An fMRI study of caring vs self-focus during induced compassion and pride

Emiliana R. Simon-Thomas,1 Jakub Godzik,2 Elizabeth Castle,3 Olga Antonenko,4 Aurelie Ponz,4 Aleksander Kogan,4 and Dacher J. Keltner4

1The Center for Compassion and Altruism Research and Education, Stanford University, Stanford, California, USA, 2Department of Medicine, Washington University, St. Louis, Missouri, USA, 3Department of Psychology, University of California, Los Angeles, Los Angeles, CA, USA, and 4Department of Psychology, University of California, Berkeley, Berkeley, California, USA

This study examined neural activation during the experience of compassion, an emotion that orients people toward vulnerable others and prompts caregiving, and pride, a self-focused emotion that signals individual strength and heightened status. Functional magnetic resonance images (fMRI) were acquired as participants viewed 55 s continuous sequences of slides to induce either compassion or pride, in alternation with sequences of neutral slides. Emotion self-report data were collected after each slide condition within the fMRI scanner. Compassion induction was associated with activation in the midbrain periaqueductal gray (PAG), a region that is activated during pain and the perception of others’ pain, and that has been implicated in parental nurturance behaviors. Pride induction engaged the posterior medial cortex, a region that has been associated with self-referent processing. Self-reports of compassion experience were correlated with increased activation in a region near the PAG, and in the right inferior frontal gyrus (IFG). Self-reports of pride experience, in contrast, were correlated with reduced activation in the IFG and the anterior insula. These results provide preliminary evidence towards understanding the neural correlates of important interpersonal dimensions of compassion and pride. Caring (compassion) and self-focus (pride) may represent core appraisals that differentiate the response profiles of many emotions.

Keywords: nurturing; self-focus; caregiving; midbrain periaqueductal gray; posterior medial cortex

INTRODUCTION
Compassion and pride represent new frontiers of inquiry in affective science. Compassion is an other-oriented emotional response to perceived suffering that involves wanting to care for those in need (Goetz et al., 2010). Pride, in contrast, is a self-oriented state responsive to actions that elevate individual or in-group status (Tracy and Robins, 2004, 2007). Empirical studies have begun to document how these two emotions are associated with different non-verbal signals (Tracy and Robins, 2004; Hertenstein et al., 2006; Simon-Thomas et al., 2009) and modes of thinking about relations between the self and others (Oveis et al., 2010). These different lines of inquiry suggest that compassion is defined by a caring orientation at the level of signal behavior and social cognition (an increased sense of similarity to others) whereas pride is defined by an individualizing, rank-related self-focus. In the present investigation we examine whether specific neural systems support the core properties of these two states: caring vs self-focus.

Compassion and caring: predictions of neural correlates
In the present research, we are guided by an appraisal tendency approach to emotion, which posits that distinct emotions are defined by core appraisals (Smith and Ellsworth, 1985; Lazarus, 1991; Lerner and Keltner, 2001). For example, anger is defined by a sense of unfairness caused by others’ intentional actions. Emotion-specific appraisals shape how emotions influence different kinds of cognition (Keltner and Lerner, 2010), and presumably the neural systems engaged during a particular emotion.

Within an appraisal framework, compassion is defined by two core processes: the empathic recognition of another’s suffering and the motivation to care for another (Lazarus, 1991; Goetz, et al., 2010). Importantly, this conceptual analysis distinguishes empathy, which involves mirroring and identifying another’s state, from compassion, which also involves caregiving motivation to reduce another’s suffering (Batson and Shaw, 1991). This conceptual analysis, as well as select neuroscience findings, set the stage for predictions concerning the neural correlates of compassion, distinguished by its core theme of caring.

The empathic response to another person’s suffering is the first stage of compassion (Batson and Shaw, 1991). By implication, empathy-related neural processes may be activated during experiences of compassion. Within the neuroscience literature several kinds of empathy have been investigated,
ranging from the feelings produced by automatic mirroring, simulation or embodiment of other peoples’ feelings (‘affective’ empathy) to understanding other peoples’ mental states (‘cognitive’ empathy), typically as the other people experience inflicted pain (Decety and Jackson, 2004; Singer, 2006; Singer and Lamm, 2009; Hooker et al., 2010; Reniers et al., 2011). These studies, while yielding variation in certain patterns of neural activation, do converge on what some consider an ‘empathy network’. This network typically includes: the insula, midbrain regions, the cingulate cortex, the anterior temporal cortex, sensorimotor cortex, the inferior frontal gyrus (IFG), the medial prefrontal cortex (mPFC), the superior temporal sulcus and gyrus and temporal parietal cortex (TPC) areas. Interestingly, individuals’ self-reported feelings of empathy, as well as the tendency to feel empathetic in daily life, predict greater activation in certain regions of the empathy network—those typically involved in representing and processing the empathizer’s own internal states (Saarela et al., 2007; Pfeifer et al., 2008; Hooker, et al. 2010). From responding with feeling to understanding what other peoples’ suffering expressions mean, regions of the empathy network are candidates of possible activation during compassion.

Beyond the initial empathic response, there is a second core process involved in compassion: the motivation to provide care to alleviate the other person’s suffering (Batson and Shaw, 1991; Eisenberg et al., 2002; Goetz et al., 2010). Hypotheses about neural correlates of this caregiving motivation are informed by neuroimaging studies of empathy toward in-group vs out-group victims: namely, certain regions of the empathy network show more robust activation in response to in-group victims likely to receive more care from empathizers. For example, Mathur and colleagues reported greater mPFC activation in response to the pain of others with in-group vs out-group membership, as well as greater altruistic motivation (Mathur et al., 2010). Another recent study showed heightened anterior cingulate, anterior insula and PAG activation during empathy toward the pain of a loved one as opposed to an unknown victim (Cheng et al., 2010). Finally, in an examination of how perceptions of responsibility affect empathy, innocent victims (who are likely to evoke greater care) led to greater activation in cingulate, anterior insula and PAG regions of the empathy network than blameworthy (less careworthy) victims (Decety et al., 2009). These studies, however, did not explicitly assess caregiving motivation, and thus do not unambiguously document a relationship between reported activation and caregiving motivation. An alternative interpretation is that the activation changes merely reflect a stronger empathic response wherein increased mirroring leads to increased engagement of social cognitive regions, i.e. more feelings and more accurate understanding. Nevertheless, a theme running across loved ones, targets of altruistic intent, and vulnerable victims, all of which known to increase activation within key regions of the empathy network, is the stronger inclination to provide care to reduce their suffering, the core appraisal that differentiates compassion from empathy (Goetz et al., 2010). If so, key regions of the empathy network may relate to a distinct process directly tied to caregiving motivation, that when engaged, produce the distinct experience of compassion.

