

Dissociating affective evaluation and social cognitive processes in the ventral medial prefrontal cortex

WOUTER VAN DEN BOS

*University of Amsterdam, Amsterdam, The Netherlands
and Princeton University, Princeton, New Jersey*

SAMUEL M. MCCLURE, LASANA T. HARRIS, AND SUSAN T. FISKE

Princeton University, Princeton, New Jersey

AND

JONATHAN D. COHEN

*Princeton University, Princeton, New Jersey
and University of Pittsburgh, Pittsburgh, Pennsylvania*

In recent studies, various regions of the ventral medial prefrontal cortex (vmPFC) have been implicated in at least two potentially different mental functions: reasoning about the minds of other people (social cognition) and processing reward related information (affective evaluation). In this study, we test whether the activation in a specific area of the vmPFC, the para-anterior cingulate cortex (PACC), correlates with the reward value of stimuli in general or is specifically associated with social cognition. Participants performed a time estimation task with trial-to-trial feedback in which reward and social context were manipulated separately. Reward was manipulated by giving either positive or negative feedback in the form of small squirts of fluid delivered orally. Social context was manipulated by instructing participants that positive and negative feedback was determined by another person or a computer. The data demonstrate a main effect of feedback, but not social context, in the PACC, suggesting that this area of the vmPFC serves a general function in evaluating and/or representing reward value. In addition, activity in a more anterior subregion of the vmPFC demonstrated reward-related sensitivity only in the social context. Another area that showed a similar interaction was the subgenual cingulate, but this region was only sensitive to negative feedback in the social condition. These findings suggest that, within the vmPFC, the PACC subserves primarily an affective function, whereas in other regions social context can modulate affective responses.

Decision making under uncertainty is central to human mental function and involves a complex process of assigning values to different options, followed by choosing the option with the highest expected value. There is a large and growing literature on the use of neuroimaging methods to identify the neural mechanisms underlying the evaluation and representation of reward information (see Montague, King-Casas, & Cohen, 2006, for a review). These studies have consistently implicated specific regions of the ventromedial prefrontal cortex (vmPFC) in reward processing. In human social interaction, the representation of reward values associated with specific outcomes is often dependent on the behavior of other individuals. Therefore, in such circumstances, it is necessary to make correct inferences about the minds of others. The ability to make inferences about mental states, such as goals and beliefs, is generally referred to as "theory of mind" (ToM; Premack & Wood-

ruff, 1978). Neuroimaging studies involving social interaction have also identified regions of the vmPFC as playing a central role in ToM processes. Some of these regions are identical to those identified in reward processing in studies that do not involve social interaction. This raises the question of the extent to which regions of the vmPFC are specifically dedicated to reward processing (affective processing) versus ToM (social cognition)—or to both.

To date, most studies have focused on either reward representation (e.g., Breiter, Aharon, Kahneman, Dale, & Shizgal, 2001; Knutson, Adams, Fong, & Hommer, 2001; O'Doherty et al., 2003) or social cognition (e.g., Castelli, Happé, Frith, & Frith, 2005; Gallagher et al., 2000; Gallagher, Jack, Roepstorff, & Frith, 2002) but have not explicitly attempted to dissociate these two processes. A review of the literature suggests that reward representation (affective/evaluative processing) and ToM (social

cognitive processing) may rely on separate, nonoverlapping networks. However, there is one area that is commonly reported to be part of both networks: the vmPFC. The vmPFC consists roughly of Brodmann areas 10 and 32, rostral to the anterior cingulate cortex and between the medial orbitofrontal cortex (area 14) and the more posterior areas of the medial frontal cortex including areas 8, 9, and 24 (Amodio & Frith, 2006). We consider this to be a geographical pointer to an area that consists of distinct anatomical and functional areas. For the purposes of our study, we are mainly interested in part of this larger area: the para-anterior cingulate cortex (PACC). In this study, we sought to test whether the PACC was involved in the processing of reward value, social cognition, or both.

Prior neuroimaging studies of reward valuation have manipulated different dimensions of rewards, such as magnitude (Knutson et al., 2001), predictability (Breiter et al., 2001), or delay (McClure, Laibson, Loewenstein, & Cohen, 2004). In these studies, a network of brain regions has been identified which is sensitive to the different reward manipulations, including the ventral striatum (vStr), orbitofrontal cortex (OFC), amygdala, and the vmPFC (Anderson et al., 2003; Baxter & Murray, 2002; Montague et al., 2006; Walter, Abler, Ciaramidaro, & Erk, 2005). In these studies, vmPFC activity scales with reward value (Blair et al., 2006; Knutson et al., 2001; O'Doherty et al., 2003). The region within vmPFC that correlates most consistently with reward processing is the PACC (see Knutson et al., 2001; McClure, Laibson, et al., 2004; Nieuwenhuis, Slagter, Alting von Geusau, Heslenfeld, & Holroyd, 2005; O'Doherty, Kringelbach, Rolls, Hornak, & Andrews, 2001; O'Doherty et al., 2003).

In a separate line of research, activity in the vmPFC has been associated with a social cognitive function referred to as "mentalizing." The verb "mentalize" is used to describe the activity of thinking about the content of other minds, or having a ToM (Premack & Woodruff, 1978). These studies typically consist of a social condition and a nonsocial control condition. The common feature of the social conditions is that participants have to think about the mental states of others. In some studies, the task involved playing cooperative or competitive games, with either another person or a computer (Gallagher et al., 2002; McCabe, Houser, Ryan, Smith, & Trouard, 2001; Rilling, Sanfey, Aronson, Nystrom, & Cohen, 2004; Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003). In other studies, participants were asked to read stories and predict the actions of people or the outcomes of physical chains of events (Fletcher et al., 1995), or to watch animations of lifelike or random moving dots on a screen. The results of these studies consistently report activation in the vmPFC associated with social cognition (for a review, see Amodio & Frith, 2006; Castelli et al., 2005; C. D. Frith & Frith, 2006; Lieberman, 2007; Saxe, Carey, & Kanwisher, 2004), as part of a larger network of structures, including the posterior cingulate cortex (PCC), temporoparietal junction (TPJ), posterior superior temporal sulcus (pSTS), and the temporal poles (U. Frith & Frith, 2003). These results have prompted suggestions that the vmPFC plays

a special dedicated role in social cognition (Amodio & Frith, 2006; U. Frith & Frith, 2003; Saxe et al., 2004), whereas the other areas of the network have more domain-general functions. However, this domain-specific view of the vmPFC is not shared by all investigators (e.g., Apperly, Samson, & Humphreys, 2005; Stone & Gerrans, 2006).

One explanation for the association of the vmPFC with both affective evaluation and social cognitive functions may be that social stimuli either elicit or modulate affective evaluative processes. Activity within the vmPFC has been observed for both primary reinforcers, such as food or juice, as well as for secondary reinforcers, such as money (Breiter et al., 2001; Knutson et al., 2001; Knutson, Fong, Bennett, Adams, & Hommer, 2003; O'Doherty et al., 2001) and cultural rewards such as music, cars, and exclusive brands (Blood, Zatorre, Bermudez, & Evans, 1999; Erk, Spitzer, Wunderlich, Galley, & Walter, 2002; McClure, Li, et al., 2004). Interestingly, social rewards such as cooperation (Rilling et al., 2004), love (Bartels & Zeki, 2004), and trust (King-Casas et al., 2005) have also been shown to produce activity within similar regions of the vmPFC, including the PACC. Thus, activity within at least some regions of the vmPFC may reflect affective evaluation—due either to intrinsic value of social stimuli or modulation of value by social context—rather than processes specific to social cognition, whereas others may be more specialized for social information processing.

