

# Social and monetary reward learning engage overlapping neural substrates

Alice Lin,<sup>1</sup> Ralph Adolphs,<sup>1,2</sup> and Antonio Rangel<sup>1,2</sup>

<sup>1</sup>California Institute of Technology, Computations and Neural Systems, MC 136-93 Pasadena and <sup>2</sup>California Institute of Technology, Division of Humanities and Social Sciences, MC 228-77 Pasadena, CA 91125, USA

**Learning to make choices that yield rewarding outcomes requires the computation of three distinct signals: stimulus values that are used to guide choices at the time of decision making, experienced utility signals that are used to evaluate the outcomes of those decisions and prediction errors that are used to update the values assigned to stimuli during reward learning. Here we investigated whether monetary and social rewards involve overlapping neural substrates during these computations. Subjects engaged in two probabilistic reward learning tasks that were identical except that rewards were either social (pictures of smiling or angry people) or monetary (gaining or losing money). We found substantial overlap between the two types of rewards for all components of the learning process: a common area of ventromedial prefrontal cortex (vmPFC) correlated with stimulus value at the time of choice and another common area of vmPFC correlated with reward magnitude and common areas in the striatum correlated with prediction errors. Taken together, the findings support the hypothesis that shared anatomical substrates are involved in the computation of both monetary and social rewards.**

**Keywords:** social reward; monetary reward; ventromedial prefrontal cortex; ventral striatum

## INTRODUCTION

The brain needs to compute several distinct signals in order for an organism to learn how to make sound decisions among alternatives. First, at the time of choice, values need to be assigned to the different stimuli associated with each choice option [which we refer to as stimulus values (SV)]; these are subsequently compared in order to choose the option with the highest value (Wallis, 2007; Rangel *et al.*, 2008; Kable and Glimcher, 2009; Rushworth *et al.*, 2009; Rangel and Hare, 2010). Stimulus value signals have been found in ventral and medial sectors of the prefrontal cortex (vmPFC) in several human fMRI (Kable and Glimcher, 2007; Plassmann *et al.*, 2007; Tom *et al.*, 2007; Hare *et al.*, 2008, 2009; Chib *et al.*, 2009; FitzGerald *et al.*, 2009; Litt *et al.*, 2009; Levy *et al.*, 2010; Plassmann *et al.*, 2010) and non-human primate electrophysiological studies (Wallis and Miller, 2003; Padoa-Schioppa and Assad, 2006, 2008; Kennerley *et al.*, 2009; Kennerley and Wallis, 2009; Padoa-Schioppa, 2009) during choices involving non-social rewards, as well as during social decisions such as donations to charities (Hare *et al.*, 2010).

Having made a choice, the brain needs to compute the reward value associated with the outcomes generated by the choice. These signals are often called reward magnitude or experienced utility (R). Several human fMRI studies have found that activity in medial regions of orbitofrontal

cortex (OFC) correlates with behavioral measures of experienced utility for a wide variety of social and non-social reward modalities (Blood and Zatorre, 2001; Small *et al.*, 2001, 2003; de Araujo *et al.*, 2003; McClure *et al.*, 2003; Kringselbach, 2005; Plassmann *et al.*, 2008; Smith *et al.*, 2010).

A third critical component is the combination of the previous two signals into a prediction-error signal (PE) that is used to update stimulus values (Schultz *et al.*, 1997). The key involvement of the ventral striatum in this third component is borne out by a sizable and rapidly growing body of human fMRI studies of reinforcement learning that have used almost exclusively non-social rewards such as monetary payments (Delgado *et al.*, 2000; Berns *et al.*, 2001; Pagnoni *et al.*, 2002; O'Doherty *et al.*, 2003b, 2004; Pessiglione *et al.*, 2006; Yacubian *et al.*, 2006; Seymour *et al.*, 2007; Hare *et al.*, 2008).

Although the findings summarized above have been replicated across species, techniques and experimental designs, the vast majority of studies have used only non-social rewards such as juice, food or money, and only a handful have directly compared social and non-social rewards. This raises a fundamental question: do the same brain regions implement reward-learning computations for social and non-social rewards? Or might the areas that encode SV, PE and R be different for social rewards, analogously to the specialized perceptual processing of social stimuli (Kanwisher and Yovel, 2006)? While a very few other studies have recently approached this issue (Izuma *et al.*, 2008; Zink *et al.*, 2008; Smith *et al.*, 2010), no study to date has investigated the question using identical tasks across the same subjects, and in a task that allows us to compare the

Received 26 October 2010; Accepted 17 January 2011

Advance Access publication 22 March 2011

This work is supported in part by grants from the Betty and Gordon Moore Foundation, an NSF IGERT (to A.L.) training grant, and a grant from NIMH (to R.A.).

Correspondence should be addressed to Antonio Rangel, Caltech HSS 228-77 Pasadena, CA 91125-7700, USA. E-mail: rangel@hss.caltech.edu

encoding of the three types of basic reward signals defined above. We undertook such an investigation here using model-based fMRI.

## METHODS

### Participants

Twenty-seven female participants from the Caltech community participated in the study (mean age = 22.4 years; range 18–28). Five were excluded from further analyses: four due to excessive head movement, one due to failure to understand task instructions. All participants were fully right-handed, had normal or corrected-to-normal vision, had no history of psychiatric or neurological disease and were not taking medications that might have interfered with BOLD-fMRI. All gave informed consent under a protocol approved by the Caltech IRB.

