

Effects of Emotional Preferences on Value-based Decision-making Are Mediated by Mentalizing and Reward Networks

Simon Evans¹, Stephen M. Fleming¹, Raymond J. Dolan¹,
and Bruno B. Averbeck^{1,2}

Abstract

■ Real-world decision-making often involves social considerations. Consequently, the social value of stimuli can induce preferences in choice behavior. However, it is unknown how financial and social values are integrated in the brain. Here, we investigated how smiling and angry face stimuli interacted with financial reward feedback in a stochastically rewarded decision-making task. Subjects reliably preferred the smiling faces despite equivalent reward feedback, demonstrating a socially driven bias. We fit a Bayesian reinforcement learning model to factor the effects of financial rewards and emotion

preferences in individual subjects, and regressed model predictions on the trial-by-trial fMRI signal. Activity in the subcallosal cingulate and the ventral striatum, both involved in reward learning, correlated with financial reward feedback, whereas the differential contribution of social value activated dorsal temporo-parietal junction and dorsal anterior cingulate cortex, previously proposed as components of a mentalizing network. We conclude that the impact of social stimuli on value-based decision processes is mediated by effects in brain regions partially separable from classical reward circuitry. ■

INTRODUCTION

Social cues are ubiquitous in day-to-day life and have substantial effects on decision-making processes. The brain mechanisms that underlie the effects of social cues on decision-making processes, however, are unclear. Much work has been done showing that financial rewards drive decision processes and that these effects are mediated by a network of anatomically interconnected areas (Haber, Kim, Maily, & Calzavara, 2006), including orbital and ventromedial prefrontal cortex and ventral striatum (Elliott, Agnew, & Deakin, 2010; Bischoff-Grethe, Hazeltine, Bergren, Ivry, & Grafton, 2009; Montague, King-Casas, & Cohen, 2006; O'Doherty et al., 2004; O'Doherty, Critchley, Deichmann, & Dolan, 2003). Other work has shown that social cues are processed by a network of areas including the temporo-parietal junction (TPJ), the temporal poles, and anterior cingulate cortex (ACC) (Van Overwalle & Baetens, 2009; Frith & Frith, 2006; Allison, Puce, & McCarthy, 2000). Social cues may also activate regions partially overlapping with classical reward circuitry, with ventromedial (Grossman et al., 2010) and orbital prefrontal cortex responding to emotional aspects of social processing (Hynes, Baird, & Grafton, 2006).

There is growing interest in the interrelationship of reward and social information processing. For example,

recent work has shown that explicit social cues, in the form of advice from a confederate, can be integrated with personal experience in learning and decision-making, an effect mediated by ACC (Behrens, Hunt, Woolrich, & Rushworth, 2008). Other work has shown that when explicit feedback is received from a confederate instead of a computer, the temporal poles are engaged (van den Bos, McClure, Harris, Fiske, & Cohen, 2007). A recent study has shown that learning is more effective when cues are emotional faces rather than cognitive stimuli, and that this effect is mediated by the amygdala (Hurlemann et al., 2010). In many cases, however, social factors are implicit, making it difficult to isolate their effects on decision-making.

We sought to examine the effect of implicit social cues on learning, using a modeling approach that allowed us to infer implicit emotional effects on decision-making directly from behavioral data. We used a task in which subjects had to determine which of two faces was being financially rewarded most often. Importantly, the faces had different emotional expressions, so the task was to associate an expression with a financial reward. Studies have shown that human subjects were more likely to cooperate with smiling partners (Scharlemann, Eckel, Kacelnik, & Wilson, 2001) and that they will forego small amounts of money to look at pictures of attractive people (Smith et al., 2010; Hayden, Parikh, Deaner, & Platt, 2007). Similarly, studies in monkeys have shown that

¹UCL Institute of Neurology, London, UK, ²National Institutes of Health, Bethesda, MD

animals will forego juice to view faces of high-ranking, but not low-ranking, monkeys (Deaner, Khera, & Platt, 2005). We have also shown previously that expressions can impact on learning, and that we can, on a subject-by-subject basis, model the relative impact of social and financial reward cues using a Bayesian model (Averbeck & Duchaine, 2009).

In principle, the behavioral effects of the facial expressions could be mediated directly by networks that process nonsocial rewards. Consistent with this hypothesis, happy expressions have been shown to activate ventromedial prefrontal cortex (O'Doherty, Winston, et al., 2003), an area which represents stimulus reward value. Moreover, damage to this region may lead to deficits in social decision-making (Grossman et al., 2010). Similarly, behavioral studies have shown that smiles can be appetitive (Murphy & Zajonc, 1993), potentially engaging Pavlovian value systems in the ventral striatum and ventromedial prefrontal cortex (Dayan, Niv, Seymour, & Daw, 2006; De Martino, Kumaran, Seymour, & Dolan, 2006; Daw, Niv, & Dayan, 2005; McClure, Laibson, Loewenstein, & Cohen, 2004). Alternatively, the effects of facial expressions on decision-making could be mediated by brain networks that underlie social processing including theory of mind (Van Overwalle & Baetens, 2009; Frith & Frith, 2006; Allison et al., 2000). Using fMRI in conjunction with our behavioral model allows us to distinguish these hypotheses.

METHODS

Task

Eighteen participants (6 women) performed a decision-making task while undergoing fMRI. All participants signed informed consent and the study was approved by the National Hospital for Neurology and Neurosurgery ethics review board. Each participant was given the following instructions beforehand: "On each trial in this task you will be presented with two faces. You will have to select one of the faces. Press the top button to select the top face, the bottom button to select the lower face. Your task is to try to figure out which face in each block has the highest probability of winning and pick that face as many times as possible. You will be told when the block switches, and at each switch the faces will be associated with new probabilities of winning."

The face stimuli and probabilities of winning were kept constant throughout each block, and each block consisted of 26 trials. Twenty-six trials are sufficient for an ideal observer to identify the correct face in 85% of blocks (Averbeck & Duchaine, 2009). Probabilities were assigned at the beginning of each block, with one face leading to a "win" 40% of the time, the other leading to a "win" 60% of the time. Probabilities were counter-balanced across blocks such that the happy face was assigned the high reward probability in two blocks and the angry face in two blocks and the order of these assign-

ments was balanced, as much as possible, across participants. Participants were given four blocks of happy and angry faces. Two male identities were used. This was interleaved with four blocks of neutral faces with different identities. Analysis was confined to the emotion blocks. All participants were paid the same amount (which was greater than their actual winnings) but they were not informed of this until after the experiment had ended.

Behavioral Data Analysis

At the beginning of each block of trials, the subjects were told that the probabilities had been reassigned and they should try to work out by trial and error which face was best. Thus, at the beginning of the block, the subjects had no evidence about which face was best and they had to begin selecting one or the other face, registering the feedback, and trying to work out which face was best. Although one face was rewarded more often than the other, the probabilities used were .6 and .4 and, as such the task, were challenging. It was possible that over short intervals the face which had a lower probability of being rewarded would be rewarded more often than the other. Therefore, in the initial analyses, we referenced the subjects' behavior to an ideal observer model which estimated, based on all of the feedback received in the current block, which face was best. By comparing each subject's choices to the ideal observer, a fraction correct could be derived, as a baseline estimate of performance. Additionally, by examining whether subjects chose the happy face more often when they should have chosen the angry face, or chose the angry face more often when they should have chosen the happy face, we could determine if their off-model decisions showed a relative preference for the happy face.

Because the outcome in each trial was either a win or a loss, the ideal observer was based upon a binomial model. The likelihood that the rewards were being generated probabilistically by an underlying probability θ_i in this model is given by:

$$p(D|\theta_i) = \theta_i^{r_i} (1 - \theta_i)^{N_i - r_i} \quad (1)$$

Here, D is the observed series of reward outcomes, θ_i is the probability that face i (angry or happy) is rewarded, r_i is the number of times face i was rewarded, and N_i is the number of times face i was selected. This equation provides the distributions over reward probabilities for each face. Specifically, as the subjects did not know the underlying probabilities, they would infer a distribution of possible probabilities, given the reward outcomes. For example, if one observed seven heads in 10 coin tosses, it would be possible that the coin was fair (i.e., $p = .5$ of heads vs. tails), but it would also be possible, in fact more likely, that the coin was unfair and had a probability of heads equal to .7. Equation 1 gives the

complete distribution over the probabilities for a given set of outcomes.

