

Published in final edited form as:

J Cogn Neurosci. 2011 December ; 23(12): 3791–3803. doi:10.1162/jocn_a_00069.

Bounded Empathy: Neural Responses to Outgroup Targets' (Mis)fortunes

Mina Cikara¹ and Susan T. Fiske²

¹MIT

²Princeton University

Abstract

The current study investigates whether mere stereotypes are sufficient to modulate empathic responses to other people's (mis)fortunes, how these modulations manifest in the brain, and whether affective and neural responses relate to endorsing harm against different outgroup targets. Participants feel least bad when misfortunes befall envied targets, and worst when misfortunes befall pitied targets, as compared to ingroup targets. Participants are also least willing to endorse harming pitied targets, despite pitied targets being outgroup members. However, those participants who exhibit increased activation in functionally-defined insula/MFG when viewing pity targets experience *positive* events not only report feeling worse about those events, but also more willing to harm pity targets in a tradeoff scenario. Similarly, increased activation in anatomically-defined bilateral anterior insula, in response to *positive* events, predicts increased willingness to harm envy targets, but decreased willingness to harm ingroup targets, above and beyond self-reported affect in response to the events. Stereotypes' specific content, and not just outgroup membership, modulates empathic responses and related behavioral consequences including harm.

A cursory reading of the emotion, empathy, and perception-action literatures might leave one with the impression that people spontaneously experience empathy in response to seeing another person in distress (Carr, Iacoboni, Dubeau, Mazziotta, & Lenzi, 2003; Preston & DeWaal, 2002; see however Decety, 2011; Jacob, 2008). Recent developments in social psychological and cognitive neuroscience research suggest otherwise: People frequently fail to empathize to the same extent with outgroup members as ingroup members (Batson & Ahmad, 2009; Chiao & Mathur, 2010; Stephan & Finlay, 1999). For example, White participants report greater empathy for, and assign less punitive punishments to White as compared to Black criminal defendants (Johnson et al., 2002); Black and White participants, at least those who exhibit greater implicit racial bias, have a stronger (empathic) sensorimotor response when watching an ingroup member's hand being pricked by a pin, compared to the hand of an outgroup member (Avenanti, Sirigu, & Aglioti, 2010); White and Asian participants demonstrate relatively less activation in the shared neural circuit for pain when viewing other-race as compared to same-race faces being pricked by a needle (Xu, Zuo, Wang, & Han, 2009).

Not all outgroups are equivalent, however. Depending on the target, people may feel not only less empathy, but also pleasure—*Schadenfreude*—in response to outgroup members' misfortunes (Cikara, Botvinick, & Fiske, 2011; Leach & Spears, 2009; Smith et al., 1996; Smith, Powell, Combs, & Schurtz, 2009): as one example, participants smile more when an envied person, compared to an ingroup target, experiences a negative event (e.g., sits in gum

on a park bench, Cikara & Fiske, in press). In contrast, there may also be specific outgroups for whom people feel even *more* empathy than ingroup members when they suffer a misfortune: college students report feeling worse when an elderly person as compared to another college student experiences a negative event (e.g., accidentally walks into a glass door, Cikara & Fiske, in press). Furthermore, no intergroup empathy study of which we are aware has assessed empathy for *positive* events, which demonstrably varies as a function of group membership. Indeed, participants report feeling worse when an envied person experiences a positive event (e.g., finds a lost sentimental possession), as compared to a pitied or ingroup target (Cikara & Fiske, in press). Thus it seems that the effect of group membership on empathy is more nuanced than merely ingroup versus outgroup responses to negative events.

One reason social psychologists and more recently cognitive neuroscientists have been interested in understanding empathy is because it is such a potent predictor of helping behavior (e.g., Batson, Van Lange, Ahmad, & Lishner, 2003; Decety & Ickes, 2009). When watching a fellow soccer fan receive an electric shock, self-reported empathic concern and increased activity in anterior insula—part of the shared neural circuit for pain—predicts participants' willingness to engage in costly helping behavior (i.e., participants volunteer to receive an electric shock in order to reduce the shock to their fellow fan; Hein, Silani, Preuschoff, Batson, & Singer, 2010). Pleasure in response to others' pain, on the other hand, may yield at best indifference to targets' suffering, and at worst facilitate active harm against them (Cikara, Bruneau, & Saxe, in press). Red Sox and Yankees fans, who show more ventral striatum activity—which is correlated with self-reported pleasure—when watching their rival fail to score, also report being more likely to actively harm the rival team's fans (Cikara, Botvinick, & Fiske, 2011). Similarly, soccer fans who exhibit ventral striatum activity, when watching a rival team's fan receive a painful electric shock, are unwilling to relieve the rival's pain by receiving half of the electric shock themselves (Hein et al. 2010).

In the current study, we examine whether specific stereotype content, absent overt category group labels (i.e., ingroup/outgroup), is sufficient to modulate empathy in response to both negative *and* positive events, as indexed by self-reported affect and neural activation. In order to investigate the effect of group membership on behavioral tendencies, we assess participants' willingness to harm ingroup and outgroup targets in a tradeoff scenario; we then examine how self-reported empathy and neural responses to targets' good and bad fortunes predict willingness to harm.