Recent theorizing and neuroscientific inquiry into a possible caregiving system provide a second platform for making predictions about neural activation associated with compassion (Mikulincer et al., 2005; Swain, 2010). Studies of non-human mammals have implicated several key regions in parental nurturance behaviors (crouching over pups, retrieval to nest, licking, prolonged nursing), which include the medial preoptic area (MPOA), the ventral bed nucleus of the stria terminalis (VBST), and areas of the midbrain including the periaqueductal gray (PAG) (Lonstein and Stern, 1997; Numan and Numan, 1997; Sturgis and Bridges, 1997; Lonstein and De Vries, 2000; Stack et al., 2002). Oxytocin (OT), a neuropeptide, which is broadly implicated in the formation of attachment bonds, caregiving behavior and pro-social orientation affects functioning within these parental nurturance regions (Insel et al., 2001; Baumgartner et al., 2008; Carter et al., 2008; Rodrigues et al., 2009; Gordon et al., 2010). A recent study tied OT in the MPOA to increased reinforcement valuation for caregiving/nurturance behaviors in lactating rats (Shahrokh et al., 2010); another showed that infusing OT into the midbrain PAG led to decreased self-focused vigilance behavior in rat dams, presumably freeing up resources to support caregiving behaviors under conditions of threat (Figueira et al., 2008). In sum, these findings highlight the midbrain-preoptic–anterior hypothalamic neural circuits that support basic parental nurturance behaviors and the formation of early attachment bonds as a candidate system for supporting the formation of attachment bonds throughout life, as well as for generating caregiving motivation toward a wide spectrum of others.

Importantly, the midbrain PAG, while implicated in caregiving, is also involved in descending pain control through the release of endogenous opioids that inhibit ascending pain signals before they reach the cortex (Heinricher et al., 2009; Lovick and Adamec, 2009). Still, PAG activation has been reported during situations that do not involve pain to the self. Alongside other regions like the orbital frontal cortex and thalamus, PAG activation has been reported in several studies of mothers viewing images of their own vs acquainted or unknown infants, or viewing video clips of their own infants exhibiting attachment-figure soliciting behaviors like smiling and crying (Bartels and Zeki, 2004; Nitschke et al., 2004; Noriuchi et al., 2008; Swain, 2008). PAG activation was also observed, along with other empathy network regions, in participants instructed to generate ‘unconditional love’ toward images of disabled people (Beauregard et al., 2009), and while participants viewed sad facial expressions with instructions to extend a
‘compassionate attitude’ toward them (Kim et al., 2009). One study that examined neural activation during ‘relived’ compassion reported activation in insula, anterior cingulate, posterior medial cortex, thalamic and midbrain areas, though not explicitly reported as the PAG (Immordino-Yang et al., 2009). Finally, Harrison et al. reported an intriguing relationship between empathetic pupillary contagion, i.e. mirroring the pupil size of sad faces, and the magnitude of midbrain PAG activation; suggesting PAG involvement in early, automatic responding to other’s sadness. Although it plays a central role in pain regulation, the PAG is also involved in empathy and caregiving motivation; as part of the human parental caregiving system (Swain, 2010), the PAG may support caregiving motivation toward victims of suffering during experiences of compassion.

Pride, self-focus and status: predictions of neural correlates
Pride is most typically experienced when an individual, or a group related to the individual’s identity, rises in social status (Tracy and Robins, 2007). Pride is defined by enhancement in the individual’s standing relative to other individuals and social standards. In this regard, pride is at its core an individuating, self-focused emotion (Tracy et al., 2007). In light of this conceptual analysis, one would expect the experience of pride to engage regions associated with self-referential processes.

Neuroimaging studies of self-referential processes frequently report activation along the cortical midline, typically focused in medial prefrontal and posterior medial cortex (mPFC and PMC, respectively). MPFC activation has been reported in participants monitoring their own physical appearance or actions (Blakemore and Frith, 2003; Sugiuira et al., 2005b), making judgments about their own qualities (Kelley et al., 2002; Kjaer et al., 2002; Blakemore and Frith, 2003; Damasio, 2003; Fossati et al., 2003; Lieberman et al., 2004; Ochsner et al., 2004; Mitchell et al., 2005b; Heatherton et al., 2006; Johnson et al., 2006; Northoff et al., 2006; D’Argembeau et al., 2007), assessing the relevance of stimuli to their own goals (Montague et al., 2006; Harris et al., 2007), and in conjunction with hippocampal structures, to retrieving autobiographical information (Macrae et al., 2004; Dasselaar et al., 2008). Activation in the PMC has been related to assessing the familiarity of stimuli, retrieving episodic emotional information and to attributing emotional quality and personal relevance to perceived stimuli (Maddock et al., 2001, 2003; Sugiuira et al., 2005a; Immordino-Yang, et al., 2009). As a result, MPFC and PMC areas are theorized to support cognitions about the self and personal goals that are characteristic of pride (D’Argembeau et al., 2009).