To make a decision, the subjects had to infer which face was better. We operationalized this decision step by assuming that subjects would compute the probability that face i was more often rewarded than face j . This was given by:

$$p(\theta_i > \theta_j) = \int_0^1 p(\theta_i | D) \int_0^{\theta_i} p(\theta_j | D) d\theta_j d\theta_i \quad (2)$$

The integral is over the posterior. For the ideal observer, the prior was flat, and as such, the posterior is just the normalized likelihood. As a decision rule, this probability can be thresholded at chance. This gives the choice of the ideal observer (\hat{f}).

$$\begin{cases} p(\theta_i > \theta_j) > 0.5 & \hat{f} = i \\ p(\theta_i > \theta_j) < 0.5 & \hat{f} = j \end{cases} \quad (3)$$

The first behavioral analyses examine the consistency between the choices of the ideal observer and the subjects by building a 2×2 contingency table, with model choice (happy or angry) as columns and subject choice (happy or angry) as rows. For the case of $p(\theta_i > \theta_j) = .5$ (i.e., when the model is equivocal between the two choices), we incremented both the happy and the angry choices of the ideal observer in the choice table by .5. The ideal observer was not used in any analyses of the fMRI data.

The ideal observer treats all choice stimuli equally and ignores the emotional expression. Human subjects, however, are known to be influenced by the emotional content of the stimuli. We added parameters to the ideal observer model to account for these potential biases and fit the models with additional parameters to the choice data from individual subjects. This Bayesian Reinforcement Learning (BRL) model contained four extra parameters which allowed us to model weighting of positive feedback, weighting of negative feedback, evidence bias (differential weighting of feedback based upon the emotional content of the face that was chosen), and a prior bias toward one of the faces.

For rewarded trials, the reward value in the model was calculated as:

$$\begin{aligned} r'_{\text{happy}}(t) &= .5 + a + c \\ r'_{\text{angry}}(t) &= .5 + a - c \end{aligned} \quad (4)$$

Whereas for unrewarded trials, it was calculated as

$$\begin{aligned} r'_{\text{happy}}(t) &= .5 - b - c \\ r'_{\text{angry}}(t) &= .5 - b + c \end{aligned} \quad (5)$$

The variables a , b , and c were fit as free parameters in the model. The parameter a is a weighting given to positive feedback. For the ideal observer, positive rewards are val-

ued at 1, so $a = 0.5$. The parameter b is a weighting given to negative feedback. For the ideal observer, negative feedback is valued at 0, so again $b = 0.5$ for the ideal observer model. Therefore, values of a and b below (above) 0.5 measure the amount that feedback is under (over) weighted and values of a and b near zero indicate that positive or negative feedback are ignored. The variable c controlled the evidence bias introduced by the expression. If c is positive, feedback is given increased weight for the happy face and decreased weight for the angry face. Therefore, positive values of c indicate that the subject picked the happy face more often than would be expected based on the feedback (evidence) alone, irrespective of the prior, discussed below. We also examined a model with separate emotion preference terms for positive and negative feedback, but results were more robust when we fit one term, c , to both.

The total reward in the block under the model for face i was then calculated as:

$$r_i = \sum_{t=1}^T r'_i(t) \quad (6)$$

Thus, the total reward for each face, up to trial T in the current block was the sum of the biased reward values.

The prior preference was modeled using a Beta distribution to model the prior disposition toward each face as:

$$p(\theta_i | \alpha_i, \beta_i) = \theta_i^{\alpha_i} (1 - \theta_i)^{\beta_i} \quad (7)$$

We constrained α_i and β_i to get a good model fit by reducing them to one degree of freedom. This was done by estimating a single parameter, d , and then computing α_i and β_i as:

$$\begin{aligned} \alpha_{\text{happy}} &= 2(0.5 + d) \\ \alpha_{\text{angry}} &= 2(0.5 - d) \\ \beta_{\text{happy}} &= 2(0.5 - d) \\ \beta_{\text{angry}} &= 2(0.5 + d) \end{aligned} \quad (8)$$

The parameters of the model were fit to individual subjects by maximizing the likelihood of the parameters, given the data. Thus, we first calculated

$$p(\theta_i | D, a, b, c, d) = \frac{p(D | \theta_i, a, b, c) p(\theta_i | d)}{p(D | a, b, c, d)} \quad (9)$$

And then calculated the belief estimate under the model using

$$p(\theta_i > \theta_j) = \int_0^1 p(\theta_i | D, a, b, c, d) \int_0^{\theta_i} p(\theta_j | D, a, b, c, d) d\theta_j d\theta_i \quad (10)$$

There were no prior distributions placed over the parameters, allowing these terms to be completely data driven. The ideal observer can be recovered by setting a and b to 0.5, c to 0, and by assuming a flat prior on θ . We next maximized the likelihood of the individual subject's sequence of decisions by adjusting the parameters a , b , c , and d . The likelihood was given by:

$$p(D^*|a, b, c, d) = \prod_{k=1}^N (p_k(\theta_i > \theta_j)l_k + (1 - p_k(\theta_i > \theta_j))(1 - l_k)) \quad (11)$$

Where $l = 1$ if the subject selected face i (e.g., the happy face) and $l = 0$ if the subject selected face j (e.g., the angry face). Here, D^* is the series of decisions of the subject, as opposed to the series of outcomes, which is collected in D in the previous equations. We maximized the likelihood using *fminsearch* in Matlab. We started from several initial conditions and examined the likelihood and parameter values for each individual subject. Most initial conditions led to the same maximum estimates, giving us confidence that these maxima were global maxima. Significance testing of model parameters was then carried out using one-sample t tests of the parameter distributions for each subject. This is a random effects approach and gives us an estimate of whether a particular term in the model is significantly different from zero in the population.

fMRI Data Analysis

Functional brain images were acquired on a 3-T Allegra scanner at a resolution of $3 \times 3 \times 3 \text{ mm}^3$, TR = 2.88 sec, TE = 65 msec. The slice angle was set to -30° and a Z-shim of $-0.4 \text{ mT/m} \cdot \text{msec}$ was applied so as to minimize signal dropout in orbito-frontal cortex and amygdala (Weiskopf, Hutton, Josephs, & Deichmann, 2006). Field maps were acquired using a dual-echo 2-D gradient-echo sequence with echoes at 10.0 and 12.46 msec. T1-weighted structural images were acquired for subject alignment at a voxel resolution of $1 \times 1 \times 1 \text{ mm}^3$. Subjects lay in the scanner with foam head-restraint pads to minimize any movement. Responses were made using a button box held in the right hand. Using SPM5, images were realigned to the first volume, normalized to a standard echo-planar imaging template, and smoothed using a 6-mm full-width half-maximum Gaussian kernel. Images were analyzed using an event-related general linear model, with the onsets of each task event represented as a delta (stick) function.

Time points of interest were the choice and feedback screens within each trial. These were modeled with two delta functions, and each had associated parametric modulators described below. Also included were several regressors of no interest representing the stimulus onsets, the button press, and motion parameters estimated

during the realignment procedure. All regressors (except the motion parameters) were convolved with a canonical hemodynamic response function and its temporal derivative. The choice screen was parametrically modulated by a single variable of interest: the probability that the chosen face was more likely to be rewarded $p(\theta_i > \theta_j)$ (Equation 10). The feedback regressor was parametrically modulated by the reward prediction error (RPE) from the Bayesian reinforcement learning model. Although the choice and feedback were separated by a fixed 4.5-sec interval (Figure 1), the parametric regressors were not strongly correlated (the regressors had $<10\%$ shared variance in the first-level model for all subjects). As they were entered simultaneously in the model, they could compete for distinct variance in the hemodynamic response.

The RPE is given by the actual reward minus the probability of the reward computed under the model. Specifically:

$$\text{RPE}(t) = \begin{cases} 1 - \frac{r_i + \alpha_i}{N_i + 2} : D^*(t) = i, r(t) = 1 \\ 0 - \frac{r_i + \alpha_i}{N_i + 2} : D^*(t) = i, r(t) = 0 \\ 1 - \frac{r_j + \alpha_j}{N_j + 2} : D^*(t) = j, r(t) = 1 \\ 0 - \frac{r_j + \alpha_j}{N_j + 2} : D^*(t) = j, r(t) = 0 \end{cases} \quad (12)$$

Note that r_i in this case reflects the biased integration of evidence (Equation 6), so it is different depending on the model parameters used to compute it. It is not simply the number of times each face has been rewarded. Additionally, the second term in the first line refers to the prediction component. More specifically, the RPE is given by $r(t) - p(r(t))$, that is, the actual reward received minus the predicted reward. Thus, the RPE can be separated into the reward received $r(t)$, which is noted as 1 or 0 in the first term of Equation 12, and the predicted reward $p(r(t))$, where the predicted reward is given by the second term, which includes both reward and prior components.

Reported results are all whole brain cluster-level corrected at $p < .05$ (height threshold = .005, uncorrected; extent threshold = 30 contiguous voxels). In addition, small-volume corrected (SVC) results are reported in areas that were hypothesized as components of the reward and mentalizing networks, including the ventral striatum and TPJ. Activations in all figures are correspondingly shown at $p < .005$, uncorrected. All significant clusters are reported in the results.

