OXFORD

Empathy and moral emotions in post-apartheid South Africa: an fMRI investigation

Melike M. Fourie,[1] Dan J. Stein,[2] Mark Solms,[3] Pumla Gobodo-Madikizela,[1] and Jean Decety[4]

[1]Studies in Historical Trauma and Transformation, Stellenbosch University, Stellenbosch, South Africa, [2]Department of Psychiatry and MRC Unit on Anxiety and Stress Disorders, University of Cape Town, Cape Town, South Africa, [3]Department of Psychology, University of Cape Town, Cape Town, South Africa, and [4]Department of Psychology and Department of Psychiatry and Behavioral Neuroscience, University of Chicago, Chicago, IL, USA

Correspondence should be addressed to Melike M. Fourie, Studies in Historical Trauma and Transformation, Stellenbosch University, Stellenbosch 7600, South Africa. E-mail: marethem@gmail.com.

Abstract

Moral emotions elicited in response to others' suffering are mediated by empathy and affect how we respond to their pain. South Africa provides a unique opportunity to study group processes given its racially divided past. The present study seeks insights into aspects of the moral brain by investigating behavioral and functional MRI responses of White and Black South Africans who lived through apartheid to in- and out-group physical and social pain. Whereas the physical pain task featured faces expressing dynamic suffering, the social pain task featured victims of apartheid violence from the South African Truth and Reconciliation Commission to elicit heartfelt emotion. Black participants' behavioral responses were suggestive of in-group favoritism, whereas White participants' responses were apparently egalitarian. However, all participants showed significant in-group biases in activation in the amygdala (physical pain), as well as areas involved in mental state representation, including the precuneus, temporoparietal junction (TPJ) and frontal pole (physical and social pain). Additionally, Black participants reacted with heightened moral indignation to own-race suffering, whereas White participants reacted with heightened shame to Black suffering, which was associated with blunted neural empathic responding. These findings provide ecologically valid insights into some behavioral and brain processes involved in complex moral situations.

Key words: empathy; moral emotion; fMRI; prejudice; group membership

Introduction

Moral emotions in response to social violations are considered powerful motivational forces (Huebner *et al.*, 2009). When we see someone suffering, for example, we may respond with concern, distress, guilt or even pleasure (depending on the context)—each resulting in different action tendencies. South Africa's history of racial discrimination during apartheid provides a tragic example.

Moral emotions evoked by witnessing the distress of others are mediated by empathy—a key moral emotional process (Tangney *et al.*, 2007). For example, empathy for a victim, combined with an awareness that a moral standard was violated, is likely to result in moral indignation. Decety and Cowell (2014) recently argued that it is critical to distinguish between the affective (experience sharing), cognitive (perspective taking) and motivational (empathic concern) components of empathy, as each uniquely affects moral responses. Equally critical, is to

881

distinguish between physical and emotional distress. Neuroimaging work on empathy has largely focused on the former (Lamm *et al.*, 2011), whereas less research has addressed the socially more relevant phenomenon of empathy for others' emotional distress—a phenomenon referred to as 'social pain' (Eisenberger, 2012). While a handful of studies have investigated brain response to stimuli depicting real emotional distress, like suffering due to natural disaster or states of grief (Immordino-Yang *et al.*, 2009; Mathur *et al.*, 2010; Cheon *et al.*, 2011), research into social pain has typically employed an exclusion game to mimic real-life social discrimination (Williams *et al.*, 2000). The latter paradigms have raised concerns about ecological validity (Risko *et al.*, 2012).

The present study seeks insights into aspects of the moral brain by investigating emotional responses to genuine social distress. Two tasks were employed to elicit empathy and moral emotions in the scanner, in White and Black individuals from the South African population. Task 1 consisted of a well-validated protocol of Black and White faces expressing dynamic suffering to elicit empathy for physical pain. Task 2 featured short video clips from the South African Truth and Reconciliation Commission (TRC) showing Black and White victims of apartheid violence in genuine emotional distress to elicit empathy for social pain. The TRC was established by the post-apartheid government to promote national reconciliation and healing through public hearings, where victims and perpetrators of apartheid atrocities recalled their experiences (Boraine *et al.*, 1997). We collected moral emotional and hemodynamic responses to TRC clips with the aim of measuring effects of racial (in/out) group membership on emotional responses to others' distress. Participants also completed behavioral measures of prejudice to enable assessment of the relationship between empathy, moral emotions, explicit prejudice and brain activity.

Study aims

First, the present study aimed to shed light on the neural correlates of empathy for ecologically valid social pain. A prominent theory in social neuroscience maintains that our nociceptive system has been co-opted by the social attachment system to detect and prevent separation (Eisenberger, 2015). However, this notion is challenged by investigators who have demonstrated that nociceptive pain and social rejection involve distinct neural representations (Cacioppo *et al.*, 2013).

Likewise, the extent to which *empathy* for physical *vs* social pain share common neural substrates remains unclear. Empathy for physical pain consistently recruits a core network consisting of the anterior mid-cingulate cortex (aMCC) and the anterior insulae (aINS) (Lamm *et al.*, 2011). These regions relate to the representation of affective distress associated with the first-hand experience of pain, although some studies suggest they are not specific to pain perception (Mouraux *et al.*, 2011; Salomons *et al.*, 2016). In contrast, studies exploring empathy for social pain have found activation in the 'mentalizing network' (Masten *et al.*, 2011; Zaki and Ochsner, 2012), although activation in areas associated with both the affective (Beeney *et al.*, 2011; Meyer *et al.*, 2013) and sensory discriminative (Novembre *et al.*, 2015) aspects of pain processing have also been observed.

Second, our study explored the effects of race on behavioral and neural measures of empathy and moral emotions in relation to in/out-group distress. Previous work suggests that people tend to favor members of social groups with which they identify (Tajfel and Turner, 1979). Moreover, several studies

have demonstrated stronger hemodynamic activation in response to others' pain for racial in- *vs* out-groups in regions associated with affective distress, as well as mentalizing [medial prefrontal cortex (MPFC), temporoparietal junction (TPJ) and precuneus] (Xu *et al.*, 2009; Mathur *et al.*, 2010; Cheon *et al.*, 2011; Azevedo *et al.*, 2013). It remains unclear, however, to what extent behavioral responses and neural activation to others' pain converge, given that in-group biases often operate outside awareness (Devine, 1989).

Finally, our study explored associations between neural activation responses, moral emotions, and explicit prejudice. Whereas the relationship between empathy and prosocial motivation has been studied (Mathur *et al.*, 2010; Masten *et al.*, 2011; Morelli *et al.*, 2014), no study to date has explored the relationships between evoked moral emotions and neural activations associated with empathy for social pain.

In light of the above, we hypothesized that neural activation responses to others' physical and social pain will both engage the aMCC/supplementary motor area (SMA) and aINS, and that social pain will recruit the mentalizing network more strongly. With regard to group-related responses, we anticipated that hemodynamic activation would be stronger for in-group than out-group members particularly in areas associated with mentalizing, rather than those associated with affective distress, given recent evidence showing aMCC and aINS activation is related to emotional salience regardless of the relationship to the person in distress (Fox *et al.*, 2013). We furthermore anticipated that behavioral empathic responses, especially for White participants, would not show this pattern of in-group bias, because of pervasive social pressures to uphold egalitarian norms in South Africa. With regard to moral emotions, we predicted that Black and White participants would be associated with different emotional overlays: given South Africa's history, we expected Black participants to experience heightened moral indignation when viewing in-group members' suffering. In contrast, we expected White participants to experience heightened guilt and shame in response to Black suffering.

Methods

Participants

Thirty-eight individuals, 19 Black-African (11 female, $M = 40.11$ years, s.d. $= 4.12$) and 19 White-Caucasian (10 female, $M = 41.47$ years, s.d. $= 5.80$) completed all study procedures and received ZAR200 compensation. Henceforth, we use the terms 'Black' and 'White' as they were defined during apartheid by the South African Population Registration Act of 1950, which divided the population into four racial groups: Whites, Blacks (Natives), Indians and Coloureds (people of mixed racial ancestry). All participants lived in South Africa during apartheid (prior to 1994) and obtained Grade 10 as a minimum level of education (Black: $M = 15.89$ years, s.d. $= 2.79$; White: $M = 15.74$ years, s.d. $= 3.04$). Participants were without previously diagnosed neurological, cardiovascular or psychiatric disorders, and none were clinically depressed (Beck *et al.*, 1996).

