Cerebral Cortex January 2016;26:225–233
doi:10.1093/cercor/bhu195
Advance Access publication September 2, 2014

Increased Moral Sensitivity for Outgroup Perpetrators Harming Ingroup Members

Pascal Molenberghs[1], Joshua Gapp[1], Bei Wang[1], Winnifred R. Louis[1] and Jean Decety[2,3]

[1]School of Psychology, The University of Queensland, Brisbane, Australia, [2]Department of Psychology and [3]Department of Psychiatry and Behavioral Neuroscience, The University of Chicago, Chicago, IL, USA

Address correspondence to P. Molenberghs, School of Psychology, McElwain Building, The University of Queensland, St Lucia, QLD 4072, Australia. Email: p.molenberghs@uq.edu.au

From an evolutionary perspective, one should be more sensitive when outgroup members attack the ingroup but less so when ingroup or outgroup members fight among themselves. Indeed, previous behavioral and neuroimaging research demonstrated that people show greater sensitivity for the suffering of ingroup compared with outgroup members. However, the question still remains whether this is always the case regardless of who is the agent causing the harm. To examine the role of agency and group membership in perception of harm, 48 participants were scanned while viewing ingroup or outgroup perpetrators intentionally harming ingroup or outgroup members. Behavioral results showed greater moral sensitivity for ingroup versus outgroup victims, but only when the perpetrator was from the outgroup. In support of this finding, fMRI data showed greater activity in left orbitofrontal cortex (OFC) for ingroup victims when they were harmed by outgroup individuals. In addition, effective connectivity analyses documented an increased coupling between left OFC and left amygdala and insula for ingroup harm, when the perpetrator was from the outgroup. Together these results indicate that we are highly sensitive to harm perpetrated by outgroup members and that increased sensitivity for ingroup victims is dependent on who is the agent of the action.

Keywords: empathy, fMRI, group processes, moral sensitivity, social neuroscience, violence, orbitofrontal cortex, amygdala, insula

Introduction

Previous neuroscience research has shown that people experience greater sensitivity for the suffering of ingroup compared with outgroup members (Xu et al. 2009; Decety et al. 2010; Gutsell and Inzlicht 2010; Hein et al. 2010; Cheon et al. 2011; Cikara and Fiske 2011; Cikara, Botvinick et al. 2011; Gutsell and Inzlicht 2012; Azevedo et al. 2013). From an evolutionary perspective, this makes sense, given the importance of one's own group for survival, wellbeing, and reproduction (Caporael 2001; Cottrell and Neuberg 2005; DeWall et al. 2011; Dunbar 2011; Jetten et al. 2012). The present research examines a potentially important limitation to this effect however, such that the group membership of the perpetrator is an important factor in how we perceive suffering for ingroup (vs. outgroup) victims.

Past research has shown that moral sensitivity is influenced not only by "who" is suffering but also by "how" the harm is caused (Young and Saxe 2009). For example, Decety et al. (2012) found that intentional harm caused to others was rated much higher on a moral sensitivity scale and elicited greater activation in neural regions implicated in moral cognition and empathy (i.e., orbitofrontal cortex (OFC), insula, and amygdala) than harm accidently caused. In addition, the distinction between intentional and accidental harm is a very fast process as demonstrated by a study using high-density EEG/ERP recordings (Decety and Cacioppo 2012). Current source density maxima in the right pSTS, as fast as 62 ms post-stimulus, first distinguished intentional versus accidental actions. Later responses in the amygdala (122 ms) and OFC (182 ms), respectively, were evoked by the perception of intentional (but not accidental) harmful actions, indicative of fast information processing associated with early stages of moral sensitivity. These studies thus suggest that we are specifically sensitive to harm caused intentionally.

However, to date no functional neuroimaging study has investigated how moral sensitivity is influenced by the group membership of the perpetrator in the context of intentional harm caused to either an ingroup or outgroup member. Therefore, the aim of this study was to investigate whether participants show increased moral sensitivity for intentional harm caused to ingroup versus outgroup members, irrespective of the group membership of the perpetrator, or whether sensitivity is affected by "who is responsible" for the suffering. Even children from the age of 3, after introduction to novel groups, predict that agents from a group are more likely to harm members of the other group rather than their own (Rhodes 2012). This suggests that moral sensitivity is influenced by group membership early on. We predicted that participants would be most sensitive to harm caused by outgroup members on ingroup members, given that this often poses an immediate threat to the survival of their own group, whereas most of the time the stakes are not as high when ingroup or outgroup members fight among themselves. In particular, it was expected that participants who show more behavioral ingroup bias (i.e., higher ratings of moral sensitivity for outgroup vs. ingroup perpetrators) would also show increased activation in brain areas involved in moral sensitivity when ingroup members are being intentionally harmed by outgroup members.

In the current study, we were particularly interested in moral sensitivity rather than complex moral reasoning. Moral sensitivity happens prior to moral reasoning in that it refers to the detection and interpretation of a moral issue or situation. This interpretative awareness is the first component of ethical decision-making in that it gives rise to the need to make a moral judgment or select a moral action (Robertson et al. 2007). The psychology of morality has typically been associated with complex abstract moral reasoning. However, automatic affective processes can be equally important in moral cognition depending on the situation (Greene and Haidt 2002; Moll et al. 2007; Greene et al. 2008; Young and Dungan 2012). As such our process of moral sensitivity is more closely related to the concept of emotional responsiveness, in which people respond to the suffering of others by vicariously sharing the affective components of their pain (Singer et al. 2004; Xu et al.

2009), rather than complex moral decision-making. However, our paradigm also differs from emotional responsiveness in that a clear moral component is involved because the harm to the victim is caused in a malicious intentional way. According to the social intuitionist model (Haidt 2001), when we witness a morally laden action, we instantly have a feeling of approval or disapproval. These feelings appear suddenly and effortlessly without much cognitive effort, and the moral reasoning is usually a post hoc construction of these quick, automatic evaluations (Haidt 2001). It is this process of fast reaction that is investigated in this study rather than the complex moral reasoning that follows.

Moral cognition as a whole is not associated with a single region (Bzdok et al. 2012). Instead, complex processes that subserve morality are distributed across the brain and are highly dependent on the modality of the stimuli and processes involved (Moll et al. 2003; Zahn et al. 2011; Young and Dungan 2012). Neuropsychological observations with patients with brain damage, as well as fMRI studies in healthy individuals, have shown that some areas are consistently involved in moral cognition (Moll et al. 2005). However, it also appears that these circuits are not unique to morality; instead, they appear to be systems and regions involved in specific motivational, cognitive, and emotional processes (Greene and Haidt 2002; Decety et al. 2012). These circuits emerge and become connected throughout development to allow for adaptive social behavior (Decety and Howard 2013).

One of the brain regions most consistently involved in moral cognition is the orbitofrontal/ventromedial prefrontal cortex (OFC; Moll et al. 2005; Blair et al. 2006; Moll et al. 2007; Decety et al. 2011). Neuroimaging studies have consistently found this region activated by tasks that involved a moral situation (Farrow et al. 2001; Moll, de Oliveira-Souza, Bramati et al. 2002; Moll, de Oliveira-Souza, Eslinger et al. 2002; Moll et al. 2006; Cikara et al. 2010; Decety and Porges 2011; Decety et al. 2012; Molenberghs et al. 2014). In addition, damage to the OFC can lead to acquired sociopathy (Eslinger and Damasio 1985; Blair and Cipolotti 2000) and an inability to recognize inappropriate social behavior (Stone et al. 1998; Ciaramelli et al. 2007). Finally, psychopaths who are characterized by a lack of empathy, remorse, and little appreciation of moral wrongdoing have functional and anatomical dysfunctions in the OFC/vmPFC (Anderson and Kiehl 2012; Decety et al. 2013).

