Self-esteem Modulates Medial Prefrontal Cortical Responses to Evaluative Social Feedback

Leah H. Somerville, William M. Kelley and Todd F. Heatherton

Department of Psychological and Brain Sciences, Center for Cognitive Neuroscience, Dartmouth College, Hanover, NH 03755, USA

Address correspondence to Leah Somerville, Weill Cornell Medical College, 1300 York Avenue, Box 140, New York, NY 10065, USA. Email: lbs2003@med.cornell.edu.

Self-esteem is a facet of personality that influences perception of social standing and modulates the salience of social acceptance and rejection. As such, self-esteem may bias neural responses to positive and negative social feedback across individuals. During functional magnetic resonance imaging scanning, participants (n = 42) engaged in a social evaluation task whereby they ostensibly received feedback from peers indicating they were liked or disliked. Results demonstrated that individuals with low self-esteem believed that they received less positive feedback from others and showed enhanced activity to positive versus negative social feedback in the ventral anterior cingulate cortex/medial prefrontal cortex (vACC/mPFC). By contrast, vACC/mPFC activity was insensitive to positive versus negative feedback in individuals with high self-esteem, and these individuals consistently overestimated the amount of positive feedback received from peers. Voxelwise analyses supported these findings; lower self-esteem predicted a linear increase in vACC/mPFC response to positive versus negative social feedback. Taken together, the present findings propose a functional role for the vACC/mPFC in representing the salience of social feedback and shaping perceptions of relative social standing.

Keywords: anterior cingulate cortex, fMRI, medial prefrontal cortex, self-esteem, social

Introduction

As a social species, humans have evolved a fundamental need to belong that encourages behaviors indicating that one is a good group member (Bowlby 1969; Baumeister and Leary 1995). Evolutionarily speaking, ejection from one’s social group would likely have resulted in death. Though the dire consequences of exclusion may no longer be inevitable in modern society, individuals still seek out social interaction, and even today a lack of social support increases a person’s risk for a number of adverse consequences, such as unhappiness, illness, and premature death (see Cacioppo et al. 2000, 2006). Therefore, it is likely that regions of the brain may function, at least in part, to monitor individuals’ inclusionary or exclusionary status based on environmental cues (Eisenberger and Lieberman 2004; Mitchell and Heatherton 2009; Heatherton 2010).

One interpersonal variable particularly relevant to monitoring inclusionary and exclusionary status is self-esteem, which is an attitude formed by self-evaluation based on positive and negative aspects of oneself (Rosenberg 1965; Brown 1993; Baumeister 1998) and which has interrelated stable (trait) and context-dependent (state) components (Heatherton and Polivy 1991). Personality and social psychological research has demonstrated that individual differences in self-esteem predict a wide variety of interpersonal and affective consequences that are relevant to health. For instance, individuals with low self-esteem are more likely to experience negative affect (e.g., Baumeister et al. 2003), show enhanced sensitivity to cues of social standing (Brodker 1983), and are more susceptible to interpersonal distress, heightened levels of depression, posttraumatic stress disorder (PTSD), social phobia (Pyszczynski and Greenberg 1987; Tennen and Affleck 1993; Izgic et al. 2004; Boscarno and Adams 2009), and substance dependence (Trzesniewski et al. 2006). However, high self-esteem is not without its drawbacks. Interpersonally, individuals with high self-esteem can be aggressive and bullying (Baumeister et al. 2003) and tend to be liked less by their peers, especially following a threat to their self-esteem (Heatherton and Vohs 2000). Given the differential salience of social feedback to high and low self-esteem individuals (Brodker 1983), we sought to test whether brain regions previously shown to be sensitive to social feedback demonstrate modulated activity based on this facet of personality.

We previously demonstrated that the ventral anterior cingulate cortex extending anteriorly into the medial prefrontal cortex (vACC/mPFC) differentiates the valence of social feedback, demonstrating a significantly enhanced response to favorable relative to unfavorable feedback from supposed peers (Somerville et al. 2006). Based on these results, it was suggested that ventral cortical midline structures including the mPFC and vACC represent affective properties of social feedback. If this interpretation is correct, then the magnitude of recruitment of this region may be subject to individual differences in social evaluation–relevant personality characteristics such as self-esteem.

To characterize the influence of self-esteem on neural responses to social feedback, we tested whether self-esteem modulates neural activity in 3 a priori defined brain regions—the vACC/mPFC, the dorsal anterior cingulate cortex (dACC) from Somerville et al. (2006), and a separate region of the dACC implicated in processing socially rejecting feedback using the Cyberball paradigm (Eisenberger et al. 2003). Focus on these regions is motivated, in part, by a lack of consensus in prior work as to which regions along the cortical midline subserve social feedback processing. The dACC has been previously suggested to be sensitive to cues of social rejection (Eisenberger et al. 2003), whereas other work has noted social feedback effects in the vACC (Somerville et al. 2006; Masten et al. 2009; Onoda et al. 2009). Still other recent experiments have implicated both dACC and vACC regions as sensitive to different aspects of social approval and disapproval signals (Burklund et al. 2007). A priori analyses were followed by a targeted voxelwise regression analysis to test the spatial specificity of cortical midline regions showing differential recruitment as a function of self-esteem.
Finally, behavioral estimates of feedback were tested for brain-behavior relationships. Specifically, participants generated retrospective estimates of how much of the feedback they received was positive and what proportion was negative. We then tested whether these estimates were modulated by self-esteem and predicted by variability in 
 vACC/mPFC activity in a similar manner to personality self-reports. Taken together, such findings would provide a brain-based account for the enhanced monitoring of social standing characteristic of low self-esteem individuals.