All participants provided informed consent. The study was approved by the University of Cape Town's Human Research Ethics Committee.

Prejudice measures

Four rating thermometers were used to assess participants' explicit attitudes toward racial groups defined during apartheid:

White, Black, Colored and Asian/Indian (Herek, 2000). White participants' motivations to behave without prejudice toward Black people were also assessed using the Internal and External Motivation to Respond Without Prejudice scales (IMS/EMS; Plant and Devine, 1998).

To reduce social desirability and priming effects, questionnaires were completed without information about the purpose of the study several weeks before the scan during a screening interview.

fMRI stimuli

Task 1:facial expressions of physical pain. Stimuli consisted of 40 video clips 2.2 s in duration showing the faces of 10 Black and 10 White males dynamically expressing either pain or no pain (neutral), validated with an independent group of 40 healthy volunteers (Decety *et al.*, 2014). Clips were presented three at a time in blocked format, separated by a 500 ms scrambled static grey image and superseded by a 3 s title indicating the nature of the block (e.g. Black Pain). No clip was repeated more than once. The interval between blocks lasted 8 s during which participants fixated at a central cross. No online rating was obtained during this task due to scan-time constraints.

Task 2:social pain. Stimuli for Task 2 consisted of short 6–9 s video clips (352 pixels × 288 pixels) showing Black/White individuals in emotional distress or neutral scenarios. Distress clips were sourced from the TRC hearings or related documentaries, whereas neutral clips were taken from interviews unrelated to the TRC, e.g. university curricula. TRC clips portrayed victims in distress due to either loss of a loved one, physical or sexual violence, and were validated by a mixed-race sample of 49 volunteers in an independent behavioral experiment (see Supplementary Material). The scanning paradigm followed the same design as Task 1: clips were presented in blocks of three, separated by a 500 ms scrambled static grey image and cued by a 3 s title. Clips in each block were similarly distressing in terms of their emotion ratings obtained in the validation study, and no clip was repeated more than once. Participants rated their empathic concern after each block of clips via button press: they indicated how 'sorry' they felt for the last person on a Likert-type rating scale from 1 (*not sorry at all*) to 4 (*extremely sorry*). The interblock intervals following each rating were 10 s.

Procedure

Participants received standardized instructions about the purpose of the study, namely, to understand how the human brain responds to others' pain. Accordingly, they learned that they would watch videos depicting people in different kinds of pain: task 1 involved people in physical pain, whereas Task 2 involved people in emotional distress taken from the TRC hearings. To ensure participants understood the context of the TRC clips, they (i) received scripted contextual information about the TRC, (ii) were shown a 2 min introduction clip about the TRC and (iii) watched video material from Task 2 paired with a contextual description of each individual that were not shown again in the scanner (e.g. *This person's son was killed by state security police*). Participants were also familiarized with the button box used in Task 2. They were instructed to empathize with all individuals in pain or distress and to stay emotionally engaged during the scan.

In the scanner, participants viewed stimuli through a mirror system mounted to the head coil. *E-Prime* software (Psychology Software Tools, Inc.) was used to display the stimuli and to record participants' behavioral responses via a button box. Participants first performed one run of Task 1, which consisted of four blocks per condition (16 in total) in pseudorandomized order. They then underwent a structural scan, followed by three runs of Task 2 counterbalanced across participants. Each run consisted of eight blocks (2 per condition), with the order of conditions randomized within a half run.

Post-scan emotion ratings

After the scan, participants rated 4 clips per condition from each task according to various scales. For Task 1, participants rated the perceived intensity of each person's pain on visual analog scales ranging from 1 (*no pain*) to 9 (*extreme pain*). For Task 2, participants reported how much they felt distressed (personal distress), sorry (empathic concern), angry (moral indignation), guilty and ashamed in response to each clip on separate 1 (*not at all*) to 9 (*extremely*) visual analog scales. Participants were then debriefed and offered the opportunity to see a counsellor, should they wish to discuss their experiences.

fMRI image acquisition and data analysis

MRI data were acquired on a 3T Allegra system (Siemens, Erlangen, Germany). The high-resolution anatomical scan was acquired with a T_1-weighted sequence (3D mprage, TR/TE = 2530/6.5 ms). Functional images covering the whole brain for Task 1 and 2 were acquired with a T_2*-weighted echo-planar (EPI) imaging sequence using blood-oxygenation-level-dependent (BOLD) contrast (TR/TE = 2000/30 ms, slice thickness = 3 mm, gap = 0.9 mm, flip angle = 90°, field of view = 240 × 240 mm). The first four volumes of each run were discarded to allow for T_1 equilibration effects.

All fMRI analyses were performed using Brain Voyager QX, version 2.8 (Brain Innovation, Maastricht, Netherlands). Preprocessing of images included correction for slice acquisition times and linear trends, temporal filtering with a high-pass filter of 2 cycles/point, and motion-correction relative to the first volume of each run with trilinear/sinc interpolation. No run exceeded 3 mm displacement/3.0° rotation. Participants' functional data sets were co-registered with their structural MRI and spatially normalized to Talaraich space.

Whole-brain group analyses were performed with a random effects analysis of variance using the general linear model (GLM) with predictors corresponding to known experimental blocks convolved by the standard hemodynamic response function. We defined predictors for the four task conditions of Black Pain/Distress, White Pain/Distress, Black Neutral, White Neutral and fixation. Task 2 also included a predictor of no interest corresponding to the emotion rating period. The six z-transformed motion correction parameters were added as predictors of no interest to reduce motion artifacts.

For both tasks, the resulting estimated beta values were entered into a second-level three-factor mixed factorial ANOVA, with the between-subjects factor Participant Race (Black *vs* White), and the within-subjects factors Pain/Distress (physical/social pain *vs* neutral) and Victim Race (same-race *vs* other-race). To evaluate task specific activations, the main effects of Pain (Task 1) and Distress (Task 2) (i.e. the contrast physical/social pain > neutral) were inspected. We report clusters that are greater than 150 contiguous voxels ($1 \times 1 \times 1$ mm^3 resolution of structural images) at $P < 0.01$, voxel-wise corrected for multiple comparisons using the false discovery rate (FDR).

To identify brain regions that responded differentially to in- *vs* out-group physical or social pain, we examined three-way interaction effects (Participant Race × Pain/Distress × Victim Race). All results are reported at $P < 0.05$, corrected for multiple comparisons using the Monte Carlo cluster threshold estimation simulation tool implemented in Brain Voyager running 1000 iterations (Forman *et al.*, 1995). Because of the exploratory nature of these interaction analyses, we applied cluster-level thresholding at an uncorrected $P < 0.05$.

Region of interest analyses were performed for all clusters that emerged in the whole-brain 3-way interaction analyses (shown in Supplementary Table S6). Random effects analysis of variance was performed on the average signal in each cluster for each participant using the GLM described above. Beta values generated by this analysis (reflecting the mean percent signal change for each condition) were analyzed by three-way ANOVA and plotted to show neural activation levels between conditions.

To enhance the power of analyses, we defined independent ROIs from the network most consistently associated with pain empathy (aMCC and aINS; Lamm *et al.*, 2011), as well as the SMA, because another meta-analysis suggested the SMA forms part of the core network of empathy, irrespective of the emotion empathized with (Fan *et al.*, 2011). We also chose ROIs for the amygdala, because previous studies report differential amygdala activation in response to racial in- *vs* out-groups (Chekroud *et al.*, 2014). All regions were defined as spheres with 10 mm radiuses centered at the peak voxel in each cluster, except for the amygdalae, where spheres had 5 mm radiuses (see Supplementary Table S2 for coordinates). ROI analyses for these independently selected regions were also conducted as described above.