An functional magnetic resonance imaging (fMRI) investigation aimed at measuring pride, in contrast to joy, presented participants with brief, status enhancing statements selected to elicit pride, e.g. ‘I won a prize at a scientific meeting’ (Takahashi et al., 2008). Compared to activation during the joy-eliciting statements, e.g. ‘I ate my favorite cake,’ pride led to activation in the posterior superior temporal sulcus (pSTS), an area implicated in perceiving social meaning, in theory of mind (TOM) and in interpreting motivation or intent from others’ actions (Pellegrino et al., 2005; Saxe and Powell, 2006; Lamm et al., 2007). The authors of this study concluded that pride involves appraisals of social meaning not inherent to joy, and attributed the lack of predicted, self-referential activation in medial cortex regions to participants’ failure to genuinely personalize the statements, and presumably, experience true self-focused pride (Takahashi, et al., 2008).

In light of theoretical claims that self-reference is central to pride and the literature on medial cortical areas and self-referential processing, we predicted that pride stimuli would selectively engage midline cortical structures, the mPFC and PMC.

Neuroimaging compassion and pride
In the present investigation, we induced compassion and pride using sequences of pre-tested picture slides found to elicit intense and relatively pure episodes of these two emotions (Oveis et al., 2010). We also collected real time subjective experience ratings for compassion, caring, pride, achievement, distress and enjoyment to validate the evocative properties of the slides and to evaluate relationships between self-reported experience and hypothesized regions of neural activation. Neural activation during compassion and pride slide conditions was contrasted with that produced by viewing sequences of neutral slides. Although the neutral slides were not expected to evoke caring or self-focus, thus providing a suitable contrast for compassion and pride slides, it is important to note that self-report data were collected in real time for all conditions. This task feature prompted ongoing self-monitoring across compassion, pride and neutral slide conditions. Thus, given the overlapping role of the mPFC in self-monitoring, social and emotional appraisal and possibly empathy-related processes (Mitchell et al., 2005a; Rudebeck et al., 2008), we note that predicted emotion-specific mPFC activation may not be observable from data acquired within this study design.

Given the present study’s design and our caregiving analysis of compassion, we predicted that compassion induction would engage regions of the empathy network that show greater activation toward in-group or vulnerable suffering such as the anterior insula, inferior frontal gyrus and the midbrain PAG, as well as other areas that support parental nurturance behaviors. In light of the self-referential core of pride, we hypothesized that pride would selectively engage self-referent cortical midline areas.

MATERIALS AND METHODS
Participants
Twenty students (11F) were recruited from a large, urban university population to participate in this study.
Participants were screened for right-handedness, current undergraduate enrollment at UC Berkeley, no history of neurological or psychiatric disorder or drug use, and negative pregnancy status. The use of human subjects was approved by the Committee for the Protection of Human Subjects at UC Berkeley and informed consent was obtained prior to participation. Participants were paid hourly for their participation.

**Study design and task**

The study design and task used here were adapted from a prior study that investigated the effects of induced compassion and pride on social judgment (Oveis et al., 2010). During fMRI acquisition, participants performed two repetitions of a blocked emotion induction, emotion self-report task as depicted in Figure 1. Instructions presented on-screen prior to the task included the following text ‘...after each sequence (of pictures), you will be asked to rate the intensity and the nature of any feelings that arose while you were viewing those pictures.’ The task consisted of 8, 55-s blocks of slide stimuli that alternated between emotionally evocative and neutral categories. A continuous sequence of 5, 11-s slides from a single category was presented during each block. Each slide block was preceded by the word ‘Ready...’ presented at central fixation for 11 s. Each block of slides was followed by a text display that prompted participants to rate the overall intensity of their emotional response to the preceding sequence of slides, and then a second display that prompted participants to rate the quality of their overall emotional experience as they viewed the preceding slides. The durations of both self-report pages were self-paced by each participant.

Emotion-inducing slide blocks progressed through compassion eliciting and pride eliciting categories, as well as pleasure and awe eliciting categories. Data from pleasure and awe slide blocks are not presented here. Neutral slide blocks were presented between each successive emotion-inducing slide block. The order of emotion-inducing slide blocks was randomized for each task repetition within, and between participants. Compassion slides featured depictions of vulnerable suffering and harm. Pride slides featured depictions of in-group, or untargeted (no pictured recipient, perspective directed toward the viewer) achievement and status including graduation scenes, home university victories, medals and trophies. The neutral slides depicted people in unexpressive, mundane contexts as well as ordinary objects like a filing cabinet (Figure 1).

The slide images used in the present study were drawn from a previous study, which had pre-tested over 150 slides, and validated a set of 15 compassion and 15 pride slides that successfully induced the target states, as well as distinct judgments of self-other similarity (Oveis et al., 2010). To refine and adapt this paradigm for use in the fMRI environment, we substituted several slides from the original set (e.g. medals and trophies replaced patriotic symbols from the original pride set) and gathered a set of neutral slides matched for social content (number of people depicted) and perceptual features (brightness, contrast). Substitute slides were drawn from the International Affective Picture System (Lang, 1999) and static images from the internet. Slide validation data from an independent sample of 36 participants who viewed the slides and reported upon their emotional experiences in response to each slide used in the present study was collected.
study showed that compassion, pride and neutral slides, respectively, produced reliably different self-reported levels of ‘compassion’ ($M = 4.3, \text{ s.e.} = 0.1$, $M = 1.9, \text{ s.e.} = 0.07$, $M = 1.6, \text{ s.e.} = 0.06$), ‘caring’ ($M = 3.9, \text{ s.e.} = 0.1$, $M = 2.0, \text{ s.e.} = 0.07$, $M = 1.7, \text{ s.e.} = 0.06$), ‘pride’ ($M = 1.3, \text{ s.e.} = 0.04$, $M = 3.2, \text{ s.e.} = 0.09$, $M = 1.3, \text{ s.e.} = 0.04$) and ‘achievement’ ($M = 1.2, \text{ s.e.} = 0.03$, $M = 3.0, \text{ s.e.} = 0.09$, $M = 1.4, \text{ s.e.} = 0.04$) (Figure 2a).