Interestingly, meta-analyses on the OFC focusing on neuroimaging and neuropsychological studies found that the medial part of the OFC is related to monitoring the reward value of many different reinforcers, whereas lateral OFC is related to the evaluation of punishers, which may lead to a change in ongoing behavior (Kringelbach and Rolls 2004; Berridge and Kringelbach 2013). So, it seems that the OFC plays an important role in linking certain behaviors and stimuli with either their associated, internalized positive (medial OFC) or negative (lateral OFC) value. This suggests that the OFC is automatically linking social-emotional responses contingent on reward and punishment, as is the case in moral sensitivity. We therefore predicted that increased activity in the lateral OFC will be associated with increased moral sensitivity to transgressions, such as intentional harm-doing.

To investigate the extent to which the neural mechanisms underpinning moral sensitivity are influenced by group membership of the victim and perpetrator, we made use of preexisting university groups. These 2 groups were the University of Queensland (UQ) and Queensland University of Technology (QUT), which are 2 competing universities in the city of Brisbane, Australia. To measure moral sensitivity, participants watched video clips of ingroup and outgroup members intentionally harming each other while undergoing fMRI. It was predicted that participants would show increased moral sensitivity, as evidenced by behavioral ratings and increased activation in the lateral OFC, for the ingroup victim when the perpetrator was from the outgroup, but not when the perpetrator was from the ingroup.

Materials and Methods

Participants

Forty-eight UQ students (24 females, mean age = 22.2 years, SD = 5.3 years) completed the experiment. Participants were paid $30 for their time. All participants gave written informed consent. The study was approved by the Ethical Review Committee of the University of Queensland.

Group Identification

Group identification was assessed by presenting participants with 2 statements: "I identify myself as a UQ student" and "I identify myself as a QUT student." The 2 items were rated on a 7-point Likert-type scale (7 = totally agree, 1 = totally disagree) with higher scores indicating greater identification. Because we wanted to make sure that people identified more with the ingroup than with the outgroup, only participants who identified more with UQ than QUT were invited to take part in the fMRI experiment. Participants who did not identify more with UQ than QUT were replaced by others.

Functional MRI Experiment

Participants were screened using an MRI safety checklist and shown a brief demonstration of the experiment on a laptop. The demonstration ensured participants correctly understood how to perform the task. The experimental stimuli were based on stimuli showing a person causing intentional harm to another person, validated by Decety et al. (2012). However, group membership was incorporated as an extra manipulation, through the use of jumper color. Half of the participants were told that the people in the video clips wearing red jumpers were from the ingroup (UQ students) whereas the people in the green jumpers were from the outgroup (QUT students). The other half were told that green was the ingroup color, so that color was counterbalanced across participants. This information was made explicit at the beginning of each run. In reality, the people in the video clips were actors that did not necessarily belong to either group. The faces of the actors were not shown to make sure participants would not recognize the actors. Sixteen series of dynamic visual stimuli, each consisting of 3 digital color images, were created for each condition (Fig. 1). They were presented in a successive manner to imply motion with each image made the same size of 600 × 480 pixels. The images depicted actions of 1 person intentionally harming another person or corresponding actions of 2 people interacting in a peaceful manner. The durations of the first, second, and third images in each series were 800, 200, and 800 ms, respectively, followed by the fixation dot for a further 700 ms. E-prime2 software (Psychology Software Tools, Inc.) was used to present the stimuli.

The experimental stimuli were categorized by perpetrator (ingroup or outgroup perpetrator) and victim (ingroup or outgroup victim) creating 4 intentional harm conditions: a perpetrator from the ingroup attacking a victim from the ingroup (PIVI), a perpetrator from the ingroup attacking a victim from the outgroup (PIVO), a perpetrator from the outgroup attacking a victim from the ingroup (POVI), and a perpetrator from the outgroup attacking a victim from the outgroup (POVO). Corresponding control conditions showing 2 people interacting peacefully were also presented: ingroup member interacting with ingroup member (CII), ingroup member interacting with outgroup

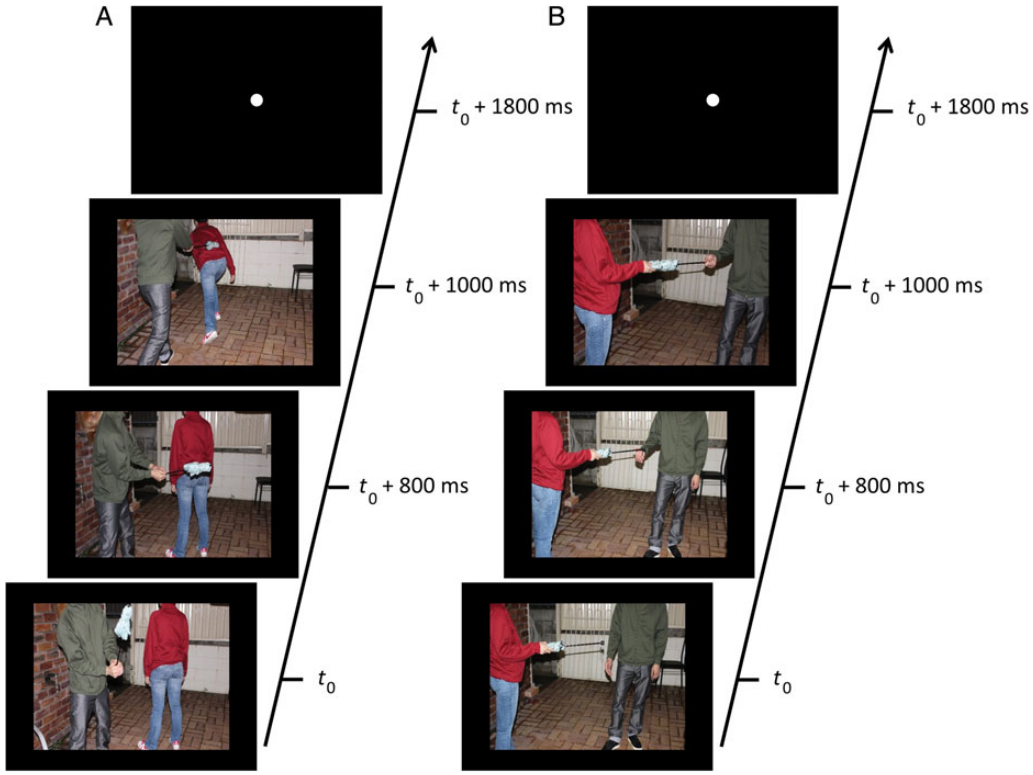

Figure 1. Schematic representation of stimuli used during the fMRI experiment. Jumper colors indicate group membership. (*A*) Intentional harm condition in which a person harms another person. (*B*) Control condition in which a person peacefully interacts with another person. The durations of the first, second, and third images in each series were 800, 200, and 800 ms, respectively, to imply motion, followed by the fixation dot for a further 700 ms.

member (CIO), outgroup member interacting with ingroup member (COI), and outgroup member interacting with outgroup member (COO). In addition, a baseline condition was included during which a white fixation dot was presented on a black background for the same length of time (10 s) as the dynamic image series.