To examine relationships between behavioral measures and brain activity, Pearson's correlations between ROI beta values for the Pain/Distress conditions and behavioral data (prejudice measures and emotion ratings) were inspected. To limit the number of correlations performed (Curtin and Schulz, 1998), ROIs included in these analyses were those previously identified as representing the core network of empathy (aMCC, aINS and SMA). In addition, we included the precuneus, derived from the whole-brain interaction analyses for Task 1 and 2, respectively, because of its apparent functional significance in responding differentially to in- *vs* out-group physical and social pain. Only correlation coefficients reflecting large effect sizes ($r \geq \pm 0.50$) were interpreted and further inspected using 95% confidence intervals (CIs) derived through bootstrapping. Finally, we conducted exploratory mediation analyses to define resulting associations (Baron and Kenny, 1986).

Results

Behavioral data

Thermometer ratings. Thermometer ratings were assessed using a 2 (Participant Race: Black *vs* White) × 4 (Thermometer Rating: White, Black, 'Coloured', Asian/Indian) mixed factorial ANOVA. The interaction was significant, $F(3,108) = 13.13$, $P < 0.001$: preference for Blacks compared to Asians/Indians, 'Coloureds', and Whites, was greater for Black than White participants ($Ps \leq 0.001$, $rs > 0.60$) (Figure 1). One-way repeated-measures ANOVAs indicated that Black participants' preference for Blacks was greater than for other racial groups ($Ps < 0.001$, $rs > 0.73$), whereas no significant preference differences existed for White participants ($P = 0.06$). Because Black participants' attitudes
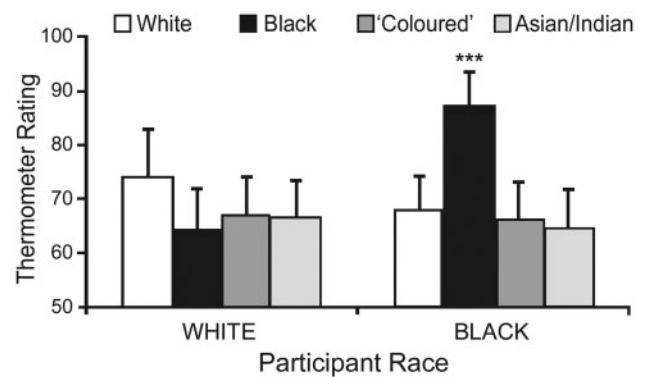

Fig. 1. Black participants demonstrated in-group favoritism: their attitudes (thermometer ratings) toward Black people were significantly more favorable than their attitudes toward other racial groups (***Ps < 0.001). Attitude ratings ranged from 0 (extremely unfavorable) to 100 (extremely favorable). Error bars indicate 95% confidence intervals.

toward out-group members were not unfavorable per se ($ts < 1.1$, $Ps > 0.30$), this phenomenon is best described as in-group favoritism (Hewstone *et al.*, 2002).

Task 1 pain ratings. Pain intensity ratings were assessed using a 2 (Participant Race: Black *vs* White) × 2 (Pain: physical pain *vs* neutral) × 2 (Victim Race: same-race *vs* other-race race) mixed factorial ANOVA. Pain intensity ratings were higher for painful than neutral facial expressions ($P < 0.001$), and Black participants' ratings were higher than those of White participants ($P < 0.01$). The main effect of victim race was not significant ($P = 0.97$), however, nor were there any significant interactions. All participants therefore rated the perceived physical pain of Black and White victims as similar (Supplementary Table S3).

Task 2 subjective emotion ratings. Within-scan emotion ratings confirmed that TRC clips ($M = 3.42$, s.d. $= 0.56$) elicited significantly greater empathic concern than neutral clips ($M = 1.40$, s.d. $= 0.53$) (Wilcoxon's signed rank: $P < 0.001$).

To examine racial biases in empathic responding, we computed an in-group empathy bias index using post-scan empathic concern ratings (see Supplementary Material). This measure showed in-group empathy bias for Black, but not White, participants ($P < 0.001$) (Supplementary Figure S1). Again, Black participants' in-group empathy bias is best described as in-group favoritism, as participant groups did not differ significantly in their empathic concern for White victims ($P = 0.16$).

Post-scan subjective emotion ratings were examined using a 2 (Participant Race: Black *vs* White) × 2 (Victim Race: same-race *vs* other-race) × 5 (Emotion: empathic concern, personal distress, moral indignation, guilt and shame) mixed factorial ANOVA (Table 1). Black participants' emotion ratings were higher than those of White participants ($P < 0.05$), and all emotion ratings were higher in response to Black than White victims ($P < 0.001$). The main effect of emotion was also significant, $F(2.79, 100.56) = 56.34$, $P < 0.001$, $\epsilon = 0.70$, with empathic concern rated as significantly higher than other emotions ($Ps < 0.001$, $rs > 0.73$).

Finally, the three-way interaction was significant, $F(4,144) = 10.43$, $P < 0.001$. Black compared to White participants' increases in emotion ratings for Black compared to White victims' distress were greater for moral indignation than any other emotion ($Ps < 0.05$, $rs > 0.36$). Black participants thus experienced heightened moral indignation when viewing in-group

Table 1. Task 2 moral emotions experienced in response to Black and White victims in distress

Emotion	Black participants ($n = 19$)		White participants ($n = 19$)	
	Black victim	White victim	Black victim	White victim
Moral indignation	7.91 (2.13)	5.62 (2.28)	4.79 (2.05)	4.21 (2.33)
Personal distress[a]	7.29 (2.16)	5.78 (2.01)	5.69 (1.84)	4.94 (1.71)
Empathic concern	9.00 (0.53)	7.62 (1.46)	7.54 (1.33)	6.92 (1.51)
Guilt	4.13 (2.80)	3.91 (2.73)	4.07 (1.88)	2.65 (1.88)
Shame	5.92 (2.78)	5.38 (2.59)	4.59 (2.19)	3.17 (1.89)

Note. Data presented are means, with standard deviations in parentheses. Ratings ranged from 1 (not at all) to 9 (extremely).
[a]Personal distress is not regarded as a moral emotion, but rather as a self-oriented, aversive reaction.
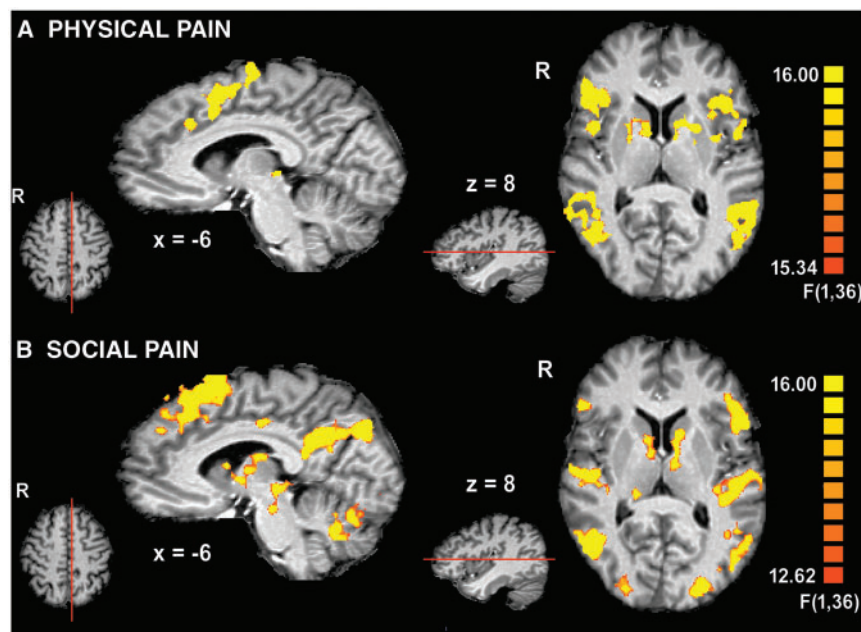

Fig. 2. Main effects of (A) perceived physical pain and (B) social pain. Contrast: physical/social pain$_{(Black + White)}$ > neutral$_{(Black + White)}$, thresholded at q(FDR) < 0.01 and 150 contiguous voxels.

members in distress, because this emotion showed the greatest difference for in-group relative to out-group victims. In contrast, White compared to Black participants' increases in emotion ratings for Black compared to White victims' distress were greater for guilt and shame than other emotions (Ps < 0.01, rs > 0.43). White participants thus experienced heightened guilt and shame when viewing out-group distress. Interestingly, White participants' guilt and shame in response to Black distress were positively correlated with their EMS scores (rs > 0.62, Ps < 0.01).

fMRI data

Whole-brain main effects of pain. The main effect of perceived physical pain (Task 1) across participants showed increased activity in areas typically associated with the perception of pain (including the affective component): aMCC, aINS, SMA, inferior frontal gyrus (IFG), precentral gyrus, thalamus and PAG (Figure 2A, Supplementary Table S4).