Materials and Methods
The present work represents a novel analysis of a data set previously used to characterize the neural representation of social feedback processing (Somerville et al. 2006). Here we incorporate self-esteem and behavioral measures not examined previously to characterize how individual differences bias neural responses to social feedback.

Participants
Fifty participants between the ages of 18 and 24 years were recruited from the local community. Eight participants were subsequently excluded due to excessive movement or insufficient variability of button responses. Of the final sample, 20 participants were enrolled in Study 1a (11 males, mean age = 20 years) and 22 were enrolled in Study 1b (10 males, mean age = 19 years). Participants reported no abnormal neurological history, were not using psychoactive medications, were native speakers of English, had normal or corrected-to-normal visual acuity, and were strongly right-handed as measured by the Edinburgh Handedness Inventory (Oldfield 1971). Participants were given course credit or cash in exchange for their participation and provided informed consent in accordance with the guidelines set by the Committee for the Protection of Human Subjects at Dartmouth College.

Stimuli
This study used a stimulus set of 400 photographs of college-aged individuals (200 males, 200 females), compiled by experimenters from the mass media. All photographs were closely cropped to depict only the head and shoulders, were of suitable image quality for clear viewing in the functional magnetic resonance imaging (fMRI) scanner, and did not contain any other faces in the background of the picture. The gender of faces was predetermined to provide positive and negative feedback in equal proportions for each participant. Face stimuli were pseudorandomized across participants such that each face was represented roughly equally in positive feedback, negative feedback, and partial trial conditions.

Pre-scan Procedure
Approximately 2 weeks prior to fMRI scanning, participants attended a brief informational session in the laboratory. Participants were given the cover story that they were participating in a multi-university study on how individuals make first impressions. They were led to believe that their photographs would be traded among the participating universities and rated by the other participants based on how likeable and desired they looked. On the day of the experiment (during fMRI scanning), the participants would have the opportunity to rate the same individuals who had rated their picture. To ensure believability, we took a photograph of each participant, which they expected to be sent to the other universities to be rated prior to their fMRI scanning session. Unbeknownst to participants, their photographs were simply deleted and all ratings and face stimuli were created and compiled pseudorandomly by experimenters.

fMRI Design and Procedure
During fMRI scanning, trials were employed that contained cue, delay, and feedback subcomponents (see Fig. 1). During the cue period, participants viewed a face centrally located on an otherwise blank screen and answered, via button press, the question, "Would I like this person?" (Study 1a) or "Would this person like me?" (Study 1b). If participants did not answer in the 3000 ms allotted, an answer (yes or no) was randomly assigned (~2% of all trials).

Following the button press or assigned response, the trial shifted to the delay portion (1000–4000 ms). The participant’s response was displayed on the left side of the screen, while they passively awaited feedback indicating whether the pictured individual ostensibly liked or disliked the participant’s photograph. Following the delay, social feedback was displayed to the right of the face for 2000 ms, which the participant passively viewed. ‘Yes’ feedback supposedly indicated that the pictured individual liked the participant, and ‘No’ feedback supposedly indicated that the pictured individual did not like the participant based on their picture.

In addition to the temporal jitter created by variable reaction times, partial trials were included so that unique estimates of the hemodynamic response could be computed for each subcomponent of the trial (Ollinger et al. 2001). As such, 20% of trials terminated after the cue period and 20% of trials terminated after the delay period. The remaining 60% of trials ran to completion. In total, each subject completed 160 partial and 240 complete trials. To mitigate confusion on the participant’s part related to early termination of the partial trials, they were instructed that if we had not received a rating from the pictured individual, then the trial would end prior to the feedback period. Partial and complete trials were randomly intermixed with periods of fixation during which participants viewed a white central fixation crosshair (2000–6000 ms).

Visual stimuli were compiled using PsyScope software (Cohen et al. 1993) and presented using an LCD projector (Epson model ELP-7000), viewable by an angled mirror mounted on top of the head coil. Stimuli were presented centrally on an otherwise black screen. Two fiber-optic key presses, one held in each hand, were used to record participants’ ‘yes’ and ‘no’ responses, which were recorded through the PsyScope button box (New Micros).