The experiment ran for a total of approximately 35 min and consisted of 5 functional runs (6 min each) and a structural scan (5 min) acquired after the third run. The 5 runs each consisted of 32 active blocks and 4 baseline blocks. During each active block, 4 stimuli series from the same condition were pseudo randomly presented (such that each stimulus only occurred once per run). This occurred 4 times per run for each condition, such that all 16 stimuli series were presented. This occurred for all 8 conditions. The series lasted 2500 ms each, meaning each block lasted 10 s. The 4 baseline blocks consisted of the fixation dot remaining on the screen for the 10-s block duration. To reduce potentially confounding neural activation associated with motor activity, no overt response was required (Decety et al. 2012). Instead, participants were instructed to watch the stimuli carefully and eye movements of participants were tracked throughout the experiment using an ASL 600 (Applied Science Laboratories) to ensure they were attending to the stimuli throughout the experiment. When participants were not focusing sufficiently on the stimuli during a particular run, this run was repeated.

fMRI Image Acquisition

fMRI data were acquired using a 3-Tesla Siemens MRI Scanner with a 32-channel head volume coil and a gradient echo planar imaging (EPI) sequence with the following parameters: repetition time (TR) 3 s; echo time (TE) 30 ms; flip angle (FA) 90°, 64 × 64 voxels at 3 mm^2 in-plane resolution. Whole-brain images were generated every 3 s, and 127 images were acquired during each functional run. The first 4 images from each functional run during which no stimuli were presented were removed from the analysis to allow for steady-state tissue magnetization. A three-dimensional high-resolution T1-weighted image covering the entire brain was acquired after the second run for anatomical reference (TR = 1900, TE = 2.32 ms, FA = 9°, 192 cubic matrix, voxel size = 0.9 cubic millimeter, slice thickness = 0.9 mm).

Post-Scan Ratings

The post-scan ratings were administered electronically using E-prime2 software. Participants were shown 64 stimuli depicting a person harming another (i.e., 16 for each condition: PIVI, PIVO, POVI, and POVO) in random order. Each scenario was followed by 6 items designed to assess the participants' feelings of moral sensitivity toward the victim. The 6 items were as follows: 1) "I feel very sad for the victim," 2) "I feel very angry about what happened," 3) "I feel very disgusted about what happened," 4) "I feel very upset about what happened," 5) "I think the perpetrator was very mean to the victim," 6) "If it was up to me, I would punish the perpetrator." The items remained on the screen until the participant responded. Post-scan items were rated on a 7-point Likert-type scale (1 = totally disagree, 4 = neutral, and 7 = totally agree) thereby, higher ratings indicated greater feeling of moral sensitivity toward the victim.

A factor analysis using principal component analysis with Varimax rotation revealed that all 6 items loaded on the same factor which explained 74.8% of the variance. Sad (0.89), angry (97), disgusted (0.91), upset (0.97), mean (0.78), and punish (0.63) all had high factor loadings. There could also have been a possibility that the effect of perpetrator group membership was not similar across all 6 items on the moral sensitivity questionnaire. For example, the difference for POVI and POVO might have been larger for "I feel very sad for the victim" but smaller for "I feel very disgusted for the victim." Therefore, to ensure that the difference between the PIVO and PIVI (perpetrator ingroup) and the POVI and POVO (perpetrator outgroup) conditions was not significantly different for the 6 questions, the difference score was calculated for PIVO and PIVI and also for POVI and POVO. These difference scores across the 6 questions for when the perpetrator was in the ingroup and for when they were in the outgroup were then run in

2 separate repeated-measures ANOVAs. These tests were non-significant for both ingroup, $F_{1,\ 47} = 0.17$, $P = 0.680$, and outgroup, $F_{1,\ 47} = 0.21$, $P = 0.652$, perpetrators. This suggests that there was no significant difference across the 6 items neither when the perpetrator was in the ingroup nor when the perpetrator was in the outgroup. Together with the results from the factor analysis, this meant the average ratings of the 6 items for each condition could be combined as 1 measure of moral sensitivity for all participants. These scores were then averaged for the 4 conditions to give a total score for each condition (PIVI, PIVO, POVI, and POVO). This was considered as the level of moral sensitivity experienced by participants in that condition.

fMRI Data Analyses

The fMRI data were analyzed using SPM8 software (Wellcome Department of Imaging Neuroscience, Institute of Neurology) implemented in Matlab (Mathworks, Inc.). For the first-level analyses, all EPI images were realigned to the first image of each run to offset any effects of head movements. The T1-weighted anatomical scan was then coregistered to the mean functional image created during realignment. Using segmentation, the coregistered anatomical scan was then normalized with a voxel size of 1 × 1 × 1 × mm to the MNI T1 standard template (Montreal Neuropsychological Institute). Following this, the same parameters were used to normalize all of the EPI images using a voxel size of 3 × 3 × 3 mm to map onto the template. The images were then smoothed with a 6-mm isotropic Gaussian kernel. Following these procedures, a general linear model was created for every participant. Regions showing significant BOLD changes were identified by a block design with a 10-s duration time-locked to the onset of the trial. This was done for all 8 conditions.

At the second level analyses, contrast images for each of the conditions minus the baseline were created and included in a 2 (harm: intentional (PIVI + PIVO + POVI + POVO)) vs. control (CII + CIO + COI + COO)) × 2 (victim: ingroup (PIVI + POVI + CII + COI) vs. outgroup (PIVO + POVO + CIO + COO)) × 2 (perpetrator: ingroup (PIVI + PIVO + CII + CIO) vs. outgroup (POVI + POVO + COI + COO)) design. Because, in this model, the control conditions did not include victims and perpetrators, additionally 2 extra 2 × 2 factorial designs were created. One analysis focused on the harm conditions: 2 (victim: ingroup (PIVI + POVI) vs. outgroup (PIVO + POVO)) × 2 (perpetrator: ingroup (PIVI + PIVO) vs. outgroup (POVI + POVO)). A second analysis focused on the control conditions: 2 (passive actor: ingroup (CII + COI) vs. outgroup (CIO + COO)) × 2 (active actor: ingroup (CII + CIO) vs. outgroup (COI + COO)). Significant main effects or interactions in any of the models were subsequently followed up with pairwise comparisons, where appropriate. Following this, a series of regression analyses were conducted between differences in brain activity between conditions and post-scan behavioral measures.

It was expected that participants who had a larger ingroup bias (i.e., those who scored higher on moral sensitivity for the ingroup vs. outgroup victim) would show greater moral sensitivity-related brain activation for this contrast. The difference score for moral sensitivity for ingroup and outgroup victims ((PIVI + POVI) − (PIVO + POVO)) was calculated such that a positive score indicated greater moral sensitivity for the ingroup. The data were then entered into a multiple regression analysis with the ingroup victim (PIVI + POVI) minus outgroup victim (PIVO + POVO) contrast using SPM8. A similar correlation analysis was run for the POVI minus POVO and PIVI minus PIVO contrast with again positive scores indicating greater moral sensitivity for the ingroup. Significant activation for all contrasts was defined at a cluster-level or voxel-level threshold with an FWE or FDR rate of $P < 0.05$ corrected for the whole brain, with clusters defined by a voxel-level probability threshold of $P < 0.001$ and a minimum cluster size of 20 voxels.

fMRI Follow-Up Analysis

Based on the significant activation in the left OFC in some of the correlation analysis, the main intentional harm minus control contrast was then further examined using Region of Interest (ROI) analysis in the left OFC (anatomically defined by the WFU PickAtlas, http://www.fmri.wfubmc.edu/cms/software). Subsequently, the percent signal change was extracted for all intentional harm conditions from each individual participant. This was extracted from a 6-mm sphere around the peak coordinate identified in the left OFC ROI analysis. The percent signal score was then used to examine the difference in percent signal change in the left OFC for ingroup compared with outgroup victims when the perpetrator was from the outgroup (POVI minus POVO) versus ingroup compared with outgroup victims when the perpetrator was from the ingroup (PIVI minus PIVO). The same was done for the control condition (COI minus COO) minus (CII minus CIO).