RESULTS

Subjects carried out a decision-making task in which they were asked to learn, in each block of trials, which of two faces was being more often rewarded, and then select that face as many times as possible (Figure 1). On each individual trial, participants were presented with two faces, one happy and one angry. The faces had the same identity and were presented pseudorandomly on either

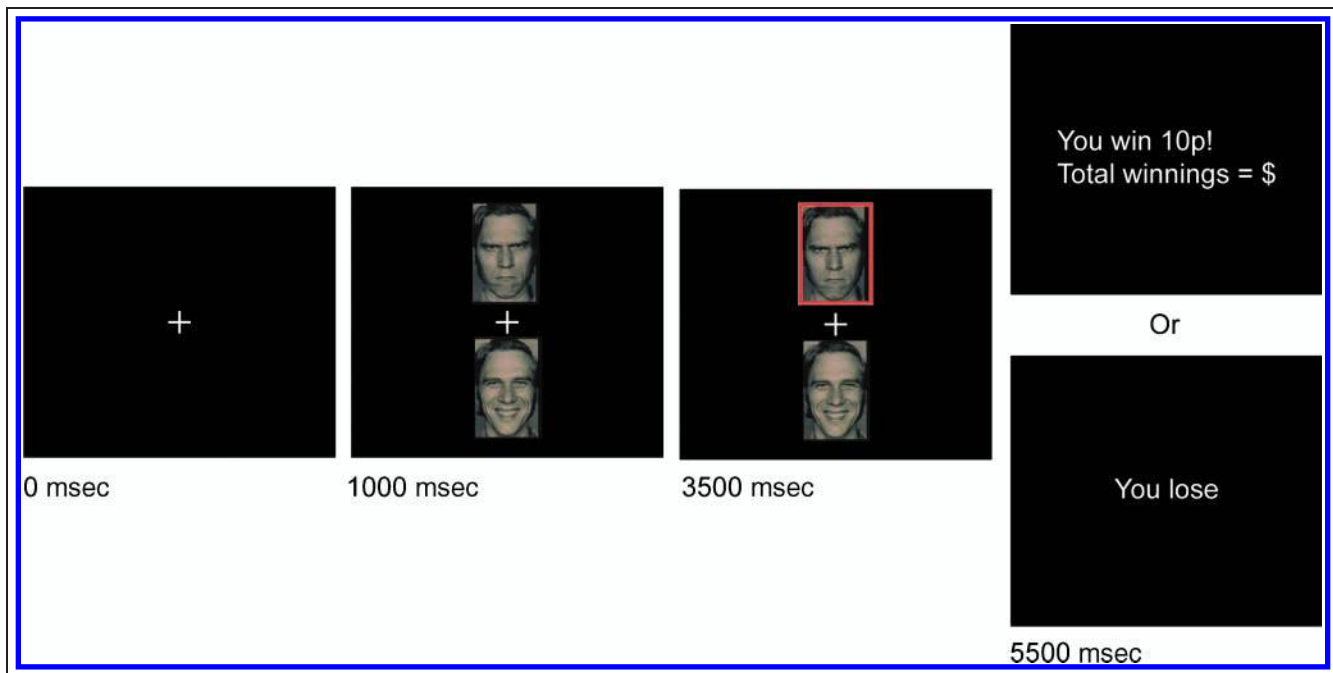


Figure 1. Task: Angry and smiling faces were presented on each trial. One of the faces was then selected using a button box, and the subject was then informed of the trial outcome (win or lose). Subjects had to integrate information across trials to determine which face had a higher probability of leading to a “win.” Time at bottom of each frame indicates the onset time.

the top or bottom of the screen. Participants were given 2.5 sec to make their decision, after which the chosen face was highlighted to confirm their decision. They were then informed as to whether they had “won” or “lost” in that trial, with a win worth 10 pence, and a loss worth nothing.

Behavior

The task was challenging because the two faces were stochastically rewarded and the difference between the probability of reward for each face was small (.6 vs. .4). To begin assessing task performance, participant behavior was compared to an ideal-observer model on a trial-by-trial basis. The ideal observer performed optimally based on the reward history up to the current trial and was not affected by the facial expression (see Methods). Therefore, deviations from this model can be used to examine the effects of the emotional expression. When referenced to this model of ideal responding, subjects were found to perform, on average, at 72.9% accuracy, which was significantly above chance [$t(17) = 8.6, p < .01$]. The ideal observer was also used to test whether participants were biased in their responding toward the happy face, averaged across the block. On average, participants chose the happy face when they should have chosen the angry face 30% of the time [$p(\text{subject choice} = \text{happy} | \text{ideal observer choice} = \text{angry}) = .30$] and they chose the angry face when they should have chosen the happy face 24% of the time [$p(\text{subject choice} = \text{angry} | \text{ideal observer choice} = \text{happy}) = .24$]. These prob-

abilities were significantly different ($p < .01$, likelihood ratio test, $df = 1$). Stated another way, subjects were biased toward choosing the happy face about 60% of the time when the model evidence was equivocal (Figure 2A; equivocal model evidence is at .5 on the x -axis). Thus, there was a preference across participants to select the happy face, even when the evidence equivocally or more strongly supported the angry face. This result replicates our previous behavioral findings that emotional expressions consistently bias learning processes (Averbeck & Duchaine, 2009).

Next, we modeled the choice behavior using a Bayesian reinforcement learning model fit to the choice data from each individual subject. The model had four parameters, as opposed to the ideal observer, which had no free parameters. These parameters allowed us to better model the individual choice behavior of each subject, and examine the relative impact of four factors on decision-making. The first two parameters, positive feedback (parameter a) and negative feedback (parameter b), modeled only the effects of the outcome the subjects received at the end of each trial. The second two parameters, a relative evidence preference for happy faces (parameter c) and a relative prior preference (or possibly an aversion) for happy faces (parameter d), modeled the effects of the expression on the decision processes, independent of how much positive or negative feedback had been given for each face. The prior effect modeled the bias toward choosing one or the other expression prior to the subjects having any evidence about which face was more rewarding. Conversely, the evidence factor modeled the expression’s

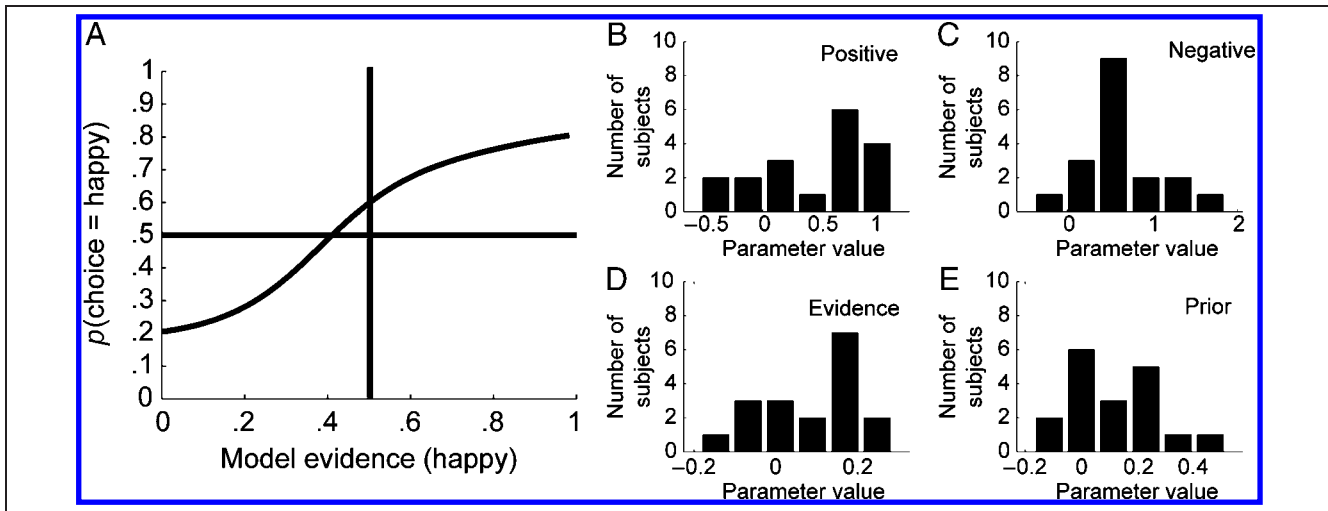


Figure 2. Behavior. (A) Ideal observer model evidence versus subject choice behavior. Shift of curve up and left indicates preference for happy face, as it is chosen more often than the feedback predicts it should be. (B) Distribution across subjects ($n = 18$) of parameters from Bayesian reinforcement learning model for positive feedback (parameter a from Methods). (C) Distribution of parameters for learning from negative feedback (parameter b). (D) Distribution of parameters for evidence bias (parameter c). (E) Distribution of parameters for prior bias (parameter d).

impact on the accumulation of evidence. For example, if a subject counted positive feedback from the happy face more than positive feedback from the angry face, they would have a positive evidence bias. Individual t tests on the four model parameters (Figure 2B–E) across subjects showed that the subjects learned from both positive [$\langle a \rangle = 0.44$, $t(17) = 3.7$, $p = .002$] and negative feedback [$\langle b \rangle = 0.671$, $t(17) = 5.6$, $p < .001$]. Additionally, the positive feedback term was significantly correlated with the fraction correct for each individual subject [$r(16) = .879$, $p < .001$], but the negative feedback parameter was not [$r(16) = .146$, $p = .562$]. Thus, the sensitivity to positive feedback more accurately characterized overall performance. Participants also had significant evidence [$\langle c \rangle = 0.09$, $t(17) = 2.8$, $p = .013$] and prior [$\langle d \rangle = 0.13$, $t(17) = 3.1$, $p = .006$] terms. Overall, the four-parameter Bayesian model provided a significantly better prediction of the behavior than the ideal observer in all 18 subjects [$\chi^2(72) = 337.8$, $p < .001$]. Thus, consistent with the results reported above, the facial expression influenced decision-making processes toward the happy face as captured by positive values of parameters c and d of the model. Furthermore, the model was able to factor these effects on learning into four components, two financial (a and b) and two social (c and d).