### TASK

Participants played two structurally identical versions of an instrumental learning task, one with monetary rewards, the second with social rewards (Figure 1A). A trial began with the display of two visually distinctive slot machines, each associated with one of three outcome distributions: mean-positive, -negative and -neutral (Figure 1B).

All participants completed one social and one monetary block of 148 trials each; block order was randomized between participants. There were two types of trials in each block. In 100 choice trials the neutral slot machine was shown paired with either the positive or negative slot machine (50/50 probability with randomized order), and participants chose one by pressing a left or right button. We refer to these as free choice trials. In 48 non-choice trials two identical copies of one of the three slot machines were shown (1/3, 1/3, 1/3 probability with randomized order), and participants merely pressed either the left or right button in order to advance the trial. We refer to these as forced choice trials. Up to 2.5 s were allowed for choice in both cases, followed by a uniformly blank screen displayed for 1–5 s (flat distribution), followed by the reward outcome displayed for 1.5 s, followed by an intertrial interval of a uniformly blank screen displayed for 1–6 s (flat distribution). Note that participants were not told the reward probabilities associated with each slot machine and had to learn them by trial and error during the task.

The forced trials provide an essential control for a potential important confound in the study. One potential concern is that the presentation of positive and aversive social outcomes might induce in the brain ‘correct’ and ‘error’ feedback signals at outcome during the social trials. This is a problem because this would suggest that the common locus of activity is not due to the activation of a social reward, but to the activation of these error feedback signals. The forced trials provide a control for this concern because when there is no free choice, there can be no error feedback regarding the correctness of the choice.

### Stimuli and rewards

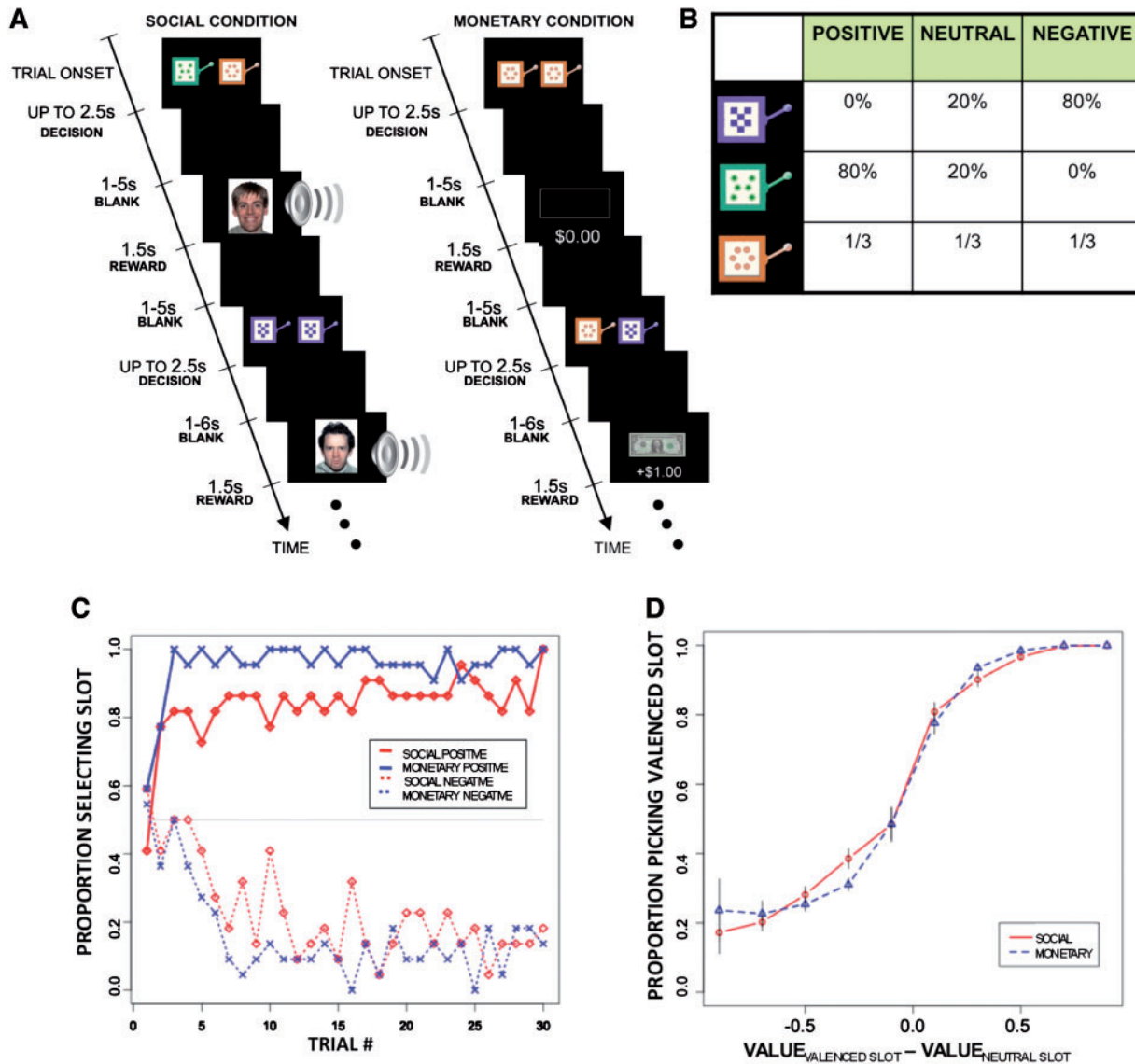
The slot machines in both conditions were represented by cartoon images of actual slot machines that varied in color and pattern (Figure 1). In the social condition, reward outcomes were color photographs of unfamiliar faces from the NimStim collection (Tottenham *et al.*, 2009) showing either an angry (negative outcome), neutral (neutral outcome) or happy (positive outcome) emotional expression, presented together with emotionally matched words played through headphones (normalized for volume and duration). Examples of positive words are excellent, bravo and fantastic. Examples of negative words are stupid, moron and wrong. Examples of neutral words are desk, paper and stapler. Extensive prior piloting had demonstrated the behavioral efficacy of these stimuli in reward learning.