In addition, effective connectivity analysis using psychophysiological interaction (PPI) was performed to estimate functional coupling between a source (seed in left OFC) and target regions of interest (left amygdala and left insula) for the POVI minus POVO contrast. Psychophysiological interaction analysis assesses the hypothesis that activity in a brain region can be explained by an interaction between a cognitive process and activity in another part of the brain. The amygdala and insula (anatomically defined by the WFU PickAtlas) in the same hemisphere were chosen a priori for the strong reciprocal connections each hold with the OFC, specifically in relation to emotions (Kringelbach and Rolls 2004). The selection of left OFC as the PPI source region was based on its significant involvement in the correlation analysis and intentional harm minus control contrast.

Activity within left OFC was used as the physiological regressor in the PPI analysis. The individual time series for the left OFC was obtained by extracting the first principal component from all raw voxel time series in a sphere (6-mm radius) centered around the peak coordinate identified in the intentional harm minus control contrast. These time series were mean-corrected and high-pass-filtered to remove low-frequency signal drifts. POVI minus POVO was the psychological regressor. The psychological variable used was a vector coding for the specific task (1 for POVI, −1 for POVO) convolved with the hemodynamic response function. A third regressor in the analysis represented the interaction between the first and second regressors. The physiological factor was multiplied with the psychological factor to constitute the interaction term.

Psychophysiological interaction analyses were carried out for each subject involving the creation of a design matrix with the interaction term, the psychological factor and the physiological factor as regressors. Subject-specific contrast images were then entered into random-effects group analyses. A PPI analysis was then conducted to identify whether left insula and left amygdala showed a significant increase in functional coupling with left OFC during POVI relative to POVO. Significant activity for the ROI analyses was defined by a voxel-level threshold with an FWE of $P < 0.05$ corrected for the size of the cluster.

Results

Group Identification

A paired-samples *t*-test indicated that participants identified more as a UQ student ($M = 6.48$, SE = 0.12) than a QUT student ($M = 1.58$, SE = 0.16), $t_{47} = 22.54$, $P < 0.001$.

Manipulation Check

To investigate whether our paradigm was effective in creating the desired states, we performed a one-sample *t*-test for all 6 post-scan measures (i.e., sad, angry, disgusted, upset, mean, and punish). If the manipulation was effective, participants' responses should indicate a score higher than 4 (i.e., neutral condition). The results indicated that sad ($M = 4.64$, SE = 0.13; $t = 7.43$), angry ($M = 4.52$, SE = 0.14; $t = 3.53$), disgusted ($M = 4.40$, SE = 0.17; $t = 2.36$), upset ($M = 4.39$, SE = 0.16; $t = 2.53$), mean ($M = 5.13$, SE = 0.11; $t = 9.81$), and punish ($M = 4.53$, SE = 0.14; $t = 3.47$) all had scores significantly higher than 4 (all $P < 0.05$).

Behavioral Results

A two-way repeated-measures ANOVA revealed no significant main effect of perpetrator in the levels of moral sensitivity

(perpetrator ingroup, PIVI and PIVO, $M = 4.64$, SE = 0.13 vs. outgroup, POVI and POVO, $M = 4.68$, SE = 0.12), $F_{1,\ 47} = 1.98$, $P = 0.166$, $\eta_p^2 = 0.04$. There was a significant main effect of victim with increased moral sensitivity for the ingroup victim (POVI and PIVI, $M = 4.71$, SE = 0.12) compared with outgroup victim (POVO and PIVO, $M = 4.61$, SE = 0.13), $F_{1,\ 47} = 5.82$, $P = 0.020$, $\eta_p^2 = 0.11$. Importantly, there was a significant interaction, $F_{1,\ 47} = 10.65$, $P = 0.002$, $\eta_p^2 = 0.19$. This point is further highlighted in Figure 2, where increased moral sensitivity for ingroup victims is only manifested when the perpetrator is from the outgroup. Follow-up paired *t*-tests revealed more moral sensitivity when the perpetrator was from the outgroup for the ingroup victim (POVI, $M = 4.78$, SE = 0.12) versus outgroup victim (POVO, $M = 4.57$, SE = 0.13), $t_{47} = -3.23$, $P = 0.002$, but equal sensitivity when the perpetrator was from the ingroup (PIVI, $M = 4.64$, SE = 0.13 vs. PIVO, $M = 4.64$, SE = 0.13), $t_{47} = 0.06$, $P = 0.952$.

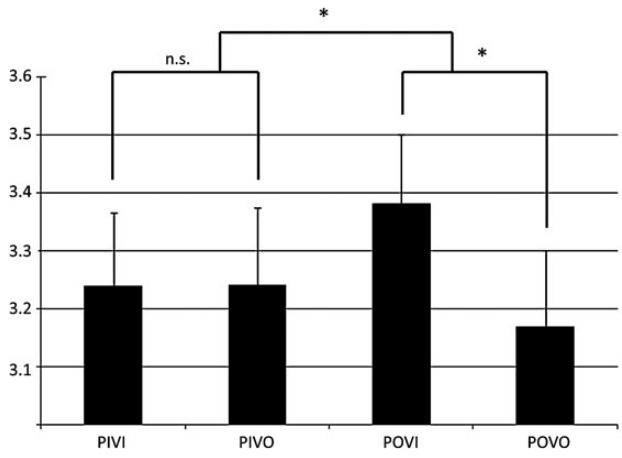

Figure 2. Mean moral sensitivity scores for the 4 intentional harm conditions. Higher scores indicate more moral sensitivity. n.s. is non-significant; *≤0.05. Error bars represent SE.

fMRI Results

A significant main effect was found for the harm factor. The intentional harm (PIVI, PIVO, POVI, and POVO) conditions produced greater activation in the action–observation network extending into occipital, parietal, and premotor areas (Caspers et al. 2010) compared with the control (CII, CIO, COI, and COO) conditions. More specifically for moral sensitivity, additional activation was also found in OFC, insula, and amygdala (Moll et al. 2005; Decety et al. 2012), which suggests that the experimental manipulation was effective (Fig. 3 and Table 1). No other main effects or interactions were significant after corrections for multiple comparisons at the whole-brain level.

fMRI Correlation Analysis

A significant positive correlation was found between differences in activation for the victim ingroup minus victim outgroup ((PIVI + POVI) − (PIVO + POVO)) contrast and corresponding moral sensitivity difference scores, in the left OFC (−36, 44, −14; $Z = 4.65$; extent = 35; *P* FWE voxel-level corrected = 0.027; Fig. 4A). A similar significant positive correlation was also found for the POVI minus POVO contrast in the left OFC (−33, 44, −14; $Z = 4.65$; extent = 92; *P* FWE voxel-level corrected = 0.042; Fig. 4*B*). No significant correlation was found for the PIVI minus PIVO contrast, indicating that differences in activation were not associated with victims' group membership when the perpetrator was from the ingroup. The same correlation analyses for the control condition ((CII + COI) − (CIO + COO), COI minus COO and CII minus CIO) revealed no significant results, even if we restricted our analysis to the left OFC (anatomically defined by the WFU PickAtlas) and used a voxel-level threshold of 0.05.