Not all Outgroups are Equivalent: Stereotype Content Model

Recent research in social cognition has firmly established that people differentiate each other not simply along an ingroup/outgroup boundary, but also by the extent to which they specifically (dis)like and (dis)respect a target or group. The Stereotype Content Model (SCM; Cuddy, Fiske, & Glick, 2007; Fiske, Cuddy, & Glick, 2007; Fiske, Cuddy, Glick, & Xu, 2002) organizes beliefs about social and cultural groups along two fundamental dimensions: perceived warmth and competence. Whether a social group is cooperative or competitive will determine if they apparently have *intent* to help or harm the culturally dominant group (or ingroup), which guides people's perceptions of that social group's *warmth*. Likewise, whether a social group does or does not have high status will determine if they apparently have *capability* to harm the ingroup, which will guide people's perceptions of the social group's *competence*. This 2 (low/high warmth) X 2 (low/high competence) mapping describes four broad stereotype categories and the emotional responses those categories elicit. Groups high on both warmth and competence (e.g., Americans, college students) elicit pride, whereas groups low on both warmth and competence (e.g., homeless people, drug addicts) elicit disgust. Groups falling in the mixed quadrants elicit ambivalent

emotions; pity is elicited by people perceived as low in competence but high in warmth (e.g., elderly people, disabled people), whereas envy is reserved for people perceived as high in competence but low in warmth (e.g., rich people, Asians, Jews, business people).

Consequences of Empathy Modulation

The prejudices predicted by the SCM matter because they have implications for behavior (i.e., helping and harm). Envied persons and groups are frequently targets of passive helping (e.g., associating), but also active harm (e.g., attack, sabotage) when conditions permit, whereas pitied targets are frequently recipients of active helping (e.g., assistance) but passive harm (e.g., neglect). Pride targets receive both active and passive help, whereas disgusting targets are the worst off because they receive both active harm and passive harm and are dehumanized (Cuddy, Fiske, & Glick, 2007; Harris & Fiske, 2006, 2009).

Empathy is a potent predictor of prosocial behavior (Batson, 2010), so it stands to reason that the extent to which people exhibit neural empathic resonance (i.e., increased ACC and AI in response to targets' negative events, and relatively decreased ACC and AI in response to positive events) will be negatively related to how much they are willing to subject those targets to harm (Hein et al., 2010). Therefore, in this study, we ask participants two weeks after scanning to report the extent to which they would be willing to subject different targets from the SCM to receive a painful electric shock, to keep four other unspecified people from receiving the shock.

Overview and Hypotheses

In the current study, participants respond to positive, negative, and neutral events happening to a variety of targets from the SCM. We predict that knowledge of a group's stereotype will spontaneously activate different kinds of prejudice, depending on the status (competence) and cooperativeness (warmth) of the target (Fiske et al., 2002, 2007). If stereotypes are sufficient to modulate behavioral and neural indicators of empathy, it suggests that groups need not have a long history of interaction to elicit variation in these empathy-related responses. Furthermore, examining the effects of the warmth and competence dimensions per se, not the groups themselves, allows predictions about responses to any social group based solely on its stereotype content.

Self-report predictions

As a validation check on the SCM, we predict that participants will report that pride and pity targets are warmer than envy and disgust targets; likewise, participants will report pride and envy targets are more competent than pity and disgust targets. Second, concerning misfortunes for each type of outgroup: we hypothesize that participants will report feeling relatively better about negative events that happen to *envied targets*, as compared to targets from the pride quadrant (the ingroup, Cuddy et al., 2008), because envy is one potent predictor of Schadenfreude (Smith et al., 1996; Van Dijk, Ouwerkerk, Goslinga, Nieweg, & Gallucci, 2006). But participants will report feeling worse when negative events befall a *pitied target* as compared to an pride/ingroup target, even though pity targets represent outgroup members. Surprisingly, previous data demonstrate that participants report feeling similarly about *disgusting targets* as compared to ingroup *pride targets*' misfortunes (see Cikara & Fiske, in press); we expect to replicate those findings here, though these ratings are subject to social desirability constraints (one argument for simultaneously collecting fMRI data). Finally, we predict the opposite pattern for positive events: that is, people will feel best when positive events happen to pitied targets, and worst when positive events happen to envied targets (with pride and disgust targets in the middle).

We are also interested in assessing whether affective or neural responses relate to individual differences in endorsement of active harm against outgroup targets, so we ask participants how likely they would be to volunteer different targets from the SCM to receive painful electric shocks in a tradeoff scenario (see methods). We predict that participants will be more willing to volunteer envy and disgust targets (who receive active harm) than pity and pride targets (who receive active help) to experience harm (Cuddy et al., 2007).

fMRI predictions

Experiencing pain is related to increased hemodynamic response in the “pain matrix,” which includes the anterior cingulate cortex (ACC) extending dorsally into supplementary motor area (SMA), the anterior insula (AI), and subcortical structures (Jackson, Rainville, & Decety, 2006). The ACC and AI are also reliably activated when observers learn of or witness other people in physical pain (e.g., listening to painfully loud noises, receiving electric shocks, making facial expressions of pain; Botvinick et al., 2005; Lamm, Batson, & Decety, 2007; Singer et al., 2004). Thus, these regions have been hypothesized to serve as a “functional bridge” between an empathic observer and a suffering target (Sommerville & Decety, 2006). ACC and AI activity, in response to perceiving a person in pain, correlate with participants’ reports of perceived pain (Saarela et al., 2006), and increase when participants are instructed to take the perspective of the victim (Lamm et al., 2007). Important for the current study, participants exhibit *reduced* ACC and AI activity (i.e., a counter-empathic response) when they perceive a racial outgroup, as compared to ingroup member, experience pain (as noted earlier).