Correlations between Model Predictions and BOLD Response at Time of Choice

Below we correlate predictions from the behavioral model with the BOLD signal on a trial-by-trial basis. To examine how emotional preferences at time of choice are captured by model parameters, we compared model predictions with subsets of model parameters set to zero (Figure 3). Choice probability estimates the probability that the subject will select each option given the feedback history, and

therefore, it is an estimate of the current subjective value of each option. Thus, if we compare a model with and without the prior term, we can see how the prior captures a preference toward the happy face that decreases as more evidence is gathered (Figure 3). Another salient point is that subjects generally do not integrate information as well as the ideal observer, so if the ideal and learning lines are compared, the choice probability is less extreme (i.e., closer to .5 which is equivocal) for the learning model. For the analyses below, we will correlate the full model with the BOLD signal, as this model best describes the subject's behavior. We then carry out between-subjects correlations

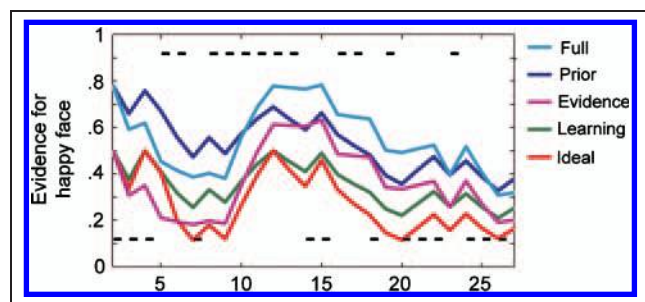


Figure 3. Choice probability predictions from Equation 10: $p(\theta_{\text{happy}} > \theta_{\text{angry}})$. Data are from an example block of trials for a single subject. Dots near .9 indicate positive feedback, dots near .1 indicate negative feedback, for the corresponding trial. Parameters were first fit to each subject to optimize the model's prediction of their choice behavior, and then these parameters were used to generate the data shown. The ideal model has all model parameters set to zero. The learning model has parameters c (evidence) and d (prior) set to zero. Thus, this model incorporates differences in how subjects learn from positive (parameter a) and negative (parameter b) feedback, but it does not incorporate any social preferences. The evidence model has parameter d (prior) set to zero and the prior model has parameter c (evidence) set to zero. The full model does not set any of the parameters to zero.

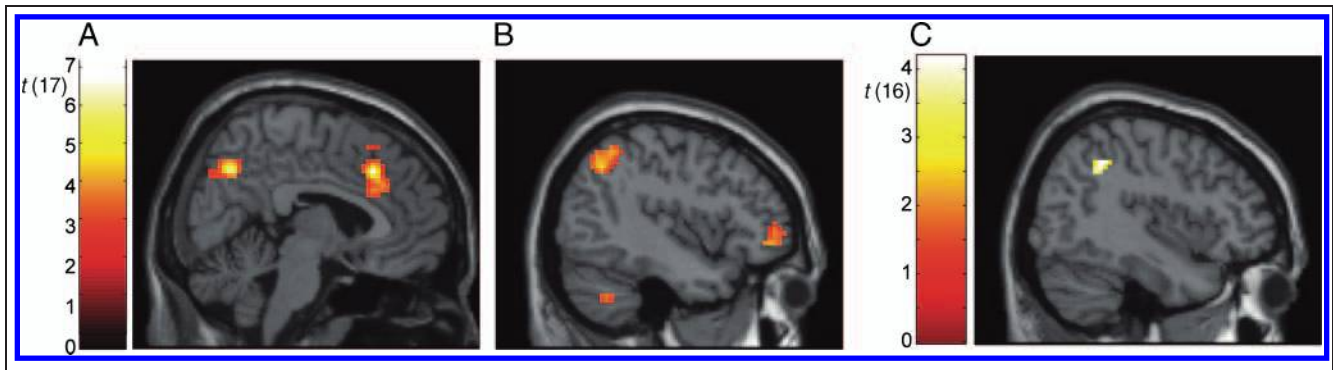


Figure 4. Significant correlations between BOLD signal and choice probability at time of choice. (A) Correlation between choice probability, $p(\theta_{\text{choice}} > \theta_{\text{alternative}})$, and BOLD response at time of choice in ACC ($x = 6, y = 24, z = 39$) and medial parietal cortex ($x = 6, y = -66, z = 39$). (B) Activation from same contrast in left parietal cortex ($x = -42, y = -60, z = 42$). (C) Between-subject correlation between learning from positive feedback (parameter a from model) and choice probability contrast in right parietal cortex ($x = 42, y = -45, z = 42$).

between model parameters and the contrast of the full model on the BOLD signal in order to identify areas that mediate individual differences in parameter strength.

When we correlated choice probability with the BOLD signal at the time of the decision, we found three significant clusters in cortex. The first was in ACC (Figure 4A; $p < .05$, cluster-level whole-brain corrected), the second was in medial parietal cortex (Figure 4A; $p < .05$, cluster-level whole-brain corrected), and the third was in left parietal cortex, dorsal to the TPJ (Figure 4B; $p < .05$, cluster-level whole-brain corrected). Thus, there were significant correlations at the time of choice with the relative value of the chosen option. We examined between-subjects correlations between learning from positive feedback (parameter a) and the contrast between choice probability and the BOLD signal. Specifically, this analysis looks for areas which show greater correlation between BOLD and choice probability in subjects who learn more from positive feedback. This analysis identified a single cluster in right parietal cortex (Figure 4C; $p < .05$, cluster-level, SVC centered at peak activation, 10 mm radius). Therefore, subjects who showed stronger corre-

lations of right parietal cortex activity with choice probability also learned more from positive feedback, leading to better performance in the task.

Correlations between Model Predictions and BOLD Response at Time of Feedback

When feedback was given, an RPE could be calculated as the difference between actual and predicted outcomes (Equation 12). The RPE also depended on the model parameters, as the model estimated an implicit expected reward. Specifically, the effects of a prior preference (Figure 5A) showed that subjects made decisions as if the happy face would be more often rewarded at the beginning of the block. This effect decreased as the block continued. The evidence preference, however, had an effect that extends throughout the block (Figure 5B). Similar to the choice probability analysis above, correlations with the BOLD signal are carried out using the full model.

The RPE was extracted trial-by-trial from the model, after the model had been fit to each individual subject.

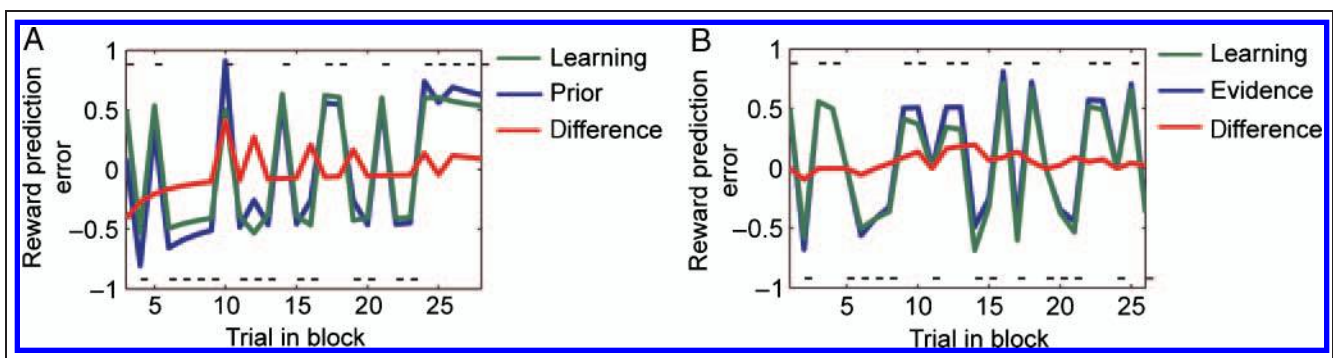


Figure 5. RPE under different models. Data are from example blocks of trials from two different subjects chosen to illustrate the effects of prior (parameter d) and evidence (parameter c) terms. Models are defined the same as in Figure 3. Difference refers to the difference between the learning model and the other model in each panel. (A) Comparison of RPE under the model with no prior and no evidence bias terms, and a model with the prior bias term. (B) Comparison of RPE under a model with no prior and no evidence bias terms and a model with the evidence term.