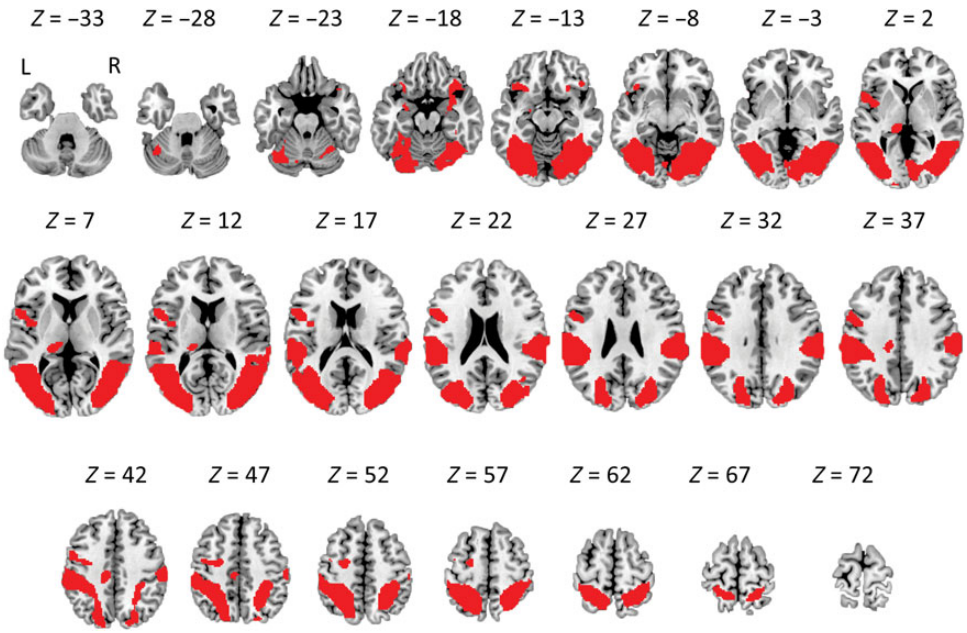

Figure 3. Significant brain activations from Table 1 for the intentional harm minus control contrast, shown on transversal slices with numbers above each slice representing *Z* coordinates in MNI space. L corresponds to left and R to the right side of the brain. Activations are displayed on a ch2better template using MRIcron software.

Table 1
Cluster size and associated peak values for the significant brain regions in the intentional harm minus control contrast

	Cluster	Peak FDR	Peak	MNI coordinates		
	Size	*P*-value	*T*-value	*X*	*Y*	*Z*
Intentional harm minus control						
Left occipitoparietal cortex	4066	<0.001	12.65	−48	−73	−2
			12.40	−48	−64	4
			11.74	−57	−25	28
Right occipitoparietal cortex	3682	<0.001	12.27	48	−73	−2
			11.61	51	−61	4
			11.16	48	−64	−8
Posterior midcingulate cortex	78	<0.001	6.12	−12	−25	40
Left premotor cortex	453	=0.001	5.45	−57	8	34
			5.33	−51	8	12
			5.26	−24	−10	52
Left thalamus	75	=0.003	5.08	−15	−28	4
Right insula extending into OFC and amygdala	120	=0.017	4.64	27	17	−17
				30	5	−20
				24	−4	−14
Left amygdala	17	=0.026	4.51	−21	−7	−14
Left OFC	59	=0.040	4.37	−36	23	−14
				−27	20	−14

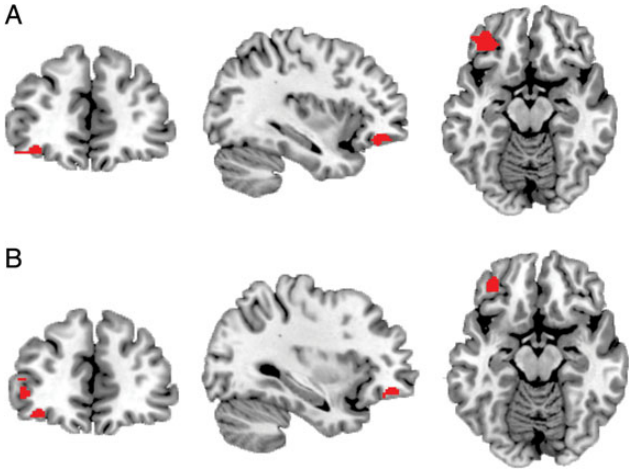

Figure 4. Significant positive correlation in left OFC between (*A*) victim ingroup minus victim outgroup ((PIVI + POVI) − (PIVO + POVO)) contrast and (*B*) victim ingroup minus victim outgroup, when perpetrator is from the outgroup (POVI minus POVO) contrast. Activations are displayed on a ch2better template using MRIcron.

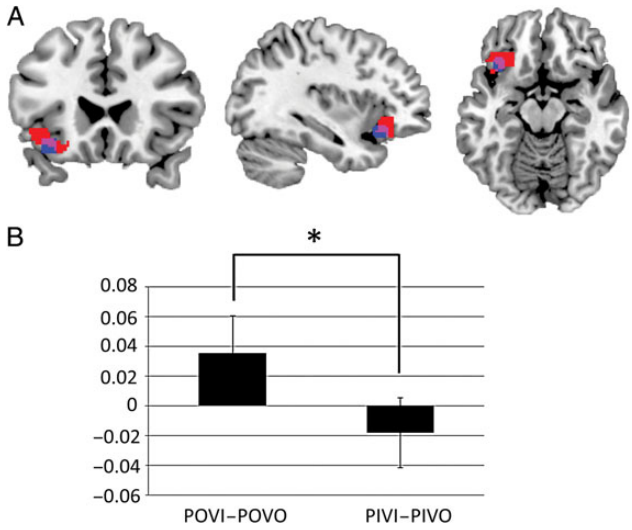

Figure 5. (*A*) Significant left OFC activation (in red) in the intentional harm minus control contrast displayed on a ch2better template using MRIcron. The blue area depicts the 6-mm sphere around the peak coordinate from which the percent signal change was extracted and which was used for the PPI analysis. (*B*) Percent signal change difference in left OFC between ingroup versus outgroup victim, when the perpetrator was from the outgroup (POVI minus POVO) compared with when the perpetrator was from the ingroup (PIVI minus PIVO). *$P < 0.05$.

Region of Interest Analysis

Considering the activation seen in the left OFC and the importance of this region in moral cognition, change in activity in this region was further explored using ROI analysis. The intentional harm conditions showed greater activation in the left OFC compared with the control conditions (−36, 23, −14; $Z = 4.31$, $k = 133$, FWE = 0.001; red in Fig. 5*A*). A follow-up paired *t*-test found a higher difference in percent signal change in this OFC region (blue in Fig. 5*A*) when the perpetrator was from the outgroup (POVI minus POVO) compared with when the perpetrator was from the ingroup (PIVI minus PIVO), $t_{47} = 2.15$, $P = 0.036$ (Fig. 5*B*). The same comparison, of the interaction effect, for the control condition (COI minus COO compared with CII minus CIO) revealed a non-significant result (47) = 0.88, $P = 0.38$.

In addition, effective connectivity analyses showed increased coupling between the left OFC and left insula (−39, −1, −11; $Z = 3.70$; extent = 146; P FWE = 0.021; Fig. 6*A*) and left amygdala (−30, −4, −17; $Z = 2.86$; extent = 22; P FWE = 0.036; Fig. 6*B*) for the POVI minus POVO contrast. No other areas showed increased connectivity with left OFC for this contrast at the whole-brain level.

