CORTEX 63 (2015) 42–54

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research report

The medial orbitofrontal cortex encodes a general unsigned value signal during anticipation of both appetitive and aversive events

*Elise Metereau and Jean-Claude Dreher**

Neuroeconomics Laboratory: Reward and Decision Making, Cognitive Neuroscience Center, Unité Mixte de Recherche 5229, Centre National pour la Recherche Scientifique (CNRS), Bron, France

ARTICLE INFO

Article history:
Received 29 January 2014
Reviewed 24 March 2014
Revised 27 May 2014
Accepted 5 August 2014
Action editor Sarah MacPherson
Published online 28 August 2014

Keywords:
Appetitive conditioning
Aversive conditioning
fMRI
Orbitofrontal cortex
Expected value

ABSTRACT

The medial orbitofrontal cortex (mOFC)/ventromedial prefrontal cortex (vmPFC) has been proposed to signal the expected value of rewards when learning stimuli-rewards associations. Yet, it is still unclear whether identical or distinct orbitofrontal cortex regions encode expected rewards and punishments at the time of the cue during appetitive and aversive classical conditioning. Moreover, it is unknown whether anticipation of different types of positive and negative reinforcers differentially influence specific orbitofrontal cortex regions. To answer these questions, this study investigated whether the human mOFC/vmPFC region encodes a general unsigned anticipatory value signal for different types of rewards and punishments (responding in a positive fashion in anticipation of both appetitive and aversive events) or a signed expected value signal (responding positively in anticipation of rewards and negatively in anticipation of punishments) when learning cue-outcomes associations. Using a model-based fMRI approach implementing a reinforcement learning model to compute the expected values of two types of rewards (pleasant juice, monetary gain) and two types of punishments (aversive juice, aversive picture), we found that mOFC/vmPFC activity correlated positively with the expected value of the cues, in anticipation of both rewards and punishments. This finding indicates that the mOFC/vmPFC encodes a general unsigned anticipatory value signal, regardless of reinforcers valence (positive/negative) and types (gustatory, visual).

1. Introduction

The orbitofrontal cortex has been proposed to signal outcome expectancies, i.e., to signal the expected characteristics and value of specific outcomes that the animal expects (Rushworth & Behrens, 2008; Schoenbaum, Roesch, Stalnaker, & Takahashi, 2009; Schoenbaum, Saddoris, & Stalnaker, 2007). Direct evidence supporting this hypothesis comes from rodents' OFC neuronal recordings directly demonstrating

* *Corresponding author*. CNRS UMR 5229, Cognitive Neuroscience Center, Reward and Decision Making Group, 67 Bd Pinel, 69675 Bron, France.
E-mail address: dreher@isc.cnrs.fr (J.-C. Dreher).
http://dx.doi.org/10.1016/j.cortex.2014.08.012

expectancy signals for appetitive and aversive outcomes (Schoenbaum et al., 2009). In humans, only a few fMRI studies specifically compared expected values during anticipation of both appetitive and aversive Pavlovian conditioning in the same experimental design. For example, expectation of a pleasant taste and of a moderately unpleasant salt taste produced activation in the lateral orbitofrontal cortex, as compared to anticipation of a neutral cue (O'Doherty, Deichmann, Critchley, & Dolan, 2002). In another study, a partial overlapping value related activity was observed within the ventromedial prefrontal cortex (vmPFC) during anticipation of juice and money reward outcomes while this region did not show increased activity during anticipation of aversive juice or of monetary losses (Kim, Shimojo, & O'Doherty, 2010).

In the current study, we focus on the Pavlovian system, which assigns values to behavior, such as approaching cues that predict reward delivery or avoiding cues that predict a punishment. During a Pavlovian conditioning procedure in which a cue predicts a reinforcer, the brain encodes three signals: the anticipatory value of the outcome that it expects to receive ("expected value"), the value of the actual outcome at the time of its reception ("outcome value") and a prediction error that measures the deviation between actual and anticipated values (Metereau & Dreher, 2013; Prevost, Pessiglione, Metereau, Clery-Melin, & Dreher, 2010; Rangel, Camerer, & Montague, 2008; Rangel & Clithero, 2012; Sescousse, Redoute, & Dreher, 2010). fMRI studies investigating Pavlovian conditioning using reinforcement learning models explored which brain system codes the prediction error when learning associations between conditioned stimuli and different types of rewards (Berns, McClure, Pagnoni, & Montague, 2001; Gottfried, O'Doherty, & Dolan, 2003; McClure, Berns, & Montague, 2003; O'Doherty et al., 2004) or punishments (Büchel, Dolan, Armony, & Friston, 1999; Delgado, Li, Schiller, & Phelps, 2008; Gottfried, Deichmann, Winston, & Dolan, 2002; Knight, Waters, King, & Bandettini, 2010; Knutson, Westdorp, Kaiser, & Hommer, 2000; LaBar, Gatenby, Gore, LeDoux, & Phelps, 1998; Nieuwenhuis, Slagter, Geusau, Heslenfeld, & Holroyd, 2005; Sarinopoulos et al., 2010; Seymour et al., 2004). However, the vast majority of these fMRI studies did not investigate the neural representation of the expected value for both positive and negative outcomes, and did not vary the type of reinforcer in different domains.

Although the engagement of the medial orbitofrontal cortex (mOFC)/ventromedial prefrontal cortex (vmPFC) in expected value representation is quite well-established for appetitive stimuli (see Boorman & Noonan, 2011 for review), it is still unknown whether this brain region also represents the expected value of aversive stimuli in humans. Moreover, it is unclear whether the mOFC/vmPFC encodes the expected value of different types of outcomes (e.g., juice *vs* money) in the anticipatory period between cue and outcome during Pavlovian learning. The current study was precisely designed to address how the brain encodes expected values at the time of the cue, both in the appetitive and aversive domains.

One fundamental question is to know whether the same mOFC/vmPFC region encodes the prediction of appetitive outcomes with increased activity and the prediction of aversive outcomes with decreased activity (signed expected values), or if it encodes the prediction of both appetitive and aversive stimuli with increased activity (general unsigned expected values). There are reasons to consider both possibilities.

On one hand, in the domain of value-based decision making, where the values assigned to the different stimuli at the time of decision are referred to as "goal values", several fMRI studies have reported mOFC/vmPFC activity that correlates positively with appetitive goal values and that correlates negatively with aversive goal values during free-bid trials for the right to avoid eating disliked foods (Plassmann, O'Doherty, & Rangel, 2007, 2010). Similarly, during a loss aversion paradigm requiring to accept or reject gambles that offered a 50% chance of gaining or losing money, the vmPFC showed increasing activity as potential gains increased and decreasing activity with potential losses (Tom, Fox, Trepel, & Poldrack, 2007).

On the other hand, animal electrophysiological findings showed that neurons from the OFC (both lateral and medial parts) respond to cues that predict the reinforcer in aversive or rewarding situations, supporting the hypothesis of a common representation of expected values regardless of valence (Hosokawa, Kato, Inoue, & Mikami, 2007; Morrison & Salzman, 2009; Schoenbaum et al., 2009; Schoenbaum, Takahashi, Liu, & McDannald, 2011). These latter studies, together with neuroimaging findings, indicate that the mOFC/vmPFC constitutes a good candidate area for mediating a common utility function during anticipation of positive and negative reinforcers.

Most of the fMRI studies investigating appetitive and aversive outcomes only concern outcome or goal values signals and have not systematically investigated the neural representation of expected values for different types of rewards and punishments during Pavlovian conditioning. Here, we used a task design simultaneously manipulating positive and negative outcomes to directly investigate whether the mOFC/vmPFC encodes a general value signal (responding positively both for anticipated rewards and punishments, regardless of their types) or whether it encodes a signed expected values (responding positively for rewards and negatively for punishments).

If the mOFC/vmPFC encodes a general 'unsigned' expected value, we would predict increased activity for cues predicting both rewards and punishments. Alternatively, if the mOFC/vmPFC encodes a signed expected value, we would predict opposite activity for cues predicting rewards and punishments, i.e., decreased activity during anticipation of punishment and increased activity during reward anticipation.