The main effect of social pain (Task 2) was also associated with increased activity in areas related to pain perception: SMA, IFG, precentral gyrus, thalamus and PAG (Figure 2B, Supplementary Table S5). In addition, significant activation was

observed in the DMPFC, precuneus, posterior cingulate cortex (PCC), amygdala, posterior superior temporal sulcus (pSTS) and temporal poles. These areas are implicated in mentalizing.

There were no other significant whole-brain main effects.

Group differences in neural responding

Whole-brain interaction effects. For Task 1, the precuneus, retrosplenial cortex (RSC), frontal pole, and TPJ showed significant 3-way interaction effects, although the latter did not survive cluster-level thresholding (Supplementary Table S6). Beta values extracted from these functionally defined ROIs and analyzed by three-way ANOVA confirmed interaction effects (Fs > 8.50, Ps < 0.01), qualified by greater increases in activation from the neutral to pain condition for same-race compared to other-race victims for both Black and White participants (Figure 3). These regions thus displayed in-group biases in activation for perceived physical pain.

For Task 2, several regions showed significant 3-way interaction effects: precuneus, TPJ, IFG, and intraparietal sulcus (IPS) (Supplementary Table S6). Analysis of beta-estimates from these functionally defined ROIs confirmed the interaction
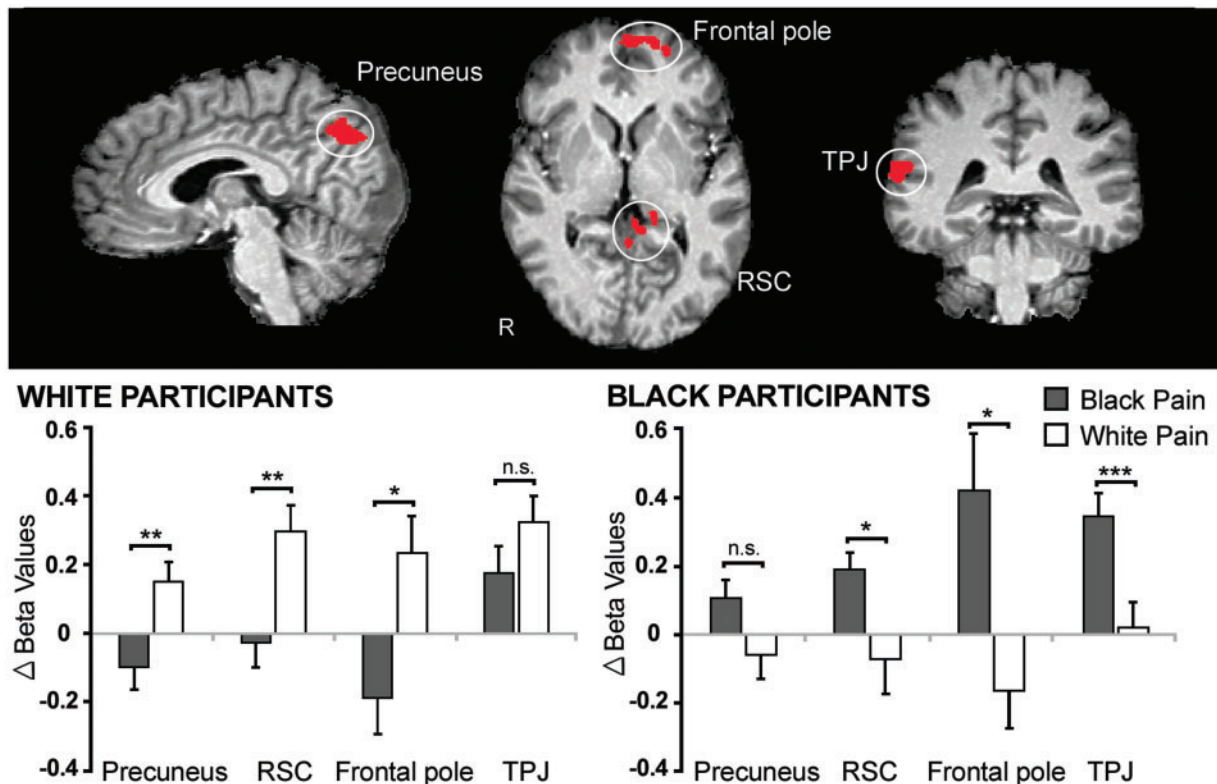
Fig. 3. Regions showing in-group biases in activation for perceived physical pain (Task 1). Activation in the precuneus, retrosplenial cortex (RSC), frontal pole and temporoparietal junction (TPJ) were significantly greater when viewing in-group compared to out-group members in pain for both Black and White participants ($P < 0.05$ corrected for multiple comparisons using Monte Carlo cluster-level thresholding). Parameter estimates (betas) plotted below reflect the change in signal (pain – neutral) for each cluster for each condition. Significance levels reflect results of paired t-tests. *$P < 0.05$. **$P < 0.01$. ***$P < 0.001$. n.s. = not significant.

effects ($Fs > 11.00$, $Ps < 0.01$), qualified by greater increases in activation from the neutral to distress condition for same-race compared to other-race victims for both Black and White participants (Figure 4). These regions thus displayed in-group biases in activation for social pain.

Independent ROI analyses. ROI analyses for Task 1 detected only significant main effects of pain for the aMCC, aINS and SMA ($Fs > 13.50$, $Ps < 0.001$), but the amygdalae revealed significant in-group biases in activation: in addition to significant main effects of pain ($Ps < 0.05$) and victim race ($Ps < 0.01$), the interaction between participant race and victim race was significant for both amygdalae ($Fs > 9.00$, $Ps < 0.01$). Extracted beta-estimates confirmed that Black and White participants' amygdala reactivity were greater for same-race than other-race individuals (Figure 5).

ROI analyses for Task 2 detected significant main effects of distress for the aMCC, aINS and SMA ($Fs > 9.00$, $Ps < 0.01$), plus significant three-way interactions in the left aINS and SMA ($Fs > 4.30$, $Ps < 0.05$), qualified by greater increases in activation from the neutral to distress conditions for same-race compared to other-race victims for both Black and White participants. ROI analyses of the amygdalae revealed only significant main effects of distress ($Ps < 0.001$).

Correlational analysis

IMS and EMS. For both tasks, we detected significant negative correlations between White participants' EMS scores and activation in the aMCC, SMA and precuneus in response to Black

physical/social pain ($rs > -0.50$, $Ps < 0.05$). IMS scores did not correlate significantly with activity in any ROIs.

In-group preference. To determine whether explicit racial attitudes were associated with neural empathic responding, we computed an in-group preference index using participants' thermometer ratings: In-group$_{Rating}$ – Out-group$_{Rating}$. Black participants' preference for same-race individuals ($M = 21.50$, s.d. = 16.46) were significantly greater than those of White participants ($M = 9.74$, s.d. = 18.82) (Mann–Whitney U: $P < 0.05$).

Black participants' in-group preference index was also positively associated with activation in the aMCC and SMA ($rs > 0.50$, $Ps < 0.05$), and the precuneus ($r = 0.65$, $P < 0.01$) during Black Distress (Task 2).

In the White group, this index did not correlate significantly with activity in any ROIs.

Task 2 subjective emotion ratings. In the White group, we detected significant negative correlations between self-reported guilt and shame in response to Black distress and activity in ROIs representing the core empathy network and the precuneus (Table 2, Figure 6).