In this study, healthy adults were asked to participate in a time estimation task in which we independently manipulated reward and social context. Neural activity was measured using fMRI following trial-to-trial feedback. We then performed both exploratory and region-of-interest (ROI) analyses to identify areas of activity associated with the reward and social manipulations and their interaction. The ROI analysis was motivated by prior studies that have implicated a subregion of the vmPFC, the PACC, in both reward processing and social cognition. We hypothesized that activation within this region is associated with domain-general evaluative processes that are sensitive but not specific to the reward value of social stimuli.

Reward was manipulated using feedback that consisted of basic reinforcement signals, positive or negative, in the form of small squirts of desirable or aversive fluid delivered orally (see also McClure, Berns, & Montague, 2003). Social context was manipulated by instructing participants that positive and negative feedback were determined either by a computer or by another person (see also Gallagher et al., 2002; McCabe et al., 2001; Rilling et al., 2004; Sanfey et al., 2003). The fact that feedback came from another person encouraged the participant to think about the communicative intent behind the feedback signals in the social condition. This design allows us to directly examine the main effects of reward and social cognition and also the interaction between these two processes. We believe this is an ecologically more valid situation, because most of our decisions in real life, including economic decisions, are made in a social context. In this experiment we used an important, but not very well studied, component of social interaction: being judged by others.

METHOD

Participants

Nineteen Princeton undergraduate and graduate students—12 men and 7 women, age range, 20.7 years ($SD = 2.5$)—participated in the study. All participants were right-handed native English speakers with no history of neurological problems. Informed consent was obtained using a consent form approved by the Institutional Review Panel at Princeton University.

Experimental Design

The experimental task consisted of a time-estimation task, which involved a decision about the time elapsed following a stimulus. The target time duration was specified before each trial (Figure 1). Participants were instructed to estimate durations that varied between 6.5 and 13 sec, and were instructed to improve their accuracy on the basis of positive and negative feedback. Each trial consisted of the following sequence of events. First, an instruction screen indicating the time to be estimated was presented. Participants pressed a button to start the trial and made a second buttonpress when they believed the instructed time had elapsed. The second buttonpress was followed by a variable 6- to 8-sec delay, followed by a 2-sec feedback period and then a 12-sec fixation intertrial interval.

Feedback consisted of reinforcing signals, positive or negative, in the form of a 2-sec visual cue and a squirt of liquid. Positive feedback consisted of the message “Good Job” coupled with the delivery of 0.5 ml of sweet fruit juice. Negative feedback consisted of the message “Wrong” coupled with the delivery of 0.5 ml of a bitter quinine hydrochloride solution (0.1 mM). The liquids were administered via two different tubes held in place by a mouthpiece. In the case of negative feedback, participants were not provided with information about the size or direction of the error. To enhance the value of the primary fluid rewards, participants were instructed not to drink any liquid three hours prior to the study.

In order to manipulate social context, participants were told that they would receive feedback from either a computer program or another participant, who was referred to as “the judge.” The human-versus-computer comparison had been successfully used in prior

studies (Gallagher et al., 2002; McCabe et al., 2001; Rilling et al., 2004; Sanfey et al., 2003). Participants were introduced to the judge (a confederate) prior to the start of the experiment. The instruction to the participant was that the feedback given by the computer was based on a threshold rule, whereas the judge was free to determine the type of feedback. To increase the subjective experience of social interaction, the participant was told that after completion of the experiment, both members of the best performing team of participant and judge would be awarded \$20. This procedure was intended to give participants the impression that there were clear incentives for the judge to assist the participant. To equalize extrinsic motivational significance in the social and the computer condition, the best behavior in the computer condition was also awarded \$20.

In reality, feedback was randomly determined for each trial in both the judge and computer conditions, resulting in an even number of trials in both conditions with either positive or negative feedback. The durations to be estimated (between 6.5 and 13 sec) varied in length so that the participants could not easily judge the correctness of their performance. The feedback was presented randomly, unless the interval between buttonpresses was more than 3 sec from the target duration, in which case negative feedback was always given.

Procedure

The four conditions (social positive, social negative, computer positive, computer negative) were presented in separate blocks of “social” and “computer” instructions, and positive and negative feedback was randomly presented in each block. In total, there were 4 social blocks, each consisting of 10 trials (5 positive and 5 negative) and 4 computer blocks, each consisting of 10 trials (5 positive and 5 negative). This resulted in 20 trials per condition. Each block lasted approximately 6 min. The presentation of social and computer blocks was counterbalanced across participants to prevent order effects. The task took 40–50 min, depending on the pace of the participant. After the experiment, the participants were questioned to confirm that the two conditions were experienced differently; subsequently, participants were informed about the actual setup of the experiment. All participants reported that they experienced the two conditions differently and that they were unaware that the feedback was presented randomly.

fMRI Data Acquisition

Imaging was performed on a 3-tesla head dedicated Siemens Allegra scanner at Princeton University. A high-resolution ($0.5 \times 0.5 \times 1.0$ mm) T1-weighted anatomical image was first acquired to enable localization of functional images. Whole-brain functional images were acquired in 30 axial slices (64×64 voxels; in-plane resolution 3×3 mm; 3-mm slices with 1-mm slice gap) approximately aligned to the AC–PC line, with a repetition time of 2 sec, an echo time of 30 msec, and a flip angle of 90° .

Using E-Prime software and a Dell computer, stimuli were projected on a screen mounted at the rear of the scanner bore, which participants viewed through a series of mirrors. A pillow and foam cushions were placed inside the head coil to minimize head movements.

fMRI Data Analysis

Data preprocessing and statistical analysis were performed using AFNI (Cox, 1996). Preprocessing consisted of slice time correction using the Fourier method, rigid body 3-D motion correction using Fourier interpolation, and spatial smoothing with a 6-mm full width at half maximum (FWHM) Gaussian kernel. The functional data were then aligned with the anatomical data, as follows: The functional data were first rotated to the same angle as the anatomical data, then shifted, using a 3-D whole-body transformation to minimize the difference from the anatomical image. Spatial normalization was performed using AFNI to transform the functional data to Talairach space (Talairach & Tournoux, 1988).