In the monetary condition, the positive outcome was a gain of one dollar (an image of a dollar bill), the negative condition was a loss of one dollar (image of a dollar bill crossed out) and the neutral condition involved no change in monetary payoff (image of an empty rectangle). Subjects were paid out the sum of their earnings at the end of the experiment.

### Computational model

We computed trial- and subject-specific values for each of the three variables described in the Introduction. The SV for every slot machine was calculated as the 10-trial moving average proportion of times that the machine was chosen when it was shown, a continuous value between 0–1. Consistent with this coding, R were assigned a value of 1 if they were positive; a value of 0.5 if they were neutral and a value of 0 if they were negative. PE at the time of outcome were calculated using a simple Rescorla–Wagner learning rule (Rescorla and Wagner, 1972) as the difference between the value of the reward outcome and the stimulus value of the machine selected for that trial:  $PE_t = R_t - SV_t$ .

Note three things about the value normalizations. First, our approach deviates from the usual practice in neuroscience studies of reinforcement learning (Pessiglione *et al.*, 2006, 2008; Seymour *et al.*, 2007; Lohrenz *et al.*, 2007; Hare *et al.*, 2008; Wunderlich *et al.*, 2009) in which it is customary to fit the values of the SV signal based on the predictions of the best fitting learning model. Here we depart from that practice because the revealed preference approach provides more accurate measures of the values computed at the time of choice (as shown in Figure 1D). Second, without loss of generality we normalize the reward outcome signals to 0 for negative outcomes and 1 for positive outcomes. Note that given the parametric nature of the general linear model specified below, this normalization does not affect the identification of areas that exhibit significant correlation with this variable. Third, we use the standard definition of prediction errors used in the literature.



**Fig. 1** Task and behavioral results. (A) Timeline of the monetary and social reward trials. Choice trials paired a neutral slot machine with a valenced slot machine. Trials were identical except for the nature of the outcomes: monetary trials had a gain/loss of +\$1, 0\$ or -\$1, whereas social trials revealed happy, neutral or angry faces accompanied with sound effects of similar emotional valence. The experiment also included no-choice trials (in which a pair of identical slot machines were shown: neutral, negative or positive) to help separate the learning and stimulus value signals. Specific slot machines were randomly assigned to specific reward outcomes at the start of the experiment for each subject, and distinct between monetary and social condition blocks. (B) Distribution of outcomes for each slot machine. First row: negative machine. Second row: positive machine. Bottom row: neutral machine. The same distribution was used in the monetary and social conditions. Actual appearance of the slot machines was randomly paired with a reward outcome distribution and distinct between monetary and social condition blocks. (C) Plot of group subject choices across trials (only the first 30 are shown). (D) Psychometric choice curve for monetary and social conditions. Bars denote standard error measures computed across subjects.

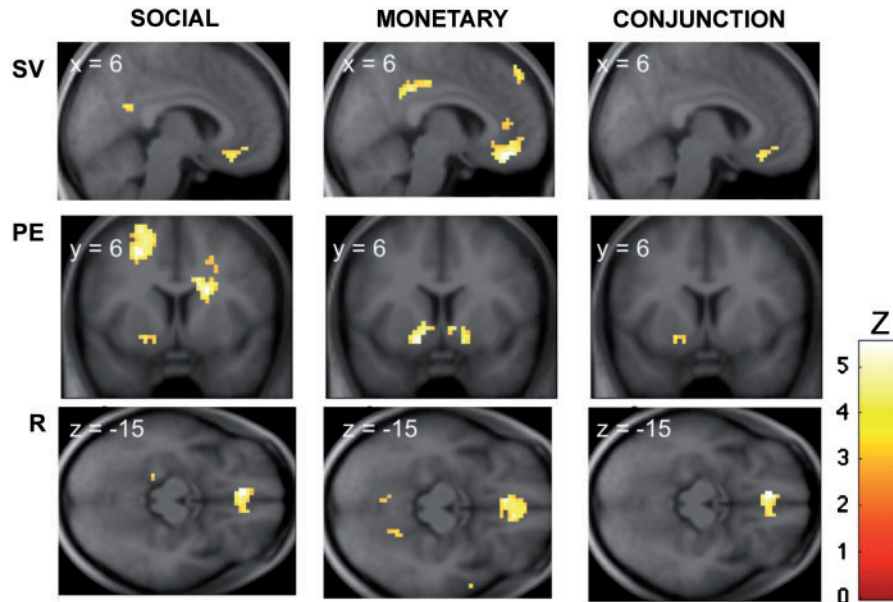
### Image acquisition

T2\*-weighted gradient-echo echo-planar (EPI) images with BOLD contrast were collected on a Siemens 3T Trio. To optimize signal in the OFC, we acquired slices in an oblique orientation of 30° to the anterior commissure–posterior commissure line (Deichmann *et al.*, 2003) and used an eight-channel phased array head coil. Each volume comprised 32 slices. Data was collected in four sessions (~12 min each). The imaging parameters were as follows: TR = 2 s, TE = 30 ms, FOV = 192 mm, 32 slices with 3 mm

thickness resulting in isotropic 3 mm voxels. Whole-brain high-resolution T1-weighted structural scans (1 × 1 × 1 mm) were co-registered with their mean T2\*-weighted images and averaged together to permit anatomical localization of the functional activations at the group level.