Participants received positive feedback for half of all full trials and negative feedback in the other half, in a pseudorandomized order across runs. The manipulation of feedback type along with the participant’s button press response (e.g., indicating ‘yes’ or ‘no’ responses) was used to sort trials into a 2 x 2 factorial design with factors of congruence and feedback type. If the participant’s choice of like or dislike matched the feedback the participant received, that was considered a “congruent” trial, while a mismatch in the 2 ratings was considered an “incongruent” trial. Because participant responses dictated whether the trial would be subsequently coded as “congruent” or “incongruent,” the number of trials per condition varied across participants. To ensure accurate estimation of the hemodynamic response for each condition, participants were only included in...
analyses if they had at least 30 full trials in each of the 4 trial types (n = 2 excluded). Following fMRI scanning, participants completed several self-report scales assessing self-esteem and anxiety, including the State-Trait Anxiety Inventory (State and Trait StAI; Spielberger et al. 1988), the State Self-Esteem Scale (SSES; Heatherton and Polivy 1991), the Rosenberg Self-Esteem Scale (RSSES; Rosenberg 1989), the Janis & Field Feelings of Inadequacy Scale (JFS; Janis, 1959; revised by Fleming and Courtney 1984), the Reaction Sensitivity: Questionnaire (RSQ; Downey and Feldman, 1996), and the Ten Item Personality Inventory (TIPI; Gosling et al. 2003). Subjects were also given an exit questionnaire, where they reported on their impression of the task and estimated the relative proportion of positive and negative feedback they received during the experiment. Although some individuals generated accurate estimates, some underestimated the proportion of positive feedback they received and still others provided overestimations relative to other participants and the objective proportion (0.5), which we refer to as an “optimism bias.” No participant reported disbelief in the cover story.

\textbf{fMRI Parameters}

Anatomical and functional whole-brain imaging was performed on a 1.5 T General Electric Signa scanner (General Electric Medical Systems). A T1-weighted high-resolution anatomical image was acquired using a 3D spoiled gradient sequence (124 sagittal slices, time echo [TE] = 6 ms, time repetition [TR] = 25 ms, flip angle = 25°, 1 x 1 x 1.2 mm voxels). Functional images were collected in 4 functional runs of 538 volumes each, using a gradient spin-echo, echo-planar sequence sensitive to blood oxygen level-dependent (BOLD) contrast (T2*) (20 axial slices per whole-brain volume, 3.75 mm in-plane resolution, 5 mm thickness, 1 mm skip, TR = 2000 ms, TE = 35 ms, flip angle = 90°).

\textbf{fMRI Analysis}

fMRI data analysis was performed in SPSS and SPM2 (Friston et al. 1995), with additional usage of in-house Matlab scripts to perform portions of region-of-interest (ROI) analyses. For functional images, data were preprocessed to remove sources of noise and artifact. Data were slice-time corrected, realigned within and across runs to correct for head movement, and coregistered with each participant’s high-resolution anatomical scan. Participants with more than 2 mm total movement in any plane were excluded (n = 6). Functional data were then spatially normalized into a standard anatomical space (resampled to 3 mm isotropic voxels) and smoothed with an 8-mm full-width at half-maximum Gaussian kernel.

For each participant, a general linear model incorporating 7 task regressors (cue, delay-like, delay-dislike, and 4 types of feedback—congruent-like, congruent-dislike, incongruent-like, incongruent-dislike) convolved with the canonical hemodynamic response function, and covariates of noninterest (session mean and linear trend for each run and 6 movement parameters derived from realignment corrections) were used to compute parameter estimates (β) and contrast images (weighted parameter estimates) for each comparison at each voxel. For participants who had more than 12 trials to which they did not register a button response (n = 11), no-response trials were modeled as an additional task regressor that was not analyzed further. All group data were collapsed across Studies 1a and 1b, as the positive and negative social feedback was delivered in an equivalent fashion across both experiments. It should be noted that self-esteem was not different between the 2 samples (RSSES, SSES P > 0.3), and including study as a covariate did not alter the significance of any of the reported effects. To test for possible interactions between the sex of the participant and the sex of the stimulus on a trial-by-trial basis, a second general linear model was conducted in which the task regressors described above were further parsed into “male” and “female” stimulus categories (see Supplementary Material).

First, we tested whether the magnitude of activation within the dACC (x = -6, y = 28, z = 32) and vACC/mPFC (x = -6, y = 49, z = -13) reported previously (Somerville et al. 2006) was differentially influenced by individual differences in self-esteem. The vACC/mPFC region was identified by a whole-brain voxelwise F test on the feedback variable (positive, negative). The dACC was not identified in the analysis of feedback effects. It is possible that the dACC was significantly more active to all feedback (regardless of valence) that was incongruent with participants’ responses. The dACC region from Eisenberger et al. (2003; 6-mm sphere centered on Montreal Neurological Institute coordinates x = -8, y = 20, z = 40) was also queried to include a region of the dACC shown by others to be sensitive to feedback cues.

To carry out analysis of a priori ROIs, each participant was characterized on self-esteem level using the total score on the SSES (Heatherton and Polivy 1991). We extracted parameter estimates from each region and performed a 2 x 2 x 2 mixed analysis of variance (ANOVA) in each region testing for effects of congruence (congruent, incongruent) and feedback type (positive, negative) as within-subjects variables and self-esteem (based on median split) as a between-subjects variable. Any significant effect of self-esteem was supplemented by targeted post hoc analyses inputting self-esteem as a continuous covariate to determine whether modulatory effects of self-esteem were linear.