For the present study, a fixed set of 40 slides (5 for each emotion and 20 neutral) was used during each task run; 8 of the 10 compassion and matched neutral slides depicted people and 5 of the 10 pride and matched neutral slides depicted people. Task runs were presented in counterbalanced order between subjects. After the final slide in each block, participants provided self-reports of their emotional response to the slides that they had just viewed guided by the following prompt: ‘While I was viewing those pictures, I experienced feelings that were: 1 = completely neutral, 4 = moderate, and 7 = very intense,’ above a row of boxes labeled 1 through 7. Participants clicked an enumerated box between 1 and 7, after which they then clicked on a box labeled ‘Next’ to continue. Participants then rated their experience of several different emotions, including ‘achievement, admiration, awe, caring, compassion, distress, enjoyment, pride and savoring’. Emotion terms appeared in a vertical column on the left side of the screen, each flanked by row of boxes enumerated 1 through 7 labeled ‘1 = not at all, 4 = moderately and 7 = very strongly’ along the top of the grid. Self-reports of ‘compassion’, ‘caring’, ‘pride’ and ‘achievement’ served as manipulation checks for compassion and pride induction, and yielded data relevant to the question of whether the experiences of compassion and pride would correlate with activation in different neural systems. Self-reported ‘distress’ and ‘enjoyment’ provided further validation of the valence-level evocative properties of the slides.

Procedure
Prior to entering the fMRI scanning chamber, participants were debriefed about the study. Specifically, experimenters told participants that they would see pictures while lying in the scanner, and then be asked to indicate how the pictures had made them feel. Participants were told that the study aimed to measure neural activation associated with their authentic emotional responses, not with their beliefs about how they ought to respond, nor with their judgments regarding what kinds of emotions the pictures represented. Participants then completed a brief practice version of the slide-viewing, self-report task that featured one mixed-category block of slides followed by the intensity-of-feeling and quality-of-emotion self-report prompts.

**Fig. 2** Mean self-reports of ‘compassion’ and ‘pride’ in response to the compassion, pride and neutral slides for (a) the behavioral pilot sample responding to each individual slide and (b) the fMRI participants responding to slide sequences. Mean self-reports of valence, ‘distress’ and ‘enjoyment,’ for the fMRI participants responding to slide sequences are depicted in (c).
Experimenters reminded participants to respond naturally to the pictures and to be honest and forthright in their emotion self-report, and suggested that disingenuous self-report would be harmful to the study.

Once inside the fMRI scanner chamber, participants were positioned in a standard transverse electromagnetic (TEM) send-and-receive radio frequency (RF) head coil with an eye-level mirror for viewing the task presentation screen. Participants held an fMRI compatible laser mouse in their right hand resting on their right thigh for making self-report ratings after each slide sequence from within the scanner chamber.

**fMRI**

Functional whole-brain images (30 axial slices, anterior commissure/posterior commissure orientation) were acquired in a 4T Varian INOVA scanner with a gradient echo T2*-weighted 1-shot gradient-echo echo-planar imaging (EPI) sequence (repetition time 2 s, echo time 28 ms, flip angle 20, field of view 224 mm with a 64 x 64 matrix size resulting in an in-plane resolution of 3.5 x 3.5 mm for each 5-mm slice). High-resolution, in-plane T1-weighted anatomical images were acquired using a gradient-echo multislice sequence (GEMS) for anatomical localization and a T1-weighted 3D MPFlash sequence was acquired for data display. Preliminary image reconstruction including ghost correction, geometric distortion correction and temporal interpolation routines was performed with ReconTools, a Python-based tool developed at UC Berkeley’s Brain Imaging Center (https://cirl.berkeley.edu/view/BIC/ReconTools).

**Data analysis**

**Self-reports of emotion**

Self-reports of compassion (‘compassion’, ‘caring’), pride (‘pride’, ‘achievement’) and valence terms (‘enjoyment’, ‘distress’) in response to compassion, pride and neutral slide blocks were averaged across two task runs. Given high Cronbach’s α coefficients between target emotion terms, we created composite measures of the experience of ‘compassion’ (‘compassion’ and ‘caring’, α = 0.95) and ‘pride’ (‘pride’ and ‘achievement’, α = 0.94) for the remainder of analyses. Though correlations were also high between self-reports of ‘distress’ and ‘compassion’ and between ‘enjoyment’ and ‘pride’ (r = 0.51 and r = 0.79, respectively), data for ‘distress’ and ‘enjoyment’ terms were analyzed independently to enable specific assessment of valence. Self-reports of ‘compassion’, ‘pride’, ‘distress’ and ‘enjoyment’ were submitted to a one-way ANOVA with slide type (compassion, pride or neutral) as the independent variable and self-report values for these four terms as the dependent variables. Self-report data for one task run from one participant were excluded due to outlying levels of ‘distress’ and ‘pride’ in response to the pride slides (‘distress’ > 2 s.d.’s from the mean, ‘pride’ < 2 s.d.’s from the mean). The participant’s own described feelings about these pride slides during debriefing indicated an unusual and specific personal interpretation of the slides (namely, that the images of military medals were connected to conspiratorial ideas about politics). Self-report data from the second task run for that individual, which were within 2 s.d.s of the mean, were included.

**fMRI analysis**

Participants’ functional images were realigned to correct for head movement, smoothed (8-mm full-width at half-maximum), normalized (parameters from MPFlash to SPM2 T1 template were applied), then submitted to a whole-brain analyses using random-effects models in Statistical Parametric Mapping (SPM) (http://www.fil.ion.ucl.ac.uk/spm) to identify brain areas specifically activated during contrasts between each emotion and neutral slide conditions. fMRI data from three participants were excluded from all group analyses due to large head size or aberrant slice definition during acquisition, factors which produced increased frontal susceptibility artifact and caused poor image normalization results (n = 17). fMRI data from one participant during the pride condition were excluded from analyses due to outlying self-report values (n = 16 for pride fMRI analyses).