### fMRI pre-processing

The imaging data was analyzed using SPM5 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK). Functional images were corrected



**Fig. 2** Basic Neuroimaging results. (Top) Activation in the vmPFC correlated with SV at the time of free choice in both monetary and social conditions. (Middle) Activation in the vStr correlated with PE at the time of outcome in both monetary and social free choice conditions (albeit the conjunction did not survive our omnibus threshold). (Bottom) Activation in the vmPFC correlated with R in both monetary and social free choice conditions. For illustration purposes only, all images are thresholded at  $P < 0.005$  uncorrected with an extent threshold of 15 voxels, except for the conjunction of PE which is  $P < 0.005$  with an extent threshold of five voxels (see Tables 1–3 for details).

for slice acquisition time within each volume, motion-corrected with realignment to the last volume, spatially normalized to the standard Montreal Neurological Institute EPI template and spatially smoothed using a Gaussian kernel with a full-width at half-maximum of 8 mm. Intensity normalization and high-pass temporal filtering (filter width = 128 s) were also applied to the data.

### fMRI data analysis

The data analysis proceeded in three steps. First, we estimated a general linear model with AR(1). This model was designed to identify regions in which BOLD activity was parametrically related to SV, R and PE. The model included the following regressors:

- (R1) An indicator function for the decision screen in free choice monetary trials.
- (R2) An indicator function for the decision screen in free choice monetary trials multiplied by the SV of the two slot machines shown in that trial (summed SV).
- (R3) An indicator function for the decision screen in free choice monetary trials multiplied by the reaction time for that trial.
- (R4–R6) Analogous indicator functions for decision screen events in free choice social trials.
- (R7) An indicator function for the decision screen in forced monetary trials.
- (R8) An indicator function for the decision screen in forced monetary trials multiplied by the SV of the slot machine displayed.
- (R9–R10) Analogous indicator functions for decision screen events in forced social trials.

(R11) A delta function for the time of response in the monetary condition.

(R12) A delta function for the time of response in the social condition.

(R13) An indicator function for the outcome screen in free monetary trials (both choice and non-choice).

(R14) An indicator function for the outcome screen in free monetary trials multiplied by the PE for the trial.

(R15) An indicator function for the outcome screen in free monetary trials multiplied by the R for the trial.

(R16–R18) Analogous indicator functions for outcome screen events in free social trials (both choice and non-choice).

We orthogonalized the modulators for the main regressors that had more than one modulator (e.g. R2 and R3). The model also included six head motion regressors, session constants and missed trials as regressors of no interest. The regressors of interest and missed trial regressor were convolved with a canonical HRF.

Second, we calculated the following first-level single-subject contrasts: (i) R2 vs baseline, (ii) R5 vs baseline, (iii) R14 vs baseline, (iv) R15 vs baseline, (v) R17 vs baseline and (vi) R18 vs baseline.

Third, we calculated second-level group contrasts using a one-sample  $t$ -test of the first level contrast statistics.

Finally, we also performed a conjunction analysis between the equivalent contrasts for the monetary and social conditions to identify areas involved in similar computations in both cases. The results are shown in Figure 2 and reported in Tables 1–3. For inference purposes we used an omnibus threshold of  $P < 0.001$  uncorrected with an extent threshold

**Table 1** Regions correlating with stimulus value at cue

Region	No. of voxels	Z-score	x	y	z
Areas correlating with SV in monetary choice trials (R2 vs baseline)					
Medial orbitofrontal cortex	214	4.53 <sup>†</sup>	0	27	-21
Frontal superior	52	4.19	-18	42	51
Mid cingulum	46	4.01	0	-30	45
Angular gyrus	61	3.91	-57	-66	30
Middle temporal gyrus	24	3.85	60	-15	-6
Areas correlating with SVs in social choice trials (R5 vs baseline)					
Medial orbitofrontal cortex	40	3.16 <sup>†</sup>	6	27	-15
Areas correlating with SVs in both monetary and social choice trials					
Medial orbitofrontal cortex	37	3.16 <sup>†</sup>	6	27	-15

Regions are significant at  $P < 0.001$  uncorrected and 15 voxels extent threshold.

<sup>†</sup>Survives  $P < 0.05$  small volume correction. Coordinates reported in MNI space.

**Table 2** Regions correlating with prediction error at outcome

Region	No. of voxels	Z-score	x	y	z
Areas correlating with PE in monetary choice trials (R13 vs baseline)					
Putamen	25	4.07 <sup>†</sup>	-15	6	-12
Caudate	22	3.75	9	9	-3
Precuneus	15	3.49	-18	-51	33
Areas correlating with PE in social choice trials (R16 vs baseline)					
-	-	-	-	-	-
Areas correlating with PE in both monetary and social choice trials					
-	-	-	-	-	-

Regions are significant at  $P < 0.001$  uncorrected and 15 voxels extent threshold.