Next, we sought to determine if individuals with low self-esteem recruited additional aspects of the mPFC to process the valence of social feedback relative to individuals with high self-esteem. The search volume was constrained to the cortical midline, defined as brain voxels between x = -20 and x = 20 and anterior to y = 5. The appropriate multiple comparisons correction to preserve alpha = 0.05 was calculated on the search volume (four thousand five hundred and fifty-five 3 x 3 x 3 voxels) using a Monte Carlo simulation carried out by the AlphSim program within AFNI (Cox 1996). Positive versus negative feedback contrast images were submitted to a regression within SPM2, including self-esteem as a continuous between-subjects regressor. All regions considered statistically significant exceeded the P value and extent threshold combination that preserves a corrected P < 0.05. This analysis is sensitive to both positive and negative relationships between feedback response and self-esteem. To facilitate cross-study comparision, we have also included a table summarizing results of whole-brain analyses at a relatively liberal P < 0.005, uncorrected, 10 voxels threshold.

As the optimism bias measure was determined to vary with self-esteem (see Results), this measure was substituted for SSES score to test whether it similarly modulated fMRI activity in the ROIs identified in the SSES analysis. We additionally performed a voxelwise between-subjects t test on positive versus negative feedback contrast estimates, inputting optimism bias as a between-subjects variable based on a median split. (Analysis based on median split was performed, rather than treating optimism bias as a continuous variable, because estimates were not normally distributed due to positive skew and leptokurtosis around the 50% mark. This property of the data also precluded the use of formal mediation analyses. Individuals in the “low” group estimated 50% or less positive feedback, and individuals in the “high” group estimated more than 50% positive feedback.)

States showing this pattern within the search volume were rendered for visualization using Caret 5.6 software (Van Essen et al. 2001). All reported coordinates have been converted to Talairach and Tournoux (1988) atlas space using a nonlinear transformation algorithm in Matlab (http://imaging.mrc-cbu.cam.ac.uk/imaging/MniTalairach). P values are 2-tailed unless otherwise noted.

\textbf{Results}

\textbf{Self-esteem Score Distribution of Participants}

SSES scores ranged from 46 to 99 (possible values: 20–100). The SSES score distribution and descriptive statistics (mean = 73, standard deviation [SD] = 11.8, median = 74.5) were comparable with published normative data on this scale (Heatherton and Polivy 1991) and included substantial variability with which to examine individual difference effects. As expected, SSES scores correlated strongly with the RSSES (r_{41} = 0.79, P < 0.001) as well as other self-esteem relevant measures including...
Male and female participants did not significantly differ on SSES received than low self-esteem participants (high SE: mean = 59.0%, SD = 11.9; low SE: mean = 48.1%, SD = 16.2; t_{40} = 2.49, P = 0.017). When compared with the actual proportion of positive feedback provided to all participants (50%), low self-esteem individuals generated a liking estimate that was not different from the actual proportion (P > 0.55), whereas high self-esteem individuals demonstrated a significant optimism bias or overestimation of the proportion of positive feedback received (t_{40} = 3.49, P = 0.002). Participant gender did not modulate liking estimates (P > 0.4).

**Task Behavior and Liking Estimates**

During the cue phase of trials, low and high self-esteem individuals endorsed a comparable proportion of “yes” and “no” responses (P > 0.15) with similar response latencies (P > 0.8). However, after fMRI scanning, when participants were asked to estimate what proportion of feedback they received was positive (e.g., “yes” feedback), individuals with high self-esteem generated a significantly larger estimate of positive feedback received than low self-esteem participants (high SE: mean = 59.0%, SD = 11.9; low SE: mean = 48.1%, SD = 16.2; t_{40} = 2.49, P = 0.017). When compared with the actual proportion of positive feedback provided to all participants (50%), low self-esteem individuals generated a liking estimate that was not different from the actual proportion (P > 0.55), whereas high self-esteem individuals demonstrated a significant optimism bias or overestimation of the proportion of positive feedback received (t_{40} = 3.49, P = 0.002). Participant gender did not modulate liking estimates (P > 0.4).

**Analyses of A Priori ROI**

The vACC/mPFC showed a significant feedback by self-esteem interaction (F_{1,40} = 9.34, P = 0.004), which replicated when SSES was treated as a continuous covariate indicating that self-esteem was linearly predictive of vACC/mPFC activity (F_{1,40} = 6.85, P = 0.012). In examining the directionality of these effects, low self-esteem individuals showed a significantly greater differentiation between positive and negative social feedback than did high self-esteem individuals (t_{40} = 3.06, P = 0.004; Fig. 2). Overall, there was no main effect of self-esteem on responding in this region (P > 0.7). This effect was not influenced by participant sex (P > 0.8). For further sex differences analyses testing for effects of receiving feedback from a same- or opposite-sex stimulus, see Supplementary Material.