If ACC and AI responses mirror the behavioral predictions, we expect that participants will exhibit greatest ACC and AI activation in response to pity targets’ misfortunes, least ACC and AI activation in response to envy target’s misfortunes, and moderate ACC and AI responses to pride/ingroup and disgust targets’ misfortunes. This hypothesis is complicated by the fact that disgust-inducing stimuli (including stigmatized individuals; Krendl, Macrae, Kelley, Fugelsang, & Heatherton, 2006) also reliably elicit AI and ACC activation (e.g., Wicker et al., 2003). Many previous studies of empathic pain use stimuli that likely disgust participants (e.g., needles pricking body parts), confounding the influence of participants’ own disgust with empathy for the target in pain. Because we include disgust targets, among other targets experiencing positive and negative events (only one of which is also inherently disgusting: “stepped in dog poo”), we can assess the relative effects of the targets (do disgust targets elicit greater ACC and AI responses as compared to other targets?), the events (do negative events elicit greater ACC and AI responses than positive events?), and the interaction of the two (e.g., do negative events only elicit more ACC and AI than positive events for pity targets?).

In contrast, *decreased* ACC and AI in response to pleasant experiences (as compared to misfortunes) constitutes an empathic response to positive events. Said another way, increased ACC and AI in response to others’ positive experiences constitutes a counter-empathic response. Indeed, participants exhibit greater ACC and AI in response to a competitor’s disadvantageous economic outcomes (Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003) and greater ACC in response to social competitors’ positive experiences (Takahashi et al., 2009). We predict that participants will exhibit least ACC and AI activation in response to pity targets experiencing good fortunes, most ACC and AI activation in response to envy targets experiencing good fortunes, and moderate ACC and AI activation in response to pride and disgust targets (though the above caveat for disgust targets applies here as well).

Correlations between self-report and neural measures

Participants' self-reported affective responses may or may not correlate reliably with the proposed patterns of activation, due to social desirability effects. Previous research has demonstrated a dissociation between physiological responses and self-reports in the context of empathy for SCM targets: though people smile more when envied targets suffer a misfortune, they report feeling neutral (Cikara & Fiske, in press). Thus, the relationship between self-report and neural responses may be weak or nonexistent, particularly in the case of negative events, because of the social undesirability of expressing diminished empathy or overt pleasure in this context.

Finally, we predict that decreased responses in ACC and AI in response to targets' misfortunes will be related to increased willingness to subject those targets to harm. Similarly, we predict that increased ACC and AI activity in response to positive events will be related to increased willingness to harm. In contrast to the self-reported affect measures, we have framed the harm question as a tradeoff scenario, which loosens the social desirability constraints: participants are able to respond in a manner that represents *both* their willingness to spare the other people in the scenario and their willingness to harm the target. The unspecified people to be spared are held constant, however, so we can attribute individual differences in harm-response variability to participants' reliance on stereotype content.

Method

Participants

Participants were 21 healthy volunteers (12 Female, $M_{\text{age}} = 20.1$, $SD = 1.55$) recruited from the Princeton University paid experiments website. All participants were right-handed, native English speakers with no history of psychiatric or neurological problems, and had normal or corrected vision. Written informed consent from each participant and experimental procedures complied with the guidelines of Princeton University's Institutional Review Board.

Stimuli

A pilot sample ($N=29$) rated the three types of events (without SCM targets) to confirm that they were perceived as negative, neutral, and positive, respectively. Examples of each event follow: "Found a \$5 bill on the street" (positive), "Tied their shoe" (neutral), "Had a stomach ache after lunch" (negative). On a scale ranging from 1 (*extremely negative*) to 10 (*extremely positive*), negative events were rated most negative ($M=3.61$, $SD=.58$), followed by neutral events ($M=6.38$, $SD=.51$; note that 5.5 is the midpoint of the scale), and positive events were rated most positive ($M=8.40$, $SD=.50$).

Each main-study participant saw and rated 27 events: 9 positive, 9 neutral, and 9 negative events. Each event was randomly paired with a picture of one person from each SCM quadrant in turn (i.e., a 'pride,' 'envy,' 'pity,' and 'disgust' target, respectively); 27 events, each paired once with each of the four quadrants, yielded 108 distinct target-event pairs total (Figure 1). Participants saw each distinct target-event pair four times over the course of the experiment, yielding 432 trials total. None of the events reflect good/bad intentions or more/less capability and therefore could not be more easily associated with some SCM targets than others.

The target images were drawn from a bank of 48, which included 12 images per quadrant. The targets were not labeled according to group membership. Not every participant saw the same sample of images because they were sampled without replacement within event type

(i.e., positive, neutral, negative). In other words, for any given participant, a specific target from a given quadrant could be randomly paired with a positive, negative, or neutral event (or some subset thereof) over the course of the study, but never 2 positive events. These pictures have been previously validated as evoking the predicted emotional responses (Harris & Fiske, 2006).

Procedure

Participants arrived to the console room, gave consent, and became acquainted with the task. Participants were told to imagine they were witnessing the event happen to the target in the photograph. Each trial consisted of a 2s exposure to a target-event pair presented simultaneously as a picture and an event phrase. After each stimulus, participants pressed a button using a touch pad to report that they had successfully seen the stimulus (this was to ensure that participants had not missed the stimuli, given their brief exposure time). Trials were separated by jittered inter-trial intervals (duration: 4 or 6 s), during which participants were instructed to fixate on a white cross-hair in the center of the black screen.