Further analyses by way of partial correlations revealed that shame ratings remained negatively correlated with activity in the aMCC, SMA and precuneus when guilt ratings were covaried ($rs > -0.40$, $Ps < 0.05$, 1-tailed). In contrast, the effects of guilt did not remain significant when shame was covaried ($Ps > 0.25$, 1-tailed). Moreover, exploratory mediation analyses indicated that shame mediated the relationship between EMS scores and activity in the aMCC, SMA and precuneus. For each region, shame predicted neural activity while controlling for EMS
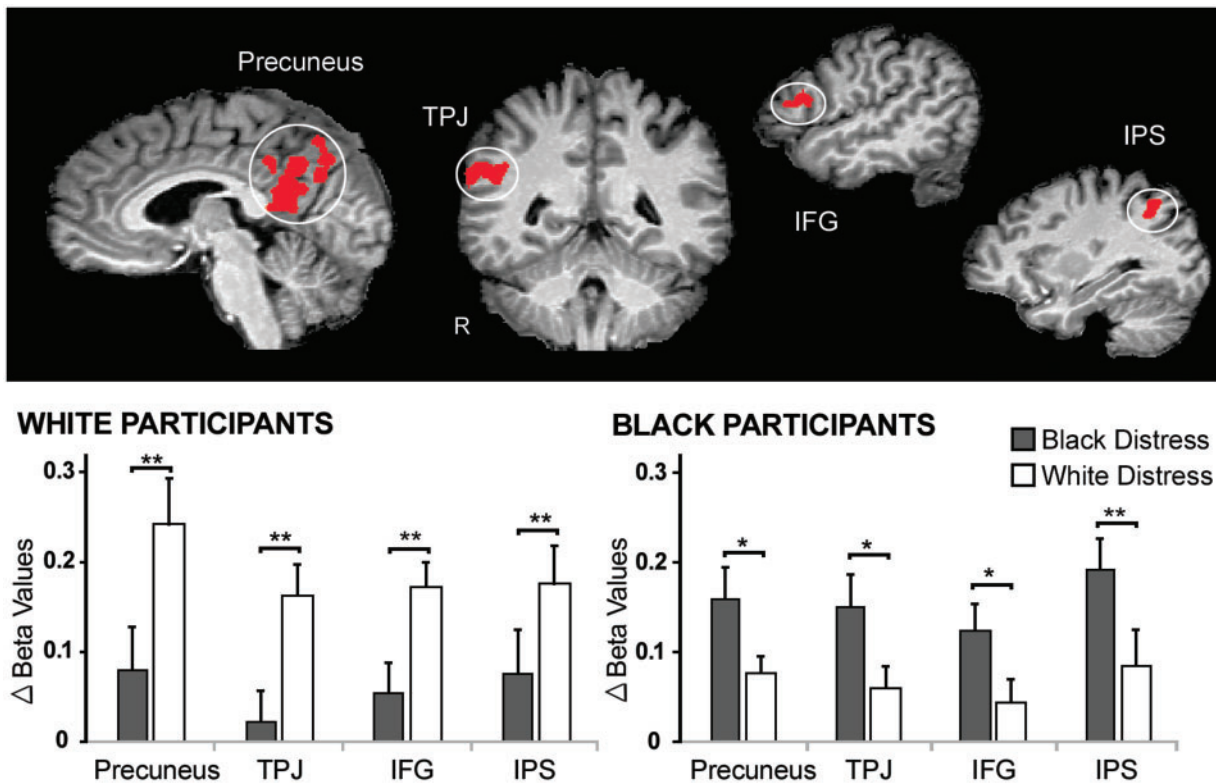
Fig. 4. Regions showing in-group biases in activation for social pain (Task 2). Activation in the precuneus, temporoparietal junction (TPJ), inferior frontal gyrus (IFG) and intraparietal sulcus (IPS) were significantly greater when viewing in-group compared to out-group victims in distress for both Black and White participants ($P < 0.05$ corrected for multiple comparisons using Monte Carlo cluster-level thresholding). Parameter estimates (betas) plotted below reflects the change in signal (distress – neutral) for each cluster for each condition. Significance levels reflect results of paired *t*-tests. *$P < 0.05$. **$P < 0.01$.
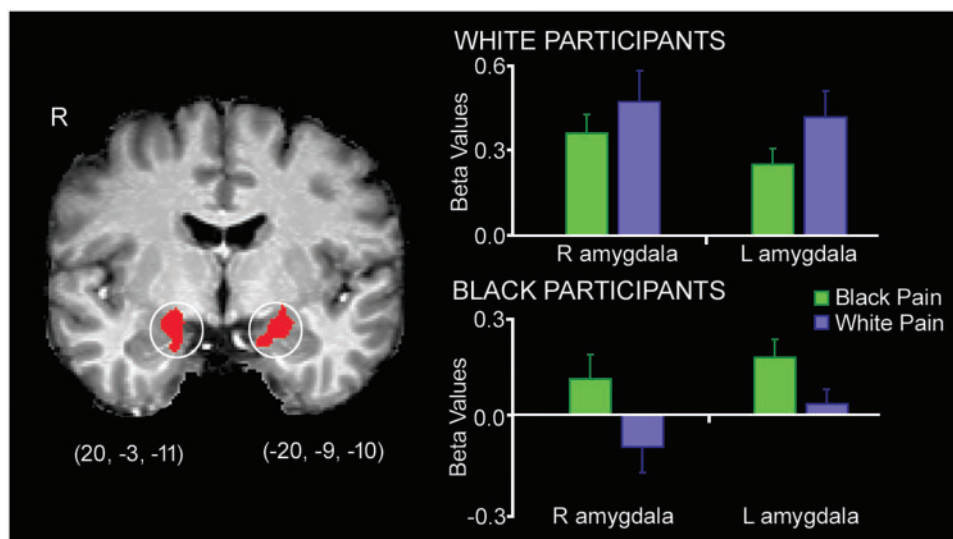

Fig. 5. In-group biases in amygdala reactivity for Task 1. Parameter estimates (betas) reflect the signal intensity in the amygdalae when participants perceived Black and White individuals in physical pain. ROI peak voxels defined based on Lamm *et al.* (2011).

scores (βs > -0.60, Ps < 0.05), while rendering the relationship between EMS scores and neural activation non-significant (βs < -0.10, Ps > 0.60).

Black participants' subjective emotion ratings did not correlate significantly with activity in any ROIs.

Discussion

Here we explored White and Black South Africans' emotional responses to each other's pain. First, empathy for ecologically valid social pain involved more complex neocortical processing than empathy for physical pain. Second, Black participants'

Table 2. Pearson's correlations between hemodynamic responses and moral emotions in response to black distress (white participants: $n = 19$)

	Region of interest					
	R aINS	L aINS	aMCC	R SMA	L SMA	Precuneus
Moral indignation	−0.06 [−0.47, 0.43]	−0.11 [−0.53, 0.44]	−0.17 [−0.56, 0.41]	−0.25 [−0.56, 0.23]	−0.16 [−0.64, 0.32]	−0.25 [−0.62, 0.37]
Personal distress	−0.06 [−0.56, 0.40]	−0.19 [−0.60, 0.45]	−0.26 [−0.70, 0.25]	−0.32 [−0.68, 0.27]	−0.22 [−0.76, 0.23]	−0.40 [−0.72, 0.15]
Empathic concern	0.05 [−0.58, 0.39]	0.04 [−0.53, 0.44]	0.00 [−0.55, 0.41]	−0.11 [−0.55, 0.41]	−0.10 [−0.61, 0.29]	−0.02 [−0.55, 0.39]
Guilt	−0.55* [−0.79, −0.22]	−0.50* [−0.78, −0.08]	−0.60** [−0.80, −0.35]	−0.65** [−0.85, −0.34]	−0.62** [−0.80, −0.37]	−0.57* [−0.77, −0.23]
Shame	−0.54* [−0.78, −0.15]	−0.50* [−0.81, −0.01]	−0.71** [−0.89, −0.39]	−0.71** [−0.86, −0.38]	−0.74** [−0.90, −0.44]	−0.65** [−0.87, −0.24]

Note. Data presented are correlation coefficients, with 95% bootstrap CIs in brackets.