Discussion

Evolution has tailored the brain to be sensitive to signs of suffering of others, and the neural mechanisms that process such signals are modulated by interpersonal relationships and group membership, with a clear bias for increased sensitivity for ingroup suffering (for reviews, see Hein and Singer 2008; Cheon et al. 2010; Chiao and Mathur 2010; Decety and Svetlova 2012; Eres and Molenberghs 2013; Molenberghs 2013). Indeed, work in social psychology and social neuroscience has demonstrated that individuals with normal empathic capacity frequently fail to respond to another's suffering if they belong

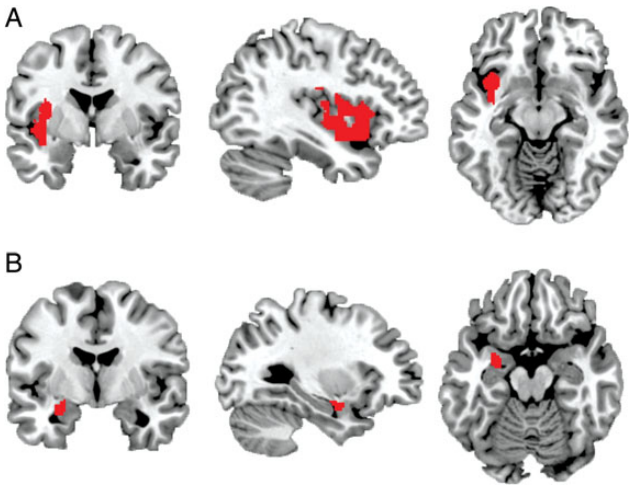

Figure 6. Psychophysiological analysis. Significant increased connectivity between left OFC and (*A*) left insula and (*B*) left amygdala for the ingroup versus outgroup victim, when the perpetrator was from the outgroup (POVI minus POVO) contrast displayed on a ch2better template using MRIcron.

to the other group and can even feel schadenfreude when experienced with the suffering of an outgroup member (Cikara, Botvinick et al. 2011; Cikara, Bruneau et al. 2011). The present study reveals an important moderator of moral sensitivity to transgressions; however, when an outgroup transgressor attacks an ingroup victim, special sensitivity may be evoked.

Specifically, in our study, we manipulated, for the first time, the group membership of both the victim and the perpetrator of harm. In line with previous work, we found increased activation in brain areas involved in moral cognition such as the OFC, amygdala, and insula when viewing actions that caused intentional harm to others (Cheng et al. 2010; Decety and Michalska 2010; Coaster et al. 2011; Decety et al. 2012; Molenberghs et al. 2014). Importantly, the behavioral results confirmed the hypothesis that participants only feel greater moral sensitivity for the suffering of an ingroup (vs. outgroup) victim when the perpetrator is from the outgroup. Indeed, when the perpetrator was from the ingroup, the level of moral sensitivity was the same for an ingroup versus outgroup victim. Previous results have shown that attacks on the ingroup by outgroup members, such as the "9/11" attacks, can result in extreme outgroup derogation (Oswald 2005). Our results show similarly that the greatest moral sensitivity is for outgroup threats to the ingroup (POVI) and that we care less for outgroup victims when they are fighting among themselves (POVO). This makes sense given that in the latter case, there is no immediate threat to the survival of the ingroup and we do not feel responsible for the violence inflicted on the outgroup victim. However, when an ingroup member is responsible for the violence inflicted on an outgroup member (PIVO), we may feel responsible for the suffering of the outgroup victim and to a similar degree as an ingroup victim (PIVI).

The fMRI results showed a similar increase in moral sensitivity as evidenced by heightened activation in left OFC for the ingroup versus outgroup victim contrast when the perpetrator was from the outgroup compared with the same contrast when the perpetrator was from the ingroup (Fig. 5). This increased response in left OFC was predicted by the amount of behavioral ingroup bias in moral sensitivity difference scores (Fig. 4). These data suggest that the lateral OFC plays a role in responding to outgroup threats to the ingroup. The important role of the OFC in moral cognition in adults is in line with our previous work that used similar intentional harm scenarios (Decety and Michalska 2010; Decety et al. 2012). The lateralization of the activation of the OFC activation also fits well with a meta-analysis on the OFC done by Berridge and Kringelbach (2013), in which they show that lateral OFC monitored learning and memory of displeasure and punishers.

No other brain areas showed a differential activation between the 4 harm conditions. However, the effective connectivity analysis, which assesses functional coupling between a particular ROI and other brain areas, did show increased connectivity between the left OFC and the insula and amygdala in the same hemisphere for the POVI minus POVO contrast. This suggest that an augmentation in hemodynamic response in OFC for the POVI minus POVO contrast is subserved by an increased coupling between the brain areas involved in more basic components of moral sensitivity such as the amygdala and insula.

Although the intentional harm condition significantly activated both the left and right OFC more than the control condition, our correlation analysis with the moral sensitivity behavioral measurers only found a significant correlation with the left OFC. It is not clear whether this laterality is specific for outgroup threats; however, previous studies have suggested that especially the left lateral OFC is important for the suppression of threats in decision-making (Bishop et al. 2004; Beer et al. 2006; Cikara et al. 2010). Given that outgroup attacks provide a clear threat for the survival of the ingroup, the left OFC might have evolved as an early suppression signal to influence automatic moral sensitivity through its connectivity with the insula and amygdala.

This OFC activation, in turn, would provide regulation and learning through negative reinforcement (Kringelbach and Rolls 2004; Berridge and Kringelbach 2013), which would allow for more efficient decision-making. Our theorizing is in line with a recent fMRI study by Decety et al. (2012) who found that older participants (compared with younger ones) showed more activity in the OFC in response to intentional harm. In addition, adults showed increased functional connectivity between the OFC and amygdala compared with the early childhood group, suggesting that negative emotional alerts are better regulated in older adults (Decety et al. 2012). This increased link between OFC and amygdala with increasing age fits well with the Integrated Emotions Systems model which states that the emotional learning system is mediated by the amygdala, which in turn provides information to a system for decision-making, on the basis of reinforcement expectations, which is mediated by the OFC (Blair 2009).

The fact the mPFC was not involved in our study, although this region is often involved in moral cognition (Moll et al. 2003, 2007), might not be that surprising given that our study focused on moral sensitivity rather than complex moral reasoning. Also, in a recent fMRI study, Cikara et al. (2014) found that participants, who showed reduced mPFC activation in response to descriptions of their own moral behaviors while competing in a group, were more willing to harm competitors. This suggests that when people act in a group situation, as in our study, reduced mPFC activation might actually lead to increased moral discomfort and a willingness to punish others.

Our study was also slightly different than some of the previous neuroimaging studies that have investigated the effect of group membership on watching others suffer, in that in our

study, the harmful behavior was clearly malicious and immoral. For example, in the study by Xu et al. (2009), participants watched images of ingroup or outgroup members receiving either needle penetration or Q-tip stimulation to the face. In this study, there was no moral component involved and as such the results were more related to emotional responsiveness. In our study, however, we used the term moral sensitivity because it was clear from the stimuli that the harm was caused to the victim in a malicious and intentional way (see for example Fig. 1). Also from the manipulation check measures, it was clear that participants thought the perpetrator was mean and that they wanted to punish him or her.