Our paradigm also allowed us to test whether activation of this brain region depends upon reinforcer type (e.g., juice *vs* money). This is a crucial question given the important differences between the reinforcers used across species in the study of reinforcement learning: human participants typically experience secondary reinforcers, such as monetary gains, while animals typically receive primary reinforcers, such as food reward. We cannot generalize across species until the influence this has on learning is resolved. Furthermore, although recent studies suggest a posterior versus anterior topographical organization of experienced value signals in the lateral OFC for different reinforcer types at outcome (Sescousse et al., 2010; Sescousse, Caldú, Segura, & Dreher,

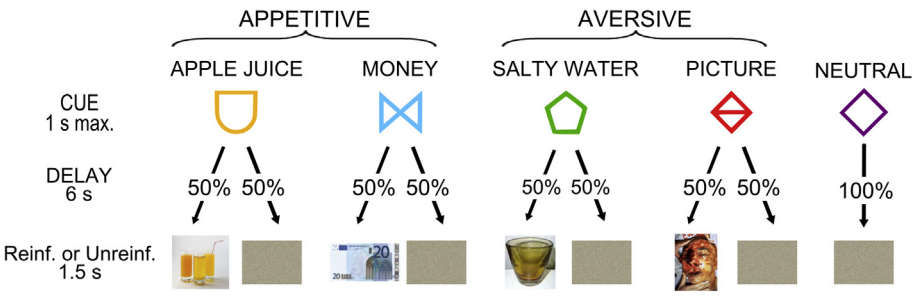

Fig. 1 – Experimental design. Subjects learned to associate various cues with 4 different types of reinforcers (2 appetitive and 2 aversive) in a classical reinforcement learning paradigm. Two types of cues were followed by positive reinforcers (apple juice, money) on 50% of occasions or by a scrambled picture (unreinforced), two other types of cues were followed by negative reinforcers (salty water, aversive picture) on 50% of occasions or by a scrambled picture (unreinforced), while some cues were always followed by a scrambled picture (neutral condition). After the cue presentation, subjects pressed a response button (<1 sec), followed by a 6 sec delay period (fixation cross) and by the positive/negative reinforcer or by a scrambled picture (Unreinf.: Unreinforced, Reinf.: Reinforced). We only investigated expected value signals during the anticipation period.

2013), it is still unclear whether similar dissociation can be observed during reinforcers anticipation (Kim et al., 2010).

2. Methods

Note that the behavioral data collected for this experiment and the fMRI prediction error results have been published previously (Metereau & Dreher, 2013). However, all the fMRI analyses and results presented in this paper, which exclusively concern the anticipatory phase and the signed and unsigned value signals are new.

2.1. *Subjects*

Twenty healthy volunteers (10 females) with no history of neurological or psychiatric illness participated in the experiment (mean age: 24.4; range: 18–33). All subjects were right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). The study was approved by the local ethics committee (Lyon) and all subjects gave informed consent.

2.2. *Materials and procedure*

The task design simultaneously manipulated 2 appetitive and 2 aversive reinforcers of different types, allowing us to directly investigate whether overlapping mPFC/mOFC regions encode expected values according to the reinforcer valence and type. The four reinforcers that we used were apple juice, salty water (.2 M of NaCl), monetary incentive and aversive picture (Fig. 1). Monetary loss was not included as reinforcer because it may not act as a primary punishment, as aversive juice or aversive pictures and because monetary gains and losses are known to weigh differently (Tom et al., 2007). A Pavlovian conditioning procedure was used with each trial divided into two phases: anticipation and reception. The anticipation phase began with an affectively neutral visual cue (geometric form) displayed until a response button was made (in less than 1 sec). This cue was followed by a delay period of 6 sec displaying a fixation cross. Then, in the reception phase, either the corresponding reinforcer or a scrambled picture was presented for 1.5 sec, with a probability of .5. Each reinforced trial was followed by real consequences: (i) in the apple juice condition, the subjects were delivered .05 ml of apple juice in the mouth while they were presented with a picture representing a glass of apple juice; (ii) in the salty water condition, the subjects were delivered .05 ml of salty water while they were presented with a picture representing a brown glass of water; (iii) in the aversive picture condition, an aversive picture was presented; (iv) in the monetary reward condition, subjects were presented with a 20 Euros bill picture and were informed that they would earn a percentage of each of these bills at the end of the experiment. A blank screen was finally used as an inter-trial interval of duration (2.5 sec–5.5 sec). In a fifth condition (neutral condition), a cue announced the neutral scrambled picture with certainty. To maintain the subjects' attention, they were asked to press a response button as soon as they saw the cue. Subjects were explicitly informed that the delivery of the reinforcer was independent of their responses and knew that the cue would disappear after 1 sec, even if they did not make a key press. The subjects were pre-trained in the same type of task one week before the scanning session, with various probabilities ranging from 0 to 1. The cues used for the pre-training session were different from the cues use in the scanning session. On the scanning day, there was no pre-training with the cues used in the scanner to ensure learning during the scanning session.

2.3. *Stimuli and reinforcers*

All the visual stimuli were back-projected on a screen located at the head of the scanner bed and presented to the subjects through an adjustable mirror located above their head. The presentation of the stimuli, as well as the juice delivery were controlled by Presentation© software (Neurobehavioral

Systems), which also recorded trigger pulses from the scanner signaling the beginning of each volume acquisition.

The aversive picture was chosen from the International Affective Picture System (Lang, Bradley, & Cuthbert, 2005) and showed a highly repelling mutilated face (picture n°3060: valence = 1.79 ± 1.56; arousal = 7.12 ± 2.09).

Subjects were not told the percentage of the 20 Euro bill that they earned on each reinforced trial of the monetary condition to avoid counting during the experiment. By the end of the experiment, each subject had seen 24 bills and earned 20€, in addition to the 50€ earned for being scanned. Thus, subjects were paid a fixed amount of 70€ for their participation.

The two liquids reinforcers were contained in two 60 ml syringes, connected to an IVAC P7000 electronic pump positioned in the scanner control room and were delivered in the subjects' mouth *via* two separate 6 m long and 1 mm wide polyethylene tubes. This small amount of liquid (2.4 ml over the whole experience) was chosen to minimize any satiety effect that could occur during the experiment. Moreover, subjects were drink-deprived for 12 h prior to scanning to ensure that they remained thirsty during the experiment. In order to reduce head movement related to swallowing, subjects were instructed to swallow only during the inter-trial interval, after the reinforcer offset and before the new trial onset.

All unreinforced trials were represented by the presentation of a unique scrambled neutral picture. It should be noted that our four reinforcers did not differ by only one factor. Indeed, the apple juice reinforcer is appetitive, gustatory, visual and immediate; the monetary incentive is appetitive, visual and delayed; the salty water reinforcer is aversive gustatory, visual and immediate and the aversive picture reinforcer is aversive, visual and immediate. However, this is not a major problem because we did not perform one to one comparison between the expected outcomes related to each reinforcer.

2.4. Experimental design

The experiment consisted of 3 scanning runs of 15 min. In each run, the five conditions were presented one time each, in blocks of 16 successive trials. The order of the condition was balanced across runs according to a Latin square design (for example: run I: 12345, run II: 43521, run III: 51432, where 1, 2, 3, 4, 5 corresponds to each condition) and different from one subject to another. In each run, a new cue was used for each condition and subjects had to learn the probabilistic association between this cue and the corresponding reinforcer. Thus, all the cues-outcome mappings were novel in each run and the probabilistic outcome association was always .5 (except for the control condition). The trials from the different conditions were not mixed between each others to avoid relative comparison between the values of the different reinforcers. Indeed, a large body of literature reports context-dependent activity in different components of the reward system (Tremblay & Schultz, 1999; Nieuwenhuis et al., 2005). This design allowed us to investigate the neural representation of expected outcome signals for different appetitive and aversive reinforcers.

2.5. Behavioral measures

First, to assess the level of thirst of the subjects, we asked them to rate their thirst on a scale ranging from 1 (not thirsty at all) to 5 (extremely thirsty), both before and after scanning.