Participants saw each of the 108 distinct target-event pairs 4 times over the course of the scanning session, viewing 432 trials total. The order in which the images appeared was randomized between participants using Analysis of Functional Neuro-Images (AFNI; Cox, 1996) 'rsfgen' program, which generates randomized stimulus timing files, creates an ideal reference function for each of files, and evaluates the quality of the experimental design using AFNI 3dDeconvolve.

Pilot testing revealed that the 2s exposure was too brief for participants to be able to encode the target *and* read the event happening to the target when the target-event pairs were completely novel. As a result, we trained participants on all of the events during the anatomical scan—which took place before the functional scans—by showing them each event (without a target) four times, for 2 s each time. Event order was randomized across the training period. This allowed participants to become acquainted with the pace of the exposures and the content of the events themselves. In the post-study debrief, all participants reported being able to process both the target and the event during the functional runs.¹

During the functional scans, the images were projected onto a screen at the rear of the bore of the magnet. Participants viewed these images via an angled mirror placed above their eyes and attached to the radio-frequency coil. Stimuli were presented using E-prime, version 1.2 (Psychology Software Tools, Inc.: <http://www.pstnet.com>), and participants responded in the scanner using a fiber-optic touchpad (Current Designs Inc.: <http://www.curdes.com/response>), which they held in their right hand.

Self-report measures—After their scanning session, participants rated each target-event pair they had seen in the scanner (“How would this make you feel?” 1 = *extremely bad*, 9 = *extremely good*). Participants were instructed to answer this question in response to the *target’s experience* in order to minimize ambiguity regarding about what they should report feeling good and bad. Order of questions was counterbalanced between trials. Participants also rated the warmth and competence of each target they had seen earlier throughout the scan: “As viewed by society, how WARM is this person?” “As viewed by society, how

¹We tested whether the training session was sufficient to allow participants to encode both the image and the event within 2s. Five judges (3 female, $M_{age} = 23$) underwent the “event-training” session: they viewed each event four times, for 2s, without a paired SCM target. Afterwards, judges viewed the 108 distinct target-event pairs. At the end of each 2s exposure judges described the target and the event to the experimenter. Target/event accuracy ranged from 99.1% (107 correct) - 100% ($M = 99.5\%$). All judges reported that it was very easy to encode both the image and the event within 2 s, having had the event-training experience.

COMPETENT is this person?" The response scale ranged from 1 (*not at all*) to 9 (*extremely*), and order of questions was counterbalanced between trials.

Follow-up questionnaire—We contacted participants 1–2 weeks after we scanned them with a websurvey. Participants were presented with the following scenario: "You are participating in a 'Fear Factor' type game show and have just won a challenge. This exempts you from the "punishment" the rest of the players face: they are all going to receive mild electric shocks, which are painful, but not lethal. The game show host gives YOU the choice to decide whether all five of the players are going to get shocks or if one person should get a stronger shock (which is again, painful, but not lethal) while you spare the other four." Afterwards participants were presented with each target from the scan and asked, one at a time, to imagine that each target was the ONE person they could shock to keep the other players from receiving their punishments. For each target, participants reported how willing they would be to volunteer the person to receive a shock so that the others could avoid the punishment (rated likelihood was 1 = *not at all*, 7 = *extremely*). Several participants did not complete the follow up; all analyses including ratings of willingness to harm include N = 17.

Framing the harm question in terms of a tradeoff is useful because it forces people to weigh alternatives that may reveal spontaneous biases, which are otherwise difficult for experimenters to detect and for participants to report (Cikara, Farnsworth, Harris, & Fiske, 2010). If instead we had asked "how willing would you personally be to administer an electric shock to this target?" we would have seen no variance, as all would refuse. The demand characteristics are loosened in a tradeoff because participants are able to respond in a manner that represents *both* their willingness to spare the other contestants and their willingness to harm the target. Because the "contestants to be spared" are held constant and unspecified, response variation reflects differences in willingness to harm particular kinds of targets. Furthermore, the game-show scenario creates a situation in which stereotypes associated with the targets constitute the only information available for participants' consideration. Participants may treat all targets equivalently, or rely on stereotype content to guide their willingness to harm individuals.

fMRI Acquisition

At the beginning of each scan session, a high-resolution T-1 weighted anatomical image (T1-MPRAGE, $0.5 \times 0.5 \times 1.0$ mm) was acquired for use in registering activity to each participant's anatomy and for spatially normalizing data across participants. Echo-planar images were acquired using a 3.0 T Siemens Allegra head-dedicated scanner (Siemens, Erlangen, Germany) with a standard "bird-cage" head coil (TR = 2000 msec, TE = 30 msec, 196mm FOV, matrix size = 64×64). Near whole-brain coverage was achieved with 32 interleaved 3.6 mm axial slices.

fMRI Preprocessing and Data Analysis

Statistical Parametric Mapping software (SPM5; Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK) preprocessed the neuroimaging data. Functional (T2*) volumes were realigned to the first volume to correct for participant head motion, using a rigid-body transformation. The mean functional volume was then registered to the anatomical (T1) volume using an affine transformation. The anatomical volume was then segmented and warped to the Montreal Neurological Institute standard space and the warp solution was applied to the functional volumes. SPM subjected the data to spatial smoothing with an 8-mm full width at half minimum Gaussian kernel. Final voxel size: $3 \times 3 \times 3$ mm.