Emotion slide blocks, neutral slide blocks, self-report intervals and six-dimension movement parameters determined during the realignment process were included as covariates in each participant’s SPM model. Statistical analyses were performed using a modified general linear model (Worsley and Friston, 1995) reference function that convolved the time series of covariates for each task condition and the movement parameters for each task run with a block-wise, idealized hemodynamic response function. Contrast images for each participant were created for compassion > neutral, pride > neutral, compassion and pride combined > neutrals from both conditions and direct comparisons between compassion and pride. Contrast images were submitted to a second level, group-wise (n = 17 for compassion > neutral, n = 16 for pride > neutral, n = 16 for across and between-emotion comparisons) t-test for each contrast (P < 0.001, uncorrected; extent: 15 voxels). Simulation using the 3D ClustSim function in AFNI (http://afni.nimh.nih.gov/pub/dist/doc/program_help/3dClustSim.html; 10,000 iterations, 64 64 30 dimensions, 3.5 3.5 5 voxels, BALL mask applied, 8-mm smoothness) indicated that an uncorrected P < 0.001 threshold and voxel cluster extent = 12.2 would provide protection against Type I error at a false discovery rate (FDR) of 0.05 (Bennett et al., 2009; Lieberman and Cunningham, 2009). Contrast images were also submitted to a second level, group-wise (n = 17 for compassion > neutral, n = 16 for pride > neutral) multiple regression with constant analysis (P < 0.001, uncorrected; extent: 15 voxels) using individuals’ emotion self-report as predictors. Specifically, multiple regression was used to examine whether ‘compassion’ or ‘distress’ self-report
values predicted activation during compassion > neutral conditions and whether ‘pride’, ‘enjoyment’ or ‘compassion’ self-report values predicted specific activation during pride > neutral conditions.

RESULTS

Self-reports of the experience of emotion

As shown in Figure 2b, the compassion, pride and neutral slide blocks (self-report values for the two neutral slide blocks are combined) elicited the predicted patterns of emotion self-report. Compassion, pride and neutral slides, respectively, produced reliably different self-reported levels of ‘compassion’ (M = 5.8, s.e. = 0.3, M = 3.1, s.e. = 0.4, M = 2.0, s.e. = 0.2) and ‘pride’ (M = 1.6, s.e. = 0.2, M = 5.1, s.e. = 0.4, M = 2.1, s.e. = 0.3). One-way ANOVAs showed that slide type (compassion, pride or neutral) significantly affected self-reported levels of ‘compassion’, F(2, 50) = 38.9, P < 0.001, and ‘pride’, F(2, 50) = 42.3, P < 0.001. Post hoc comparisons between specific conditions indicated that each condition elicited: (i) higher ratings for target emotions (e.g. higher ‘compassion’ for the compassion slides than for the pride, or neutral slides, P < 0.001 for all comparisons) and (ii) low ratings for non-target emotions (e.g. low ‘pride’ for the compassion and neutral slides, P < 0.001 for all comparisons).

Though ‘compassion’ was significantly lower for the pride than for the compassion slide condition (P < 0.001), it was greater for pride than for the neutral slide condition (P < 0.05), suggesting that the lower than moderate self-rated level of ‘compassion’ in response to the pride slides was still greater than the very low ‘compassion’ ratings in response to neutral slides (self-reported ‘compassion’ is thus incorporated into regression analyses involving the pride > neutral fMRI contrast). Finally, no differences were observed between non-target ‘pride’ ratings for compassion and neutral slide conditions (P < 0.001 for all comparisons).

As evident from Figure 2c, compassion, pride and neutral slides also produced different self-report levels of affective valence measured as ‘distress’ (M = 5.2, s.e. = 0.3, M = 2.1, s.e. = 0.3, M = 1.7, s.e. = 0.19, respectively) and ‘enjoyment’ (M = 1.3, s.e. = 0.53, M = 4.8, s.e. = 0.36, M = 1.9, s.e. = 0.14). There was a main effect of task condition on self-reported ‘distress’, F(2, 50) = 44.9, P < 0.001; post hoc comparisons indicated significantly higher ‘distress’ for compassion than for pride or neutral slide conditions, P < 0.001 for both comparisons. There was also a main effect of task condition on self-reported ‘enjoyment’, F(2, 50) = 65.5, P < 0.001; post hoc comparisons indicated higher ‘enjoyment’ for pride than for compassion or neutral conditions, P < 0.001 for both comparisons. No differences between ‘distress’ ratings for pride and neutral conditions or between ‘enjoyment’ ratings for compassion and neutral conditions were observed.

fMRI data

Compared with the activation produced by viewing blocks of neutral slides, viewing blocks of compassion slides was associated with activation in the midbrain PAG. No significant activation was observed in any other voxel clusters for this contrast, even in predicted ‘empathy network’ or parental nurturance regions, at designated thresholds. During pride compared to neutral slide blocks, the posterior medial cortex showed increased activation as predicted; no significant activation was observed in any other voxel clusters. No significant activation was observed for the combination of compassion and pride vs neutral conditions at designated thresholds (P < 0.001, uncorrected; extent: 15 voxels) (Figure 3 and Table 1).

In direct comparisons between the activation associated with compassion and pride conditions, no voxel clusters were observed to be significantly more active during compassion than pride. Several regions typically implicated in self-reflective processing including the PMC, parahippocampal and inferior temporal regions showed greater activation during pride in contrast with compassion (Table 1.)

We next ran multiple regression analyses to examine whether the self-reported experiences of ‘compassion’ and ‘pride’ predicted neural activation during target emotion conditions. More specifically, we ran several multiple

![Fig. 3](image-url)
regression analyses to examine whether ‘compassion’ and ‘distress,’ both characteristic responses to suffering (Eisenberg et al., 1989; Goetz, et al., 2010), predicted distinct activation during the compassion > neutral condition contrast. These analyses revealed that (i) ‘compassion’ in combination with ‘distress’ predicted increased activation in a near-PAG region and (ii) ‘compassion’ controlling for ‘distress’ predicted increased activation within the right inferior frontal gyrus (IFG). Though the cluster foci are distinct, the near-PAG region predicted by greater ‘compassion’ and ‘distress’ overlapped with the PAG region activated during the group-wise compassion > neutral condition contrast (Figure 4). There were no negative relationships between ‘compassion’ and/or ‘distress’ self-reports and neural activation associated with the compassion > neutral condition contrast.