[a]MCC, anterior midcingulate cortex; aINS, anterior insula; SMA, supplementary motor area.
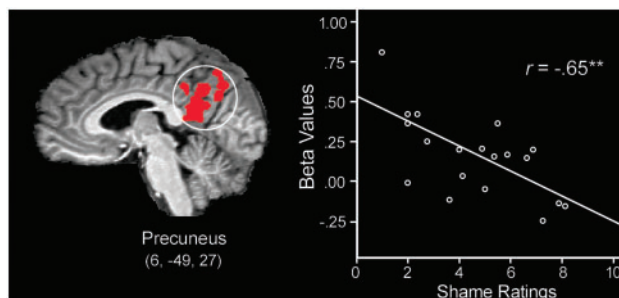
*$P < 0.05$; ** $P < 0.01$.

Fig. 6. White participants' mean activation level (betas) observed within the precuneus and self-reported shame in response to Black Distress (Task 2) is negatively correlated. The line represents the linear best fit.

behavioral responses were suggestive of in-group favoritism, whereas White participants' were more egalitarian. However, all participants showed significant in-group biases in the amygdala, as well as areas that allow for mental state representation. Furthermore, Black participants reacted with heightened moral indignation to own-race suffering, whereas White participants reacted with heightened shame to Black suffering. Finally, neural activation was modulated significantly by a priori racial attitudes, motivations for unprejudiced behavior and shame. These data reveal enduring, possibly implicit, effects of pervasive racial discrimination in terms of inter-group empathy and moral emotional responding.

Empathy for social pain

Like empathy for physical pain, empathy for social pain resulted in increased activity in areas associated with affective distress, including the aMCC and aINS (ROI analysis), as well as the SMA (whole-brain analysis), which may be important in activating motor programs (Shackman *et al.*, 2011). The genuine emotional distress conveyed by TRC victims probably accounts for activation in these areas, as well as those in the amygdala and PAG, even though victims were strangers. Previous studies have mostly detected activation in affective pain areas when those in distress were emotionally close (Beeney *et al.*, 2011; Meyer *et al.*, 2013). Our data suggest that it is not emotional closeness per se, but rather emotional salience that lead to activation in these

areas. Likewise, recent theorizing on the functional significance of aMCC and aINS activation in pain studies suggest that these structures form part of a broader multimodal network that responds to salient sensory stimuli regardless of the perception of pain (Legrain *et al.*, 2011; Iannetti *et al.*, 2013).

Of significance is that witnessing others in emotional distress recruited areas associated with mentalizing, including the DMPFC, pSTS, temporal poles and precuneus/PCC (Schurz *et al.*, 2014). However, accumulative evidence links a similar network also to self-referential processing and the default mode of brain function (Northoff *et al.*, 2006; Mars *et al.*, 2012). Moreover, episodic recall and reflecting upon future events (one's own and others') engage a similar network, consisting mainly of midline frontal regions, medial and lateral parietal regions, and medial temporal lobe structures (Buckner and Carroll, 2007; Mitchell, 2009). Hence, it has been suggested that these structures facilitate mentally projecting ourselves into the subjective reality of others, enabling the sharing of their state based upon personal experiences.

The discussion above highlights two important aspects of participants' empathic responses to the social distress task (i.e. perceiving victims of gross human rights violations), which may differ qualitatively and quantitatively from previous studies examining emotional responses to social pain. First, unlike most previous studies, activation in the aMCC and aINS, as well as the amygdala and PAG, in response to strangers in distress testifies to the emotional poignancy of participants' responses. Second, extensive activation in the DMPFC and other mentalizing regions point to the cognitively demanding nature of our task. That is, it could be argued that our task required more cognitive processing necessary to understand the distress of a person suffering from human rights abuses, which also taps into memory and other self-reflective processes, than, for example, witnessing someone being excluded in a computer game. Importantly, several participants reported that viewing the TRC material 'triggered memories' and 'opened wounds again'. Moreover, reliving personal social pain has recently been linked to increased DMPFC activity (Meyer *et al.*, 2015).

Group and individual differences

Intergroup bias has been shown to impact emotional responding toward others when it matters most—when they are in

pain. Yet response differences across social categories are not inevitable. Recent work demonstrates the impact of culturally-acquired prejudices (Mathur *et al.*, 2010; Cheon *et al.*, 2011; Azevedo *et al.*, 2013; Zuo and Han, 2013), and inter-racial conflict (Gobodo-Madikizela, 2008, 2015) in shaping empathic responses in intergroup contexts. Given these findings, we wanted to understand what factors influence the way in which Black and White South Africans respond (implicitly) to each other's pain.

In our study, behavioral (explicit) and imaging (implicit) responses differed in interesting ways. Whereas behavioral empathy ratings for Task 1 were not indicative of in-group biases, empathic concern ratings for Task 2 were characterized by significant in-group biases in Black (but not White) participants. All participants showed striking differences at the neural level, however, with several areas showing in-group biases in activation: amygdala (Task 1), aINS and SMA (Task 2), and areas involved in mentalizing (Tasks 1 and 2).

The behavioral in-group serving tendency of Black participants (thermometer and empathic concern ratings), took the form of in-group favoritism rather than out-group derogation (Hewstone *et al.*, 2002). While this conclusion cannot be confirmed definitively without a control group baseline (Cikara *et al.*, 2014), no differences existed between Black and White participants' general attitudes toward Whites, nor their empathic concern for White distress. At the neural level, Black participants' in-group preference was also positively correlated with neural activation in the core network of empathy, and the precuneus, in response to Black distress.

Black participants' in-group favoritism may reflect greater group identification than White participants. The Rejection–Identification model posits that pervasive discrimination against members of disadvantaged groups leads to increased in-group identification, which may alleviate psychological distress resulting from societal dehumanization (Branscombe *et al.*, 1999; Schmitt and Branscombe, 2002). In-group identification and intergroup bias have long been positively related (Hewstone *et al.*, 2002). More recently, greater identification with one's group has also been associated with extraordinary empathy for in-group members (Mathur *et al.*, 2010).

In contrast with Black participants, White participants' explicit and implicit responses diverged on both tasks. This pattern of results may be indicative of social pressure to uphold egalitarian norms and not necessarily of overt racial prejudice. Several studies have shown that in-group biases can operate implicitly. That is, while people may declare egalitarian views, they may simultaneously harbor racial biases of which they are unaware (Phelps *et al.*, 2000; Ito and Bartholow, 2009).

Our data furthermore show that White participants' neural responses toward Black others in distress were influenced by motivations to respond without prejudice. People may be (internally) motivated by sincere changes in their personal attitude, or by (external) social pressures (Plant and Devine, 1998). Data from both tasks suggested that higher EMS scores were associated with reduced activation in core empathy areas, and the precuneus, in response to Black distress. Several previous studies have demonstrated the moderating effects of internal/external motivations to respond without prejudice on implicit measures of prejudice (Devine *et al.*, 2002; Amodio *et al.*, 2008). In particular, despite egalitarian self-reports, high-EMS individuals appear to be just as ineffective in controlling automatic prejudiced responses on implicit measures as individuals who are explicitly racist. Our results extend these findings, suggesting that high-EMS individuals also have dampened neural empathic responses to out-group members.

Both tasks showed in-group biases in areas associated with mentalizing and the default mode network, including the precuneus and RSC, TPJ and frontal pole. As described above, areas in these networks are implicated in self-referential processing, episodic memory retrieval, and thinking about other minds, and may thus allow for a richer representation of another's physical/psychological pain (Saxe and Wexler, 2005; Buckner and Carroll, 2007). It should be noted that the TPJ, IFG and IPS also participate in attention networks dedicated to processing salient, behaviorally relevant events (Vossel *et al.*, 2014).

The precuneus may function as a critical hub in the network that distinguishes neural responses when we empathize with 'us' *vs* 'them'. Despite relatively little attention in functional accounts, converging evidence suggests that the precuneus contributes mental imagery to represent the perspective of another (Schurz *et al.*, 2014) . In our study, precuneus activity was consistently greater for in-group members. Moreover, activity in this area appeared sensitive to top-down influences, like motivation to respond without prejudice and in-group preference. Precuneus activity was evidently modulated by the perceiver's convictions.