Task-related activity was measured using a window of 2 s, during which target-event pairs appeared on the screen. For each participant, AFNI computed mean BOLD signal in each condition within the 2 s in the response window. For statistical analysis, each stimulus time series was convolved with a hemodynamic response function to create a unique regressor for each of the 12 conditions: positive, negative, and neutral events, each happening to disgust, envy, pity, and pride targets. In addition, regressors of noninterest were included in the multiple regression model to factor out variance associated with participant head motion, and mean, linear, and quadratic trends in each run.

Whole brain exploratory analyses were performed with a voxel-wise significance threshold of $p < .005$. AlphaSim (included in AFNI) corrected for multiple comparisons. A Monte Carlo simulation determined a minimum cluster size of 99 voxels to achieve corrected significance of $p < .05$ for whole-brain contrasts, with a voxel-wise threshold of $p < .005$.

Whole-brain contrasts—The contrasts between parameter estimates for different experimental conditions within each participant were submitted to a group analysis that treated the variability between participants as a random effect. Statistical parametric maps were derived from the resulting t values associated with each voxel. AFNI's 3dANOVA3 program conducted a 3 (negative/neutral/positive events) \times 4 (disgust/envy/pity/pride target) \times 21 (participants) ANOVA to examine the main effects of event type and target, and the event \times target interaction. We created masks of all clusters that surpassed the minimum threshold designated by rsfgen, and calculated the mean coefficients for each participant, for each condition in that cluster. Neither harm nor affect ratings were included in the GLM used to define these regions.

Anatomical region of interest (ROI) analyses—Based on previous findings regarding the empathic pain network—ACC and AI (see, Jackson et al., 2006 for meta-analysis)—we drew anatomical masks for each subject individually (ACC masks drawn according to landmarks from Paus et al., 2006; AI masks drawn according to Crespo-Facorro et al., 2000). We then extracted average parameter estimates in each person's mask, in response to each of the 12 conditions, respectively. Using anatomically defined, a priori ROIs ensures that any analyses examining the relationship between neural activity in these ROIs, affect, and harm are independent (Vul, Harris, Winkielman, & Pashler, 2009).

Results

Self-report Results

We did not observe any participant gender differences in the self-reported affect, harm ratings, fMRI, or fMRI/harm regression results, so all analyses collapse over gender. All p -values correspond to two-tailed tests unless we note otherwise.

Warmth and competence ratings—In line with our hypotheses, participants rated pride ($M = 6.29$, $SD = 0.58$) and pity targets ($M = 6.54$, $SD = 0.76$) as significantly warmer than envy ($M = 6.83$, $SD = 1.18$) and disgust targets ($M = 3.33$, $SD = 0.85$), $F_{\text{warmth}}(1,20) = 229.15$, $p < .001$, $\eta^2_p = .92$. Participants also rated pride ($M = 7.43$, $SD = 0.65$) and envy targets ($M = 6.77$, $SD = 1.16$) as significantly more competent than pity ($M = 4.58$, $SD = 1.02$) and disgust targets ($M = 2.24$, $SD = 0.54$), $F_{\text{competence}}(1,20) = 415.12$, $p < .001$, $\eta^2_p = .95$.

Affect ratings—A 4 (disgust/envy/pity/pride target) \times 3 (positive/neutral/negative event) within-subjects ANOVA predicted participants' affective ratings. Both the main effect of target, $F_{\text{target}}(3, 60) = 4.53$, $p < .05$, $\eta^2_p = .19$, and the main effect of event, $F_{\text{event}}(1,09)$,

21.82) = 86.69, $p < .01$, $\eta^2 = .81$, were significant (df values are adjusted where the sphericity assumption is violated; we report Greenhouse-Geisser corrected values). The main effects were qualified by a significant interaction, $F_{\text{target} \times \text{event}}(2.03, 40.60) = 13.02$, $p < .01$, $\eta^2 = .39$: As predicted, participants reported that they felt worst when negative events befell pity targets ($M = 2.81$, $SD = 1.02$), surprisingly followed by disgust ($M = 3.06$, $SD = 0.92$), then pride targets ($M = 3.60$, $SD = 1.17$), and least bad (almost neutral) when negative events befell envy targets ($M = 4.20$, $SD = 1.31$; 5 is the midpoint of the scale), all pair-wise $t(20) > 3.2$, $p < .01$, except disgust-pity, $t(20) = 1.86$, $p = .07$ (Figure 2).

Participants also reported that they felt best when positive events happened to pity targets ($M = 7.11$, $SD = 0.93$), followed by pride ($M = 6.89$, $SD = 0.79$), then disgust ($M = 6.82$, $SD = 0.94$), and felt *least* good when positive events happened to envy targets ($M = 6.25$, $SD = 1.09$). Envy means were significantly lower than pride, pity, and disgust in response to positive events, $t(20) > 2.2$, $p < .05$; none of the other targets differed significantly from one another (Figure 2).