The left OFC showed significantly more activation in the intentional harm versus control condition at the whole-brain level. However, the same region did not show more activation for the ingroup victim or a significant interaction at the whole-brain level. There was, however, a significant interaction effect in ROI analysis focused on the left OFC, such that in line with the behavioral results, ingroup victims (compared with outgroup victims) suffering caused more activation in this region, but only when the perpetrator was from the outgroup (Fig. 5). More interestingly, there were strong significant results at the whole-brain level in the same region when we took inter-individual differences in moral sensitivity into account (Fig. 4). This is not surprising given that not all participants would have been influenced to the same degree by the group membership of the perpetrator and victim. In our study, the outgroup member was a neighboring university student, and these outgroup members were probably not disliked by everyone. As such, previous review papers have highlighted the importance at looking at inter-individual difference in empathy-related measures (Hein and Singer 2008). We expect that in more extreme ingroup versus outgroup situations such as for example between avid rival sporting teams (Hein et al. 2010; Cikara, Botvinick et al. 2011), where the outgroup is typically more disliked, the ingroup bias would be more similar across participants.

Overall, our results support the growing body of evidence that group membership influences a wide range of psychological and neural mechanisms across various modalities (e.g., face perception, action, emotions, etc.) which in turn can result in ingroup bias (Amodio 2008; Batson and Ahmad 2009; Ito and Bartholow 2009; Kubota et al. 2012; Derks et al. 2013; Eres and Molenberghs 2013; Molenberghs 2013; Molenberghs et al. 2013; Cikara and Van Bavel 2014). We demonstrate that group members are vigilant to ingroup bias: intentional harm-doing by an outgroup member to an ingroup victim evokes increased reactivity, compared with harm-doing by ingroup members. Moreover, the special sensitivity to ingroup members' suffering, compared with outgroup members', may be eliminated when the suffering is intentionally caused by ingroup transgressors. A better understanding of these underlying mechanisms is important and will lead to new insights into complex social problems such as violence between groups, racism, ingroup bias, and war.

To conclude, our data show that people have an increased sensitivity for outgroup threats toward the ingroup. These data provide novel insights into the existing literature on this topic by showing the importance of transgressor group membership and specifically that increased sensitivity for the suffering of an ingroup victim (Xu et al. 2009; Decety et al. 2010; Hein et al. 2010) is not absolute but is influenced by the group membership of the perpetrator.

Funding

This work was supported by an ARC Early Career Research Award (DE130100120) awarded to PM, an ARC Discovery Grant (DP130100559) awarded to PM and JD, and an ARC Discovery Grant (DP1092490) awarded to WL.

Notes

Conflict of Interest: None declared.

References

Amodio DM. 2008. The social neuroscience of intergroup relations. Eur Rev Soc Psychol. 19:1–54.

Anderson NE, Kiehl KA. 2012. The psychopath magnetized: insights from brain imaging. Trends Cogn Sci. 16:52–60.

Azevedo RT, Macaluso E, Avenanti A, Santangelo V, Cazzato V, Aglioti SM. 2013. Their pain is not our pain: brain and autonomic correlates of empathic resonance with the pain of same and different race individuals. Hum Brain Mapp. 34:3168–3181.

Batson CD, Ahmad NY. 2009. Using empathy to improve intergroup attitudes and relations. Soc Issues Policy Rev. 3:141–177.

Beer JS, Knight RT, D'Esposito M. 2006. Controlling the integration of emotion and cognition: the role of frontal cortex in distinguishing helpful from hurtful emotional information. Psychol Sci. 17:448–453.

Berridge KC, Kringelbach ML. 2013. Neuroscience of affect: brain mechanisms of pleasure and displeasure. Curr Opin Neurobiol. 23:294–303.

Bishop S, Duncan J, Brett M, Lawrence AD. 2004. Prefrontal cortical function and anxiety: controlling attention to threat-related stimuli. Nat Neurosc. 7:184–188.

Blair J. 2009. Neuro-cognitive systems involved in moral reasoning. In: Verplaetse J, de Schrijver J, Vanneste S, Braeckman J, editors. The Moral Brain. New York: Springer. p 87–105.

Blair J, Marsh AA, Finger E, Blair KS, Luo J. 2006. Neuro-cognitive systems involved in morality. Philos Explor. 9:13–27.

Blair RJR, Cipolotti L. 2000. Impaired social response reversal: a case of 'acquired sociopathy'. Brain. 123:1122–1141.

Bzdok D, Schilbach L, Vogeley K, Schneider K, Laird AR, Langner R, Eickhoff SB. 2012. Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. Brain Struct Funct. 217:783–796.

Caporael LR. 2001. Parts and wholes: the evolutionary importance of groups. In: Sedikides C, Brewer MB, editors. Individual Self, Relational Self, Collective Self. New York: Psychology Press. p 241–258.

Caspers S, Zilles K, Laird AR, Eickhoff SB. 2010. ALE meta-analysis of action observation and imitation in the human brain. NeuroImage. 50:1148–1167.

Cheng Y, Chen C, Lin C-P, Chou K-H, Decety J. 2010. Love hurts: an fMRI study. Neuroimage. 51:923–929.

Cheon BK, Im D-m, Harada T, Kim J-S, Mathur VA, Scimeca JM, Parrish TB, Park HW, Chiao JY. 2011. Cultural influences on neural basis of intergroup empathy. NeuroImage. 57:642–650.

Cheon BK, Mathur VA, Chiao JY. 2010. Empathy as cultural process: insights from the cultural neuroscience of empathy. World Cult Psychiatry Res Rev. 5:32–42.

Chiao JY, Mathur VA. 2010. Intergroup empathy: how does race affect empathic neural responses? Curr Biol. 20:R478–R480.

Ciaramelli E, Muccioli M, Ladavas E, di Pellegrino G. 2007. Selective deficit in personal moral judgment following damage to ventromedial prefrontal cortex. Soc Cogn Affect Neur. 2:84–92.

Cikara M, Botvinick MM, Fiske ST. 2011. Us versus them social identity shapes neural responses to intergroup competition and harm. Psychol Sci. 22:306–313.

Cikara M, Bruneau EG, Saxe RR. 2011. Us and them intergroup failures of empathy. Curr Dir Psychol. 20:149–153.