Guided by a similar logic, we ran a second set of multiple regression analyses to examine whether self-reported levels of ‘pride,’ ‘enjoyment,’ and ‘compassion’ predicted activation associated with the pride > neutral condition contrast. No positive correlation was observed between self-report levels and pride > neutral contrast activation. Notable negative relationships, however, were observed (i) between all three predictors, ‘pride,’ ‘enjoyment,’ and ‘compassion,’ and the right IFG (Brodmann area 47) and a closely neighboring ventral aspect of the right lateral prefrontal cortex (Brodmann area 46), (ii) between ‘pride’ and ‘enjoyment’ controlling for ‘compassion’ and the same right frontal regions, as well as the left anterior insula and (iii) between ‘enjoyment’ controlling for ‘pride’ and ‘compassion’ and the left anterior insula. In contrast to the results for ‘compassion’ controlling for ‘distress,’ participants who reported greater feelings of ‘pride’ and ‘enjoyment’ independent of ‘compassion’ for pride slides showed reduced activation in right IFG/prefrontal regions (Figure 5 and Table 2).

In a final set of multiple regression analyses, we examined whether self-reports of ‘compassion’ and ‘pride’ across the compassion and pride conditions predicted specific patterns of activation in a contrast that combined compassion and pride > neutral condition contrasts. No significant positive or negative relationships were observed in this analysis at designated thresholds (P < 0.001, uncorrected; extent: 15 voxels).

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Brain region</th>
<th>MNI coordinates</th>
<th>t-statistic</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compassion &gt; neutral</td>
<td>Midbrain PAG</td>
<td>0, -32, -8</td>
<td>5.7</td>
<td>20</td>
</tr>
<tr>
<td>Pride &gt; neutral</td>
<td>Posterior Medial Cortex</td>
<td>2, -64, 16</td>
<td>4.17</td>
<td>25</td>
</tr>
<tr>
<td>Compassion and pride &gt; neutral</td>
<td>L cerebellum</td>
<td>-28, -42, -28</td>
<td>5.75</td>
<td>46</td>
</tr>
<tr>
<td>Compassion &gt; pride</td>
<td>L Inferior temporal gyrus</td>
<td>8, -72, 10</td>
<td>5.35</td>
<td>43</td>
</tr>
<tr>
<td>Pride &gt; compassion</td>
<td>Posterior medial cortex</td>
<td>23/31, -8, -68, 24</td>
<td>4.74</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>L Parahippocampal</td>
<td>-22, -20, 4</td>
<td>4.4</td>
<td>18</td>
</tr>
</tbody>
</table>

The MNI coordinates of the maximally active focus within each structure are reported. BA = Brodmann’s Area, L = Left, R = Right, Nsv = no significant voxels. P < 0.001, uncorrected; extent: 15 voxels.

**DISCUSSION**

The present research sought to document patterns of neural activation associated with compassion, an emotion defined by empathic response and caregiving motivation, and pride, a self-focused emotion focused on elevated status. Toward this end, we used validated emotionally evocative slides to elicit relatively intense and pure experiences of compassion and pride.

Guided by appraisal analyses of the two emotions of interest, we predicted that experiences of compassion would be associated with increased activation in ‘empathy network’ regions (particularly regions responsive to a greater sense of care), as well as regions involved in parental nurturance behaviors. The latter prediction was supported by the data from the present study: laboratory induction of compassion relative to a neutral condition was associated with increased activation in the midbrain PAG, an area implicated across empathy and parental care-taking behaviors. Further regression analyses showed that self-reported experiences of ‘compassion’ and ‘distress,’ two core processes involved in pro-social responses to suffering (Batson and Shaw, 1991), also predicted activation in a near-PAG region.

These findings are in keeping recent studies that report PAG activation during heightened empathy for other people’s pain and during attachment-related states, namely maternal love and unconditional love. They also dovetail with a non-human literature that has implicated the PAG, via OT connections with MPOA and VBST regions, in motivating, reinforcing and freeing up resources for parental nurturance
behaviors (Figueira et al., 2008; Rosen et al., 2008; Lovick and Adamec, 2009; Shahrokh et al., 2010).

The simple conclusion that the PAG supports the caregiving tendencies that define compassion is clouded by the PAG’s well-known role in pain modulation. In the present study, it may have been that the PAG was activated due to the participant’s personal, simulated experience of pain when viewing the images of suffering (Singer and Frith, 2005).

Fig. 4 Group wise main effects activation in the PAG during the compassion > neutral (yellow) condition contrast and a near-PAG cluster predicted by self-reported ‘compassion’ and ‘distress’ for the compassion > neutral contrast (cyan). Activation is displayed at \( P < 0.005 \) (uncorrected; extent: 15 voxels) to show the extent of activation and is displayed on a mean anatomical image calculated across fMRI participants. Self-reported ‘compassion’ and ‘distress’ for the compassion > neutral condition contrast are plotted against mean activation across voxels in the near-PAG cluster \( P < 0.001 \) for correlation r values unless otherwise labeled.