Second, in order to investigate whether the subjects have a real preference for cues explicitly associated with positive reinforcers relative to cues associated with negative reinforcers, we asked them to perform a two-alternative forced-choice preference task before the scanning session. In this task, on each trial (72 trials total), two out of four possible cues explicitly indicating both the type of reinforcer and the chance to be reinforced ($p = .5$), were presented side-by-side. Subjects had to choose which one they preferred by a left or right key press. There were 6 different pairs of cues, repeated 12 times each, and the cues were randomly assigned to the left or right side of the screen. All the cues used for this task were different from the cues presented during the scanning session. After the decision, the chosen reinforcer was effectively delivered with a probability of $p = .5$ (that is, the cue predicting apple juice or salty water were followed in 50% of the trials by the simultaneous presentation of the corresponding glass and by the delivery of .05 ml of liquid in subjects' mouth). For each reinforcer, we computed a preference score (percent chosen if available) as the number of times one cue was chosen divided by the number of times this cue was presented. Moreover, after the scanning session, we checked that the valence of our reinforcers was well perceived by the subjects by asking them to provide pleasantness ratings for each reinforcer on a scale ranging from −2 (very unpleasant) to 2 (very pleasant).

Finally, to assess whether the probabilities of the cue-reinforcer association were explicitly learnt, during scanning subjects were asked to evaluate the probability of the cue-reinforcer association of each block just after its last trial, by positioning a cursor on a continuous scale from 0 to 1. Percent accuracy were used to assess whether subjects paid attention to the cues. Moreover, subjects' response times (RTs) were compared across conditions with a one-way repeated-measures ANOVA.

The a priori significance level was defined at $p < .05$ for all the behavioral tests.

2.6. fMRI data acquisition and pre-processes

fMRI data were acquired on a 1.5 Tesla Siemens MRI scanner. Blood-oxygenation-level-dependent (BOLD) signal was measured with gradient echo T2* weighted echo-planar images (EPI). Twenty-six interleaved slices parallel to the AC-PC line were acquired per volume (matrix 64 × 64, voxel size = 3.4 × 3.4 × 4 mm). In total 410 volumes were acquired continuously every 2.5 sec for each of the three runs. The first 4 volumes of each run were discarded to allow the BOLD signal to reach a steady state. A T1-weighted structural image (1 × 1 × 1 mm) was also acquired for each subject at the end of the experiment.

Data were preprocessed using the SPM5 software package. First, outlier scans (>1.5% variation in global intensity or >.5 mm/TR scan-to-scan motion) were detected using the ArtRepair SPM toolbox http://cibsr.stanford.edu/tools/human-brain-project/artrepair-software.html. Since less

than 5% of outlier scans were detected per subject, no repair was performed. Then, images were corrected for slice timing and spatially realigned to the first image from the first run. They were normalized to SPM5's EPI template in Montreal Neurological Institute space with a resampled voxel size of $3 \times 3 \times 3$ mm and spatially smoothed using a Gaussian kernel with full-width at half maximum (FWHM) of 8 mm. The T1-weighted structural scan of each subject was normalized to a standard T1 template in Montreal Neurological Institute space with a resampled voxel size of $1 \times 1 \times 1$ mm.

2.7. Computational model

Here, we used a reinforcement learning framework that has been used extensively to describe animal learning in reward-driven tasks and the generation of phasic dopaminergic firing patterns. In this framework, learning occurs through updating expectations in proportion to prediction error (i.e., discrepancy between expected outcome and actual outcome), so that across trials, the expected outcome value converges to the actual outcome value.

We computed the expected values and the prediction errors for each type of cue separately and for each subject according to the sequence of stimuli they received, providing a regressor for the fMRI data. The use of a probabilistic reinforcement strategy, in which the cues are only 50% predictive of their outcomes, ensures constant learning and updating of the expected values and generates both positive and negative prediction error throughout the course of the experiment.

Expected values and prediction error values were calculated trial-by-trial by using a Rescorla–Wagner rule (Rescorla & Wagner, 1972). For each trial t a prediction error $\delta(t)$ was computed as the difference between the actual outcome value $R(t)$ and its expected value $V(t)$ on that trial (Eq. (1)):

$$\delta(t) = R(t) - V(t) \quad (1)$$

Then, the expected value of the next trial $V(t+1)$ was updated by adding the prediction error $\delta(t)$ weighted by a learning rate α (Eq. (2)):

$$V(t+1) = V(t) + \alpha\delta(t) \quad (2)$$

The outcome value $R(t)$ was set to 1 when a reinforcer (either a reward or a punishment) was delivered and to 0 when a scrambled picture was displayed. $V(t)$ was initialized to 0. The learning rate α was derived from subjects' response times (RTs) to the cue. RTs have been shown to be good indicators of conditioning (Critchley, Mathias, & Dolan, 2002; Gottfried et al., 2003) and to be correlated with the prediction $V(t)$ estimated by a reinforcement learning model (Seymour et al., 2004). Several fMRI studies have used RTs to estimate the learning rate of reinforcement learning models during tasks in which the buttons presses were irrelevant to receive the reward (Bray & O'Doherty, 2007; Seymour et al., 2005). First, RTs were normalized to allow analysis across subjects. Then, the first trial of each cues-outcome association was discarded because the RT of this trial was strongly slower due to the novelty effect. We derived the prediction $V(t)$ for each subject based on their individual conditioning histories for a range of learning rates (ranging from .1 to .5). Then, trial-by-trial RTs across subjects were fitted to a regression model that included the prediction $V(t)$. The best fit yielded a learning rate of .1, which is close to the value used in other studies (Bray & O'Doherty, 2007; Jensen et al., 2007; O'Doherty, Buchanan, Seymour, & Dolan, 2006). Fig. 2A shows an example of the successive values taken by $V(t)$, as calculated by this model, during 16 consecutive trials for one of the 4 reinforcers (converging towards .5). The values learnt were positive for the appetitive cues and negative for the aversive cues.

2.8. fMRI data analysis

Statistical analysis was performed using the general linear model (GLM). For each of the five conditions (apple juice, monetary reward, salty water, aversive picture and neutral) two phases (anticipation and reception) were modeled, resulting in the creation of 10 regressors. The anticipation phase was modeled as an epoch, time-locked to the onset time of the cue, with a duration equal to RT + anticipatory period

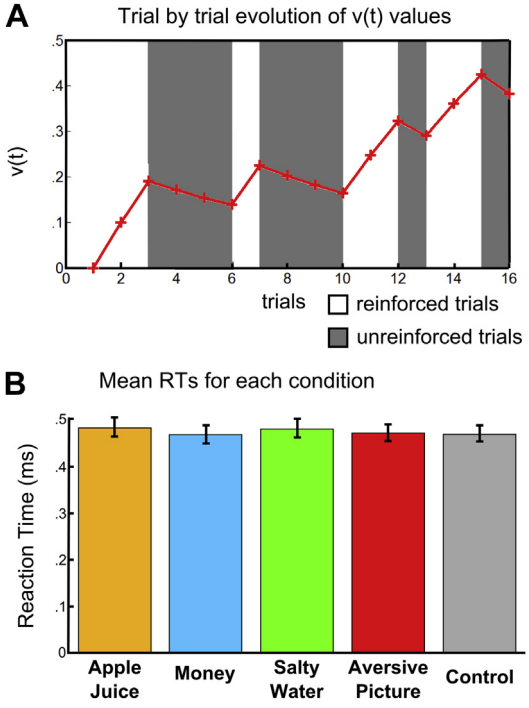

Fig. 2 – Expected value and Response Times. (A) Example of trial by trial evolution of the expected value V(t), as calculated by the Rescorla–Wagner rule with the outcome value R(t) set to 1 when a reinforcer (either a reward or a punishment) is delivered and to 0 otherwise, during 16 consecutive trials for one of the cue. Reinforced trials are represented by a white background while unreinforced trials are displayed on a gray background. (B) Response times (RTs) averaged across subjects, for the detection of the cue in the apple juice condition (yellow), the monetary reward condition (blue), the salty water condition (green), the aversive picture condition (red) and the control condition (gray). Error bar indicate SEM. A one-way repeated-measures ANOVA on these RTs failed to revealed any significant difference at $p < .05$.

Table 1 – MNI coordinates and statistic *t* for regions in which the cerebral activity is modulated by the unsigned expected values (EV).