Harm ratings—A $2(\text{warmth}) \times 2(\text{competence})$ within-subjects ANOVA demonstrated a main effect of warmth on willingness to harm, $F_{\text{warmth}}(1,16) = 23.78$, $p < .001$, $\eta^2 = .60$, and a main effect of competence, $F_{\text{competence}}(1,16) = 41.09$, $p < .001$, $\eta^2 = .72$; these were qualified by a significant interaction, $F_{\text{warmth} \times \text{competence}}(1,16) = 11.02$, $p < .01$, $\eta^2 = .41$. As predicted, participants reported that they would be more willing to volunteer an envy target ($M = 3.97$, $SD = 1.22$) to receive an electric shock, than pity ($M = 1.42$, $SD = 0.42$), $t_{\text{envy-pity}}(16) = 8.60$, $p < .001$, disgust ($M = 3.31$, $SD = 1.35$), $t_{\text{envy-disgust}}(16) = 2.09$, $p = .05$, and pride targets ($M = 3.40$, $SD = 1.14$), though this last comparison was only significant with a one-tailed test, $t_{\text{envy-pride}}(16) = 1.74$, $p = .05$, one-tailed. Participants were also less likely to volunteer a pity target as compared to disgust, $t_{\text{disgust-pity}}(16) = 5.90$, $p < .001$, and pride targets, $t_{\text{pride-pity}}(16) = 7.64$, $p < .001$, to receive an electric shock. Willingness to harm pride and disgust targets was statistically equivalent, $t_{\text{pride-disgust}}(16) = .25$, ns (Figure 3).

fMRI Results

Whole-brain analyses—All clusters reported here are thresholded, $t(20) = 3.15$, $p < .005$, with a minimum volume of 99 contiguous voxels (designated by rsfgen, see methods) unless otherwise noted; MNI coordinates and volume are reported in Table 1. Neither the main effect of event type, nor the main effect of target yielded clusters of activation that surpassed the designated threshold for whole-brain exploratory analyses. The interaction between event and target types identified two significant clusters centered on right insula (including right middle frontal gyrus, MFG) and right superior parietal lobule (SPL), respectively.

To unpack the interactions, we computed mean coefficients for each of the 12 conditions, in each participant, within each of the clusters separately (Figure 4). Activation in right insula/MFG in response to negative events did not vary as a function of target, all pairwise $t(20) < 1.5$, ns . During neutral events, disgust targets were associated with the greatest response in right insula/MFG ($M = .0084$), followed by envy ($M = .0057$) and pride ($M = .0046$), which were statistically equivalent, followed by pity targets ($M = .0015$). During positive events, however, envy targets ($M = .0072$) elicited the strongest response compared to the other three targets (note that the envy > pity comparison is only significant with a one-tailed test, $t(20) = 1.76$, $p < .05$, one-tailed; Figure 4, top). Insula/MFG responses to disgust, envy, and pride targets during positive events were statistically equivalent.²

²Insula/MFG responses during envy targets experiencing positive events ($M = .0072$) were statistically equivalent to responses during disgust and envy targets experiencing neutral events ($M = .0084$ and $.0057$, respectively), but greater than pity and pride targets experiencing neutral events ($M = .0015$ and $.0046$, respectively).

In order to better understand the relative effects of the events for each target type, we examined how insula/MFG responses changed when targets were paired with positive and negative events, respectively, as compared to neutral events. For disgust targets, insula/MFG responses were significantly lower when targets were paired with *negative* as compared to neutral events, $t(20) = -3.74$, $p < .01$. In contrast, for pity targets, insula/MFG responses were significantly higher when targets were paired with negative as compared to neutral events, $t(20) = 2.30$, $p < .05$. The differences between responses to negative and neutral events were not significant for envy or pride targets, $t_s < 1.3$, *ns*. We ran parallel analyses for positive events, which yielded a similar pattern. For disgust targets, insula/MFG responses were significantly lower when targets were paired with *positive* as compared to neutral events, $t(20) = -3.43$, $p < .01$. In contrast, for pity targets, insula/MFG responses were significantly higher when targets were paired with positive as compared to neutral events, $t(20) = 2.01$, $p = .06$. The differences between responses to positive and neutral events were not significant for envy or pride targets, $t_s < 1.3$, *ns*.

To address our hypothesis, we examined insula/MFG responses to positive as compared to negative events across the four target types: in line with our predictions, insula/MFG responses were significantly higher for envy targets paired with positive events as compared to negative events, $t(20) = 2.58$, $p < .05$. The positive-negative comparisons were not significant for the other target types, $t_s < 1.6$, *ns*.

Activation in right SPL demonstrated a similar pattern of results: responses to negative events did not vary as a function of target, all pairwise $t_s(20) < 1.4$, *ns*. During neutral events, disgust targets were associated with the greatest response in right superior parietal lobule ($M = .0135$), followed by envy, pride, and finally pity targets. During positive events, however, envy targets elicited the strongest response compared to the other 3 targets (Figure 4, bottom). SPL responses to disgust, pity, envy, and pride targets during positive events were statistically equivalent.³

We also examined how SPL responses changed when targets were paired with positive and negative events, respectively, as compared to neutral events. For disgust targets, SPL responses were significantly lower when targets were paired with *negative* as compared to neutral events, $t(20) = -3.25$, $p < .01$. In contrast, for pity targets, SPL responses were significantly higher when targets were paired with negative as compared to neutral events, though this difference was only significant with a one-tailed test, $t(20) = 1.75$, $p = .05$, one-tailed. The differences between responses to negative and neutral events were not significant for envy or pride targets, $t_s < 1.4$, *ns*. We ran parallel analyses for positive events, which yielded a slightly different pattern. For disgust targets, SPL responses were significantly lower when targets were paired with *positive* as compared to neutral events, $t(20) = -2.98$, $p < .01$. For envy targets, SPL responses were significantly higher when targets were paired with positive as compared to neutral events, $t(20) = 2.38$, $p < .05$. The differences between responses to positive and neutral events were not significant for pity or pride targets, $t_s < 1.5$, *ns*.