The dACC region identified in Somerville et al. (2006) showed no main effect of feedback type, no main effect of self-esteem, and no feedback by self-esteem interaction (P > 0.4). In addition, the dACC region identified in Eisenberger et al. (2003) also showed no main effect of feedback type (P > 0.4), no main effect of self-esteem (P > 0.2), and no feedback by self-esteem interaction (P > 0.9). Participant sex did not modulate these effects (P > 0.2). Thus, the dACC regions tested were insensitive to the feedback manipulations and further did not modulate in activity based on self-esteem.

**Specificity of Self-esteem Effects**

Following the identification of a significant feedback by self-esteem interaction in the vACC/mPFC, we tested the specificity of vACC/mPFC modulation to self-esteem-related personality variables. To argue that self-esteem biases the neural responses to social feedback, the observed effects should show some degree of specificity to self-esteem at the exclusion of other personality variables. To test this, we performed additional ANOVA analyses on the a priori vACC/mPFC ROI (within factor: BOLD response to positive feedback, BOLD response to negative feedback) and substituted scores on the JFS, RSES, RSQ, STAI, and TIPI as between-subjects factors.

As would be expected, use of the RSES generated the same pattern of results (t_{40} = 2.84, P = 0.007) as did the JFS (t_{40} = 2.13, P = 0.04), both of which are trait self-esteem measures. We also observed marginal effects for the Spielberger State (t_{40} = 2.05, P = 0.05) and Trait (t_{40} = 1.82, P = 0.08) Anxiety Inventories, with greater anxiety corresponding to greater differentiation of vACC/mPFC response to positive versus negative social feedback. This is likely due to shared variance between the constructs of anxiety and self-esteem, which has been documented previously (Heatherton and Polivy 1991). However, effects observed with the SSES remained significant when controlling for anxiety (State: F_{1,39} = 4.75, P = 0.035; Trait: F_{1,39} = 6.23, P = 0.017), whereas anxiety effects were no longer significant when controlling for self-esteem (P > 0.15), an indication that self-esteem was a more robust predictor of vACC/mPFC response than was anxiety. The RSQ, a measure of rejection sensitivity, did not show significant relationships with vACC/mPFC activity (P > 0.5). We also did not observe any modulation of vACC/mPFC responses by scores on a measure of the “Big 5” personality traits (Gosling et al. 2003): Extraversion, Neuroticism, Openness, Conscientiousness, or Agreeableness (P > 0.25).

**Voxelwise Analysis of Cortical Midline**

The results focusing on a priori ROIs suggest that individuals with lower self-esteem show greater differential response to positive than negative feedback along the ventral cortical midline. We followed up this targeted analysis with a voxelwise regression aimed to identify brain regions whose magnitude of response to positive versus negative feedback was predicted by self-esteem.

Again, we observed a region of the mPFC, contiguous with the a priori ROI and extending anteriorly into Brodmann area (BA) 10 that showed a linear change in recruitment as a function of self-esteem (x = −9, y = 57, z = −15, 45 voxels, t_{41} = 3.54, P < 0.05, corrected, see Fig. 3, Table 1). Specifically, lower self-esteem predicted a greater differentiation in mPFC response to positive relative to negative social feedback. The same pattern of activity was also observed in a second cortical midline region at the confluence of the vACC and the medial...
Evaluating Positive Feedback Estimates as a Predictor of vACC/mPFC Response

Finally, we tested whether individual differences in an optimism bias measure predicted vACC/mPFC responses to positive relative to negative social feedback. To do so, we compared contrast estimates for responses to positive relative to negative feedback as a function of optimism bias in 3 regions defined previously: the a priori vACC/mPFC region and the 2 vACC/mPFC regions identified in the voxelwise self-esteem analysis. Individuals with lower optimism bias showed a significantly greater vACC/mPFC response differentiating positive relative to negative social feedback in all regions tested relative to those demonstrating a greater optimism bias (a priori vACC/mPFC: $t_{40} = 1.65, P = 0.05$ [1-tailed]; voxelwise mPFC: $t_{40} = 2.54, P = 0.02$; voxelwise vACC: $t_{40} = 2.21, P = 0.05$). To facilitate cross-study comparison, results of an exploratory whole-brain analysis comparing neural responses to positive relative to negative social feedback in individuals high and low in optimism bias are reported in Table 2.

Discussion

Across 2 studies, we measured fMRI and behavioral responses while participants received positive and negative social feedback from purported peers. We previously demonstrated that the vACC/mPFC shows differential activity to feedback as a function of valence, with positive feedback generating a significantly greater response than negative feedback. Here we demonstrate that the magnitude of vACC/mPFC engagement while processing social feedback is modulated by self-esteem such that low self-esteem individuals show a greater polarization of vACC/mPFC response to social feedback based on valence. This effect appears to be relatively specific to the personality dimension of self-esteem, as a general tendency to experience negative affect (e.g., neuroticism), and other personality variables did not similarly predict vACC/mPFC responsivity. Enhanced sensitivity of the vACC/mPFC also predicted accuracy in estimating one’s relative degree of social acceptance, whereas vACC/mPFC insensitivity correlated with overestimations in judging how well-liked the subject was during the experiment.