Regions	Laterality	Nb. of voxels	x	y	z	t
a. (C1+) Localization of activities positively modulated by apple juice EV ($p < .001$ unc., $k \geq 25$)						
Medial orbital gyrus[a]	L	37	−3	30	−18	5.75
Rolandic operculum	R	384	48	−9	18	5.68
Insula	L	54	−36	−9	12	4.77
Inferior parietal lobule	L	38	−60	−21	45	3.92
b. (C2+) Localization of activities positively modulated by monetary EV ($p < .001$ unc., $k \geq 25$)						
Medial orbital gyrus[a]	R	41	15	33	−12	4.09
c. (C3+) Localization of activities positively modulated by salty water EV ($p < .001$ unc., $k \geq 25$)						
Rolandic operculum	R	1216	60	−3	12	8.14
	L	793	−48	−12	15	6.45
Medial orbital gyrus[a]	L	69	−3	36	−12	4.56
d. (C4+) Localization of activities positively modulated by aversive picture EV ($p < .001$ unc., $k \geq 25$)						
Cuneus	R	225	21	−90	39	6.41
Precentral gyrus	R	122	42	−18	69	5.11
SMA	R	107	12	−21	57	5.01
Medial orbital gyrus[a]	R	25	3	42	−18	4.12
Rolandic operculum	R	56	45	−18	27	4.04
	R	36	33	12	57	4.66
	R	28	45	51	−3	4.03
e. Global unsigned EV: conjunction of C1+ & C2+ & C3+ & C4+ ($p < .005$ unc., $k \geq 5$)						
Medial orbital gyrus[a]	L	9	−3	36	−18	3.26

[a] Survives small volume correction (SVC) at FWE $p < .05$ within an 10-mm sphere centered on MNI coordinates from previous work: [x = −2, y = 28, z = −18].

(=6 sec). The reception phase was modeled as a box-car of 1.5 sec duration. For each reinforced condition (i.e., all conditions except the neutral condition), expected values V(*t*) and prediction error δ(*t*), generated by the Rescorla–Wagner model, were used as parametric modulators of the anticipation and reception regressors, respectively. All of these 18 regressors were convolved with a canonical hemodynamic response function. In addition, the six on-going motion parameters estimated during realignment were included as regressors of no interest. A high-pass filter with a cut-off of 128 sec was applied to the time series to remove low-frequency noise and baseline drifts and an AR(1) model plus white noise was used to correct for temporal autocorrelation.

For the second-level analysis, we constructed a one-way flexible factorial design including the images of the parameter estimates (betas) from the expected value [V(*t*)] parametric modulation for each reinforced condition (apple juice, money, salty water and aversive picture), as well as a subject factor accounting for between-subject variability. This model allowed us to test several contrasts to identify areas in which activity is modulated by the expected value during anticipation of the four reinforcers.

We first searched for brain regions in which the activity would be positively modulated by the anticipation of the apple juice condition (contrast C1+) and the monetary condition (C2+). To investigate whether anticipation of aversive reinforcers modulates brain activity positively (unsigned expected value hypothesis) or negatively (signed expected value hypothesis), we tested both possibilities with the contrasts C3+ and C3− for the salty water condition and the contrasts C4+ and C4− for the aversive picture condition. Conjunctions of these contrasts were used to test for the signed expected value hypothesis versus the unsigned expected value hypothesis. That is, we tested for brain activities commonly modulated by the signed expected value during the anticipation of all the reinforcers, regardless of their valence or modality, with a conjunction of the contrasts C1+ & C2+ & C3− & C4−. Likewise, brain regions commonly modulated by the unsigned expected value were identified with a conjunction of the contrasts C1+ & C2+ & C3+ & C4+. All the conjunctions were based on the conjunction null hypothesis, as implemented in SPM5 (Nichols, Brett, Andersson, Wager, & Poline, 2005).

Post hoc examinations of how activity in the OFC/vmPFC scaled with the expected value were conducted using a region of interest (ROI) approach for each subject. First we used the Marsbar toolbox for SPM (http://marsbar.sourceforge.net/) to create a ROI made of the 9 voxels significantly activated in the unsigned expected value conjunction analysis with a threshold of $p < .005$ uncorrected. We then estimated a new GLM in this region for each subject. This model was similar to the main one except that the cue onsets of the trials with low, middle and high expected value were separated in three distinct regressors, allowing us to estimate a beta parameter for each of these three categories, for each condition. Finally, individual beta parameters were averaged across subjects and plotted for each condition to illustrate the correlation between the BOLD signal and the expected value.

2.9. *Activations localization and reported statistics*

Anatomic labeling of activated regions was done using the SPM Anatomy toolbox (http://www.fz-juelich.de/inm/inm-1/DE/Forschung/_docs/SPMAnatomyToolbox/SPMAnatomyToolbox_node.html) and the probabilistic atlas of Hammers et al. (2003). Reported coordinates conform to the Montreal Neurological

Table 2 – MNI coordinates and statistic t for regions in which the cerebral activity is modulated by the signed expected values (EV).

Regions	Laterality	Nb. of voxels	x	y	z	t
a. (C3−) Localization of activities negatively modulated by salty water EV ($p < .001$ unc., $k \geq 25$)						
Precuneus	L	120	−3	−51	75	6.66
Angular gyrus	R	321	30	−60	42	5.66
Inferior occipital gyrus	R	128	51	−66	−12	5.36
	L	761	−42	−66	−15	5.33
Lingual gyrus	R	56	6	−39	3	4.30
Superior parietal gyrus	L	107	−24	−57	66	4.29
Fusiform gyrus	R	25	36	−54	−12	4.16
b. (C4−) Localization of activities negatively modulated by aversive picture EV ($p < .001$ unc., $k \geq 25$)						
Middle occipital gyrus	R	177	30	−96	0	7.21
	L	188	−21	−99	3	6.32
Inferior parietal lobule	L	410	−45	−45	60	5.79
Angular gyrus	R	680	33	−60	48	5.60
Precentral gyrus	L	55	−57	9	36	4.70
Middle frontal gyrus	R	255	48	36	33	5.34
	R	36	33	12	57	4.66
	R	28	45	51	−3	4.03
c. Global signed EV: conjunction of C1+ & C2+ & C3− & C4− ($p < .005$ unc., $k \geq 5$)						
None						

Institute (MNI) space. Given our specific a priori in the vmPFC, a small volume correction (SVC) was performed in this region. This SVC was performed with a Family Error Wise (FEW) corrected significance threshold of $p < .05$, in a 10-mm radius sphere centered on the coordinates [x,y,z = −2, 28, −18] taken from a meta-analysis reporting a peak of activity in the vmPFC for subjective value computation (Clithero & Rangel, 2013). For completeness, Tables 1 and 2 list all the regions displaying an effect with a voxel-wise significant threshold of $p < .001$ uncorrected and a cluster size ≥25 voxels, except for conjunction analyses which are reported with a threshold of $p < .005$ uncorrected and a cluster size ≥5 voxels, since conjunctions are very conservative (Friston, Penny, & Glaser, 2005).

3. Results

3.1. Behavioral results

First, subjects' self reports indicated that they were drink-deprived for an average of 12 h 24 min ± 2 h 16 min prior to scanning, therefore respecting their instructions to be drink-deprived for 12 h. No significant difference was observed between the ratings of the thirst sensation performed before and after scanning (before: 3.16 ± 1.01, after: 2.74 ± .99; paired t-test: $t(18) = -1.41$, $p = .180$), supporting that the motivation to drink was stable over the course of the experiment.

Second, the preference scores showed that the cues announcing positive reinforcers (apple juice and money) were preferred to those announcing negative reinforcers (salty water and aversive picture) ($F_{(1,19)} = 2684.94$, $p = 10^{-6}$). Moreover, post-scan subjective ratings confirmed that subjects perceived positive reinforcers as more pleasant than negative reinforcers ($F_{(1,19)} = 180.91$, $p = 10^{-6}$).

Finally, during scanning, the percent accuracy for detecting the cues was 97%, confirming that subjects paid attention to the cues. No significant difference in the RTs was found according to the reinforcer type at the time of the cue ($F_{(4,72)} = 2.0242$, $p = .1$; Fig 2B). In addition, the probabilities of the association between cues and reinforcers were explicitly learnt in a valence-incentive manner since for each of the 4 conditions having a probability of $p = .5$, the rating of the estimated probability that a cue led to a specific reinforcer was non-significantly different from .5. Moreover, for the neutral condition ($p = 0$), this estimated probability did not differ from 0.