Fig. 5 Activation in the right IFG predicted by greater self-reported ‘compassion’ (red) for the compassion > neutral condition contrast, and by lower self-reported ‘pride’ and ‘enjoyment’ (yellow) for the pride > neutral condition contrast. Activation is displayed at \( P < 0.005 \) (uncorrected; extent: 15 voxels) to show the extent of activation and is displayed on a mean anatomical image calculated across fMRI participants. Self-reported ‘pride’ and ‘enjoyment’ for the pride > neutral condition contrast, and ‘compassion’ for the compassion > neutral condition contrast are plotted against mean activation across voxels in the right IFG regions \( P < 0.001 \) for correlation r values unless otherwise labeled.
Indeed, participants reported high levels of ‘distress’ alongside ‘compassion’ while viewing the images of suffering. Multiple regression analyses of ‘compassion’ and ‘distress’ showed activation in a cluster of voxels partially overlapping with the PAG during the group-wise compassion > neutral contrast. This combination of findings suggests that PAG activation during compassion might be related to several compassion-related processes, including the empathic mirroring of another’s suffering, felt personal pain or distress in response to perceived suffering, and/or caregiving motivation. Future studies that (i) acquire additional self-report to identify appraisals related to distress, e.g. enduring personal threat or pain vs concern toward other person’s suffering and (ii) directly contrast conditions more purely evocative of empathized personal pain or distress, i.e. induced empathy with implied threat to the self, with explicit prevention to caregiving, with no manipulation and with support for caregiving could more definitively tease apart the relative involvement of the PAG in pain, empathy, distress and caregiving processes.

The lack of observed activation in other parental nurturing regions like the MPOA and VBST in the compassion > neutral fMRI contrast analyses is noteworthy. Lack of engagement of other parental nurturing regions may relate to the complex valence properties of compassion. As elicited here, compassion was associated with self-reported ‘distress’, and not with ‘enjoyment’. Participants did not have the opportunity to help or act on their inclination to care for the targets of compassion here, which may have attenuated any anticipated reward associated with predicted social benefits of caregiving. While OT function in the MPOA has been associated with reinforcing fundamentally neutral or mildly pleasant nurturance behaviors, OT/PAG functions may figure more prominently in caregiving motivation particularly in response to pain and suffering, key antecedents of compassion. Direct examination of how parental nurturance systems respond to care evoking situations of different valence, and with or without the opportunity to intervene, could clarify these concerns.

In addition to the PAG findings, multiple regression analyses showed that self-reported levels of ‘compassion’ controlling for self-reported ‘distress’ predicted activation in an anterior region of the right IFG—pars triangularis. This finding is interesting in light of findings relating right IFG activation to emotion inference, regulation, as well as empathic mirroring of facial expression (Carr et al., 2003; Hooker et al., 2008; Berkman and Lieberman, 2009). A recent study of patients with IFG lesions reported impaired affective empathy and emotion recognition; the authors theorized that Brodmann area 44 of the IFG, the pars opercularis, the human analog to ‘mirror neuron’ primate area F5 (slightly caudal to the IFG regions observed here), plays a principal role in an emotion simulation system (Shamay-Tsoory et al., 2009). Increased right IFG activation predicted by self-reported ‘compassion’ here suggests that greater compassion experience may relate to stronger mirroring of expressions shown by the targets of compassion. The IFG is also implicated in cognitive control of memory processes (Badre and Wagner, 2007); IFG activation during compassion may also relate to controlled memory processes engaged in response to the contents of evocative slides.

Contrary to expectations generated our appraisal analysis of compassion, activation was not observed in other empathy network regions, such as the insula, mPFC or the TPC during the group-wise compassion > neutral contrast at the designated threshold (P < 0.001, uncorrected; extent: 15 voxels).
This could, in part, have arisen as a function of our emotion induction approach. The task required that participants report on their emotions after each block in all of the conditions, which likely prompted ongoing self-monitoring, self-evaluation and appraisal, processes which typically engage the mPFC (Hutcherson et al., 2005; Northoff et al., 2006). This design feature likely engaged self-evaluative and affective simulation regions like the insula and mPFC during neutral, as well as emotion slide conditions, thereby attenuating observable differences in contrasts between conditions. Similarly, use of slides that depicted people in social settings across neutral and emotion-inducing conditions likely engaged areas of TPC involved in perception of social meaning across conditions (Pelphrey et al., 2005; Zahn et al., 2007). With regards to other empathy network regions, particularly those that have been shown to be sensitive to caregiving orientation, at a speculative, exploratory threshold (P < 0.005, uncorrected; extent: 10 voxels), activation clusters in the insula (36, 6, 22; 12 voxels), anterior temporal (−32, 2, −32; 21 voxels) and left prefrontal regions (32, 50, 10; 14 voxels) were observed. Compassion may indeed engage additional empathy network areas; evidence from future studies of the compassionate response will be more definitive with respect to this possibility.

In keeping with our second set of hypotheses, activation was observed in the posterior medial cortex (PMC), an area implicated in self-reflection, during the pride > neutral contrast (Daselaar et al., 2001; Maddock et al., 2001; Johnson et al., 2006; D’Argembeau et al., 2008). Directly compared to compassion eliciting slides, the pride slides, which included images of symbolic success like medals and trophies, also elicited activation in left inferior temporal and parahippocampal regions. The engagement of autobiographical memory processes is an important component of self-focus, was probably evoked by the pride slides (e.g. ‘I remember winning a medal like this one at a swim competition.’) and likely accounts for the parahippocampal findings observed here. Future investigations of pride resulting from actual increases in personal strength or status will clarify the role of hippocampal activation and memory-related processes in the specific experience of pride.

As was the case with compassion, activation in the mPFC, a region strongly associated with self-appraisal, was not observed during pride. This finding could have been due to the self-reflective demands required of the neutral condition, as discussed earlier (Ochsner et al., 2004). Overlapping, task-related self-referential processes across pride and neutral conditions may have yielded a ceiling effect for mPFC activation, minimizing emotion-specific observable differences between conditions.