Statistical analyses were performed using a general linear model (GLM), in which individual events were modeled with a canonical

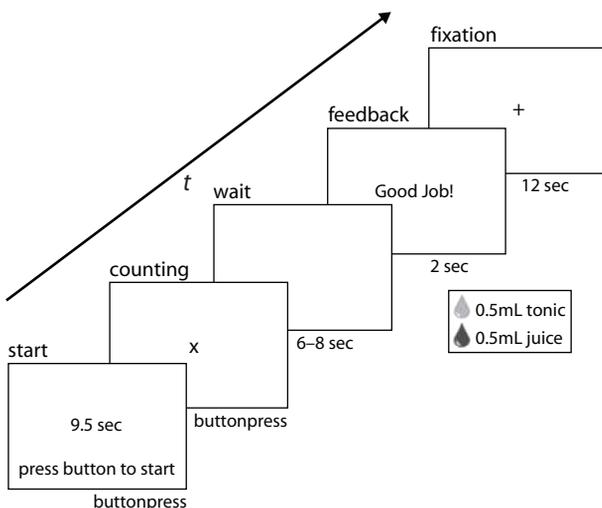


Figure 1. Each trial started with a display presenting the duration to be estimated. When the participant pressed the button, this screen was followed by a screen with an “x” in the center to indicate that the estimation had started. The second time a button was pressed this screen disappeared, and a blank screen followed for 6–8 sec, representing the judging period. Next, the type of feedback was displayed and liquid was administered. The trial was always followed by a fixation cross for 12 sec.

Table 1
Absolute Adjustment Effects (abs[RT(*n*) – RT(*n*–1)])
on the Subsequent Trial, in Milliseconds,
for Each of the Four Conditions

Context	Feedback	<i>M</i>	<i>SE</i>
Computer	Positive	639	27
	Negative	672	40
Social	Positive	688	40
	Negative	795	32

hemodynamic response function. The GLM consisted of 13 regressors: two types of feedback in the social and computer conditions (4 regressors), starting and stopping of the trial (2 regressors), counting (1 regressor modeling the response during the estimation interval) and 6 regressors to factor out effects due to minor head movements (as estimated from motion correction of functional brain images).

This analysis was individually performed on each participant's data set. The contrast between parameter estimates for different within-participants events was submitted to an ANOVA, which treated the variability between participants as a random effect. We tested directly for interactions at the voxelwise level via an ANOVA *F* test. Statistical parametric maps were derived from the resulting *t* value associated with each voxel, and results were reported for brain areas composed of 10 or more contiguous voxels (120 mm²), each significant at an uncorrected *p* < .001. This combination of *p* value and extent thresholding reduces the effective per voxels false positive rate to *p* < .05 (corrected; Forman et al., 1995; estimated using AFNI programs 3dFWHM and AlphaSim). The locations of peak activity of the reported clusters are in Talairach coordinates.

ROIs

Mean event-related time series were calculated for each identified brain region. BOLD data from each brain voxel were converted to percent signal change and then were averaged over all voxels in the relevant region. Since events could occur at any point relative to the onset of scans, event-related signals were calculated by linear interpolation, beginning 2 sec prior to the choice and continuing at 2-sec intervals for 12 sec, using AFNI's tent function.

RESULTS

Behavioral Data

We calculated the absolute RT difference between each trial *n* and subsequent trial *n* + 1, then we analyzed that difference as a function of whether participants had received correct or error feedback on trial *n*. Replicating previous findings (e.g., van Veen, Holroyd, Cohen, Stenger, & Carter, 2004), participants changed their time estimates more following error feedback than following correct feedback [$F(1,246) = 234.3, p < .001$] (Table 1). These results strongly suggest that participants performed the task as expected; they used the feedback to guide and adjust their behavior. In addition, the adjustments were also larger in the social context than in the nonsocial context [$F(1,246) = 273.94, p < .001$]. Although social feedback was associated with a greater adjustment effect, it did not result in higher accuracy. A comparison of RTs revealed that there were no significant differences in the accuracy of estimation times for the two conditions [$t(1,759) = -1.16, p = .24$].

Whole-Brain Comparisons

As expected, the positive > negative feedback contrast revealed activation in bilateral ventral striatum, including nucleus accumbens, ventral caudate, and putamen (see Table 2 and Figure 2). In contrast, the negative > positive contrast revealed activation in a number of regions, including the rostral cingulate zone (RCZ), the right anterior insula, cerebellum, and subgenual anterior cingulate cortex (sgACC; Figure 3). Next, we tested for main effects associated with the social-versus-computer condition. The social > computer comparison revealed activation in the bilateral temporal poles, an area often implicated in social cognition. When the significance threshold for this contrast was lowered to *p* < .005, activation was also

Table 2
Peak Activations of Whole-Brain Analysis Reported for Each Region
in Talairach Coordinates

Anatomical Label	Right/Left	<i>x</i>	<i>y</i>	<i>z</i>	Max <i>t</i> Value	Max <i>F</i> Value
Positive Feedback > Negative Feedback						
Ventral striatum						
Ventral caudate	R	8	8	0	6.266	
	L	-10	8	-1	4.791	
Nucleus accumbens	R	13	8	-10	4.362	
Putamen	L	-17	9	-7	4.511	
Negative Feedback > Positive Feedback						
Subgenual cingulate	R	2	19	-7	4.271	
Rostral cingulate zone	R	2	16	32	4.367	
Anterior insula	R	38	3	10	5.017	
Cerebellum	R	1	-48	-19	4.455	
Social > Computer						
Temporal poles	R	55	-2	-24	4.562	
	L	-54	8	-24	4.482	
Interaction (Feedback × Condition)						
Anterior vmPFC	L	-0	55	29		6.665
Subgenual cingulate	R	2	19	-7		7.760

Note—All regions are *p* < .001 uncorrected. Minimal voxel contiguity is 120 mm² (10 voxels).