3.2. fMRI results

3.2.1. Brain activity positively modulated by the expected value of each anticipated reinforcer

We first performed 4 whole-brain comparisons to identify areas in which the activity is positively modulated by the expected value during the anticipation of each reinforcer: apple juice (C1+), monetary reward (C2+), salty water (C3+) and the aversive picture (C4+). Each of these analyses led to an activation in the mOFC/vmPFC: apple juice expected value: [x = −3, y = 30, z = −18], t = 5.75, FWE $p < .05$, small-volume-corrected; monetary expected value [x = 15, y = 33, z = −12], t = 4.09, FWE $p < .05$, small-volume-corrected (Fig. 3A, Table 1a and b); salty water expected value: [x = −3, y = 36, z = −12], t = 4.56, FWE $p < .05$, small-volume-corrected; aversive picture expected value: [x = 3, y = 42, z = −18], t = 4.12, FWE $p < .05$, small-volume-corrected (Fig. 3B, Table 1c and d). Other foci revealed by these four analyses are reported in Table 1a, b, c and d respectively. Thus, the mOFC/vmPFC does not respond in opposite fashion for cues predicting rewards and punishments. Instead, both cues predicting these two types of outcomes engaged similar mOFC/vmPFC region.

3.2.2. *Brain activity negatively modulated by the expected value during the anticipation of the aversive reinforcers*

In the same way, two whole-brain comparisons allowed us to identify areas in which the activity is negatively modulated by the expected value during the anticipation of the salty water (C3−) and the aversive picture (C4−). These analyses revealed activities mainly in the parietal and the occipital lobes, among other regions (see Table 2a and b).

3.2.3. *Brain activities commonly modulated by the unsigned expected value for all the reinforcers*

We identified the brain activities commonly modulated by the unsigned expected values during the anticipation of all the reinforcers, regardless of their valence or modality, with a conjunction of the contrasts C1+ & C2+ & C3+ & C4+. This analysis revealed that the mOFC/vmPFC is commonly activated by the unsigned expected value for all types of reinforcers ([$x = -3$, $y = 36$, $z = -18$], $t = 3.26$, $k = 9$, with $p < .005$ uncorrected) (Fig. 4 upper part). This analysis survived the FWE $p < .05$ small volume correction within a 10-mm radius sphere centered on [$x = -2$, $y = 28$, $z = -18$]. To illustrate this effect, we used a ROI approach in the mOFC/vmPFC cluster revealed by this conjunction analysis, in which we estimated a new model with three distinct regressors for the trials with low, middle and high expected value for each subject. As expected, the plot of the averaged beta parameters across subject showed that the higher the expected value, the higher the BOLD signal is for each condition (Fig. 4, lower part). No other area of the brain showed significant correlations in the whole brain conjunction analysis at $p < .005$ uncorrected (Table 1e). Together, these results provide evidence for a common representation of the expected value in the mOFC/vmPFC, which was not found anywhere else in the brain.

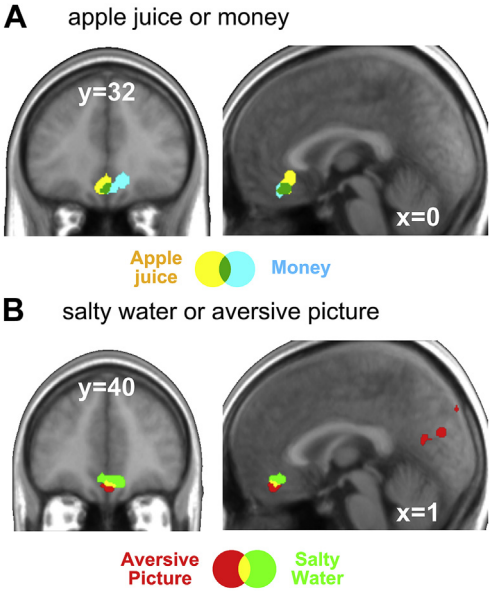

Fig. 3 − Activity in OFC is positively modulated by the expected value for each reinforcer. (A) The activity of the mOFC/vmPFC is positively modulated by the expected value V(t) during anticipation of apple juice (yellow) and money (blue). The overlap between these two conditions is represented in green. The statistical maps of the two conditions were thresholded at $p < .001$, uncorrected with a cluster size ≥25 voxels. (B) The activity of the mOFC/vmPFC is positively modulated by the expected value V(t) during anticipation of salty water (light green) and of aversive pictures (red). The overlap between these two conditions is represented in yellow. For illustration, the statistical maps of the two conditions were thresholded at $p < .001$, uncorrected with a cluster size ≥25 voxels.

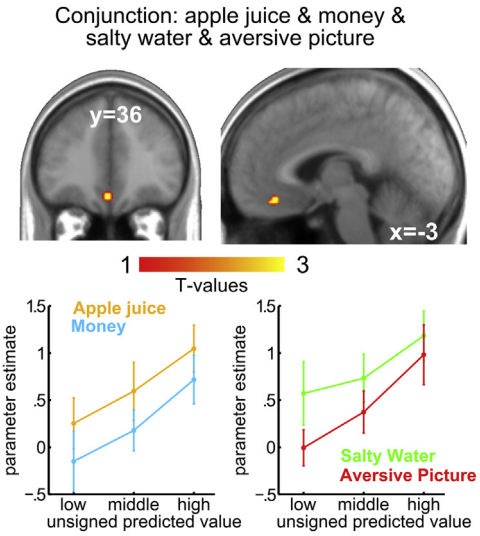

Fig. 4 − Activity in OFC is commonly modulated by the unsigned expected value for all the reinforcers. Activation resulting from the conjunction (logical AND) of the positive correlation between the expected value and the mOFC/vmPFC activity for the four conditions. Color bars represent t values. Plots illustrate the positive correlation between the BOLD signal and the unsigned expected value for each condition. A new model separating the cues onset of the trials with a low, middle and high expected value was used to estimate the parameter estimate (beta) for each of these three categories. This model was estimated in the mOFC/vmPFC region identified by the unsigned expected value conjunction analysis ($p < .005$ unc.) for each subject and the betas were then averaged across subjects. Error bars indicate SEM.

3.2.4. *Brain activities commonly modulated by the signed expected value for all the reinforcers*

To test for the signed expected value hypothesis, we performed the conjunction of the contrasts C1+ & C2+ & C3− & C4−. Even at the very liberal uncorrected threshold of $p < .1$, there were no significant voxels activated by this conjunction (Table 2c).

4. Discussion

This study used a model-based fMRI approach to show that the mOFC/vmPFC encodes a general expected value signal for different types of appetitive and aversive reinforcers during

Pavlovian learning. One important implication of our observed positive correlation between mOFC/vmPFC activity and expected value signal regardless of outcome valence is that it did not encode the signed value of outcomes. Instead, our data indicate that this brain region is responding positively to the predictiveness of an outcome independently of its valence. This finding extends reports from discrimination learning tasks, showing that neurons in the rat OFC initially fire in response to either the rewarding reinforcer or the aversive reinforce (Schoenbaum et al., 2009). After a number of trials, these neurons fire in anticipation of the reinforcer and then fire in response to cues that predict the reinforcer. Interestingly, recent monkey single neurons recordings reported regionally distinct vmPFC for processing rewards and punishments (Monosov & Hikosaka, 2012). These previous findings are not necessarily contradictory with our current results because, given the spatial limitation of fMRI, we cannot rule out the possibility that the mOFC/vmPFC region we identified contains neurons that encode separately for the expected value of appetitive and aversive cues. In fact, the mOFC/vmPFC activity we observed with the unsigned expected value signal for different types of rewards and punishments is likely to result from distributed populations of neurons with different tuning properties. Indeed, positive and negative value coding neurons, respectively increasing and decreasing their firing rates with increasing value, coexist in the OFC (Kennerley, Dahmubed, Lara, & Wallis, 2009; Kobayashi, Pinto de Carvalho, & Schultz, 2010; Morrison & Salzman, 2009; Padoa-Schioppa & Assad, 2006). fMRI studies using multivariate pattern classification analyses (MVPA) can now be used to interpret overlapping functional activations from independent contrasts and distinguish between a common-coding interpretation – a shared region is thought to contain neurons that are engaged in a common computational process – from a functional independence interpretation – two overlapping but functionally independent neural populations are thought to be engaged within the common region – (Peelen & Downing, 2007). Future fMRI experiments designed for MVPA should help to test whether the vmPFC activation reflects shared neural processing for expected value of appetitive and aversive cues.