Therefore, this RPE reflects the combined effects of all four model parameters and was optimized to fit the choices of each individual subject. When the RPE was correlated with the BOLD response at the time of feedback, we found three areas that were significant after whole-brain correction. One in the anterior subcallosal cingulate (Figure 6A and B; $p < .05$ cluster-level, whole-brain corrected), one in the posterior cingulate (Figure 6A; $p < .05$, cluster-level, whole-brain corrected), and one in visual cortex (Figure 6A; $p < .05$, cluster-level, whole-brain corrected). Activation was also seen in the ventral striatum that survived small-volume correction [SVC] (Figure 6C; $p < .05$, cluster-level, SVC centered at peak activation, 10 mm radius).

The model factors the overall effects of the decision process into four components, positive feedback, negative feedback, and evidence and prior preferences for the emotional expressions. To examine which brain areas mediate different components of the decision process, we correlated model parameters between subjects with contrast estimates which assessed correlations between the RPE and the BOLD signal. Thus, this analysis looked for areas which had a stronger (or weaker) modulation of the BOLD signal for subjects which learned more from positive feedback. We first examined correlations between learning from positive feedback (parameter a) and found three areas that were significant. Specifically, one cluster in the right ventral striatum (Figure 7A; peak activation at $x = 12, y = 12, z = -9, p < .05$, cluster-level whole-brain corrected), one in ACC (Figure 7B; peak activation at $x = 3, y = 39, z = 15; p < .05$, cluster-level whole-brain corrected), and one in dorsolateral prefrontal cortex (not shown; $p < .05$, cluster-level whole-brain corrected, peak activation at $x = 39, y = 42, z = 27$). Additionally, the correlation in the left ventral striatum was significant after SVC (Figure 7A; $p = .001$ cluster-level, SVC centered at peak activation, 10 mm radius). There were no clusters that exceeded chance for learning from negative feedback (parameter b), which is unsurprising, given that this parameter did not correlate with overall performance in the task.

When the same analysis was carried out for correlations with prior preference (parameter d), two significant clusters were found. The first was in the caudal anterior cingulate (Figure 7C; $p < .05$, cluster-level whole-brain corrected) and the second was in the right dorsal TPJ (Figure 7D; $p < .05$, cluster-level whole-brain corrected). Finally, when correlations with the evidence bias (parameter c) were examined, a single cluster in the right TPJ reached significance (Figure 7E; $p < .05$, cluster-level whole-brain corrected). Interestingly, while the TPJ clusters correlating with the evidence and prior parameters were very near each other, there was only minor overlap between activations (2 common voxels). Examination of these effects at the peak voxel of each significant cluster showed no significant correlations with the alternative parameter (Figure 7, right column). Thus, correlations between model parameters related to learning and brain activation were found in areas commonly implicated in reward processing, and correlations between model parameters related to social preferences were found in components of the mentalizing network.

The correlations between individual differences in learning from positive feedback and activation in the ventral striatum (Figure 7A) were negative. This suggests that larger prediction errors lead to larger BOLD responses in subjects that learned less (given by the learning from positive feedback parameter, which is strongly correlated with overall performance as reported above). This finding was somewhat counterintuitive. To examine this in more detail, we carried out an additional analysis, in which we separated the RPE into the reward (i.e., “You win” vs. “You lose”) and the prediction (i.e., the subject’s prediction of reward, based on the model) components, and simultaneously regressed these on the BOLD response (see Methods for an additional discussion of the reward and prediction components of the RPE). As the outcomes were highly stochastic (60%/40% splits in every block), reward and reward prediction were relatively independent of each other. When we carried out this analysis, we first

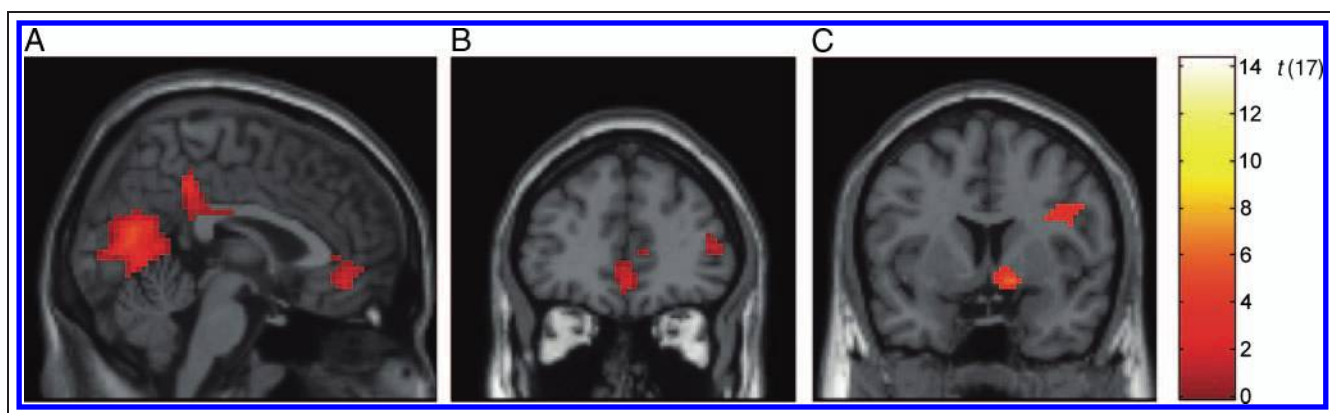
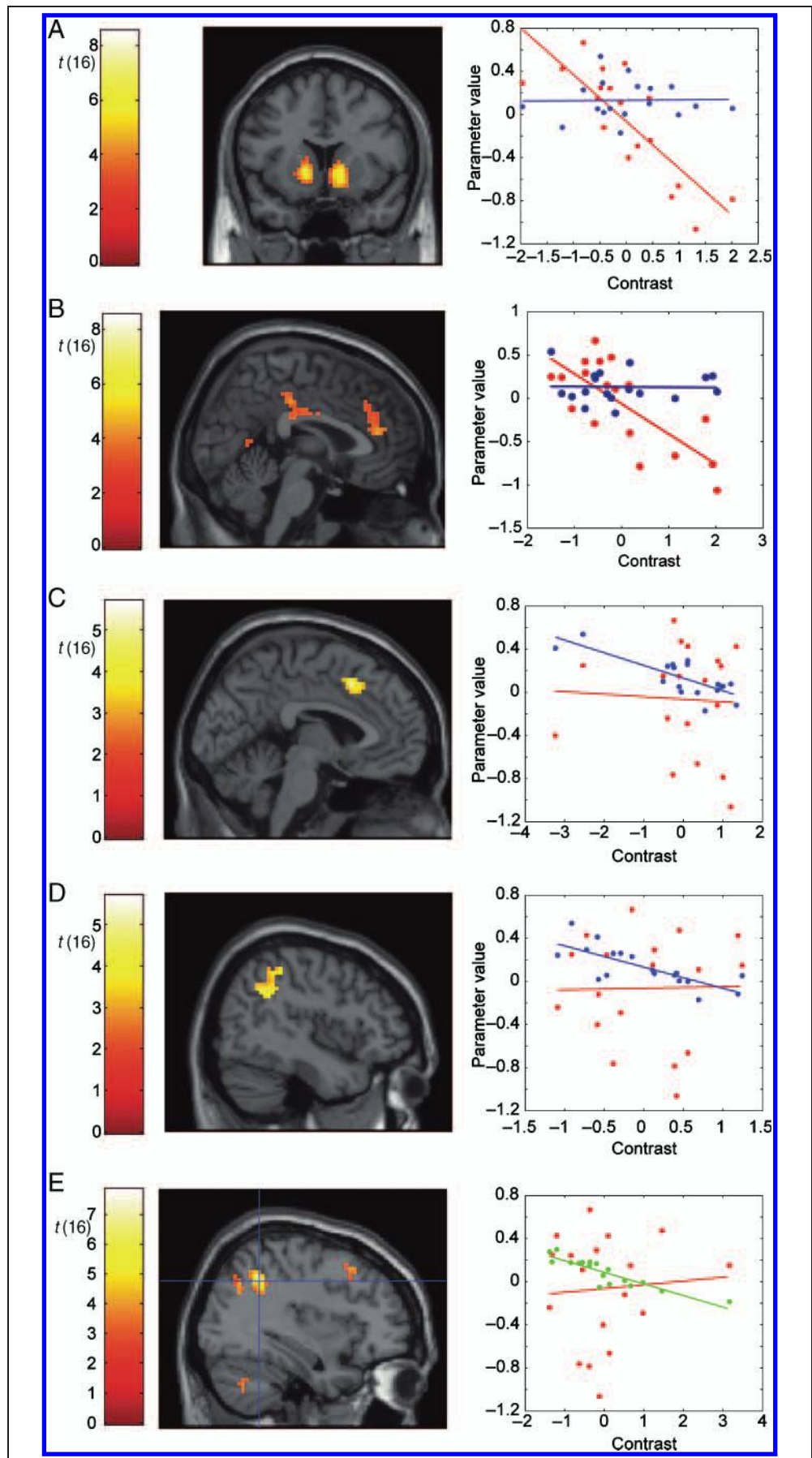


Figure 6. Significant correlations with RPE (Equation 12) at time of feedback. (A) RPE correlation in anterior ($x = 3, y = 42, z = -9$) and posterior ($x = 3, y = -39, z = 39$) cingulate cortex and visual cortex ($x = -9, y = -81, z = 0$). (B) RPE correlation shown in coronal section through ACC. (C) RPE correlation in the right ventral striatum ($x = 12, y = 6, z = -9$).