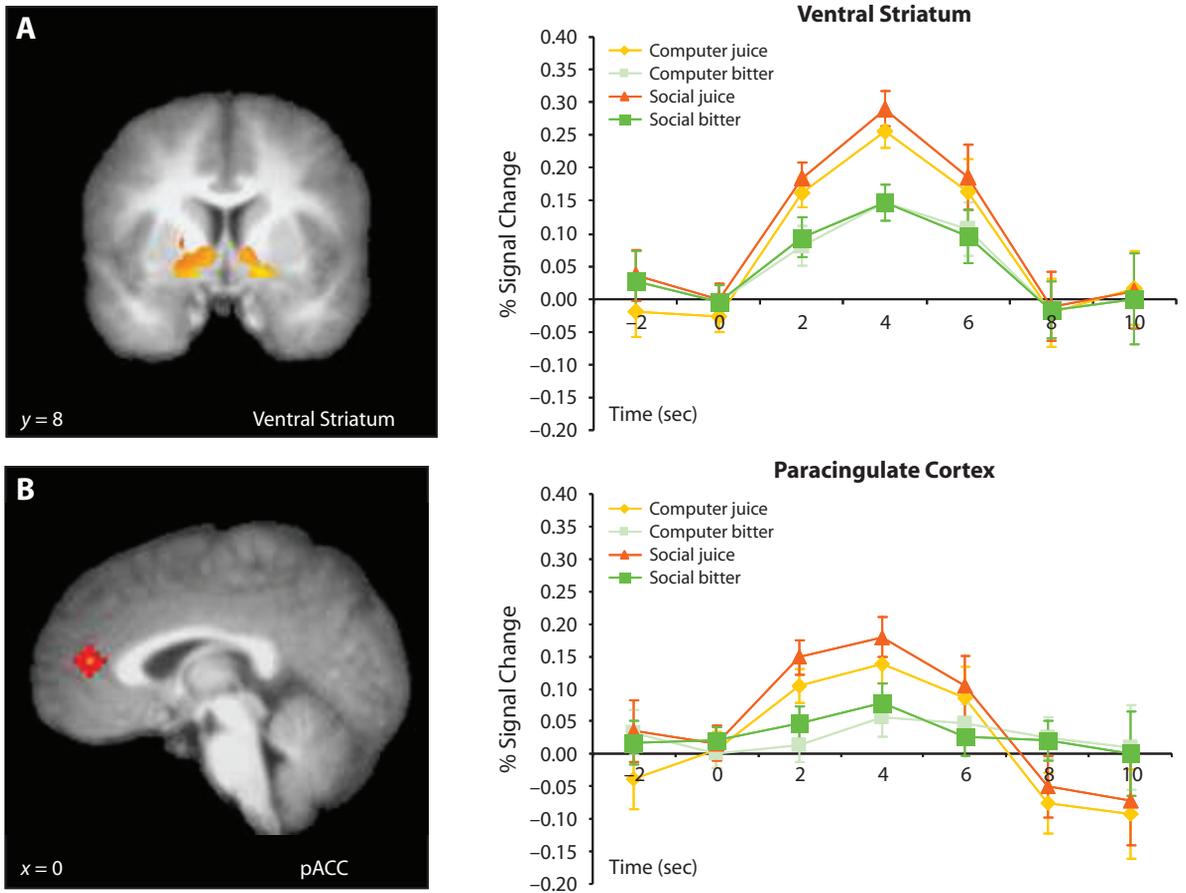


Figure 2. Brain regions that are preferentially activated for positive feedback versus negative feedback in both contexts. (A) A random effects general linear model analysis revealed several areas of the ventral striatum which are significantly more activated by positive feedback (at $p < .001$, uncorrected; 10 contiguous voxels). (B) ROI analysis showed that the PACC (Talairach coordinates of central voxel: $x = 0, y = 44, z = 12$) was more active following positive than following negative feedback [$F(1,18) = 7.99, p < .01$]. The graphs represent mean event-related time courses for each region, starting at 2 sec before feedback is received. Error bars represent standard errors of the mean; $N = 19$ participants.

observed in the pSTS (Figure 4). The computer > social comparison did not result in any significant activation.

We next tested for the interaction between the reward and social context manipulations to determine whether the difference in neural activation following positive versus negative feedback was different for the social and computer conditions. This analysis revealed two foci within the vmPFC (Figure 5). The first was located in the anteriormost part of the vmPFC (Talairach coordinates $x = 0, y = 55, z = 29$), and exhibited greater activity to positive than to negative feedback, but in the social, not the computer, condition. This area is one that has commonly been associated with the processing of social information (Amodio & Frith, 2006; U. Frith & Frith, 2003). The second region showing a significant effect of the interaction between the reward and social manipulations was within the sgACC. Like the anterior region of the vmPFC, the area within the sgACC showed a differential response to feedback in the social but not in the computer condition; however, it showed the reverse response to feedback, with greater activity for negative than for positive feedback

in the social condition. No brain areas were found that showed greater responses to rewards (vs. punishments) in the computer (vs. social) condition. Contrary to expectations, the whole-brain analysis did not reveal significant effects within the PACC region of the vmPFC.

ROI for PACC

A planned ROI analysis was performed on a sphere of voxels within the PACC to test our a priori hypotheses about the function of this region (Figure 2). The placement of the ROI was based on previous findings of overlapping areas of activity within the vmPFC, reported in a number of reward-processing studies (Knutson et al., 2001; McClure, Laibson, et al., 2004; Nieuwenhuis et al., 2005; O’Doherty et al., 2001; O’Doherty et al., 2003; Stone & Gerrans, 2006) and social cognition (Gallagher et al., 2000; McCabe et al., 2001; Mitchell, Banaji, & Macrae, 2005a; Rilling et al., 2004; Sanfey et al., 2003). The central voxel of the spherical region (consisting of 33 voxels) was located at Talairach coordinates ($x = 0, y = 44, z = 12$). Consistent with our predictions, the ROI analysis

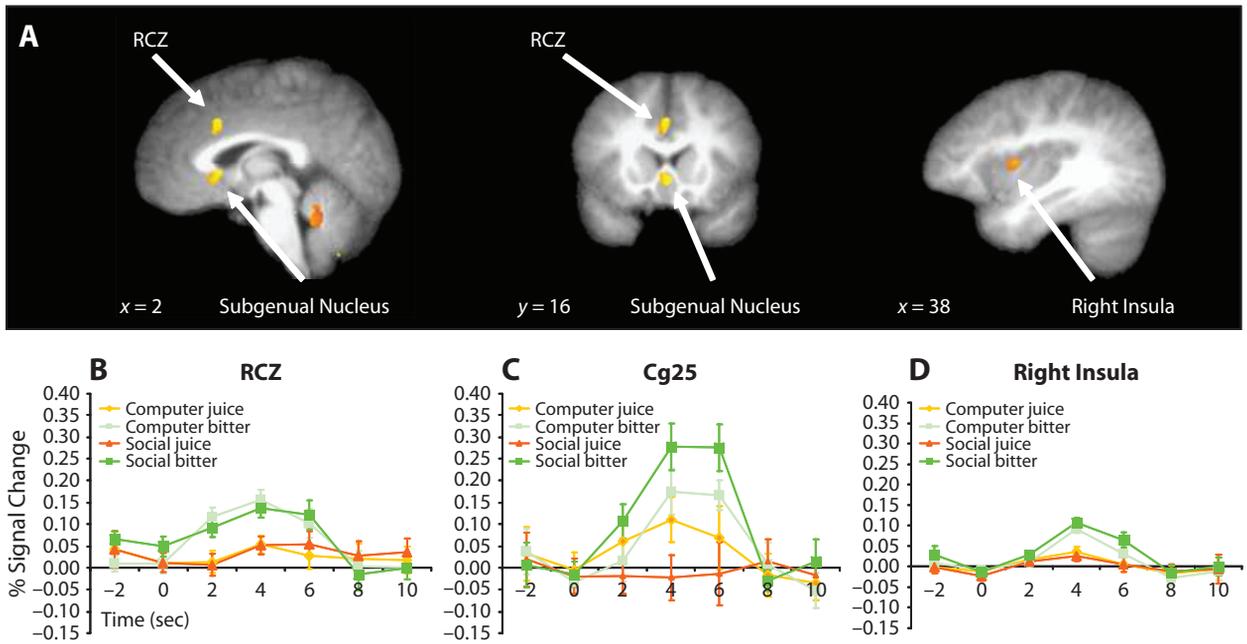


Figure 3. Brain regions that are preferentially activated for negative feedback versus positive feedback in both contexts. (A) A random effects general linear model analysis revealed several areas which are significantly more activated by negative feedback (at $p < .001$, uncorrected; 10 contiguous voxels), including rostral cingulate zone, subgenual nucleus, and the insula. The graphs represent mean event-related time courses for each region, starting at 2 sec before feedback is received. (B) Rostral cingulate zone. (C) Subgenual nucleus. (D) The insula. Error bars represent standard errors of the mean; $N = 19$ participants.

showed that the PACC was more active following positive than following negative feedback [$F(1,18) = 7.99$, $p < .01$]. This effect was observed in both the computer condition [$F(1,18) = 6.74$, $p < .02$] and the social condition [$F(1,18) = 5.98$, $p < .05$], and did not differ between the latter two (interaction of condition \times feedback) [$F(1,18) < 1$, $p = .91$].