<sup>†</sup>Survives  $P < 0.05$  small volume correction. Coordinates reported in MNI space.

of 15 voxels. However, given the strong priors from the previous literature about the role of the vmPFC in encoding stimulus value and reward outcome signals, as well as the role of the ventral striatum in encoding prediction errors, we also report activity in these two areas if they survive small volume corrections (SVC) at  $P < 0.05$ . The mask for the SVC in vmPFC at choice was taken using a sphere of 10-mm radius defined around the peak activation coordinates that correlated with stimulus values in Rolls *et al.* (Rolls *et al.*, 2008). The mask for the vmPFC SVC at reward outcome was given by a sphere of 10-mm radius defined around the peak coordinates that correlated with the magnitude of reward outcome in O'Doherty *et al.* (O'Doherty *et al.*, 2002). The mask for the SVC in ventral striatum was taken using a sphere of 10-mm radius defined around the peak activation coordinates that correlated with prediction errors in Pessiglione *et al.* (Pessiglione *et al.*, 2006). For display purposes only activity in selected SPMS is reported at  $P < 0.005$  uncorrected with an extent threshold of five voxels. Anatomical localizations were performed by overlaying the *t*-maps on a normalized structural image averaged across subjects, and with reference to an anatomical atlas (Duvernoy, 1999).

**Table 3** Regions correlating with reward at outcome

Region	No. of Voxels	Z score	x	y	z
Areas correlating with R in monetary choice trials (R14 vs baseline)					
Occipital	124	4.74	21	-75	15
Insula	125	4.68	-33	3	12
Inferior parietal	116	4.43	-51	-36	27
Occipital	59	4.29	-6	87	18
Insula	33	4.23	39	-18	18
Cingulum	52	3.99	-6	9	36
Medial frontal gyrus	86	3.96	-15	-6	57
Inferior parietal	78	3.95	51	-33	30
Medial orbitofrontal cortex	136	3.88 <sup>†</sup>	6	33	-12
Superior frontal gyrus	26	3.84	-18	27	57
Superior frontal gyrus	20	3.66	-30	36	33
Rolandic operculum	18	3.66	57	0	12
Heschl gyrus	21	3.63	-39	-24	3
Inferior parietal	21	3.61	-36	-27	24
Calcarine	15	3.42	-18	-72	9
Areas correlating with R in social choice trials (R17 vs baseline)					
Medial orbitofrontal cortex	29	4.16 <sup>†</sup>	-6	36	-15
Areas correlating with R in both monetary and social choice trials					
Medial orbitofrontal cortex	129	4.16 <sup>†</sup>	-6	36	-15

Regions are significant at  $P < 0.001$  uncorrected and 15 voxels extent threshold.

<sup>†</sup>Survives  $P < 0.05$  small volume correction. Coordinates reported in MNI space.

## RESULTS

### Behavioral results

Participants reliably learned to select the slot machine associated with the highest probability of a positive valenced outcome within a few choice trials for both social and non-social rewards (Figure 1C). The figure also reveals two additional interesting patterns about the learning process. First, participants were somewhat slower at learning to discriminate between social rewards than between monetary rewards. For example, by the 10th exposure, the positive monetary machine was chosen with 92% whereas the social positive machine was chosen with 72% frequency ( $P < 0.001$ ). Second, participants were slower in learning to avoid the negative slot machines than in choosing the positive ones. For example, by the tenth presentation the positive slot machines were chosen 85% of the time, whereas the negative ones were avoided only 68% of the time ( $P < 0.001$ ). Both differences were not significant on the last third of the learning trials, which suggests that they are related to the speed of learning, and not to the ability to ultimately learn the value of the stimuli.

Figure 1D shows the psychometric choice curves for the social and monetary conditions based on their SV. Note several things about the curves. First, when the values of valenced and neutral slot machines were identical, participants exhibited no choice bias (0.5 on the *y*-axis corresponds to 0.0 on the *x*-axis). Second, the choice curves are not significantly different from each other (greatest difference at  $x = 0.25$  had  $P = 0.32$  with Bonferroni correction). Third, the choice curve is asymmetric: whereas participants chose

the valenced slot machine over the neutral slot machine with probability close to one when its relative stimulus value was sufficiently positive (far right side of curve), subjects chose the neutral slot machine only 80% of the time even when it was the most favorable (far left side of curve).

### Neural correlates of stimulus values

We estimated a parametric general linear model of the BOLD signal to identify areas in which activation correlated with SV at the time of choice, and with PE and R at outcome, during free choice trials (see 'Methods' section for details). In the free choice monetary task, activation in the vmPFC correlated with SV of the slot machines. SV signals were additionally found in the mid-cingulum, the superior frontal gyrus and the angular gyrus (Table 1 and Figure 2). In the free choice social task, activation correlating with SV was also found in a similar region of vmPFC. A conjunction analysis showed that activation in a common area of vmPFC correlated with SV in both social and monetary conditions.

### Neural correlates of prediction errors

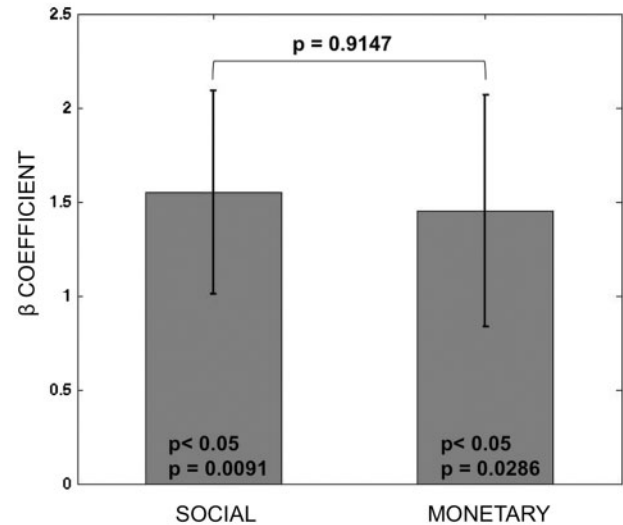
In the free choice monetary task, PE correlated with activation in the caudate and putamen (Table 2 and Figure 2). In the free choice social task, PE did not exhibit any correlations at our omnibus threshold ( $P < 0.001$  uncorrected, 15 voxels). However, for completeness we show areas of the striatum that correlate with PE in the social free choice condition at  $P < 0.005$  uncorrected, as well as the resulting conjunction results using this lower threshold.