Finally, we examined SPL responses to positive as compared to negative events across the four target types: SPL responses were significantly higher for envy targets paired with positive events as compared to negative events, $t(20) = 3.12$, $p < .01$. Again, the positive-negative comparisons were not significant for the other target types, $t_s < 1.6$, *ns*.

³SPL responses during envy targets experiencing positive events ($M = .0145$) were statistically equivalent to responses during disgust targets experiencing neutral events ($M = .0135$), but greater than envy, pity, and pride targets experiencing neutral events ($M = .0112$, $.0051$, and $.0099$, respectively).

As suspected, self-reported affect ratings of target-event pairs did not correlate with insula/MFG or SPL activation during corresponding stimuli, with one exception: participants who exhibited less right insula/MFG activity while watching pity targets experience *positive* events also reported feeling relatively *better* about those events, $r(19) = -.60, p < .05$. Again, these analyses are independent because the subjective affect ratings were not included in the GLM that defined the interaction contrast (Vul et al., 2009).

We also examined the correlation between harm ratings and activity in insula/MFG and SPL while viewing event pairs. Harm ratings did not correlate with insula/MFG or SPL activation during corresponding stimuli, again, with one exception: participants who exhibited increased right insula/MFG activity while watching pity targets experience *positive* events also reported being *more willing* to harm pity targets, $r(15) = .56, p < .05$. Said another way, people who exhibited less insula/MFG activation when watching pity targets experience positive events (i.e., an empathic response) not only reported feeling more positive affect about those events, but also reported being less willing to harm pity targets in a tradeoff scenario.

ROI analyses—We calculated the mean parameter estimates in ACC and bilateral AI for each participant, for each of the 12 conditions, respectively. The group level analyses only trended in the predicted direction; we report the findings in a footnote.⁴ We were interested in examining whether activation in ACC and AI in response to positive and negative events could predict harm ratings above and beyond self-reported affect. We ran a separate regression for each target type: we entered self-reported affect ratings in response to positive and negative events in the first step, average parameter estimates in ACC or AI in response to neutral events in the second step, and average parameter estimates in ACC or AI in response to positive and negative events in the final step. ACC activity in response to positive and negative events did not predict willingness to harm any of the four target types. Similar to the finding for pity targets reported above, increased AI activation, when watching an envy target experience a positive event, was related to increased willingness to harm envy targets ($\beta = .59, t(15) = 2.48, p < .05$); in contrast, increased AI activation, when watching a pride/ingroup target experience a positive event, was related to *decreased* willingness to harm ingroup targets ($\beta = -.54, t(15) = -2.47, p < .05$; Table 3).

Discussion

The current study demonstrates that participants felt best about negative events, and worst about positive events when they happened to envy targets as compared to other targets. In contrast people felt worse about negative events when they happened to pity targets as compared to admired, ingroup targets, despite pity targets representing outgroups. The current study also assessed the extent to which participants were willing to volunteer

⁴In order to examine the effect of positive and negative events on ACC and AI activation as a function of target (accounting for effect of the targets' pictures) we computed a difference score in each ROI subtracting responses during neutral events from responses during positive and negative events, respectively. In other words, we calculated 'negative events minus neutral events' and 'positive events minus neutral events' responses for disgust, envy, pity, and pride targets in ACC and AI. Target type did not have an effect on ACC activity in response to negative events, $F_{\text{target}}(3, 60) = 1.03, ns$, nor in response to positive events, $F_{\text{target}}(3, 60) = 1.94, ns$. In contrast, target type had a marginally significant effect on bilateral AI in response to negative events, $F_{\text{target}}(3, 60) = 2.06, p = .11$. Insula activity in response to negative (minus neutral) events was lowest for the disgust targets, followed by envy, then pride, and highest for pity targets. Paired t-tests revealed that the means for disgust and envy were statistically equivalent $t_{\text{envy-disgust}}(20) = 0.24, ns$, as were the means for pity and pride, $t_{\text{pity-pride}}(20) = 0.63, ns$. In contrast, the means for disgust were significantly different from pity, $t_{\text{pity-disgust}}(20) = 2.21, p < .05$, and pride, $t_{\text{pride-disgust}}(20) = 2.18, p < .05$. The means for envy were only marginally different from pity, $t_{\text{pity-envy}}(20) = 1.64, p = .12$, and not significantly different from pride, $t_{\text{pride-envy}}(20) = 1.27, ns$. Similarly, target type had a marginally significant effect on bilateral AI in response to positive events, $F_{\text{target}}(3, 60) = 2.33, p = .08$. Paired t-tests revealed that this effect was driven by the difference between the pity and disgust means: AI response to positive events for disgust targets were significantly lower than for pity targets, $t_{\text{pity-disgust}}(20) = 2.43, p < .05$. None of the other targets were significantly different from one another, all $t(20) < 1.43, ns$.

different targets for painful electric shocks: People were least willing to harm pity targets. This accords with recent evidence that people are more willing to subject more agentic targets (even good agents, such as Mother Teresa) to pain than less agentic targets (e.g., an orphan, Gray & Wegner, 2009). As illustrated in Figure 3, participants' self-reported harm data did not support the hypothesis that participants would be more likely to harm disgust as compared to pride targets. The extant literature reports dehumanized perception (e.g., Harris & Fiske, 2006; 2009) as well as active and passive harm tendencies against disgust targets (Cuddy et al., 2007). Our paradigm differs from previous studies, however, because we place the targets in a context (i.e., experiencing positive or negative events). We will need to run follow-up studies to understand better the conditions under which disgust targets elicit protective responses.