Though no positive relationship was observed between self-reported ‘pride’, ‘enjoyment’ or ‘compassion’ in response to the pride slides and neural activation during pride (vs neutral) conditions, lower ratings on pride-related terms (‘pride’, ‘enjoyment’) predicted greater activation in the right IFG, a closely neighboring ventral aspect of the right prefrontal cortex, and anterior insula regions. Anterior insula and IFG activation are routinely observed in emotion expression identification and empathy eliciting paradigms, are sensitive to degree of care toward a social target, and are posited to play key roles in emotion sharing, and by some accounts, empathic accuracy (Carr et al., 2003; Iacoboni et al., 2005; Saarela et al., 2007). Less activation in the anterior insula and IFG with greater pride experience may relate to self-focus, less concern toward others, and by extension, fewer resources dedicated to processing emotional information from others. This finding is in keeping with recent studies finding that pride is associated with a reduced sense of similarity to, and in turn, caregiving sentiment toward others (Oveis et al., 2010).

Examination of activation across combined compassion and pride vs neutral conditions revealed no systematic activation at designated thresholds. At a speculative, exploratory threshold (P < 0.005, uncorrected; extent: 10 voxels), activation was observed during this combined analysis in empathy network regions associated with perception and interpretation of social meaning or intent including the right superior temporal sulcus (STS), rIFG and fusiform gyrus (Pelphrey, et al., 2005; Saxe and Wexler 2005; Hooker et al., 2008). This trend may reflect processing of social meaning across compassionate and pride evoking vs neutral slides. Multiple regression analyses to examine whether the combination of self-reported ‘compassion’ and ‘pride’ across the compassion and pride slide conditions predicted specific activation also showed no significant result. In effect, these analyses asked whether individuals prone to the experience of compassion or pride show distinct activation in response to emotionally evocative slides. Although the present study was not designed to explore this possibility (and the low levels of ‘pride’ felt during the compassion slides are problematic), several recent studies have found that individual differences in emotion guide specific appraisals of broad classes of stimuli, even those unrelated to the emotion of interest (Gross et al., 1998; Oveis et al., 2009). This literature would suggest that individuals prone to compassion, as assessed in our self-report data, would show compassion-related patterns of activation even in response to the pride slides, and that complementary results would be observed with individuals who report high levels of pride even in response to the compassion slides. It may be more fruitful to look for ‘emotion-proneness’-related patterns of neural activation by studying responses to prototypical stimuli or a broader range of emotion categories.

Inferences drawn from the present findings are constrained by certain features of the study design. First, for the purposes of gathering validation data, our emotion elicitation procedures required that participants report on their own emotional experiences, likely engaging participants in sustained monitoring of their own states. Requiring participants to monitor themselves may have interfered with...
emotion-related neural activity (Northoff et al., 2004; Hutcherson et al., 2005). Some brain regions hypothesized to support monitoring and self-reporting of emotions such as the mPFC, were also expected to be active during experiences of compassion and pride. Future research would be well served by studying the patterns of neural activation associated with compassion and pride in the absence of anticipated self-report procedures.

Second, our protocol featured unusually long fMRI blocks. Acquiring long blocks of fMRI data biases observations toward regions that show persistent activation over the entire interval (54 s blocks vs typical 15–30 s blocks). Longer induction intervals are likely to produce stronger experiences of emotion and a greater likelihood of activating core appraisals associated with the emotion, such as caregiving motivation in the case of compassion. Longer block designs also risk greater overlap low frequency artifact inherent to fMRI data (D’Esposito, 2000). Here, each participant’s data was carefully inspected for quality, and high pass filtered at 220 s (4 x 55 s block length) to minimize the contribution of low frequency artifact.

Third, it is fair to raise doubts about the ecological validity of our emotion-eliciting stimuli—static slides of social situations not directly relevant to the participant’s concerns (Coan and Allen, 2007). Slides are ideally suited to the environment of fMRI data acquisition, and in the case of compassion and pride, control for significant behavioral differences of the two emotions (e.g. the approach-related tendency of compassion vs the social disengagement associated with pride). Nevertheless, the use of slides introduced certain limitations with the present study’s findings. Pride, for example, is conceptualized as a response to heightened status; our slides were limited in their ability to directly convey changes in status. Further, pride slides included images of students from the home university enthusiasm for celebration, which likely elicited collective, in-group pride. These concerns are mitigated, somewhat, by the self-report data gathered in the present study, and the fact that self-reports of ‘compassion’ and ‘pride’ predicted theoretically relevant regions of activation. Still, greater confidence will be gained if the present research is extended to studies using different emotion elicitation techniques. Studies of neural activation during emotional states elicited using alternative techniques, like emotional imagery, reliving and guided autobiographical recall (Damasio et al., 2000; Immordino-Yang et al., 2009), point to intriguing extensions of the present work on the neural correlates of elicited emotions.

Finally, patterns of significant activation in only one voxel cluster for the compassion > neutral and pride > neutral contrasts, or none at all in the case of the combined compassion and pride analysis, confers some degree of ambiguity to interpretation of the results reported here. Generally, complex social emotions like compassion and pride are expected to engage parallel brain networks supporting concurrent affect and appraisal processes. The lack of multiple regional activation here, may, in part be related to the long fMRI block approach used here, which may average out fluctuating engagement of multiple regions, leaving only signals from regions making the most enduring, tonic contribution to the processes. The lack of significant activation for combined compassion and pride may be related to variability in subjective experience of these complex emotions, or to variance in temporal properties of regional activation supporting fluctuating, complex affective and cognitive processes across the long emotion elicitation blocks. Further studies with different time constraints, specifically designs that independently distinctly evoke specific affect and appraisal-related component processes that factor into compassion and pride will clarify interpretation of the results reported here.

The work presented here provides a first within-subjects examination of neural activation during laboratory induction of compassion and pride. The present findings point to the possible involvement of the midbrain PAG, insula and IFG regions vs posterior midline structures in states characterized by other-oriented, caregiving motivation vs status-based self-focus, respectively. As this literature on compassion and pride grows, several opportunities await: studies of gender differences in the neural correlates of these potentially gendered emotions; trait and state-level variations in these emotions and their correlates; effects of social group status on the emergence of these emotions; and potential linkages between the central nervous system processing and possible peripheral correlates of compassion and pride. This kind of work will help delineate the physiological processes associated with compassion and pride, two emotions that serve important social functions.

REFERENCES