Figure 7. Between-subject correlations between contrast estimates for the RPE at time of feedback and model parameters from the Bayesian reinforcement learning model. (A) Significant correlation between positive feedback (parameter a) and contrast in the ventral striatum ($x = 12, y = 12, z = -9$). (B) Significant correlation between positive feedback and contrast in ACC ($3, 39, 15$). (C) Significant correlation between prior (parameter d) and dorsal ACC ($x = -3, y = 15, z = 45$). (D) Significant correlation between prior and TPJ ($x = 45, y = -54, z = 36$). (E) Significant correlation between evidence bias (parameter c) and TPJ ($x = 36, y = -51, z = 39$). The right-hand column shows the correlations between the peak voxel in these contrasts and model parameters. We note that these plots are biased toward showing a strong correlation with the parameter used to select the voxel; they are included to illustrate the specificity of the relationship between separable model parameters and brain activity. Blue lines and dots indicate the relation between prior (parameter d , Equation 8) and contrast; red line and dots indicate the relation between learning from positive feedback (parameter a , Equation 4) and contrast. Green line and dots indicate the relation between activity and the evidence bias (parameter c , Equations 4 and 5).



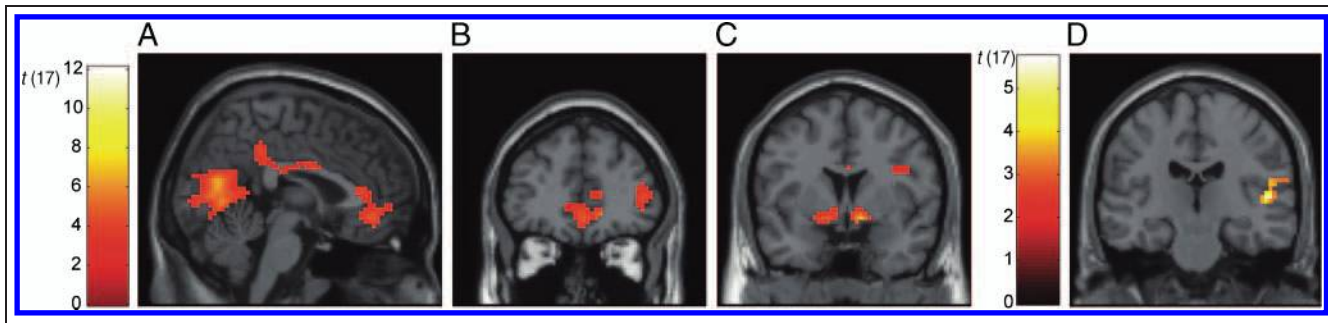


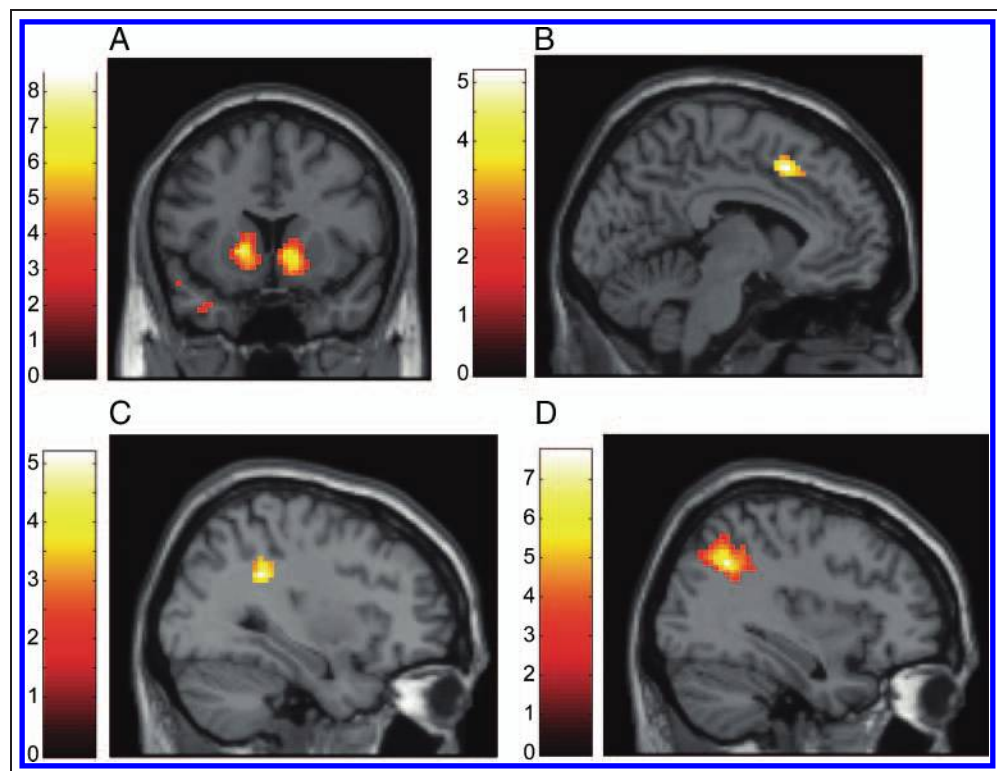
Figure 8. Separate reward (i.e., win vs. lose) and prediction correlations (see Methods for a description of the separate components of the RPE). (A–C) The network correlated with reward is highly similar to the network correlated with the RPE, shown in Figure 6. (D) Specific correlations with reward prediction were found in the right TPJ ($x = 51, y = -21, z = 6$).

found that the reward component correlated with a network nearly identical to that which correlated with the RPE. Specifically, significant clusters were found in subcallosal ACC (Figure 8A and B; $p < .05$, cluster-level whole-brain corrected) and in the ventral striatum (Figure 8C; $p < .05$, cluster-level, SVC centered at peak activation, 10 mm radius). Thus, the correlation in these structures with the RPE appears to be largely driven by reward versus no-reward outcomes. When we examined correlations with the prediction component of the RPE, a cluster in the TPJ was significant (Figure 8D; cluster-level whole-brain corrected). There were, however, no significant clusters or even suprathreshold voxels in the subcallosal cingulate or the ventral striatum, further supporting the hypothesis that these correlations are driven by reward and not prediction. We note that the TPJ cluster, which correlated with prediction, did not appear in the original analysis with

compound RPE. One potential explanation for this discrepancy is that the variance in the RPE is dominated by reward versus no-reward outcomes (Figure 5). Thus, the TPJ seems to specifically correlate with prediction of reward association with the faces rather than responding to reward feedback itself.

In the final analysis, we again carried out between-subject correlations, here between the separate reward and prediction components and the factors from the behavioral model to examine which aspect of the feedback, the reward or the prediction, better correlated with the parameters from the model identified in Figure 7. In this analysis, we found a negative correlation in the ventral striatum between reward outcome and learning from positive feedback (Figure 9A; $p < .05$, cluster-level whole-brain corrected). No suprathreshold voxels were found for the correlation with prediction. Similarly, the prior

Figure 9. Correlations between reward and prediction contrast and model parameters. (A) Significant negative correlation between reward contrast and learning from positive feedback (left side: $x = -15, y = 15, z = 0$; right side: $x = 9, y = 6, z = -12$). (B) Significant negative correlation between reward contrast and prior ($x = -6, y = 15, z = 45$). (C) Significant negative correlation between reward contrast and prior in the TPJ ($x = 33, y = -42, z = 30$). (D) Significant positive correlation between prediction contrast and learning from positive feedback ($x = 36, y = -57, z = 36$).



correlated with reward in caudal ACC (Figure 9B; $p < .05$, cluster-level whole-brain corrected) and the dorsal TPJ (Figure 9C; $p < .05$, cluster-level whole-brain corrected) but not with prediction (no suprathreshold voxels). One cluster showed a positive correlation between prediction and learning from positive feedback in the dorsal TPJ (Figure 9D; $p < .05$, cluster-level whole-brain corrected). Thus, the correlations in Figure 7 can be explained by correlations with reward feedback and not with prediction.