### Neural correlates reward magnitude

In the free choice monetary task, reward outcome correlated with activation in vmPFC, insula, occipital cortex, cingulate gyrus and superior frontal gyrus (Table 3 and Figure 2). In the free choice social task, reward outcome correlated with activation in vmPFC. A conjunction analysis revealed that activation in a common area of the vmPFC correlated with reward magnitude in the social and non-social conditions.

### Ruling out a potential confound

A non-trivial potential confound is that the happy and angry faces might activate 'correct' and 'error' feedback signals in the brain regarding the adequacy of choice, and that the areas of co-activation might be due to the presence of these error signals, and not the computation of social rewards. In fact, these types of stimuli have previously been used just for that purpose (Cools *et al.*, 2007). Fortunately, the forced choice trials provide a control that allows us to test if the previous results are driven by this potential confound. Figure 3 describes the strength of the correlation between outcome reward signals and BOLD activity in the area of vmPFC identified by the conjunction of outcome rewards in both conditions. It shows that the strength of the correlation in the social and monetary trials is of similar



**Fig. 3** ROI analysis of outcome reward signals in vmPFC during forced choice trials. Average beta plots for activity during reward outcome in forced choice trials. The functional mask of vmPFC is given by the area that exhibits correlation with reward outcomes in social and monetary free choice trials at  $P < 0.05$  SVC. The  $P$ -values inside the bars are for  $t$ -tests vs zero.

magnitude and not statistically different ( $P = 0.91$ , two-sided paired  $t$ -test) even in the absence of error feedback. This implies that the signal in the vmPFC during social outcomes cannot be attributed to error feedback, and that the concern about the potential confound in this task was unfounded.

## DISCUSSION

A fundamental open question in behavioral and social neuroscience is whether common regions of the brain encode the value signals that are necessary to make sound decisions for both social and non-social rewards. Prior evidence suggested that there might be such an overlap. In the case of stimulus values, a recent paper found that the values of charities at the time of decision making were encoded in areas of the vmPFC that overlap with those that have been found for private rewards (Hare *et al.*, 2010). In the case of experienced utility for social rewards, several studies found that activity in the OFC correlates with the perceived attractiveness of faces (Aharon *et al.*, 2001; O'Doherty *et al.*, 2003a; Cloutier *et al.*, 2008; Smith *et al.*, 2010). Finally, in the case of prediction errors, studies have found that activity in the ventral striatum correlates with prediction error-like signals in a task involving the receipt of anticipated social rewards (Spreckelmeyer *et al.*, 2009) and in tasks involving social reputation and status (Izuma *et al.*, 2008; Zink *et al.*, 2008). These latter two studies in particular, compared both social and monetary rewards, as we did in the present study, and provided strong initial evidence for the idea that neural representations for these two types of rewards are at least partly overlapping. What has been missing to date is a study that compared social and non-social rewards

across tasks whose basic structure and reward probabilities are matched for the two types of rewards, and in which the three basic computations associated with reward learning (SV, PE and R) are at work.

We addressed this open question by asking subjects to perform an otherwise identical simple probabilistic learning decision-making task in which stimuli were associated with either monetary or social rewards. We found evidence for common signals in all cases: a common area of vmPFC correlated with SV, a common area of vmPFC correlated with R, and common areas of ventral striatum correlated with PE, albeit in the later case only at a relatively low threshold of  $P < 0.005$  unc. Together with other recent findings (Izuma *et al.*, 2008; Zink *et al.*, 2008; Chib *et al.*, 2009; Hare *et al.*, 2010), our results provide increasing support that overlapping areas of vmPFC and ventral striatum encode value signals for both types of rewards (Montague and Berns, 2002; Rangel, 2008).

Behaviorally, our subjects were slower to learn the value of social and negative stimuli. Since the type of reinforcement learning models that have been successfully used to account for the behavioral data do not predict such asymmetries (Rescola and Wagner, 1972; Sutton and Barto, 1998; Montague and Berns, 2002; Niv and Montague, 2008), this raises an apparent puzzle. However, there are two potential explanations for this aspect of the findings. First, the reward magnitude of both types of stimuli might not have been perfectly matched in our population (so that, for example, subjects found the \$1 outcome more rewarding than the positive social stimuli). Second, individuals stop selecting the negative slot machine after a while, which means that learning stops and subjects might not get sufficient negative reinforcement to learn the full extent of the negative outcomes associated with these machines.

We emphasize that the existence of areas involved in the encoding of reward in social and non-social situations does not mean that the full network involved in processing both types of rewards is identical. For example, it is known that areas involved in theory of mind computations are more likely to become active during social decisions than during choices among non-social rewards (Saxe and Kanwisher, 2003; Saxe, 2006; Krach *et al.*, 2010).

It is important to highlight two limitations of our results. First, given the limited spatial resolution of fMRI we cannot rule out the possibility that there might be neuronal subpopulations within the vmPFC and ventral striatum specialized in valuing certain types of rewards. Future studies using fMRI adaptation designs, or direct electrophysiological recordings within these regions, will have to address this issue before the existence of a common valuation currency can be definitely established.