DISCUSSION

The goal of this study was to test the sensitivity of brain regions previously implicated in social and affective neuroscience studies in response to performance feedback and social context. Two brain networks were sensitive only to the valence of feedback, whether the feedback was delivered in the social or the computer condition. The first network involved the PACC and the striatum; these regions were more active following positive than negative feedback. The second network included the rostral cingulate zone and the insula; these regions were more active following negative than positive feedback. Two other regions were only responsive to feedback valence in the social condition, but not in the computer condition. These regions included the anterior vmPFC, which was more active following positive than negative feedback in the social condition, and sgACC, which was more active following negative than positive feedback in the social condition. Thus, we were able to distinguish between brain regions that were sensitive to general reward/punishment information, regardless of source and brain regions modulated by social context. In addition to the patterns of brain activation, the behavioral

data independently show that the adjustment effects are sensitive to feedback type and type of context.

The Role of PACC in Affective and Social Cognition

As expected, the PACC of the vmPFC was sensitive to positive feedback, consistent with prior studies implicating this region in reward processing (Knutson et al., 2001; McClure, Laibson, et al., 2004; Nieuwenhuis et al., 2005; O'Doherty et al., 2003; Walter et al., 2005; Walter et al., 2004). However, contrary to prior reports, this effect was noted regardless of whether the positive feedback was presented in the social or the computer condition. In prior studies, PACC had been implicated in social processes such as mentalizing and ToM (Castelli et al., 2005; Gallagher et al., 2002; Rilling et al., 2004; Sanfey et al., 2003). Because the activity in this region, as opposed to other areas in the mPFC, was not modulated by social context, our findings suggest that this effect may reflect the intrinsically rewarding value of social stimuli, and the operation of a domain-general reward-processing mechanism.

A possible explanation for vmPFC activity reported in other studies on social cognition is that social stimuli have an intrinsic rewarding property not inherent in the control stimuli. The studies that have reported anterior vmPFC activation typically involved thinking about friends, like-minded people, or self-referential/self-relevant topics (Mitchell, Banaji, & Macrae, 2005b; Mitchell, Macrae, & Banaji, 2006). The studies share stimuli with positive dimensions; for instance, it is known that individuals are more likely to endorse positive information as

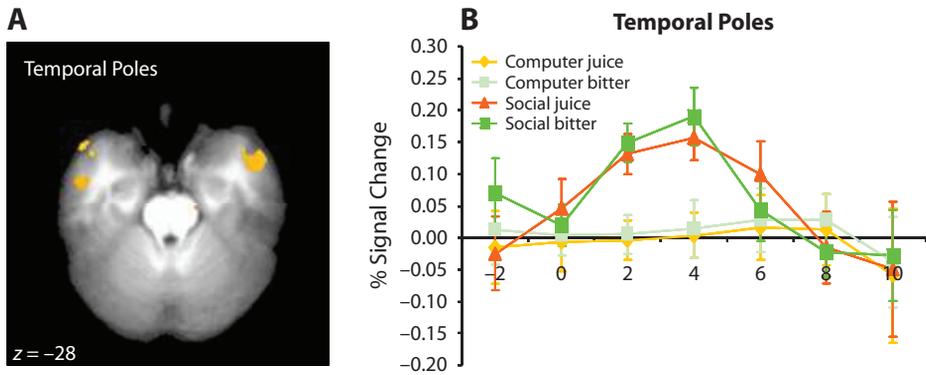


Figure 4. Brain regions that are preferentially activated for social versus computer context. (A) A random effects general linear model analysis revealed the bilateral temporal poles which are significantly more activated by positive feedback (at $p < .001$, uncorrected; 10 contiguous voxels). (B) The graphs represent mean event-related time courses for each region, starting at 2 sec before feedback is received. Error bars represent standard errors of the mean; $N = 19$ participants.

self-descriptive (Taylor & Brown, 1988). In addition, the activity in the vmPFC reported by Mitchell et al. (2006) associated with thinking about like-minded people also correlated with how positive the ratings given to these people were.

It may not always be possible to explain increased activation in the vmPFC in terms of the inherent positive value of social stimuli. For instance, studies in which there are two social conditions, neither of which is *prima facie* more positive than the other, have also reported increased activation in the vmPFC for social stimuli (Grèzes, Frith, & Passingham, 2004; Walter et al., 2004). However, whereas these conditions may appear to be primarily social, it is possible that the involvement of the vmPFC can still be explained in terms of affective processing. The social conditions in these studies typically involve thinking about goals or motivations versus beliefs or knowledge, a distinction similar to the affective versus cognitive dissociation previously suggested by Völlm et al. (2006), in which study it was suggested that making inferences about the affective states of others (e.g., feelings, desires, and motivations) and making inferences about more dispassionate cognitive states (e.g., beliefs and knowledge) are represented in different areas of the brain. In particular, they suggest a functional dissociation in the mPFC, where the vmPFC is engaged in the evaluation and representation of affective valence and the dorsal mPFC is involved in more abstract and/or deliberative forms of mentalizing (see also Saxe et al., 2004).

The activation of part of the valuation network in these kinds of studies is consistent with inferences that can be drawn based on simulation theory (Gallese & Goldman, 1998; Gordon, 1992). To predict the future behavior of others, it is necessary to infer what motivates them (e.g., “he is thirsty”) and what knowledge/beliefs they possess (“there is a café around the corner”). Next, assuming that the other person is not too different from oneself, a prediction is made (“he will quickly go to the café”), on the basis of self-knowledge. The last assumption is captured by “simulation” theories of social cognition; people infer

the mental states of others by “simulating” their own experience as if they were in the situation of the other, on the assumption that others would experience the same thing (Gallese & Goldman, 1998; Gordon, 1992). Positive reinforcements can be seen as very basic motivational forces that guide our behavior. When we think about the future actions of others, we think about what motivates them by simulation, thus activating our own valuation network, of which the PACC is part. Importantly, these explanations of activity in the PACC in studies on social cognition are not mutually exclusive. The activity in the PACC can in some experiments be caused by the difference in intrinsic value of the stimuli and in other experiments by the “simulation” of motivational states of others.

Context-Dependent Sensitivity to Positive and Negative Feedback

Consistent with prior studies, positive performance feedback resulted in activation in the PACC and the striatum, whereas negative performance feedback resulted in activation in the RCZ and the insula. These findings support the possibility of a positive–negative dissociation between ventral (PACC) and dorsal/posterior (RCZ) parts of the medial PFC (Blair et al., 2006). The pattern of increased activity associated with positive relative to negative feedback in the striatum is in accord with previous evidence that the ventral striatum is involved in reward expectation and delivery in humans (Breiter et al., 2001; Delgado, Nystrom, Fissell, Noll, & Fiez, 2000; Knutson et al., 2003; McClure et al., 2003; O’Doherty et al., 2003; Pagnoni, Zink, Montague, & Berns, 2002; Ramnani, Elliott, Athwal, & Passingham, 2004), and nonhuman species (Schultz, Dayan, & Montague, 1997). The time series of the ventral striatum reveal that there is also an increase in activity following negative feedback (Figure 2). Previous studies have also found an increased BOLD signal to aversive stimuli in the ventral striatum (Jensen et al., 2003; Seymour et al., 2004).