Amygdala responses suggested that participants experienced heightened arousal for own- *vs* other-race individuals in physical pain. In the prejudice literature, amygdala activation typically shows the opposite pattern: heightened activation in response to racial out-group members, which may reflect a threat response (Chekroud *et al.*, 2014). In our study, enhanced amygdala activation toward same-race others likely reflects approach-related motivation and attention in line with task demands (Amodio, 2014). The fact that amygdala in-group biases were observed for physical and not social pain suggests that implicit prejudices affecting emotional arousal operated more freely in this condition—consistent with the notion that empathy for physical pain relies on more reflexive processes (Eisenberger, 2012).

Finally, the frontal pole showed preferential activity for in-group members in physical pain. Because the MPFC is strongly associated with mentalizing (Van Overwalle, 2009), *lack* of MPFC activity (including the frontal pole) in response to a social target may indicate diminished mental state attribution—a form of prejudice characterized by dehumanization and therefore, impaired empathy (Harris and Fiske, 2006; Cikara *et al.*, 2011). Frontal pole in-group biases during Task 1 suggest that mental state attribution toward out-group members is less readily engaged during more automatic experiences when stereotypes may prevail.

Moral emotions

Empathy-mediated moral emotions distinguished Black and White participants' responses in important ways.

White participants' responses to Black others in distress were characterized by heightened guilt and shame, which appeared to diminish out-group empathic responding: heightened guilt, and especially shame, were associated with reduced activation in the core network of empathy, and the precuneus. Moreover, mediation analyses suggested that shame mediated the relationship between external motivation to respond without prejudice and neural activation responses to Black distress. These analyses suggest strongly that White participants' feelings of shame were associated with blunted out-group empathic responding.

Guilt and shame are not always experienced in relation to personal transgressions. To the extent that group identification

defines who we are, it is possible to construe the behavior of in-group members as reflecting on the self (Salice and Montes Sánchez, 2016). Research on collective guilt and shame (feelings in response to others' transgressions), suggests that these phenomena also parallel personal guilt and shame. One notable example is guilt and shame's differing relationships to empathy: whereas guilt is associated with other-oriented concern and motivation to make amends, shame involves an egocentric focus on one's own distress, which derails the empathic process (Tangney *et al.*, 2007). With shame, the global self is perceived as fundamentally flawed, causing feelings of worthlessness and incompetence (Lewis, 1971). Thus unlike guilt, shame does not motivate reparative behaviors, but rather facilitates reactions such as hiding, denial, or escape of the shame-inducing situation (Tangney, 1991; Lickel *et al.*, 2005). Speculatively, the inverse relationship observed between shame and several areas that appear pivotal for empathy may thus reflect a kind of disengagement, such that the shamed person is less able to focus cognitive and emotional resources on the wounded other.

With regard to Black participants, it is plausible that the observed heightened moral indignation to in-group distress was in fact directed toward White perpetrators who are still, to some extent, responsible for social inequality in South Africa. Indeed, during debriefing sessions, high levels of current indignation amongst many Black participants surfaced toward White people.

Conclusions

Empathy for genuine social pain involves significant cognitive complexity and may require additional mental effort (and time) for introspective processing of culturally shaped knowledge—an attribute that may limit full experience thereof, particularly toward out-group members. Our neuroimaging findings indicate that, irrespective of explicit self-reports, group membership affects how readily we project ourselves into the reality of another to share/understand their psychological state. Moreover, group membership profoundly impacts moral emotional responding, and thus, behavior. These findings highlight the complexities of inter-group emotional responding in a context characterized by racial inequality. The intergroup landscape is not uniformly bleak, however. Several studies suggest that the more people become aware of their implicit racial biases, the more they are able to regulate them (Molenberghs, 2013). Our data stress the limitations of external motivations for egalitarian behavior, as well as feelings of shame, which undermine genuine, cross-racial empathy.

Supplementary data

Supplementary data are available at *SCAN* online.

Conflict of interest. None declared.

Funding

References

Amodio, D.M. (2014). The neuroscience of prejudice and stereotyping. *Nature Reviews Neuroscience*, **15**, 670–82.

Amodio, D.M., Devine, P.G., Harmon-Jones, E. (2008). Individual differences in the regulation of intergroup bias: the role of conflict monitoring and neural signals for control. *Journal of Personality and Social Psychology*, **94**, 60–74.

Azevedo, R.T., Macaluso, E., Avenanti, A., Santangelo, V., Cazzato, V., Aglioti, S.M. (2013). Their pain is not our pain: brain and autonomic correlates of empathic resonance with the pain of same and different race individuals. *Human Brain Mapping*, **34**, 3168–81.

Baron, R.M., Kenny, D.A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, **51**, 1173–82.

Beck, A.T., Steer, R.A., Brown, G.K. (1996). *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation.

Beeney, J.E., Franklin, R.G., Jr., Levy, K.N., Adams, R.B. Jr. (2011). I feel your pain: emotional closeness modulates neural responses to empathically experienced rejection. *Social Neuroscience*, **6**, 369–76.

Boraine, A., Levy, J., Scheffer, R. (1997). *Dealing with the Past: Truth and Reconciliation in South Africa*, 2nd edn, Cape Town: IDASA.

Branscombe, N.R., Schmitt, M.T., Harvey, R.D. (1999). Perceiving pervasive discrimination among African Americans: implications for group identification and well-being. *Journal of Personality and Social Psychology*, **77**, 135–49.

Buckner, R.L., Carroll, D.C. (2007). Self-projection and the brain. *Trends in Cognitive Sciences*, **11**, 49–57.

Cacioppo, S., Frum, C., Asp, E., Weiss, R.M., Lewis, J.W., Cacioppo, J.T. (2013). A quantitative meta-analysis of functional imaging studies of social rejection. *Scientific Reports*, **3**, 2027.

Chekroud, A.M., Everett, J.A.C., Bridge, H., Hewstone, M. (2014). A review of neuroimaging studies of race-related prejudice: does amygdala response reflect threat? *Frontiers in Human Neuroscience*, **8**, 179.

Cheon, B.K., Im, D.M., Harada, T., *et al.* (2011). Cultural influences on neural basis of intergroup empathy. *NeuroImage*, **57**, 642–50.

Cikara, M., Bruneau, E., Van Bavel, J.J., Saxe, R. (2014). Their pain gives us pleasure: how intergroup dynamics shape empathic failures and counter-empathic responses. *Journal of Experimental Social Psychology*, **55**, 110–25.

Cikara, M., Eberhardt, J.L., Fiske, S.T. (2011). From agents to objects: sexist attitudes and neural responses to sexualized targets. *Journal of Cognitive Neuroscience*, **23**, 540–51.

Curtin, F., Schulz, P. (1998). Multiple correlations and Bonferroni's correction. *Biological Psychiatry*, **44**, 775–7.

Decety, J., Cowell, J.M. (2014). Friends or foes: is empathy necessary for moral behavior? *Perspectives on Psychological Science*, **9**, 525–37.

Decety, J., Skelly, L., Yoder, K.J., Kiehl, K.A. (2014). Neural processing of dynamic emotional facial expressions in psychopaths. *Social Neuroscience*, **9**, 36–49.

Devine, P.G. (1989). Stereotypes and prejudice: their automatic and controlled components. *Journal of Personality and Social Psychology*, **56**, 5–18.

Devine, P.G., Plant, E.A., Amodio, D.M., Harmon-Jones, E., Vance, S.L. (2002). The regulation of explicit and implicit race bias: the role of motivations to respond without prejudice. *Journal of Personality and Social Psychology*, **82**, 835–48.

Eisenberger, N.I. (2012). The pain of social disconnection: examining the shared neural underpinnings of physical and social pain. *Nature Reviews Neuroscience*, **13**, 421–34.

Eisenberger, N.I. (2015). Social pain and the brain: controversies, questions, and where to go from here. *Annual Review of Psychology*, **66**, 601–29.