DISCUSSION

Decision-making processes are frequently studied in the context of reward, as decisions are assumed to maximize a subjective utility function, where utility is mapped to reward or punishment. In our task, we found that both the financial reward and the emotional expression drove decision-making processes. Based on behavior alone, one might hypothesize that the smiling face was appetitive and the angry face was aversive, and that viewing them gave some small additional increment to the reward circuitry which drove the decision process (Montague & Berns, 2002). This would be consistent with the assumption that decisions are driven by rewards and that all types of reward, for example, social and monetary, are processed through a common circuitry (Chib, Rangel, Shimojo, & O'Doherty, 2009). At some point in the process, this premise has to be true, as the decision is ultimately revealed by a motor action. The question is whether all cues which drive decisions are processed by a single system or whether, in our case, biases in learning induced by emotional expression are mediated by a partially separate network which processes social information (Van Overwalle, 2009; Frith & Frith, 2006). Importantly, the prior and evidence parameters of the model used in our study captured learning biases driven by the emotional expression of the face, independent of the financial feedback. With respect to this division, we found that, across subjects, the ventral striatum mediated learning from positive financial feedback, whereas the caudal anterior cingulate and the dorsal TPJ mediated prior- and evidence-related preferences for the happy face.

Whether financial and social cues are processed by one or multiple systems cannot be distinguished easily using behavioral studies. However, this questions can be addressed using fMRI, as networks mediating the effects of rewards tend to differ from networks that process social information (van den Bos et al., 2007; Walter, Abler, Ciaramidaro, & Erk, 2005), although there is no complete separation of these networks. For example, orbital and ventromedial prefrontal cortex have been implicated in both reward (Elliott & Deakin, 2005; O'Doherty, Critchley, et al., 2003) and social processing (Grossman et al., 2010; Hynes et al., 2006; O'Doherty, Winston, et al., 2003). It is likely that the distinction between social and emotional processing becomes quite important when considering

ventromedial and orbital–prefrontal cortex. For example, both rewards and social cues can have positive and negative affective value, and it may be this affective component, as opposed to the specifically social component, that is engaging these areas. Indeed, previous studies have shown that orbital and ventromedial prefrontal areas are preferentially engaged by contrasts over affective aspects of the social tasks (Grossman et al., 2010; Hynes et al., 2006). Interestingly, in our task, we did not find differential activation in orbital prefrontal cortex for rewards or social processing, although we found activation in ventromedial prefrontal cortex for financial reward.

Previous work has addressed related but different questions. For example, Behrens et al. (2008) found that when subjects were integrating feedback and explicit advice from a confederate about which stimulus to choose, a social predication error activated the caudal anterior cingulate as well as the TPJ. These results are consistent with ours, despite the fact that our task utilized an implicit social cue that subjects could ignore, whereas Behrens et al. utilized an explicit social cue. Other studies have examined the impact of feedback that was given by either a computer or a confederate (van den Bos et al., 2007), and found activation in the temporal poles when contrasting the effects of confederate versus computer feedback. Inferring second-order effects of one's own actions on the actions of an opponent in a strategic game also activates anterior cingulate and temporo-parietal areas (Hampton, Bossaerts, & O'Doherty, 2008). All of these tasks, particularly the latter, differ from more commonly used theory of mind tasks, but, as in the present study, appear to automatically engage components of the same network. In general, theory of mind refers to the mental process of inferring the thoughts or intentions of others. This mental process is often engaged by asking subjects to infer knowledge about participants in various scenarios.

Although our study is related to these previous studies, it differs in important ways. First, our subjects were required to learn the value of a face, where the faces differed in their emotional expressions. Previous studies focused on learning the value of an abstract image (Behrens et al., 2008) or the value of an abstract choice (Hampton et al., 2008), or focused on task performance (van den Bos et al., 2007). Second, there was no social feedback per se in our experiment. The social effects were mediated by implicit effects of emotional expressions. Third, using a Bayesian reinforcement learning model, we were able to take advantage of the fact that expressions have prior intrinsic value, and factor the effect of the expressions into prior and evidence terms. This is an important distinction, given that previous work has tended to focus on social versus nonsocial *evidence* and has not examined effects of stimuli which have intrinsic social information, unlike in our task. Recent work has shown that the intrinsic value of faces can be more effective than abstract stimuli when used as a reinforcer and further, that this value is specifically sensitive to oxytocin (Hurlemann et al., 2010). In

accordance with the present data, the results of Hurlemann et al. suggest that feedback information provided by social stimuli is processed by different networks and is affected by different pharmacological mechanisms than nonemotional cognitive feedback. We have also found, in related work, that oxytocin can specifically affect the dispreference for the angry face in our task, without affecting the way reward feedback is processed (Evans, Shergill, & Averbeck, 2010).

We further found that subjects who did not learn as well had larger RPE responses in the ventral striatum than subjects who learned better. When we split RPE into reward and prediction components, the correlations between reward and behavioral parameters were similar to the correlations between RPE and behavioral parameters, demonstrating that these effects are predominantly driven by the reward component of the RPE and not by the prediction component. This potentially counterintuitive finding can be explained by noting that we found a positive correlation between the choice probability contrast and learning from positive feedback at the time of choice in parietal cortex. Subjects who learned better showed stronger correlations between probability that they were picking the best face and BOLD signal at the time of choice. Thus, subjects who learned poorly appeared to respond to the reward (increased striatal response), but they did not appropriately use this reward information to bias brain activity at the time of choice. The correlation in the TPJ between reward prediction and learning from positive feedback further corroborated this interpretation, as this correlation was positive. An analogous result is that smokers computed a fictitious learning prediction error, but they did not appropriately integrate it (Chiu, Lohrenz, & Montague, 2008). Here we show a similar effect, although our finding is within a group of normal subjects who vary in their learning rates. One other study has found that the dorsal striatum showed increased modulation with the RPE in good learners versus bad learners (Schonberg, Daw, Joel, & O'Doherty, 2007). We did not observe a dorsal striatal response in the present study, but this null result could be due to differences in task structure and/or subjects' learning performance. Further work is needed to identify the causal role of reinforcement-related brain activity on individual differences in performance.

We found that caudal cingulate activity correlated with the RPE, an area often seen to be activated in interactive behavioral economics experiments (Tomlin et al., 2006). For example, during economic exchanges, when investors see the trustee's response, there is increased activation in the caudal portion of the cingulate, just dorsal to the corpus callosum. Activation has also been seen in this area when subjects viewed faces of opponents with whom they were interacting in prisoner's dilemma games (Rilling, Sanfey, Aronson, Nystrom, & Cohen, 2004), as well as in mentalizing experiments (Lombardo et al., 2010; Abraham, Werning, Rakoczy, von Cramon, & Schubotz, 2008; Gobbini, Koralek, Bryan, Montgomery, & Haxby,

2007). Thus, activation in this area may be related to associating an outcome with another agent in a social interaction. This is consistent with our hypothesis that the preference effects we see in our task are mediated by mentalizing networks.

Several contrasts identified nonoverlapping but spatially adjacent anterior cingulate and ventromedial prefrontal areas. Specifically, choice probability (caudal, ventral ACC), reward feedback (anterior, ventromedial prefrontal), and a prior preference for happy faces (anterior ventral ACC) all activated nearby areas. The reward activation in ventromedial prefrontal cortex is consistent with a role of ventromedial prefrontal cortex in processing of reward feedback (O'Doherty, Critchley, et al., 2003). The correlations between prior social preferences and ACC activation, and choice probability and ACC activation, are consistent with a role of this area in processing social information relevant to decisions (Behrens et al., 2008; Hampton et al., 2008). This heterogeneity of activation is consistent with a role for ACC in guiding behavior based on multiple sources of value (Rushworth, Buckley, Behrens, Walton, & Bannerman, 2007). Thus, the separable components of our task appeared to engage nearby but separate components of the network mediating value-based decisions.

A consistent finding across a number of contrasts was that the TPJ and the adjacent parietal cortex mediated aspects of learning. A positive correlation at the time of choice was found between learning from positive feedback and the dependence of choice probability on BOLD activation. Thus, subjects who learned better from positive feedback had a stronger correlation between their BOLD response and choice probability in parietal cortex. There were also correlations at the time of feedback between prior and evidence parameters and the correlation between the RPE and the BOLD signal in this region. In this case, subjects who had larger prior and evidence biases had a lower dependence of the RPE on the BOLD signal. Thus, subjects whose decisions were more strongly driven by emotional preferences showed a weaker modulation of the BOLD signal in this area by reward. Finally, when the RPE was split into a reward and a prediction component, the TPJ was the only area that showed a correlation between learning from positive feedback and reward prediction. Thus, this area appears to play a key role in assigning value to the faces in this task. More specifically, the correlation of TPJ activity with reward prediction mediates aspects of the decision process that are updated dynamically, trial-by-trial, whereas the dorsal TPJ correlation with the prior mediates aspects of the decision process that are static and predetermined by each subject's reaction to the face stimuli. Some authors have argued that the TPJ is more important for attention than social processes (Corbetta & Shulman, 2002). However, value and attention are closely related (Maunsell, 2004), and it is not surprising that learning would modulate a network which may also be involved in attention. In our case, however, it is specifically learning the value of a socially relevant image that engages the TPJ.