Activation associated with negative feedback in the RCZ and in the right anterior insula is consistent with

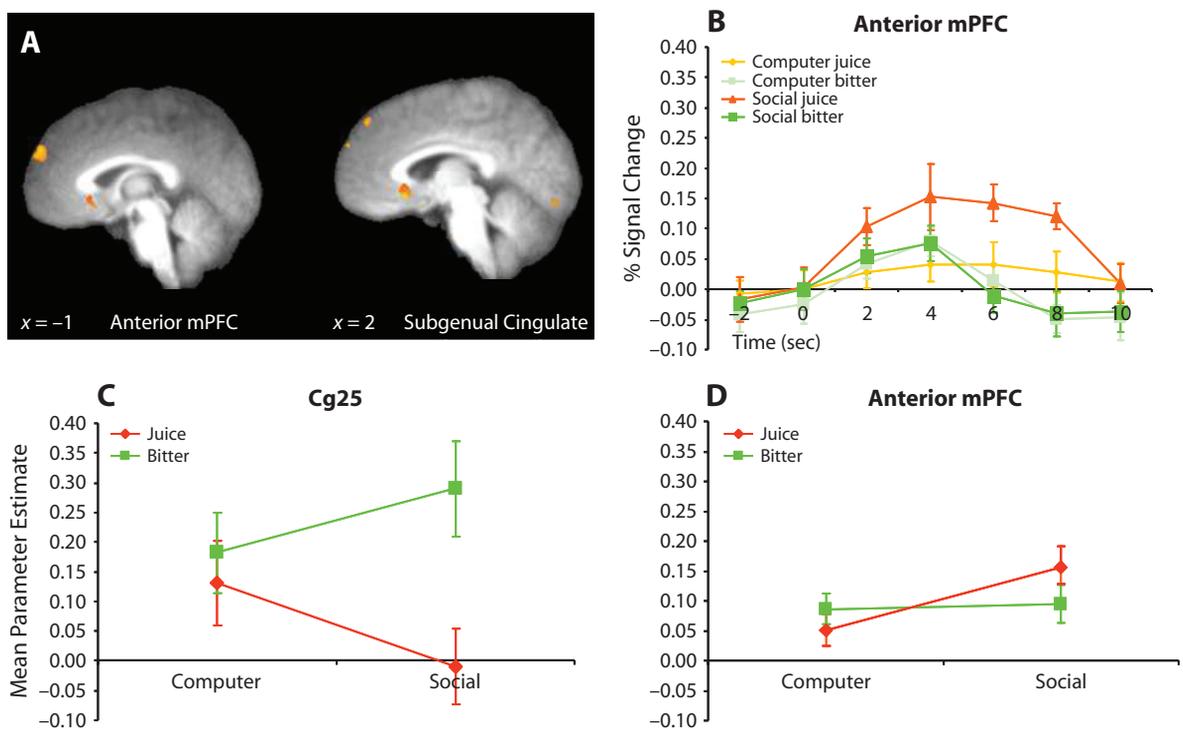


Figure 5. (A) Brain regions that are differentially activated for the positive versus negative feedback for the computer relative to the social condition, including the subgenual nucleus (Cg25) and the anterior vmPFC (at $p < .001$, uncorrected; 10 contiguous voxels). (B) The graphs represent mean event-related time courses for the anterior vmPFC, starting at 2 sec before feedback is received. (C) Neural responses in the Cg25 were associated with more activation for negative than for positive feedback in the social condition, whereas this difference was absent in the computer condition. (D) The anterior vmPFC region was more sensitive to positive than negative performance feedback, only in the social condition and not in the computer condition. Error bars represent standard errors of the mean; $N = 19$ participants.

prior studies suggesting that these regions are involved in processing negative feedback (Holroyd et al., 2004; Mars et al., 2005; Ullsperger & Von Cramon, 2004; for a review, see Ridderinkhof, Ullsperger, Croen, & Nieuwenhuis 2004) and experiencing disgust (e.g., Schäfer, Schienle, & Vaitl, 2005; Wicker et al., 2003). More generally, these findings are consistent with the hypothesis that these regions code for negative utility (McClure, Botvinick, Yeung, Greene, & Cohen, 2007). However, previous studies have reported negative effects in the RCZ using a different timing estimation task; this was interpreted as problematic for certain error-processing theories of anterior cingulate function (Nieuwenhuis et al., 2005; van Veen et al., 2004). The design of this study is similar in the sense that the outcome of the feedback is unpredictable; therefore, activity in the RCZ correlated with negative feedback is most likely not due to expectancy violation. These results are consistent with previous theories that the RCZ is more generally involved in conflict monitoring, of which error detection is a special case (Yeung, Cohen, & Botvinick, 2004).

Although the PACC was not specifically sensitive to social context, two other regions demonstrated a specific sensitivity to social manipulations. First, an anterior sub-region of the vmPFC showed greater neural activity for positive feedback than for negative feedback in the social

context, but not in the computer context. This finding suggests that this area of the vmPFC is an affective region specifically sensitive to social cues. The result is in accord with a study by Harris, McClure, van den Bos, Fiske, and Cohen (2007) in which a very similar area shows greater activity for positive than for negative stimuli in the social domain, but not in the nonsocial object domain. It is possible that the more complex nature of positive feedback in the social context is represented in more anterior parts of the vmPFC (Adolphs, 2003).

Second, the sgACC showed sensitivity to negative relative to positive feedback in the social context, but not in the computer context. These findings indicate that when negative feedback comes from another person rather than from a computer, the response of this area is amplified. This pattern of activity is in accord with the report that hyperactivity in this area is strongly implicated in major depression, and that reducing this activity by deep brain stimulation can ameliorate the symptoms of depression (Mayberg et al., 2005). Although these studies suggest that the sgACC is sensitive to socially mediated negative information, others have reported *increased* activity for positive and *decreased* activity in the sgACC for self-referential terms (Moran, Macrae, Heatherton, Wyland, & Kelley, 2006). This difference in findings may reflect functional specialization within subregions of the sgACC

as far as the processing of negatively and positively valenced stimuli (Gotlib et al., 2005) is concerned. Independent of the direction of the valence effect, the results reported by Moran and colleagues are consistent with our findings in showing that the sgACC plays an important role in processing the valence of *social* judgments. Additional studies are needed to determine the extent to which this sensitivity is specifically associated with positive or negative information.

