 2013; 37: 681–96.
Sescousse G, Li Y, Dreher J-C. A common currency for the computation of motivational values in the human striatum. Soc Cogn Affect Neurosci 2015; 10: 467–73.
Sescousse G, Redouté J, Dreher J-C. The architecture of reward value coding in the human orbitofrontal cortex. J Neurosci 2010; 30: 13095–104.
Shuler MG, Bear MF. Reward timing in the primary visual cortex. Science 2006; 311: 1606–9.
Stopper CM, Green EB, Floresco SB. Selective involvement by the medial orbitofrontal cortex in biasing risky, but not impulsive, choice. Cereb Cortex 2014; 24: 154–62.
Sugam JA, Day JJ, Wightman RM, Carelli RM. Phasic nucleus accumbens dopamine encodes risk-based decision-making behavior. Biol Psychiatry 2012; 71: 199–205.
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. Nat Neurosci 2011; 14: 1590–7.
Talairach J, Bancaud J. Stereotaxic approach to epilepsy. Methodology of anatomo-functional stereotaxic investigations. Prog Neurol Surg 1973; 5: 297–354.
Thomas J, Vanni-Mercier G, Dreher J-C. Neural dynamics of reward probability coding: a Magnetoencephalographic study in humans. Front Neurosci 2013; 7: 214.

Tobler PN, Christopoulos GI, O'Doherty JP, Dolan RJ, Schultz W. Risk-dependent reward value signal in human prefrontal cortex. Proc Natl Acad Sci USA 2009; 106: 7185–90.

Tobler PN, O'Doherty JP, Dolan RJ, Schultz W. Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. J Neurophysiol 2007; 97: 1621–32.

Vanni-Mercier G, Mauguiere F, Isnard J, Dreher JC. The hippocampus codes the uncertainty of cue-outcome associations: an intracranial electrophysiological study in humans. J Neurosci 2009; 29: 5287–94.

Von Neumann J, Morgenstern O. Theory of games and economic behavior. Bull AmMath Soc 1945; 51: 498–504.

Wallis JD. Cross-species studies of orbitofrontal cortex and value-based decision-making. Nat Neurosci 2012; 15: 13–19.

Wright ND, Symmonds M, Hodgson K, Fitzgerald TH, Crawford B, Dolan RJ. Approach-avoidance processes contribute to dissociable impacts of risk and loss on choice. J Neurosci 2012; 32: 7009–20.

Zaghloul KA, Blanco JA, Weidemann CT, McGill K, Jaggi JL, Baltuch GH, et al. Human substantia nigra neurons encode unexpected financial rewards. Science 2009; 323: 1496–9.

Zald DH, McHugo M, Ray KL, Glahn DC, Eickhoff SB, Laird AR. Meta-analytic connectivity modeling reveals differential functional connectivity of the medial and lateral orbitofrontal cortex. Cereb Cortex 2014; 24: 232–48.