In the current study, we observed mOFC/vmPFC activity during the anticipatory period responding to the predictiveness of an outcome independently of its valence. Such mOFC/vmPFC activity may reflect a signal encoding acquired salience, increasing with learning of a partially rewarded or punished cue and being similar for different valence and different types of outcomes (Ogawa et al., 2013). At the time of outcome, we previously searched for brain response correlating with a salient prediction error for appetitive and aversive reinforcers, that is, for the set of brain regions showing high BOLD signal when both rewards and punishments are delivered and low BOLD signal for their omission (Metereau & Dreher, 2013). This brain network encompassed the striatum, anterior insula and anterior cingulate cortex for appetitive and aversive juice. However, no vmPFC activity was observed with the salient prediction error, potentially reflecting the fact that learning of the stimuli-outcome associations could not be fully established since the cue-outcome association had maximal uncertainty (reinforcer probability = .5). Confirming this hypothesis, a number of previous fMRI studies investigating anticipatory and experienced value signals reported vmPFC activity in both phases when no stimulus-outcome learning was required (in these studies the cue indicated explicitly the reward probability) (Kahnt, Heinzle, Park, & Haynes, 2010; Sescousse et al., 2010; Sescousse, Caldú et al., 2013; Sescousse, Barbalat, Domenech, & Dreher, 2013; Sescousse, Li, & Dreher, 2014). Another fMRI study recently reported that such similarity between vmPFC patterns for the reward value of predicted and actual outcomes was also present after, but not before, learning cue-outcome associations with 100% contingency, further suggesting that the representations of predicted and actual outcomes become similar during learning (Kahnt, Heinzle, Park, & Haynes, 2011).

Only a few previous fMRI studies investigated anticipatory signals, and rarely combined model-based fMRI approach of anticipatory value signals and investigation of different types of rewards and punishments (Jensen et al., 2003, 2007; Kim et al., 2010; O'Doherty et al., 2002). A partially overlapping response for the expectation of apple juice and monetary rewards was recently found in the vmPFC (Kim et al., 2010). However, this study did not report vmPFC activity during anticipation of aversive reinforcers (i.e., monetary loss and salty tea). Yet, our results are difficult to compare with those of this previous study, because it did not investigate whether the vmPFC codes an unsigned expected value signal. Moreover, this previous study did not model the cue-outcome learning process and the same types of cues (specific shapes) were always associated (outcome probability = 1) with the same outcomes, while in the current design learning was maximal (outcome probability = .5).

The view that the mOFC/vmPFC codes a general unsigned expected value signal for both appetitive and aversive reinforcers is consistent with direct electrophysiological recordings showing that neurons in macaque OFC respond to both a predicting stimulus of an electrical shock and a reward-predicting stimulus (Hosokawa et al., 2007). Neurons in the OFC also code relative preference of both reward and aversive outcomes (Hosokawa et al., 2007; Morrison & Salzman, 2009; Schoenbaum et al., 2009) and a cardinal-like valuation scale (Padoa-Schioppa & Assad, 2006, 2008). Moreover, rodent's OFC recordings indicate that this brain region tends to anticipate the event instead of being triggered by it (Schoenbaum et al., 2011) and monkey electrophysiological studies demonstrated that vmPFC/mOFC neuronal activity is modulated by factors that determine reward's value, such as satiety levels and the animals' motivational state (Bouret & Richmond, 2010). All these studies point to a general (appetitive and aversive) expected outcome signal of the vmPFC during outcome anticipation.

Anticipatory firing of expected rewards has also been observed in other areas, such as midbrain dopaminergic neurons (Fiorillo, Tobler, & Schultz, 2003), but recent electrophysiological findings indicate that they may occur first in the OFC and be sent to dopaminergic neurons afterward (Takahashi et al., 2011). Thus, the anticipatory outcome signal of the mOFC may be a key signal for midbrain dopamine neurons, which encode not only a reward prediction error (Schultz, 1998), but also a salient prediction error signal responding positively to both appetitive and aversive stimuli

(Matsumoto & Hikosaka, 2009; Bromberg-Martin, Matsumoto, & Hikosaka, 2010). In turn, these two midbrain dopamine signals may be important for encoding the unsigned expected value signal in the vmPFC and may contribute to the learning of both aversive and appetitive cue-outcomes associations.

Our current results are in contrast with two fMRI studies investigating goal value or outcome value signals (Plassmann et al., 2010; Tom et al., 2007), which reported a signed value signal in the mOFC (i.e., positive correlation of the activity with appetitive outcomes and negative correlation with negative outcomes). One reason for this apparent discrepancy is that the goal value and outcome value signals investigated in these previous studies require distinct neural computations than the expected value signal from the current study. Another reason is that monetary loss and placing bids may not engage an anticipatory salient signal for aversive outcomes in the same way than a physical punishment, such as an aversive juice really experienced in the scanner.

Aversive reinforcers, such as monetary losses, unattractive faces and aversive odors have sometimes been found to elicit activity in the lateral OFC (Elliott, Agnew, & Deakin, 2010; Liu et al., 2007; O'Doherty, Kringelbach, Rolls, Hornak, & Andrews, 2001; Ursu & Carter, 2005). Yet, these results have mostly been reported at the time of outcome and not during anticipation of the reinforcers. For instance, O'Doherty, Critchley, Deichmann, and Dolan (2003) found that lateral OFC was engaged after detection of contingency changes after a monetary loss rather than by a monetary loss *per se*. In the same line of research, Elliott et al. (2010) recently reported that a right lateral OFC region responding to punishment also responded to cues for behavioral change whereas a more ventral and anterior bilateral OFC region responded to cues for behavioral maintenance. These fMRI results support the view that punishments and shift cues are associated with a similar right lateral OFC region, suggesting a connection between emotional response to negative reinforcement and use of negative information to cue behavioral change. However, a number of fMRI studies found no evidence of a medio-lateral OFC dissociation between appetitive and aversive reinforcers for outcome or decision value signals (Breiter, Aharon, Kahneman, Dale, & Shizgal, 2001; Noonan, Mars, & Rushworth, 2011; Tom et al., 2007).

Most of previous fMRI studies using classical conditioning paradigms have focused either on prediction error-related activity (D'Ardenne, McClure, Nystrom, & Cohen, 2008; Menon et al., 2007; Metereau & Dreher, 2013; McClure et al., 2003; O'Doherty, Dayan, Friston, Critchley, & Dolan, 2003; O'Doherty et al., 2004; Seymour et al., 2004; Tobler, O'doherty, Dolan, & Schultz, 2006) or on outcome-related activity (Francis et al., 1999; O'Doherty, Rolls, Francis, Bowtell, & McGlone, 2001; Sescousse et al., 2010; Small et al., 1999; Zald, Lee, Fluegel, & Pardo, 1998). These fMRI studies often reported ventral striatal activity to prediction error and/or to salient prediction error (i.e., responding positively to both appetitive and aversive reinforcers) (D'Ardenne et al., 2008; McClure et al., 2003; Menon et al., 2007; Metereau & Dreher, 2013; O'Doherty, Dayan, et al., 2003; O'Doherty et al., 2004; Sescousse et al., 2010; Tobler et al., 2006) and mOFC/vmPFC activity at the time of rewarded outcome (experienced value) relative to either neutral or negative outcomes (Elliott et al., 2010; Sescousse et al., 2010). Indeed, the mOFC/vmPFC carries information regarding the outcome value of rewards, responding to the reception of a number of rewards regardless of their types, while the anterior and posterior parts of the lateral OFC are differentially recruited for different reward types (Sescousse et al., 2010, Sescousse, Caldú, et al., 2013). That is, the anterior part of the lateral OFC is more engaged with secondary rewards while the posterior lateral OFC is more engaged with primary rewards. This anterior-posterior lateral OFC dissociation has recently been confirmed for different primary rewards in a large quantitative meta-analysis of neuroimaging studies (Sescousse, Caldú, et al., 2013).