Two possible caveats should be noted with regard to the experimental design of the study reported here. First, reward sensitivity of the PACC was observed in the ROI analysis but not in the whole-brain analysis, suggesting that this effect was not as strong as had been observed in prior studies. Possibly, efforts to induce mild dehydration prior to the experiment caused the negative feedback to be rewarding (despite its bitter taste), thereby reducing the contrast between the positive and negative primary reinforcers. This explanation is also consistent with the finding that negative feedback led to activity above baseline, even in the PACC. A second caveat concerns the social manipulation. In the social condition, the “judge” could also receive an extra \$20. Thus, it could be claimed that this extra monetary reward, although it was not received by the participant, influenced the experience of reward in the social context. However, this concern is at least partly mitigated by the observation that it did not produce a significant increase in activity in an area that is known to scale with reward value (the PACC). Nevertheless, these issues should be explored in future research. A possible way to address this issue is to do an experiment with a very similar design that in addition to context and type of feedback also manipulates the magnitude of feedback.

To conclude, our results confirm that PACC is sensitive to the evaluation and/or representation of reward information, and disconfirm the hypothesis that this region is uniquely or even primarily sensitive to social information. Whereas there are areas in the mPFC that are modulated by social context, the PACC is not one of them. Taken together, these results favor the hypothesis that the PACC serves a more general affective function, as part of the valuation network of the brain. Social stimuli may elicit responses within this region, either by being intrinsically more rewarding than other types of stimuli or through the internal “simulation” of the desires or motivations of others. In either case, by directly examining the interaction between reward processing and social context, the results of this study contribute to the identification of functionally distinguishable subregions of the mPFC that may help inform future social neuroscientific and neuroeconomic studies concerning the interaction between reward processing and social context.

AUTHOR NOTE

We thank Eveline Crone, Richard Ridderinkhof, and Birte Forstmann for their valuable comments. S.M.M. is now in the Department of Psychology at Stanford University; L.T.H., in the Department of Psychology at New York University. Correspondence concerning this article should be addressed to W. van den Bos, Department of Psychology, Leiden

University, 2333 AK Leiden, The Netherlands (e-mail: wbos@fsw.leidenuniv.nl).

REFERENCES

- ADOLPHS, R. (2003). Cognitive neuroscience of human social behaviour. *Nature Reviews Neuroscience*, *4*, 165-178.
- AMODIO, D. M., & FRITH, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, *7*, 268-277.
- ANDERSON, A. K., CHRISTOFF, K., STAPPEN, I., PANITZ, D., GHAREMANI, D. G., GLOVER, G., ET AL. (2003). Dissociated neural representations of intensity and valence in human olfaction. *Nature Neuroscience*, *6*, 196-202.
- APPERLY, I. A., SAMSON, D., & HUMPHREYS, G. W. (2005). Domain-specificity and theory of mind: Evaluating neuropsychological evidence. *Trends in Cognitive Sciences*, *9*, 572-577.
- BARTELS, A., & ZEKI, S. (2004). The neural correlates of maternal and romantic love. *NeuroImage*, *21*, 1155-1166.
- BAXTER, M. G., & MURRAY, E. A. (2002). The amygdala and reward. *Nature Reviews Neuroscience*, *3*, 563-573.
- BLAIR, K., MARSH, A. A., MORTON, J., VYTHILINGAM, M., JONES, M., MONDILLO, K., ET AL. (2006). Choosing the lesser of two evils, the better of two goods: Specifying the roles of ventromedial prefrontal cortex and dorsal anterior cingulate in object choice. *Journal of Neuroscience*, *26*, 11379-11386.
- BLOOD, A. J., ZATORRE, R. J., BERMUDEZ, P., & EVANS, A. C. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nature Neuroscience*, *2*, 382-387.
- BREITER, H. C., AHARON, I., KAHNEMAN, D., DALE, A., & SHIZGAL, P. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron*, *30*, 619-639.
- CASTELLI, F., HAPPE, F., FRITH, U., & FRITH, C. (2005). Movement and mind: A functional imaging study of perception and interpretation of complex intentional movement patterns. In J. T. Cacioppo & G. G. Berntson (Eds.), *Social neuroscience: Key readings* (pp. 155-169). New York: Psychology Press.
- COX, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers & Biomedical Research*, *29*, 162-173.
- DELGADO, M. R., NYSTROM, L. E., FISELL, C., NOLL, D. C., & FIEZ, J. A. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *Journal of Neurophysiology*, *84*, 3072-3077.
- ERK, S., SPITZER, M., WUNDERLICH, A. P., GALLEY, L., & WALTER, H. (2002). Cultural objects modulate reward circuitry. *NeuroReport*, *13*, 2499-2503.
- FLETCHER, P. C., HAPPE, F., FRITH, U., BAKER, S. C., DOLAN, R. J., FRACKOWIAK, R. S. J., & FRITH, C. D. (1995). Other minds in the brain: A functional imaging study of “theory of mind” in story comprehension. *Cognition*, *57*, 109-128.
- FORMAN, S. D., COHEN, J. D., FITZGERALD, M., EDDY, W. F., MINTUN, M. A., & NOLL, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. *Magnetic Resonance in Medicine*, *33*, 636-647.
- FRITH, C. D., & FRITH, U. (2006). The neural basis of mentalizing. *Neuron*, *50*, 531-534.
- FRITH, U., & FRITH, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society B*, *358*, 459-473.
- GALLAGHER, H. L., HAPPE, F., BRUNSWICK, N., FLETCHER, P. C., FRITH, U., & FRITH, C. D. (2000). Reading the mind in cartoons and stories: An fMRI study of “theory of mind” in verbal and nonverbal tasks. *Neuropsychologia*, *38*, 11-21.
- GALLAGHER, H. L., JACK, A. I., ROEPSTORFF, A., & FRITH, C. D. (2002). Imaging the intentional stance in a competitive game. *NeuroImage*, *16*, 814-821.
- GALLESE, V., & GOLDMAN, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, *2*, 493-501.
- GORDON, R. M. (1992). Folk psychology as simulation. *Mind & Language*, *1*, 158-171.
- GOTLIB, I. H., SIVERS, H., GABRIELI, J. D. E., WHITFIELD-GABRIELI, S.,