Table 2

<table>
<thead>
<tr>
<th>Brain region</th>
<th>BA</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>t</th>
<th>Voxel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low &gt; High optimism bias</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right mPFC</td>
<td>11</td>
<td>18</td>
<td>23</td>
<td>-9</td>
<td>3.79</td>
<td>30</td>
</tr>
<tr>
<td>Right mPFC</td>
<td>11</td>
<td>3</td>
<td>40</td>
<td>-25</td>
<td>3.43</td>
<td>30</td>
</tr>
<tr>
<td>Left dorsomedial prefrontal cortex</td>
<td>8</td>
<td>-6</td>
<td>42</td>
<td>31</td>
<td>3.60</td>
<td>13</td>
</tr>
<tr>
<td>Left posterior cingulate gyrus</td>
<td>31</td>
<td>-12</td>
<td>48</td>
<td>25</td>
<td>3.55</td>
<td>36</td>
</tr>
<tr>
<td>Right middle temporal gyrus</td>
<td>19</td>
<td>24</td>
<td>-66</td>
<td>12</td>
<td>3.21</td>
<td>12</td>
</tr>
<tr>
<td>Right superior occipital gyrus</td>
<td>19</td>
<td>24</td>
<td>-66</td>
<td>12</td>
<td>3.21</td>
<td>12</td>
</tr>
<tr>
<td>High &gt; Low optimism bias</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left medial frontal gyrus</td>
<td>6</td>
<td>-6</td>
<td>21</td>
<td>45</td>
<td>3.58</td>
<td>17</td>
</tr>
<tr>
<td>Right temporal pole</td>
<td>21</td>
<td>60</td>
<td>-6</td>
<td>3</td>
<td>3.48</td>
<td>12</td>
</tr>
<tr>
<td>Right postcentral gyrus</td>
<td>1/3</td>
<td>30</td>
<td>36</td>
<td>57</td>
<td>3.39</td>
<td>11</td>
</tr>
<tr>
<td>Hippocampus</td>
<td>14</td>
<td>-27</td>
<td>-12</td>
<td>-12</td>
<td>3.39</td>
<td>11</td>
</tr>
</tbody>
</table>

Note: Regions listed exceeded threshold of $P < 0.005$, uncorrected, with at least 10 contiguous voxels. Coordinates represent spatial location of activation in Talairach and Tournoux (1988) atlas space. Regions denoted with “*” were also identified in a targeted analysis of mPFC effects at $P < 0.05$, corrected for multiple comparisons.
objective accuracy should be considered preliminary as it is possible that participants used theory-based reasoning to generate liking estimates (e.g., a low self-esteem individual believing that they are generally liked by others about half the time, irrespective of the proportion of negative feedback given during the task; Cronbach 1955). Future work varying the proportion of positive and negative feedback may test whether the correspondence between liking estimates and absolute task-based feedback rates is meaningful.

Irrespective, these findings reliably demonstrate that high self-esteem is associated with inflated perceptions of liking relative to low self-esteem individuals. This observation converges with prior work that has identified a host of biases observed in the behavior of high self-esteem individuals following negative interpersonal feedback. For example, high self-esteem individuals perceive themselves to have been accepted even after being outwardly rejected by peers (Nezlek et al. 1997) and, when faced with negative interpersonal feedback, tend to emphasize their positive features while minimizing their own negative features and derogating others (Vohs and Heatherton 2004). Following threat, those with high self-esteem become more self-focused and boastful and behave in ways that lead them to be evaluated negatively by their interaction partners (Heatherton and Vohs 2000; Vohs and Heatherton 2001). By stark contrast, those with low self-esteem behave in ways that increase how much others like them, perhaps in attempts to stave off potential rejection (Heatherton and Vohs 2000). These behavioral findings are consistent with the ideas behind "sociometer theory" (Leary et al. 1995), which proposes that changes in self-esteem may facilitate motivation to engage in behaviors to preserve their status as group members. The present findings support the general conclusion that those with low self-esteem are more sensitive to social feedback than those with high self-esteem.

**Differential fMRI Responses to Social Feedback Based on Self-esteem**

Individuals with lower self-esteem demonstrated a robust vACC/mPFC response to positive versus negative social feedback compared with those with high self-esteem. Additionally, lower self-esteem predicted a linear increase in activity differentiating feedback valence extending anteriorly to the mPFC (BA 10) and overlapping with what is sometimes termed the orbitofrontal cortex. It is important to note that individuals with low self-esteem did not show generalized increases in neural responses to social feedback, rather it is the coding of feedback "valence" that appears to be exaggerated in low self-esteem individuals. The current findings are consistent with the psychological consequence of low self-esteem in sensitizing individuals to monitoring cues related to social acceptance and rejection (Leary et al. 1995). This interpretation is also in line with prior work demonstrating that low self-esteem individuals show exaggerated physiological responses to cues of rejection relative to their high self-esteem counterparts (Gyurak and Ayduk 2007). We have identified a possible neural mechanism for this heightened sensitivity in the fMRI response in the vACC/mPFC when encountering cues indicative of social standing.