Second, previous experiments suggest that males and females process some types of social rewards differently (Spreckelmeyer *et al.*, 2009), which opens the possibility that there might be a gender difference in the extent to

which common circuitry is used in the social and non-social domains to carry out basic reward computations. Unfortunately, we cannot resolve this issue with this data set since only females participated in the experiment.

### Conflict of Interest

None declared.

### REFERENCES

- Aharon, I., Etcoff, N., Ariely, D., Chabris, C.F., O'Connor, E., Breiter, H.C. (2001). Beautiful faces have variable reward value: fMRI and behavioral evidence. *Neuron*, 32, 537–51.
- Berns, G.S., McClure, S.M., Pagnoni, G., Montague, P.R. (2001). Predictability modulates human brain response to reward. *Journal of Neuroscience*, 21, 2793–8.
- Blood, A.J., Zatorre, R.J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences USA*, 98, 11818–23.
- 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. *Journal of Neuroscience*, 29, 12315–20.
- Cloutier, J., Heatherton, T.F., Whalen, P.J., Kelley, W.M. (2008). Are attractive people rewarding? Sex differences in the neural substrates of facial attractiveness. *Journal of Cognitive Neuroscience*, 20, 941–51.
- Cools, R., Lewis, S.J., Clark, L., Barker, R.A., Robbins, T.W. (2007). L-DOPA disrupts activity in the nucleus accumbens during reversal learning in Parkinson's disease. *Neuropsychopharmacology*, 32, 180–9.
- de Araujo, I.E., Rolls, E.T., Kringelbach, M.L., McGlone, F., Phillips, N. (2003). Taste-olfactory convergence, and the representation of the pleasantness of flavour, in the human brain. *European Journal of Neuroscience*, 18, 2059–68.
- Deichmann, R., Gottfried, J.A., Hutton, C., Turner, R. (2003). Optimized EPI for fMRI studies of the orbitofrontal cortex. *Neuroimage*, 19, 430–41.
- Delgado, M.R., Nystrom, L.E., Fissell, C., Noll, D.C., Fiez, J.A. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *Journal of Neurophysiology*, 84, 3072–7.
- Duvernoy, H.M. (1999). *The Human Brain: Surface, Three-Dimensional Sectional Anatomy with MRI, and Blood Supply*. Berlin: Springer.
- FitzGerald, T.H., Seymour, B., Dolan, R.J. (2009). The role of human orbitofrontal cortex in value comparison for incommensurable objects. *Journal of Neuroscience*, 29, 8388–95.
- Hare, T., Camerer, C., Rangel, A. (2009). Self-control in decision-making involves modulation of the vmPFC valuation system. *Science*, 324, 646–8.
- Hare, T.A., Camerer, C.F., Knoepfle, D.T., Rangel, A. (2010). Value computations in ventral medial prefrontal cortex during charitable decision making incorporate input from regions involved in social cognition. *Journal of Neuroscience*, 30, 583–90.
- 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. *Journal of Neuroscience*, 28, 5623–30.
- Izuma, K., Saito, D.N., Sadato, N. (2008). Processing of social and monetary rewards in the human striatum. *Neuron*, 58, 284–94.
- Kable, J.W., Glimcher, P.W. (2007). The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience*, 10, 1625–33.
- Kable, J.W., Glimcher, P.W. (2009). The neurobiology of decision: consensus and controversy. *Neuron*, 63, 733–45.
- Kanwisher, N., Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of The Royal Society London B: Biological Sciences*, 361, 2109–28.
- Kennerly, 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, 1162–78.