Our current results suggest that encoding a general expected value signal in the mOFC/vmPFC may preside encoding of value signals computed subsequently, such as outcome value and decision value signals. This may explain why this brain region is engaged in a variety of situations, such as during the anticipatory period between a cue and reinforcer in Pavlovian conditioning, at the time of outcome or during decision making. This may also explain why the value of many different categories of primary and secondary reinforcers-money, sexual images, food and nonfood consumables-received after making different types of response or after choosing between visual stimuli, all appear to be represented in the mOFC/vmPFC (Behrens, Hunt, Woolrich, & Rushworth, 2008; Chib, Rangel, Shimojo, & O'Doherty, 2009; Gläscher, Hampton, & O'Doherty, 2009; Hare, O'Doherty, Camerer, Schultz, & Rangel, 2008; Hare, Schultz, Camerer, O'Doherty, & Rangel, 2011; Lim, O'Doherty, & Rangel, 2011; Lin, Adolphs, & Rangel, 2012; Litt, Plassmann, Shiv, & Rangel, 2011; Noonan et al., 2011; Plassmann et al., 2010, 2007; Prevost et al., 2010; Sescousse et al., 2010, 2014; Tanaka et al., 2004; Wunderlich, Rangel, & O'Doherty, 2010).

A remaining question relates to the comparative nature of the mOFC/vmPFC signal and its relationship to the process of making a decision. This value comparison function is compatible with the expected value signal from the mOFC/vmPFC reported in the current study. Indeed, the vmPFC may integrate information from other brain regions into a value signal computing a difference-based comparator operation (Basten, Biele, Heekeren, & Fiebach, 2010) or a sum of the stimulus values (Hare et al., 2011). The mOFC/vmPFC signal may reflect the difference in value between the choice that is taken and the choice that is rejected (Boorman & Rushworth, 2009; FitzGerald, Seymour, & Dolan, 2009), both when the choices are made between stimuli indicating the probability and magnitude of money that might be won or when they are made between bundles of consumer items. Further studies will need to address whether economic choices are implemented in the human brain as a two-stage process in which the mOFC/vmPFC decision value signal drives a sequential sampling decision process implemented somewhere else in the brain.

Our study reveals the contributions of the mOFC/vmPFC in the computations underlying Pavlovian learning of different types of rewards and punishments. Our observed mOFC/vmPFC activity coding positively an expected value signal for both positive and aversive reinforcers bridges the gap between the appetitive and aversive conditioning literatures (Jensen

et al., 2007; Kim et al., 2010; Menon et al., 2007; Sarinopoulos et al., 2010; Seymour et al., 2004). In conclusion, our findings advance our understanding of the neurobiological mechanisms underlying the ability to anticipate different types of rewards or punishments. Our model-based fMRI approach pinpoint the role of the mOFC/vmPFC in encoding a general unsigned expected value signal, regardless of valence and type of reinforcers (gustatory, visual, monetary). It finally provides direct empirical evidence for formal learning theories that posit a critical role for the expected value signal during Pavlovian conditioning in humans.

Acknowledgments

This work was funded by the ANR EMCO-HEIDI 2011 and a Fyssen Foundation grant to J-C D. This work was also supported by the Laboratory of Excellence LABEX CORTEX (ANR-11-LABEX-0042), within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR). We thank the CERMEP staff for help during scanning.

REFERENCES

Basten, U., Biele, G., Heekeren, H. R., & Fiebach, C. J. (2010). How the brain integrates costs and benefits during decision making. *Proceedings of the National Academy of Sciences of the United States of America, 107*, 21767–21772.

Behrens, T. E. J., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. S. (2008). Associative learning of social value. *Nature, 456*, 245–249.

Berns, G. S., McClure, S. M., Pagnoni, G., & Montague, P. R. (2001). Predictability modulates human brain response to reward. *The Journal of Neuroscience, 21*, 2793–2798.

Boorman, E. D., & Noonan, M. P. (2011). Contributions of ventromedial prefrontal and frontal polar cortex to reinforcement learning and value-based choice. In R. B. Mars, J. Sallet, M. F. S. Rushworth, & N. Yeung (Eds.), *Neural basis of motivational and cognitive control* (pp. 54–74). The MIT Press.

Boorman, E. D., & Rushworth, M. F. S. (2009). Conceptual representation and the making of new decisions. *Neuron, 63*, 721–723.

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

Bray, S., & O'Doherty, J. (2007). Neural coding of reward-prediction error signals during classical conditioning with attractive faces. *Journal of Neurophysiology, 97*, 3036–3045.

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.

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

Büchel, C., Dolan, R. J., Armony, J. L., & Friston, K. J. (1999). Amygdala-hippocampal involvement in human aversive trace conditioning revealed through event-related functional magnetic resonance imaging. *The Journal of Neuroscience, 19*, 10869–10876.

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. *The Journal of Neuroscience, 29*, 12315–12320.

Clithero, J. A., & Rangel, A. (2013). Informatic parcellation of the network involved in the computation of subjective value. *Social Cognitive and Affective Neuroscience, 9*, 1289–1302.

Critchley, H. D., Mathias, C. J., & Dolan, R. J. (2002). Fear conditioning in humans: the influence of awareness and autonomic arousal on functional neuroanatomy. *Neuron, 33*, 653–663.

D'Ardenne, K., McClure, S. M., Nystrom, L. E., & Cohen, J. D. (2008). BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science, 319*, 1264–1267.

Delgado, M. R., Li, J., Schiller, D., & Phelps, E. A. (2008). The role of the striatum in aversive learning and aversive prediction errors. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 363*, 3787–3800.

Elliott, R., Agnew, Z., & Deakin, J. F. W. (2010). Hedonic and informational functions of the human orbitofrontal cortex. *Cerebral Cortex, 20*, 198–204.

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

FitzGerald, T. H. B., Seymour, B., & Dolan, R. J. (2009). The role of human orbitofrontal cortex in value comparison for incommensurable objects. *The Journal of Neuroscience, 29*, 8388–8395.

Francis, S., Rolls, E. T., Bowtell, R., McGlone, F., O'Doherty, J., Browning, A., et al. (1999). The representation of pleasant touch in the brain and its relationship with taste and olfactory areas. *NeuroReport, 10*, 453–459.

Friston, K. J., Penny, W. D., & Glaser, D. E. (2005). Conjunction revisited. *NeuroImage, 25*, 661–667.

Gläscher, J., Hampton, A. N., & O'Doherty, J. P. (2009). Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. *Cerebral Cortex, 19*, 483–495.

Gottfried, J. A., Deichmann, R., Winston, J. S., & Dolan, R. J. (2002). Functional heterogeneity in human olfactory cortex: an event-related functional magnetic resonance imaging study. *The Journal of Neuroscience, 22*, 10819–10828.

Gottfried, J. A., O'Doherty, J., & Dolan, R. J. (2003). Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science, 301*, 1104–1107.

Hammers, A., Allom, R., Koepp, M. J., Free, S. L., Myers, R., Lemieux, L., et al. (2003). Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. *Human Brain Mapping, 19*, 224–247.

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

Hare, T. A., Schultz, W., Camerer, C. F., O'Doherty, J. P., & Rangel, A. (2011). Transformation of stimulus value signals into motor commands during simple choice. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 18120–18125.

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

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.

Jensen, J., Smith, A. J., Willeit, M., Crawley, A. P., Mikulis, D. J., Vitcu, I., et al. (2007). Separate brain regions code for salience