Such a role for the vACC/mPFC is consistent with other work documenting a role of the mPFC in processing self-relevant information and the vACC in processing cues of affective significance. The mPFC is more active, for example, when people report on their personality traits (Craik et al. 1999; Johnson et al. 2002; Kelley et al. 2002; Heatherton et al. 2006; Moran et al. 2006; Schmitz and Johnson, 2006), make self-relevant judgments about pictures (Ochsner et al. 2004), and retrieve autobiographical memories of past events (Macrae et al. 2004). Particularly relevant are the results of Moran et al. (2006), showing significantly greater activity in the vACC when judging positive relative to negative traits. Importantly, Moran et al. (2006) only observed this response profile in the vACC when the personality traits were also judged to be self-relevant; vACC activity did not code the valence of traits judged to be unlike oneself. In the present experimental context, positive and negative social feedback is necessarily self-relevant and, taken together with the work of Moran et al., suggests a putative role for the vACC in monitoring affective aspects of self-relevant information.

Social psychological research has demonstrated a robust relationship between self-esteem and one’s self-concept (Jones 1973). When judging trait adjectives in relation to oneself, individuals with lower self-esteem tend to choose intermediate values and do so with lower internal consistency than individuals with high self-esteem (Sande et al. 1988). This has led some to posit that low self-esteem individuals have a “fuzzier” self-concept that is molded by external feedback (Campbell 1990), whereas the self-concept of high self-esteem individuals is more stable and resistant to outside influence. Though not directly tested, the present findings are consistent with this conceptualization, given a role for the vACC/mPFC in representing the salience of social feedback cues. The enhanced differential activity in low self-esteem individuals may reflect the greater significance of this external information for their self-concepts (Brockner 1983). Future research testing modulatory effects of self-esteem in self-referential processing tasks may test whether mPFC responses to trait adjectives are polarized by low self-esteem in a similar manner as the externally generated feedback used in the present experiment. Confirmatory findings would provide additional evidence implicating the mPFC in representing the relevance of information contributing to ideas about the self.

The mPFC has been implicated more generally in representing the value of rewards and punishments (Elliott et al. 2000) with particular emphasis on rewards (O’Doherty 2004) and using this information to guide subsequent behavior (Rolls 2000). More recently, the mPFC has been identified as a common locus of activity for several varieties of social information processing (Mitchell 2009). The current findings, implicating the mPFC in representing neural responses to positive and negative social feedback (and biased activity with low self-esteem), may lie at the intersection of these research domains. Indeed, a review by Amodio and Frith (2006) delineating the function of subcomponents of the mPFC implicates the region observed in this experiment in monitoring the outcomes of social exchanges, with a particular sensitivity to processing feedback.

Results of the present study are consistent with this interpretation and extend the functioning of this region to show differential modulation based on individual difference variables that bias the salience of such outcomes. These results also converge with lesion work demonstrating that damage to this region of the brain often results in a deficiency in incorporating feedback from others (and social norms) to...
make appropriate behavioral choices in social contexts, resulting in social disinhibition and inappropriate approach behavior toward other individuals (Beer et al. 2003, 2006). The current findings support the notion that vACC/mPFC activity may serve, in part, to shape perceptions of social standing and modulate behavior accordingly.

In addition to the vACC/mPFC, we tested whether activity in the dACC, previously implicated in processing social rejection cues, demonstrated modulated activity based on self-esteem. In a prior report using this same task (Somerville et al. 2006), we observed that dACC activity did not differ as a function of feedback valence in either direction (positive > negative or negative > positive). In the present data set, the dACC did not demonstrate modulated activity based on individual differences in self-esteem, as evidenced by both a priori ROI and voxelwise regression analyses. Moreover, we tested the specific dACC location obtained by Eisenberger et al. (2003) in their study identifying brain regions active in contexts of social rejection. In the current experimental context, this region did not show task-modulated activity or differential recruitment with variability in self-esteem. There are numerous differences between the 2 tasks that may explain these divergent effects, and it is hoped that future research will provide additional evidence for the specificity of cortical midline responses to social feedback cues.

At face value, generating an accurate perception of one’s social standing may appear to be the ideal mode of operation when contrasted with a bias toward inaccurate self-enhancement. If true, the present data suggest that individuals with the strongest vACC/mPFC coding of social feedback should have optimal functioning. However, the assumption that accurate self-evaluation is most adaptive has been challenged for decades (Taylor and Brown 1988). In fact, individuals who possess moderately inflated self-views tend to be happier (Diener E and Diener M 1995; Furnham and Cheng, 2000; Baumeister et al. 2003) and may be protected against negative affect and depression (Greenberg et al. 1992; Roberts and Monroe 1994; Bonanno et al. 2002). Thus, a relative insensitivity in vACC/mPFC coding of social feedback cues may actually be more adaptive when compared with the exaggerated responses to valenced social feedback in low self-esteem individuals.