Conclusion

We found that emotional expressions were able to influence decision processes, and that the brain network underlying this effect differs from that characterized for reward processing. Specifically, although the financial reward component of the task was mediated by the ventral striatum and the subcallosal anterior cingulate, the differential effect of the emotion component was mediated by the mid-anterior cingulate and the dorsal TPJ. Thus, in the context of our task, when social information biases reward-based decision processes, it does so by engaging regions implicated in social processing, rather than regions involved in processing primary reward. It is important to point out, however, that the preferences induced by the faces may not be solely driven by social considerations. Further experiments could separate emotional and social valence, as well as perhaps considering other appetitive and aversive stimuli to ask whether they engage similar or different networks. Overall, this study has important implications for understanding how information from different sources is integrated in the brain for real-world decision processes.

Acknowledgments

We thank all the volunteers who participated in this study. This work was supported by a Wellcome Trust Programme Grant to R. J. D., MRC funding within the UCL 4-Year PhD in Neuroscience to S. M. F. This work was supported in part by the Intramural Program of the NIH, National Institute of Mental Health.

Reprint requests should be sent to Bruno B. Averbeck, Laboratory of Neuropsychology, NIMH/NIH, Building 49 Room 1B80, 49 Convent Drive MSC 4415, Bethesda, MD 20892-4415, or via e-mail: bruno.averbeck@nih.gov.

REFERENCES

- Abraham, A., Werning, M., Rakoczy, H., von Cramon, D. Y., & Schubotz, R. I. (2008). Minds, persons, and space: An fMRI investigation into the relational complexity of higher-order intentionality. *Consciousness and Cognition, 17*, 438–450.
- Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences, 4*, 267–278.
- Averbeck, B. B., & Duchaine, B. (2009). Integration of social and utilitarian factors in decision making. *Emotion, 9*, 599–608.
- Behrens, T. E., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. (2008). Associative learning of social value. *Nature, 456*, 245–249.
- Bischoff-Grethe, A., Hazeltine, E., Bergren, L., Ivry, R. B., & Grafton, S. T. (2009). The influence of feedback valence in associative learning. *Neuroimage, 44*, 243–251.
- 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*, 12315–12320.
- Chiu, P. H., Lohrenz, T. M., & Montague, P. R. (2008). Smokers' brains compute, but ignore, a fictive error signal in a sequential investment task. *Nature Neuroscience, 11*, 514–520.
- Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*, 201–215.
- Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience, 8*, 1704–1711.
- Dayan, P., Niv, Y., Seymour, B., & Daw, N. D. (2006). The misbehavior of value and the discipline of the will. *Neural Networks, 19*, 1153–1160.
- De Martino, B., Kumaran, D., Seymour, B., & Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science, 313*, 684–687.
- Deaner, R. O., Khera, A. V., & Platt, M. L. (2005). Monkeys pay per view: Adaptive valuation of social images by rhesus macaques. *Current Biology, 15*, 543–548.
- Elliott, R., Agnew, Z., & Deakin, J. F. (2010). Hedonic and informational functions of the human orbitofrontal cortex. *Cerebral Cortex, 20*, 198–204.
- Elliott, R., & Deakin, B. (2005). Role of the orbitofrontal cortex in reinforcement processing and inhibitory control: Evidence from functional magnetic resonance imaging studies in healthy human subjects. *International Review of Neurobiology, 65*, 89–116.
- Evans, S., Shergill, S. S., & Averbeck, B. B. (2010). Oxytocin decreases aversion to angry faces in an associative learning task. *Neuropsychopharmacology, 35*, 2502–2509.
- Frith, C. D., & Frith, U. (2006). The neural basis of mentalizing. *Neuron, 50*, 531–534.
- Gobbini, M. I., Koralek, A. C., Bryan, R. E., Montgomery, K. J., & Haxby, J. V. (2007). Two takes on the social brain: A comparison of theory of mind tasks. *Journal of Cognitive Neuroscience, 19*, 1803–1814.
- Grossman, M., Eslinger, P. J., Troiani, V., Andeson, C., Avants, B., Gee, J. C., et al. (2010). The role of ventral medial prefrontal cortex in social decisions: Converging evidence from fMRI and frontotemporal lobar degeneration. *Neuropsychologia, 48*, 3505–3512.
- Haber, S. N., Kim, K. S., Maily, P., & Calzavara, R. (2006). Reward-related cortical inputs define a large striatal region in primates that interface with associative cortical connections, providing a substrate for incentive-based learning. *Journal of Neuroscience, 26*, 8368–8376.
- Hampton, A. N., Bossaerts, P., & O'Doherty, J. P. (2008). Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proceedings of the National Academy of Sciences, U.S.A., 105*, 6741–6746.
- Hayden, B. Y., Parikh, P. C., Deaner, R. O., & Platt, M. L. (2007). Economic principles motivating social attention in humans. *Proceedings of the Royal Society of London, Series B, Biological Sciences, 274*, 1751–1756.
- Hurlemann, R., Patin, A., Onur, O. A., Cohen, M. X., Baumgartner, T., Metzler, S., et al. (2010). Oxytocin enhances amygdala-dependent, socially reinforced learning and emotional empathy in humans. *Journal of Neuroscience, 30*, 4999–5007.
- Hynes, C. A., Baird, A. A., & Grafton, S. T. (2006). Differential role of the orbital frontal lobe in emotional versus cognitive perspective-taking. *Neuropsychologia, 44*, 374–383.
- Lombardo, M. V., Chakrabarti, B., Bullmore, E. T., Wheelwright, S. J., Sadek, S. A., Suckling, J., et al. (2010). Shared neural circuits for mentalizing about the self and others. *Journal of Cognitive Neuroscience, 22*, 1623–1635.
- Maunsell, J. H. (2004). Neuronal representations of cognitive state: Reward or attention? *Trends in Cognitive Sciences, 8*, 261–265.

- McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science*, *306*, 503–507.
- Montague, P. R., & Berns, G. S. (2002). Neural economics and the biological substrates of valuation. *Neuron*, *36*, 265–284.
- Montague, P. R., King-Casas, B., & Cohen, J. D. (2006). Imaging valuation models in human choice. *Annual Review of Neuroscience*, *29*, 417–448.
- Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology*, *64*, 723–739.
- O'Doherty, J., Critchley, H., Deichmann, R., & Dolan, R. J. (2003). Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. *Journal of Neuroscience*, *23*, 7931–7939.
- 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., Winston, J., Critchley, H., Perrett, D., Burt, D. M., & Dolan, R. J. (2003). Beauty in a smile: The role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia*, *41*, 147–155.
- Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2004). The neural correlates of theory of mind within interpersonal interactions. *Neuroimage*, *22*, 1694–1703.
- Rushworth, M. F., Buckley, M. J., Behrens, T. E., Walton, M. E., & Bannerman, D. M. (2007). Functional organization of the medial frontal cortex. *Current Opinion in Neurobiology*, *17*, 220–227.
- Scharlemann, J. P. W., Eckel, C. C., Kacelnik, A., & Wilson, R. K. (2001). The value of a smile: Game theory with a human face. *Journal of Economic Psychology*, *22*, 617–640.
- Schonberg, T., Daw, N. D., Joel, D., & O'Doherty, J. P. (2007). Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. *Journal of Neuroscience*, *27*, 12860–12867.
- 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*, 2490–2495.
- Tomlin, D., Kayali, M. A., King-Casas, B., Anen, C., Camerer, C. F., Quartz, S. R., et al. (2006). Agent-specific responses in the cingulate cortex during economic exchanges. *Science*, *312*, 1047–1050.
- van den Bos, W., McClure, S. M., Harris, L. T., Fiske, S. T., & Cohen, J. D. (2007). Dissociating affective evaluation and social cognitive processes in the ventral medial prefrontal cortex. *Cognitive Affective & Behavioral Neuroscience*, *7*, 337–346.
- Van Overwalle, F. (2009). Social cognition and the brain: A meta-analysis. *Human Brain Mapping*, *30*, 829–858.
- Van Overwalle, F., & Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: A meta-analysis. *Neuroimage*, *48*, 564–584.
- Walter, H., Abler, B., Ciaramidaro, A., & Erk, S. (2005). Motivating forces of human actions. Neuroimaging reward and social interaction. *Brain Research Bulletin*, *67*, 368–381.
- Weiskopf, N., Hutton, C., Josephs, O., & Deichmann, R. (2006). Optimal EPI parameters for reduction of susceptibility-induced BOLD sensitivity losses: A whole-brain analysis at 3 T and 1.5 T. *Neuroimage*, *33*, 493–504.

This article has been cited by:

1. Atbin Djamshidian, Sean S. O'Sullivan, Bianca C. Wittmann, Andrew J. Lees, Bruno B. Averbeck. 2011. Novelty seeking behaviour in Parkinson's disease. *Neuropsychologia* . [[CrossRef](#)]