Fan, Y., Duncan, N.W., de Greck, M., Northoff, G. (2011). Is there a core neural network in empathy? An fMRI based quantitative meta-analysis. *Neuroscience and Biobehavioral Reviews*, **35**, 903–11.

Forman, S.D., Cohen, J.D., Fitzgerald, M., Eddy, W.F., Mintun, M.A., Noll, D.C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magnetic Resonance in Medicine*, **33**, 636–47.

Fox, G.R., Sobhani, M., Aziz-Zadeh, L. (2013). Witnessing hateful people in pain modulates brain activity in regions associated with physical pain and reward. *Frontiers in Psychology*, **4**, 772.

Gobodo-Madikizela, P. (2008). Empathetic repair after mass trauma. When vengeance is arrested. *European Journal of Social Theory*, **11**, 331–50.

Gobodo-Madikizela, P. (2015). Psychological repair. The intersubjective dialogue of remorse and forgiveness in the aftermath of gross human rights violations. *Journal of the American Psychoanalytic Association*, **63**, 1085–123.

Harris, L.T., Fiske, S.T. (2006). Dehumanizing the lowest of the low: neuroimaging responses to extreme out-groups. *Psychological Science*, **17**, 847–53.

Herek, G.M. (2000). Sexual prejudice and gender: do heterosexuals' attitudes toward lesbians and gay men differ? *Journal of Social Issues*, **56**, 251–66.

Hewstone, M., Rubin, M., Willis, H. (2002). Intergroup bias. *Annual Review of Psychology*, **53**, 575–604.

Huebner, B., Dwyer, S., Hauser, M. (2009). The role of emotion in moral psychology. *Trends in Cognitive Sciences*, **13**, 1–6.

Iannetti, G.D., Salomons, T.V., Moayedi, M., Mouraux, A., Davis, K.D. (2013). Beyond metaphor: contrasting mechanisms of social and physical pain. *Trends in Cognitive Sciences*, **17**, 371–8.

Immordino-Yang, M.H., McColl, A., Damasio, H., Damasio, A. (2009). Neural correlates of admiration and compassion. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 8021–6.

Ito, T.A., Bartholow, B.D. (2009). The neural correlates of race. *Trends in Cognitive Sciences*, **13**, 524–31.

Lamm, C., Decety, J., Singer, T. (2011). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *NeuroImage*, **54**, 2492–502.

Legrain, V., Iannetti, G.D., Plaghki, L., Mouraux, A. (2011). The pain matrix reloaded: a salience detection system for the body. *Progress in Neurobiology*, **93**, 111–24.

Lewis, H.B. (1971). *Shame and Guilt in Neurosis*. New York: International Universities Press.

Lickel, B., Schmader, T., Curtis, M., Scarnier, M., Ames, D.R. (2005). Vicarious shame and guilt. *Group Processes & Intergroup Relations*, **8**, 145–57.

Mars, R.B., Neubert, F.X., Noonan, M.P., Sallet, J., Toni, I., Rushworth, M.F.S. (2012). On the relationship between the 'default mode network' and the 'social brain'. *Frontiers in Human Neuroscience*, **6**, 189.

Masten, C.L., Morelli, S.A., Eisenberger, N.I. (2011). An fMRI investigation of empathy for 'social pain' and subsequent prosocial behavior. *NeuroImage*, **55**, 381–8.

Mathur, V.A., Harada, T., Lipke, T., Chiao, J.Y. (2010). Neural basis of extraordinary empathy and altruistic motivation. *NeuroImage*, **51**, 1468–75.

Meyer, M.L., Masten, C.L., Ma, Y., *et al.* (2013). Empathy for the social suffering of friends and strangers recruits distinct patterns of brain activation. *Social Cognitive and Affective Neuroscience*, **8**, 446–54.

Meyer, M.L., Williams, K.D., Eisenberger, N.I. (2015). Why social pain can live on: different neural mechanisms are associated with reliving social and physical pain. *PLoS One*, **10**, e0128294.

Mitchell, J.P. (2009). Inferences about mental states. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences*, **364**, 1309–16.

Molenberghs, P. (2013). The neuroscience of in-group bias. *Neuroscience & Biobehavioral Reviews*, **37**, 1530–6.

Morelli, S.A., Rameson, L.T., Lieberman, M.D. (2014). The neural components of empathy: predicting daily prosocial behavior. *Social Cognitive and Affective Neuroscience*, **9**, 39–47.

Mouraux, A., Diukova, A., Lee, M.C., Wise, R.G., Iannetti, G.D. (2011). A multisensory investigation of the functional significance of the "pain matrix". *NeuroImage*, **54**, 2237–49.

Northoff, G., Heinzel, A., de Greck, M., Bermpohl, F., Dobrowolny, H., Panksepp, J. (2006). Self-referential processing in our brain—a meta-analysis of imaging studies on the self. *NeuroImage*, **31**, 440–57.

Novembre, G., Zanon, M., Silani, G. (2015). Empathy for social exclusion involves the sensory-discriminative component of pain: a within-subject fMRI study. *Social Cognitive and Affective Neuroscience*, **10**, 153–64.

Phelps, E.A., O'Connor, K.J., Cunningham, W.A., *et al.* (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience*, **12**, 729–38.

Plant, E.A., Devine, P.G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, **75**, 811–32.

Risko, E.F., Laidlaw, K., Freeth, M., Foulsham, T., Kingstone, A. (2012). Social attention with real versus reel stimuli: toward an empirical approach to concerns about ecological validity. *Frontiers in Human Neuroscience*, **6**, 143.

Salice, A., Montes Sánchez, A. (2016). Pride, shame, and group identification. *Frontiers in Psychology*, **7**, 557.

Salomons, T.V., Iannetti, G., Liang, M., Wood, J.N. (2016). The "pain matrix" in pain-free individuals. *JAMA Neurology*, **73**, 755–6.

Saxe, R., Wexler, A. (2005). Making sense of another mind: the role of the right temporo-parietal junction. *Neuropsychologia*, **43**, 1391–9.

Schmitt, M.T., Branscombe, N.R. (2002). The meaning and consequences of perceived discrimination in disadvantaged and privileged social groups. *European Review of Social Psychology*, **12**, 167–99.

Schurz, M., Radua, J., Aichhorn, M., Richlan, F., Perner, J. (2014). Fractionating theory of mind: a meta-analysis of functional brain imaging studies. *Neuroscience and Biobehavioral Reviews*, **42**, 9–34.

Shackman, A.J., Salomons, T.V., Slagter, H.A., Fox, A.S., Winter, J.J., Davidson, R.J. (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nature Reviews Neuroscience*, **12**, 154–67.

Tajfel, H., Turner, J. (1979). An integrative theory of intergroup conflict. In: Austen, W. G., Worchel, S., editors. *The Social Psychology of Intergroup Relations*. Monterey, CA: Brooks-Cole.

Tangney, J.P. (1991). Moral affect: the good, the bad, and the ugly. *Journal of Personality and Social Psychology*, **61**, 598–607.

Tangney, J.P., Struewig, J., Mashek, D.J. (2007). Moral emotions and moral behavior. *Annual Review of Psychology*, **58**, 345–72.

Van Overwalle, F. (2009). Social cognition and the brain: a meta-analysis. *Human Brain Mapping*, **30**, 829–58.

Vossel, S., Geng, J.J., Fink, G.R. (2014). Dorsal and ventral attention systems: distinct neural circuits but collaborative roles. *Neuroscientist*, **20**, 150–9.

Williams, K.D., Cheung, C.K., Choi, W. (2000). Cyberostracism: effects of being ignored over the Internet. *Journal of Personality and Social Psychology*, **79**, 748–62.

Xu, X., Zuo, X., Wang, X., Han, S. (2009). Do you feel my pain? Racial group membership modulates empathic neural responses. *Journal of Neuroscience*, **29**, 8525–9.

Zaki, J., Ochsner, K.N. (2012). The neuroscience of empathy: progress, pitfalls and promise. *Nature Neuroscience*, **15**, 675–80.

Zuo, X., Han, S. (2013). Cultural experiences reduce racial bias in neural responses to others' suffering. *Culture and Brain*, **1**, 34–46.