- Kennerley, S.W., Wallis, J.D. (2009). Evaluating choices by single neurons in the frontal lobe: outcome value encoded across multiple decision variables. *European Journal of Neuroscience*, 29, 2061–73.
- Krach, S., Paulus, F.M., Bodden, M., Kircher, T. (2010). The rewarding nature of social interactions. *Frontiers in Behavioural Neuroscience*, 4, 22.
- Kringelbach, M.L. (2005). The human orbitofrontal cortex: linking reward to hedonic experience. *Nature Reviews Neuroscience*, 6, 691–702.
- Levy, I., Snell, J., Nelson, A.J., Rustichini, A., Glimcher, P.W. (2010). The neural representation of subjective value under risk and ambiguity. *Journal of Neurophysiology*, 103, 1036–47.
- Litt, A., Plassmann, H., Shiv, B., Rangel, A. (2011). Dissociating valuation and saliency signals during decision-making. *Cerebral Cortex*, 21, 95–102.
- Lohrenz, T., McCabe, K., Camerer, C.F., Montague, P.R. (2007). Neural signature of fictive learning signals in a sequential investment task. *Proceedings of The National Academy of Sciences USA*, 104, 9493–8.
- McClure, S.M., Berns, G.S., Montague, P.R. (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron*, 38, 339–46.
- Montague, P.R., Berns, G.S. (2002). Neural economics and the biological substrates of valuation. *Neuron*, 36, 265–84.
- Niv, Y., Montague, P.R. (2008). Theoretical and empirical studies of learning. In: Glimcher, P.W., Fehr, E., Camerer, C., Poldrack, R.A., editors. *Neuroeconomics: Decision-Making and the Brain*. New York: Elsevier.
- 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, 329–37.
- O'Doherty, J.P., Deichmann, R., Critchley, H., Dolan, R.J. (2002). Neural responses during anticipation of a primary taste reward. *Neuron*, 33, 815–26.
- 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, 452–4.
- O'Doherty, J., Winston, J., Critchley, H., Perrett, D., Burt, D.M., Dolan, R.J. (2003a). Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia*, 41, 147–55.
- Padoa-Schioppa, C. (2009). Range-adapting representation of economic value in the orbitofrontal cortex. *Journal of Neuroscience*, 29, 14004–14.
- Padoa-Schioppa, C., Assad, J.A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature*, 441, 223–6.
- 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.
- Pagnoni, G., Zink, C.F., Montague, P.R., Berns, G.S. (2002). Activity in human ventral striatum locked to errors of reward prediction. *Nature Neuroscience*, 5, 97–8.
- Pessiglione, M., Petrovic, P., Daunizeau, J., Palminteri, S., Dolan, R.J., Frith, C.D. (2008). Subliminal instrumental conditioning demonstrated in the human brain. *Neuron*, 59, 561–7.
- Pessiglione, M., Seymour, B., Flandin, G., Dolan, R.J., Frith, C.D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, 442, 1042–5.
- Plassmann, H., O'Doherty, J., Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *Journal of Neuroscience*, 27, 9984–8.
- Plassmann, H., O'Doherty, J., Rangel, A. (2010). Appetitive and aversive goal values are encoded in the medial orbitofrontal cortex at the time of decision making. *Journal of Neuroscience*, 30, 10799–808.
- Plassmann, H., O'Doherty, J., Shiv, B., Rangel, A. (2008). Marketing actions can modulate neural representations of experienced pleasantness. *Proceedings of The National Academy of Sciences USA*, 105, 1050–4.
- Rangel, A. (2008). The computation and comparison of value in goal-directed choice. In: Glimcher, P.W., Camerer, C.F., Fehr, E., Poldrack, R.A., editors. *Neuroeconomics: Decision Making and the Brain*. New York: Elsevier.
- Rangel, A., Camerer, C., Montague, P.R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, 9, 545–56.
- Rangel, A., Hare, T. (2010). Neural computations associated with goal-directed choice. *Current Opinion in Neurobiology*, 20, 262–70.
- Rescola, R.A., Wagner, A.R. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and non-reinforcement. In: Black, A.H., Prokasy, W.F., editors. *Classical Conditioning II: Current Research and Theory*. New York, NY: Appleton Century Crofts, pp. 406–12.
- Rolls, E.T., McCabe, C., Redoute, J. (2008). Expected value, reward outcome, and temporal difference error representations in a probabilistic decision task. *Cerebral Cortex*, 18, 652–63.
- Rushworth, M.F., Mars, R.B., Summerfield, C. (2009). General mechanisms for making decisions? *Current Opinion Neurobiology*, 19, 75–83.
- Saxe, R. (2006). Uniquely human social cognition. *Current Opinion in Neurobiology*, 16, 235–9.
- Saxe, R., Kanswisher, N. (2003). People thinking about thinking people: the role of the temporo-parietal junction in “theory of mind”. *Neuroimage*, 19, 1835–42.
- Schultz, W., Dayan, P., Montague, P.R. (1997). A neural substrate of prediction and reward. *Science*, 275, 1593–9.
- Seymour, B., Daw, N., Dayan, P., Singer, T., Dolan, R. (2007). Differential encoding of losses and gains in the human striatum. *Journal of Neuroscience*, 27, 4826–31.
- Small, D.M., Gregory, M.D., Mak, Y.E., Gitelman, D., Mesulam, M.M., Parrish, T. (2003). Dissociation of neural representation of intensity and affective valuation in human gustation. *Neuron*, 39, 701–11.
- Small, D.M., Zatorre, R.J., Dagher, A., Evans, A.C., Jones-Gotman, M. (2001). Changes in brain activity related to eating chocolate: from pleasure to aversion. *Brain*, 124, 1720–33.
- Smith, D.V., Hayden, B.Y., Truong, T.K., Song, A.W., Platt, M.L., Huettel, S.A. (2010). Distinct value signals in anterior and posterior ventromedial prefrontal cortex. *Journal of Neuroscience*, 30, 2490–5.
- Spreckelmeyer, K.N., Krach, S., Kohls, G., et al. (2009). Anticipation of monetary and social reward differently activates mesolimbic brain structures in men and women. *Social Cognitive and Affective neuroscience*, 4, 158–65.
- Sutton, R.S., Barto, A.G. (1998). *Reinforcement Learning: An Introduction*. Cambridge: MIT Press.
- 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, 515–8.
- Tottenham, N., Tanaka, J.W., Leon, A.C., et al. (2009). The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Research*, 168, 242–9.
- Wallis, J.D. (2007). Orbitofrontal cortex and its contribution to decision-making. *Annual Review of Neuroscience*, 30, 31–56.
- Wallis, J.D., Miller, E.K. (2003). Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *European Journal of Neuroscience*, 18, 2069–81.
- Wunderlich, K., Rangel, A., O'Doherty, J.P. (2009). Neural computations underlying action-based decision making in the human brain. *Proceedings of The National Academy of Sciences USA*, 106, 17199–204.
- Yacubian, J., Glascher, J., Schroeder, K., Sommer, T., Braus, D.F., Büchel, C. (2006). Dissociable systems for gain- and loss-related value predictions and errors of prediction in the human brain. *Journal of Neuroscience*, 26, 9530–7.
- Zink, C.F., Tong, Y., Chen, Q., Bassett, D.S., Stein, J.L., Meyer-Lindenberg, A. (2008). Know your place: neural processing of social hierarchy in humans. *Neuron*, 58, 273–283.