Cikara M, Farnsworth RA, Harris LT, Fiske ST. 2010. On the wrong side of the trolley track: neural correlates of relative social valuation. Soc Cogn Affect Neur. 5:404–413.
Cikara M, Fiske ST. 2011. Bounded empathy: neural responses to outgroup targets' (mis)fortunes. J Cognitive Neurosci. 23:3791–3803.
Cikara M, Jenkins A, Dufour N, Saxe R. 2014. Reduced self-referential neural response during intergroup competition predicts competitor harm. NeuroImage. 96:36–43.
Cikara M, Van Bavel JJ. 2014. The neuroscience of intergroup relations: an integrative review. Perspect Psychol Sci. 9:245–274.
Coaster M, Rogers BP, Jones OD, Viscusi WK, Merkle KL, Zald DH, Gore JC. 2011. Variables influencing the neural correlates of perceived risk of physical harm. Cogn Affect Behav Ne. 11: 494–507.
Cottrell CA, Neuberg SL. 2005. Different emotional reactions to different groups: a sociofunctional threat-based approach to "prejudice". J Pers Soc Psychol. 88:770–789.
Decety J, Cacioppo S. 2012. The speed of morality: a high-density electrical neuroimaging study. J Neurophysiol. 108:3068–3072.
Decety J, Chen C, Harenski C, Kiehl KA. 2013. An fMRI study of affective perspective taking in individuals with psychopathy: imagining another in pain does not evoke empathy. Front Hum Neurosci. 7:489.
Decety J, Echols S, Correll J. 2010. The blame game: the effect of responsibility and social stigma on empathy for pain. J Cognitive Neurosci. 22:985–997.
Decety J, Howard L. 2013. The role of affect in the neurodevelopment of morality. Child Dev Perspect. 7:49–54.
Decety J, Michalska KJ. 2010. Neurodevelopmental changes in the circuits underlying empathy and sympathy from childhood to adulthood. Develop Sci. 13:886–899.
Decety J, Michalska KJ, Kinzler KD. 2012. The contribution of emotion and cognition to moral sensitivity: a neurodevelopmental study. Cereb Cortex. 22:209–220.
Decety J, Michalska KJ, Kinzler KD. 2011. The developmental neuroscience of moral sensitivity. Emotion Rev. 3:305–307.
Decety J, Porges EC. 2011. Imagining being the agent of actions that carry different moral consequences: an fMRI study. Neuropsychologia. 49:2994–3001.
Decety J, Svetlova M. 2012. Putting together phylogenetic and ontogenetic perspectives on empathy. Dev Cogn Neurosci. 2:1–24.
Derks B, Scheepers D, Ellemers N, editors. 2013. The neuroscience of prejudice and intergroup relations. New York: Psychology Press.
DeWall CN, Deckman T, Pond RS, Bonser I. 2011. Belongingness as a core personality trait: How social exclusion influences social functioning and personality expression. J Pers. 79:1281–1314.
Dunbar R. 2011. Evolutionary basis of the social brain. In: Decety J, Cacioppo JT, editors. The Oxford Handbook of Social Neuroscience. New York: Oxford University Press. p 28–38.
Eres R, Molenberghs P. 2013. The influence of group membership on the neural correlates involved in empathy. Front Hum Neurosci. 7:176.
Eslinger PJ, Damasio AR. 1985. Severe disturbance of higher cognition after bilateral frontal lobe ablation Patient EVR. Neurology. 35:1731.
Farrow TF, Zheng Y, Wilkinson ID, Spence SA, Deakin JW, Tarrier N, Griffiths PD, Woodruff PW. 2001. Investigating the functional anatomy of empathy and forgiveness. Neuroreport. 12:2433–2438.
Greene JD, Haidt J. 2002. How (and where) does moral judgment work? Trends Cogn Sci. 6:517–523.
Greene JD, Morelli SA, Lowenberg K, Nystrom LE, Cohen JD. 2008. Cognitive load selectively interferes with utilitarian moral judgment. Cognition. 107:1144–1154.
Gutsell JN, Inzlicht M. 2010. Empathy constrained: prejudice predicts reduced mental simulation of actions during observation of outgroups. J Exp Soc Psychol. 46:841–845.
Gutsell JN, Inzlicht M. 2012. Intergroup differences in the sharing of emotive states: neural evidence of an empathy gap. Soc Cogn Affect Neur. 7:596–603.
Haidt J. 2001. The emotional dog and its rational tail: a social intuitionist approach to moral judgment. Psychol Rev. 108:814–834.
Hein G, Silani G, Preuschoff K, Batson CD, Singer T. 2010. Neural responses to ingroup and outgroup members' suffering predict individual differences in costly helping. Neuron. 68:149–160.
Hein G, Singer T. 2008. I feel how you feel but not always: the empathic brain and its modulation. Curr Opin Neurobiol. 18:153–158.
Ito TA, Bartholow BD. 2009. The neural correlates of race. Trends Cogn Sci. 13:524–531.
Jetten J, Haslam C, Haslam SA. 2012. The social cure: Identity, health and well-being. East Sussex: Psychology Press.
Kringelbach ML, Rolls ET. 2004. The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. Prog Neurobiol. 72:341–372.
Kubota JT, Banaji MR, Phelps EA. 2012. The neuroscience of race. Nat Neurosci. 15:940–948.
Molenberghs P. 2013. The neuroscience of in-group bias. Neurosci Biobehav R. 37:1530–1536.
Molenberghs P, Bosworth R, Nott Z, Louis WR, Smith JR, Amiot CE, Vohs KD, Decety J. 2014. The influence of group membership and individual differences in psychopathy and perspective taking on neural responses when punishing and rewarding others. Hum Brain Mapp. Advance online publication. doi: 10.1002/hbm.22527.
Molenberghs P, Halász V, Mattingley JB, Vanman EJ, Cunnington R. 2013. Seeing is believing: Neural mechanisms of action perception are biased by team membership. Hum Brain Mapp. 34:2055–2068.
Moll J, de Oliveira-Souza R, Bramati IE, Grafman J. 2002. Functional networks in emotional moral and nonmoral social judgments. Neuroimage. 16:696–703.
Moll J, de Oliveira-Souza R, Eslinger PJ. 2003. Morals and the human brain: a working model. Neuroreport. 14:299–305.
Moll J, de Oliveira-Souza R, Eslinger PJ, Bramati IE, Mourão-Miranda Jn, Andreiuolo PA, Pessoa L. 2002. The neural correlates of moral sensitivity: a functional magnetic resonance imaging investigation of basic and moral emotions. J Neurosci. 22:2730–2736.
Moll J, Krueger F, Zahn R, Pardini M, de Oliveira-Souza R, Grafman J. 2006. Human fronto–mesolimbic networks guide decisions about charitable donation. P Nat A Sci. 103:15623–15628.
Moll J, Oliveira-Souza Rd, Garrido GJ, Bramati IE, Caparelli-Daquer EM, Paiva ML, Zahn R, Grafman J. 2007. The self as a moral agent: linking the neural bases of social agency and moral sensitivity. Soc Neurosci. 2:336–352.
Moll J, Zahn R, de Oliveira-Souza R, Krueger F, Grafman J. 2005. The neural basis of human moral cognition. Nat Rev Neurosci. 6:799–809.
Oswald DL. 2005. Understanding anti-arab reactions post-9/11: the role of threats, social categories, and personal ideologies. J Appl Soc Psychol. 35:1775–1799.
Rhodes M. 2012. Naïve theories of social groups. Child Dev. 83: 1900–1916.
Robertson D, Snarey J, Ousley O, Harenski K, Bowman FD, Gilkey R, Kilts C. 2007. The neural processing of moral sensitivity to issues of justice and care. Neuropsychologia. 45:755–766.
Singer T, Seymour B, O'Doherty J, Kaube H, Dolan RJ, Frith CD. 2004. Empathy for pain involves the affective but not sensory components of pain. Science. 303:1157–1162.
Stone VE, Baron-Cohen S, Knight RT. 1998. Frontal lobe contributions to theory of mind. J Cognitive Neurosci. 10:640–656.
Xu X, Zuo X, Wang X, Han S. 2009. Do you feel my pain? Racial group membership modulates empathic neural responses. J Neurosci. 29:8525–8529.
Young L, Dungan J. 2012. Where in the brain is morality? Everywhere and maybe nowhere. Soc Neurosci. 7:1–10.
Young L, Saxe R. 2009. Innocent intentions: a correlation between forgiveness for accidental harm and neural activity. Neuropsychologia. 47:2065–2072.
Zahn R, De Oliveira-Souza R, Moll J. 2011. The neuroscience of moral cognition and emotion. In: Decety J, Cacioppo JT, editors. The Oxford Handbook of Social Neurosciene. New York: Oxford University Press. p 477–490.