- GOLDIN, P., MINOR, K. L., & CANLI, T. (2005). Subgenual anterior cingulate activation to valenced emotional stimuli in major depression. *NeuroReport*, *16*, 1731-1734.
- GRÈZES, J., FRITH, C., & PASSINGHAM, R. E. (2004). Brain mechanisms for inferring deceit in the actions of others. *Journal of Neuroscience*, *24*, 5500-5505.
- HARRIS, L. T., MCCLURE, S., VAN DEN BOS, W., COHEN, J., & FISKE, S. T. (2007). Regions of MPFC differentially turned to social and nonsocial affective evaluation. *Cognitive, Affective, & Behavioral Neuroscience*, *7*, 309-316.
- HOLROYD, C. B., NIEUWENHUIS, S., YEUNG, N., NYSTROM, L., MARS, R. B., COLES, M. G. H., & COHEN, J. D. (2004). Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. *Nature Neuroscience*, *7*, 497-498.
- JENSEN, J., MCINTOSH, A. R., CRAWLEY, A. P., MIKULIS, D. J., REMINGTON, G., & KAPUR, S. (2003). Direct activation of the ventral striatum in anticipation of aversive stimuli. *Neuron*, *40*, 1251-1257.
- KING-CASAS, B., TOMLIN, D., ANEN, C., CAMERER, C. F., QUARTZ, S. R., & MONTAGUE, P. R. (2005). Getting to know you: Reputation and trust in a two-person economic exchange. *Science*, *308*, 78-83.
- KNUTSON, B., ADAMS, C. M., FONG, G. W., & HOMMER, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *Journal of Neuroscience*, *21*, 1-5.
- KNUTSON, B., FONG, G. W., BENNETT, S. M., ADAMS, C. M., & HOMMER, D. (2003). A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: Characterization with rapid event-related fMRI. *NeuroImage*, *18*, 263-272.
- LIEBERMAN, M. D. (2007). Social cognitive neuroscience: A review of core processes. *Annual Review of Psychology*, *58*, 259-289.
- MARS, R. B., COLES, M. G. H., GROL, M. J., HOLROYD, C. B., NIEUWENHUIS, S., HULSTIJN, W., & TONI, I. (2005). Neural dynamics of error processing in medial frontal cortex. *NeuroImage*, *28*, 1007-1013.
- MAYBERG, H. S., LOZANO, A. M., VOON, V., MCNEELY, H. E., SEMINOWICZ, D., HAMANI, C., ET AL. (2005). Deep brain stimulation for treatment-resistant depression. *Neuron*, *45*, 651-660.
- MCCABE, K., HOUSER, D., RYAN, L., SMITH, V., & TROUARD, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of Sciences*, *98*, 11832-11835.
- MCCLURE, S. M., BERNIS, G. S., & MONTAGUE, P. R. (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron*, *38*, 339-346.
- MCCLURE, S. M., BOTVINICK, M. M., YEUNG, N., GREENE, J. D., & COHEN, J. D. (2007). Conflict monitoring in cognition-emotion competition. In J. J. Gross (Ed.), *Handbook of emotion regulation* (pp. 204-228). New York: Guilford.
- 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.
- MCCLURE, S. M., LI, J., TOMLIN, D., CYPERT, K. S., MONTAGUE, L. M., & MONTAGUE, R. M. (2004). Neural correlates of behavioral preference for culturally familiar drinks. *Neuron*, *44*, 379-387.
- MITCHELL, J. P., BANAJI, M. R., & MACRAE, C. N. (2005a). General and specific contributions of the medial prefrontal cortex to knowledge about mental states. *NeuroImage*, *28*, 757-762.
- MITCHELL, J. P., BANAJI, M. R., & MACRAE, C. N. (2005b). The link between social cognition and self-referential thought in the medial prefrontal cortex. *Journal of Cognitive Neuroscience*, *17*, 1306-1315.
- MITCHELL, J. P., MACRAE, C. N., & BANAJI, M. R. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron*, *50*, 655-663.
- MONTAGUE, P. R., KING-CASAS, B., & COHEN, J. D. (2006). Imaging valuation models in human choice. *Annual Review of Neuroscience*, *29*, 417-448.
- MORAN, J. M., MACRAE, C. N., HEATHERTON, T. F., WYLAND, C. L., & KELLEY, W. M. (2006). Neuroanatomical evidence for distinct cognitive and affective components of self. *Journal of Cognitive Neuroscience*, *18*, 1586-1594.
- NIEUWENHUIS, S., SLAGTER, H. A., ALTING VON GEUSAU, N. J., HESLENFELD, D. J., & HOLROYD, C. B. (2005). Knowing good from bad: Differential activation of human cortical areas by positive and negative outcomes. *European Journal of Neuroscience*, *21*, 3161-3168.
- O'DOHERTY, J., KRINGELBACH, M. L., ROLLS, E. T., HORNAK, J., & ANDREWS, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience*, *4*, 95-102.
- 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.
- PAGNONI, G., ZINK, C. F., MONTAGUE, P. R., & BERNIS, G. S. (2002). Activity in human ventral striatum locked to errors of reward prediction. *Nature Neuroscience*, *5*, 97-98.
- PREMACK, D., & WOODRUFF, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral & Brain Sciences*, *1*, 515-526.
- RAMNANI, N., ELLIOTT, R., ATHWAL, B. S., & PASSINGHAM, R. E. (2004). Prediction error for free monetary reward in the human prefrontal cortex. *NeuroImage*, *23*, 777-786.
- RIDDERINKHOF, K. R., ULLSPERGER, M., CRONE, E. A., & NIEUWENHUIS, S. (2004). The role of the medial frontal cortex in cognitive control. *Science*, *306*, 443-447.
- 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.
- SANFEY, A. G., RILLING, J. K., ARONSON, J. A., NYSTROM, L. E., & COHEN, J. D. (2003). The neural basis of economic decision making in the ultimatum game. *Science*, *300*, 1755-1757.
- SAXE, R., CAREY, S., & KANWISHER, N. (2004). Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annual Review of Psychology*, *55*, 87-124.
- SCHÄFER, A., SCHIENLE, A., & VAITL, D. (2005). Stimulus type and design influence hemodynamic responses towards visual disgust and fear elicitors. *International Journal of Psychophysiology*, *57*, 53-59.
- SCHULTZ, W., DAYAN, P., & MONTAGUE, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*, 1593-1599.
- SEYMOUR, B., O'DOHERTY, J. P., DAYAN, P., KOLTZENBURG, M., JONES, A. K., DOLAN, R. J., ET AL. (2004). Temporal difference models describe higher-order learning in humans. *Nature*, *429*, 664-667.
- STONE, V. E., & GERRANS, P. (2006). Does the normal brain have a theory of mind? *Trends in Cognitive Sciences*, *10*, 3-4.
- TALAIRACH, J., & TOURNOUX, P. (1988). *Co-planar stereotaxic atlas of the human brain: An approach to medical cerebral imaging*. Stuttgart: Thieme.
- TAYLOR, S. E., & BROWN, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, *103*, 193-210.
- ULLSPERGER, M., & VON CRAMON, D. Y. (2004). Neuroimaging of performance monitoring: Error detection and beyond. *Cortex*, *40*, 593-604.
- VAN VEEN, V., HOLROYD, C. B., COHEN, J. D., STENGER, V. A., & CARTER, C. S. (2004). Errors without conflict: Implications for performance monitoring theories of anterior cingulate cortex. *Brain & Cognition*, *56*, 267-276.
- VÖLLM, B. A., TAYLOR, A. N. W., RICHARDSON, P., CORCORAN, R., STIRLING, J., MCKIE, S., ET AL. (2006). Neuronal correlates of theory of mind and empathy: A functional magnetic resonance imaging study in a nonverbal task. *NeuroImage*, *29*, 90-98.
- 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.
- WALTER, H., ADENZATO, M., CIARAMIDARO, A., ENRICI, I., PIA, L., & BARA, B. G. (2004). Understanding intentions in social interaction: The role of the anterior paracingulate cortex. *Journal of Cognitive Neuroscience*, *16*, 1854-1863.
- WICKER, B., KEYSERS, C., PLAILLY, J., ROYET, J. P., GALLESE, V., & RIZZOLATTI, G. (2003). Both of us disgusted in my insula: The common neural basis of seeing and feeling disgust. *Neuron*, *40*, 655-664.
- YEUNG, N., COHEN, J. D., & BOTVINICK, M. M. (2004). The neural basis of error detection: Conflict monitoring and the error-related negativity. *Psychological Review*, *111*, 931-959.