This conceptualization converges with linkages between self-esteem and psychiatric health. It has long been recognized that low self-esteem is a risk factor for major depression and other psychiatric illnesses such as PTSD and social phobia (Pyszczynski and Greenberg, 1987; Tennen and Affleck, 1993; Boccarino and Adams, 2009). Moreover, both clinical depression and PTSD are associated with abnormal properties of the vACC including reduced volume (e.g., Drevets et al. 1997; Tang et al. 2007; although see Mayberg et al. 1999) and dysregulated functional responses to affective cues. Reports of PTSD patients showed decreased vACC activity to trauma script imagery (Shin et al. 2001; Frewen et al. 2008), while Gotlib et al. (2005) have reported altered vACC responses to facial expressions of emotion in depressed individuals relative to controls. Functional properties of the vACC are also predictive of clinical outcomes, with activity in vACC observed during the processing of emotional cues related to overall number of symptoms in PTSD (Shin et al. 2005) as well as predicting which patients will respond to cognitive behavioral therapy (Bryant et al. 2008).

Characterizing the neurobiological correlates of risk factors for psychiatric illnesses (such as low self-esteem) may shed light on the pathophysiology of these illnesses. The present findings suggest that healthy individuals with enhanced risk for developing these psychiatric illnesses may show some degree of similar functional dysregulation in midline cortical structures when processing interpersonal feedback. Although low self-esteem has been identified as a risk factor for depression, we did not explicitly collect measures of depression symptomatology. However, future work may explicitly test the role of depression symptomatology in contributing to, or modulating, the observed effects by using a more heterogeneous sample.

In this study, we sought to characterize responses to receiving feedback relevant to one's likeability. This dimension was chosen because processing feedback based on likeability is arguably more universal than processing other social judgments that involve more complex scenarios, such as those that contain sexual dimensions (such as appearance-based measures) or other more elaborate inferential processes related to mental states (such as whether the subject is ethical). Feedback regarding likeability can readily be received both within and across genders, irrespective of sexual preference, and receipt of like/dislike feedback requires little by way of elaborative processing. It could be argued that like/dislike feedback is overly simplistic relative to the types of social feedback experienced in the real world. However, our objective was to establish biased behavioral and neural responses to this relatively simplistic form of social feedback, in hopes that it lays the groundwork for future experimentation on more complex and ecologically valid social evaluations.

An open question is whether it is stable or context-dependent aspects of self-esteem that predict biased neural responding. According to the sociometer theory (Leary et al. 1995), a drop in self-esteem based on environmental cues (in this case, feedback in the experiment) may serve to enhance sensitivity to valenced social feedback—which is consistent with findings being observed using a measure of state self-esteem, the SSes (Heatherton and Polivy 1991). However, the SSes consistently shows high correlations with trait measures of self-esteem—as it did in this study—and similar results were observed when substituting trait measurements. That said, the possibility that responses to the trait scales were influenced by the task cannot be ruled out given their administration postscanning. Future work may determine the respective roles of stable versus context-dependent facets of self-esteem by deconfounding state and trait self-esteem using experimental manipulations (e.g., Volhs and Heatherton 2001) or administering self-report scales before and after the experimental task. Finally, the possibility of an unmeasured third variable modulating both self-esteem scores and neural responses to feedback (such as self-monitoring or attention to others) should be acknowledged.

Conclusion
In this experiment, participants received favorable and unfavorable feedback during a social evaluation task. Prior work has demonstrated that people with low self-esteem are especially sensitive to social cues that provide information about their inclusionary/exclusionary status within social groups. The present findings provide a neurobiological mechanism for this enhanced sensitivity, with low self-esteem individuals generating heightened activity in the vACC/mPFC to positive versus negative social feedback. These findings
support a role for the vACC/mPFC in representing the affective or reward value of environmental social cues, and enhanced recruitment with lower self-esteem may reflect the magnified salience of cues of social standing in these individuals. In addition, a greater differentiation in vACC/mPFC activity in response to positive relative to negative feedback predicted an accurate retrospective account of their relative social standing. Taken together, the present findings propose one functional role of the vACC/mPFC in differentially processing social feedback as a function of its salience and shaping perceptions of relative social standing.

Supplementary Material
Supplementary material can be found at: http://www.cercor.oxfordjournals.org/.

Funding
National Institutes of Health (MH059282 to T.H.); National Institute of Drug Abuse (DA007274) postdoctoral support for LS; National Science Foundation (BCS-0746220 to W.K., Graduate Research Fellowship to LS); Dartmouth Brain Imaging Center.

Notes
We would like to thank Nicole Magoon and Amy Rosenblum for assistance in compiling the stimulus materials, Joe Moran for assistance with fMRI analyses, George Wolford for statistical advice, and the anonymous reviewers who provided useful feedback on a previous version of this manuscript. Conflict of Interest: None declared.

References
Cronbach LJ. 1955. Processes affecting scores on ‘‘understanding others’’ and ‘‘assumed similarity’’. Psychol Bull. 52:177-193.