vs. valence during reward prediction in humans. *Human Brain Mapping, 28*, 294–302.
Kahnt, T., Heinzle, J., Park, S. Q., & Haynes, J.-D. (2010). The neural code of reward anticipation in human orbitofrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America, 107*, 6010–6015.
Kahnt, T., Heinzle, J., Park, S. Q., & Haynes, J.-D. (2011). Decoding the formation of reward predictions across learning. *The Journal of Neuroscience, 31*, 14624–14630.
Kennerley, S. W., Dahmubed, A. F., Lara, A. H., & Wallis, J. D. (2009). Neurons in the frontal lobe encode the value of multiple decision variables. *Journal of Cognitive Neuroscience, 21*, 1162–1178.
Kim, H., Shimojo, S., & O'Doherty, J. P. (2010). Overlapping responses for the expectation of juice and money rewards in human ventromedial prefrontal cortex. *Cerebral Cortex, 21*, 769–776.
Knight, D. C., Waters, N. S., King, M. K., & Bandettini, P. A. (2010). Learning-related diminution of unconditioned SCR and fMRI signal responses. *NeuroImage, 49*, 843–848.
Knutson, B., Westdorp, A., Kaiser, E., & Hommer, D. (2000). FMRI visualization of brain activity during a monetary incentive delay task. *NeuroImage, 12*, 20–27.
Kobayashi, S., Pinto de Carvalho, O., & Schultz, W. (2010). Adaptation of reward sensitivity in orbitofrontal neurons. *The Journal of Neuroscience, 30*, 534–544.
LaBar, K. S., Gatenby, J. C., Gore, J. C., LeDoux, J. E., & Phelps, E. A. (1998). Human amygdala activation during conditioned fear acquisition and extinction: a mixed-trial fMRI study. *Neuron, 20*, 937–945.
Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*. Gainesville, FL: University of Florida.
Lim, S.-L., O'Doherty, J. P., & Rangel, A. (2011). The decision value computations in the vmPFC and striatum use a relative value code that is guided by visual attention. *The Journal of Neuroscience, 31*, 13214–13223.
Lin, A., Adolphs, R., & Rangel, A. (2012). Social and monetary reward learning engage overlapping neural substrates. *Social Cognitive and Affective Neuroscience, 7*, 274–281.
Litt, A., Plassmann, H., Shiv, B., & Rangel, A. (2011). Dissociating valuation and saliency signals during decision-making. *Cerebral Cortex, 21*, 95–102.
Liu, X., Powell, D. K., Wang, H., Gold, B. T., Corbly, C. R., & Joseph, J. E. (2007). Functional dissociation in frontal and striatal areas for processing of positive and negative reward information. *The Journal of Neuroscience, 27*, 4587–4597.
Matsumoto, M., & Hikosaka, O. (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature, 459*, 837–841.
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.
Menon, M., Jensen, J., Vitcu, I., Graff-Guerrero, A., Crawley, A., Smith, M. A., et al. (2007). Temporal difference modeling of the blood-oxygen level dependent response during aversive conditioning in humans: effects of dopaminergic modulation. *Biological Psychiatry, 62*, 765–772.
Metereau, E., & Dreher, J.-C. (2013). Cerebral correlates of salient prediction error for different rewards and punishments. *Cerebral Cortex, 23*, 477–487.
Monosov, I. E., & Hikosaka, O. (2012). Regionally distinct processing of rewards and punishments by the primate ventromedial prefrontal cortex. *The Journal of Neuroscience, 32*, 10318–10330.
Morrison, S. E., & Salzman, C. D. (2009). The convergence of information about rewarding and aversive stimuli in single neurons. *The Journal of Neuroscience, 29*, 11471–11483.
Nichols, T., Brett, M., Andersson, J., Wager, T., & Poline, J.-B. (2005). Valid conjunction inference with the minimum statistic. *NeuroImage, 25*, 653–660.
Nieuwenhuis, S., Slagter, H. A., von Geusau, N. J. A., Heslenfeld, D. J., & Holroyd, C. B. (2005). Knowing good from bad: differential activation of human cortical areas by positive and negative outcomes. *European Journal of Neuroscience, 21*, 3161–3168.
Noonan, M. P., Mars, R. B., & Rushworth, M. F. S. (2011). Distinct roles of three frontal cortical areas in reward-guided behavior. *The Journal of Neuroscience, 31*, 14399–14412.
O'Doherty, J. P., Buchanan, T. W., Seymour, B., & Dolan, R. J. (2006). Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum. *Neuron, 49*, 157–166.
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. *The Journal of Neuroscience, 23*, 7931–7939.
O'Doherty, J. P., Deichmann, R., Critchley, H. D., & Dolan, R. J. (2002). Neural responses during anticipation of a primary taste reward. *Neuron, 33*, 815–826.
O'Doherty, J. P., 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., 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.
O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001a). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience, 4*, 95–102.
O'Doherty, J., Rolls, E. T., Francis, S., Bowtell, R., & McGlone, F. (2001b). Representation of pleasant and aversive taste in the human brain. *Journal of Neurophysiology, 85*, 1315–1321.
Ogawa, M., van der Meer, M. A. A., Esber, G. R., Cerri, D. H., Stalnaker, T. A., & Schoenbaum, G. (2013). Risk-responsive orbitofrontal neurons track acquired salience. *Neuron, 77*, 251–258.
Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia, 9*, 97–113.
Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature, 441*, 223–226.
Padoa-Schioppa, C., & Assad, J. A. (2008). The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nature Neuroscience, 11*, 95–102.
Peelen, M. V., & Downing, P. E. (2007). Using multi-voxel pattern analysis of fMRI data to interpret overlapping functional activations. *Trends in Cognitive Sciences, 11*, 4–5.
Plassmann, H., O'Doherty, J., & Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *The Journal of Neuroscience, 27*, 9984–9988.
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. *The Journal of Neuroscience, 30*, 10799–10808.
Prevost, C., Pessiglione, M., Metereau, E., Clery-Melin, M.-L., & Dreher, J.-C. (2010). Separate valuation subsystems for delay and effort decision costs. *The Journal of Neuroscience, 30*, 14080–14090.
Rushworth, M. F. S., & Behrens, T. E. J. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature Neuroscience, 11*, 389–397.
Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience, 9*, 545–556.

Rangel, A., & Clithero, J. A. (2012). Value normalization in decision making: theory and evidence. *Current Opinion in Neurobiology, 22*, 970–981.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black, & W. F. Prokasy (Eds.), *Classical conditioning ii: Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.

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

Sescousse, G., Barbalat, G., Domenech, P., & Dreher, J.-C. (2013a). Imbalance in the sensitivity to different types of rewards in pathological gambling. *Brain, 136*, 2527–2538.

Sescousse, G., Caldú, X., Segura, B., & Dreher, J.-C. (2013b). Processing of primary and secondary rewards: a quantitative meta-analysis and review of human functional neuroimaging studies. *Neuroscience & Biobehavioral Reviews, 37*, 681–696.

Sarinopoulos, I., Grupe, D. W., Mackiewicz, K. L., Herrington, J. D., Lor, M., Steege, E. E., et al. (2010). Uncertainty during anticipation modulates neural responses to aversion in human insula and amygdala. *Cerebral Cortex, 20*, 929–940.

Sescousse, G., Li, Y., & Dreher, J.-C. (2014). A common currency for the computation of motivational values in the human striatum. *Social Cognitive and Affective Neuroscience* (Epub ahead of print).

Seymour, B., O'Doherty, J. P., Dayan, P., Koltzenburg, M., Jones, A. K., Dolan, R. J., et al. (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., et al. (2005). Opponent appetitive-aversive neural processes underlie predictive learning of pain relief. *Nature Neuroscience, 8*, 1234–1240.

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

Schoenbaum, G., Roesch, M. R., Stalnaker, T. A., & Takahashi, Y. K. (2009). A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. *Nature Reviews Neuroscience, 10*, 885–892.

Schoenbaum, G., Saddoris, M. P., & Stalnaker, T. A. (2007). Reconciling the roles of orbitofrontal cortex in reversal learning and the encoding of outcome expectancies. *Annual Meeting of the New York Academy of Sciences, 1121*, 320–335.

Schoenbaum, G., Takahashi, Y., Liu, T.-L., & McDannald, M. A. (2011). Does the orbitofrontal cortex signal value? *Annual Meeting of the New York Academy of Sciences, 1239*, 87–99.

Small, D. M., Zald, D. H., Jones-Gotman, M., Zatorre, R. J., Pardo, J. V., Frey, S., et al. (1999). Human cortical gustatory areas: a review of functional neuroimaging data. *NeuroReport, 10*, 7–14.

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. *Nature Neuroscience, 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.

Tobler, P. N., O'doherty, J. P., Dolan, R. J., & Schultz, W. (2006). Human neural learning depends on reward prediction errors in the blocking paradigm. *Journal of Neurophysiology, 95*, 301–310.

Takahashi, Y. K., Roesch, M. R., Wilson, R. C., Toreson, K., O'Donnell, P., Niv, Y., et al. (2011). Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. *The Journal of Neuroscience, 14*, 1590–1597.

Tremblay, L., & Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex. *Nature, 398*, 704–708.

Ursu, S., & Carter, C. S. (2005). Outcome representations, counterfactual comparisons and the human orbitofrontal cortex: implications for neuroimaging studies of decision-making. *Brain Research. Cognitive Brain Research, 23*, 51–60.

Wunderlich, K., Rangel, A., & O'Doherty, J. P. (2010). Economic choices can be made using only stimulus values. *Proceedings of the National Academy of Sciences of the United States of America, 107*, 15005–15010.

Zald, D. H., Lee, J. T., Fluegel, K. W., & Pardo, J. V. (1998). Aversive gustatory stimulation activates limbic circuits in humans. *Brain, 121*(Pt 6), 1143–1154.