Functionally-defined insula responses varied as a function of event and target type. In line with our hypotheses, insula/MFG responses were significantly higher when envy targets were paired with positive as compared to negative events (i.e., a counter-empathic response). In contrast to our hypotheses, however, we did not observe the opposite pattern for pity or pride targets. Insula/MFG responses were not related to individual differences in affect or harm ratings for envy targets; however, we observed a different pattern for pity targets: relatively less insula/MFG activation when viewing pity targets paired with *positive* events was related to feeling better about the events, and also related to decreased willingness to submit pity targets to receive electric shocks in the harm tradeoff scenario. To the extent that *decreased* AI in response to *positive* events at the individual level represents an empathic response (or increased AI in response to positive events represents a counter-empathic response), these data suggest that increased empathic resonance is related to greater empathic affect and decreased aggressive behavioral responses to pity targets. These findings also fit with studies of unfairness in social interactions: the less AI activation participants exhibit in response to unfair offers from another person in an ultimatum game, the less likely participants are to reject the offer (Sanfey et al., 2003).

Responses in functionally-defined SPL also varied as a function of target and event type: SPL responses were significantly higher when envy targets were paired with positive as compared to negative events. In contrast to the findings for insula/MFG, SPL responses were not related to affect or harm ratings for any of the targets. Understanding the precise role of the SPL in participants' responses to ingroup and outgroup members' good and bad fortunes is beyond the scope of the current study. That said, recent research on the neural representations of social hierarchies offers an interesting possibility: judgments of both numerical magnitude and relative social rank elicit activity in bilateral intraparietal sulci and adjacent regions of parietal cortex (Chiao, in press; Chiao et al., 2009). Moreover, activity in these regions is greater when the comparison is between two close, as compared to two distant targets. Social comparisons are essentially contrast effects (Wedell, 1994) and abundant evidence suggests that the self constitutes a particularly pervasive standard in evaluations of others (e.g., Dunning & Hayes, 1996). Thus it is possible that our participants were most likely to engage in social comparison when confronted with an envy target. This may initially seem to contradict the SCM's predictions; however, evidence demonstrates that average participants perceive envy targets as being similar to themselves, familiar, and likely associates (Harris, Cikara, & Fiske, 2008). Indeed, although people compare downward to feel better (Wills, 1981), people compare upward to otherwise-similar others in order to feel inspired (see Fiske, 2011, for context). This link is purely speculative, however, and does not address why the SPL response is so high for disgust targets experiencing neutral events (see Footnote 3, Figure 4). Future studies should directly test how participants rank themselves relative to specific SCM targets and why the effects of social rank may shift as a function of the context in which the target is judged (i.e., experiencing a positive or negative event).

Our data did not support our ACC and AI hypotheses at the group-level. Responses in anatomically-defined ACC did not change as a function of target type paired with positive or negative events (controlling for responses during neutral events). Responses in anatomically-defined AI in the context of negative events (controlling for responses during neutral events) were lower for disgust as compared to pity and pride targets, but equivalent to envy targets, which were significantly lower than pity, but not pride targets. Finally, AI responses in the context of positive events were significantly lower for disgust as compared to pity targets, but none of the other contrasts were significant.

The most interesting finding was that individual differences analyses in anatomically-defined AI related to willingness to harm ingroup versus outgroup targets. Participants who exhibited decreased AI activation when viewing *positive* events (i.e., an empathic response) paired with *envy* targets also reported being *less* willing to harm envy targets (similar to the negative relationship between insula responses to pity targets with positive events and harm). In contrast, participants who exhibited greater AI activation (i.e., perhaps arousal) when viewing positive events happen to ingroup targets also reported being more willing to harm ingroup targets. Thus, at the individual level, ingroup/outgroup distinctions moderate the relationship between insula activation in response to positive events and willingness to harm targets.

Our data did not support the hypothesis that individuals who exhibited less ACC and AI activation (i.e., pain matrix network) in response to *negative* events happening to outgroup targets' would be related to increased willingness to harm said outgroups. We suspect that this due to the nature of the stimuli: the targets in all of the images had neutral to smiling facial expressions, which are congruent with experiencing neutral and positive events, but incongruent with experiencing negative events (e.g., the furthest right panel in Figure 1). This incongruence may have made it more difficult for participants to imagine the target's experience of negative events in the brief time allotted (2s). Future studies should either increase stimulus exposure time, ensure that all faces are neutral, or find targets whose expressions match the affect a person is expected to feel in response to positive, neutral, and negative events, respectively.

General discussion

The implications of the current findings are that group-based modulations of empathy are not as simple as more empathy for the ingroup (Mathur, Harada, Lipke, & Chiao, 2010) or relatively less empathy for the outgroup (Avenanti et al., 2010; Xu et al., 2009): specific outgroups in question, and their associated stereotypes matter. Furthermore, examining affective and neural responses to both good *and* bad fortunes (with an appropriate neutral baseline for comparison) is important because neural responses to positive events are often overlooked in neuroimaging studies of empathic resonance. Modulation of empathic responses has serious consequences, including decreased helping: people who attribute fewer uniquely human emotions (e.g., anguish, mourning) to opposite-race Katrina victims are also less willing to volunteer for relief efforts to help those victims (Cuddy, Rock, & Norton, 2007). In the current study, individual differences in AI responses to *positive* events are potent predictors of willingness to harm outgroup targets.

We employed a variety of methods to assess the effects of stereotype content on affective and neural responses to a variety of social groups' good and bad fortunes. Because expressing counter-empathic responses is socially undesirable, people may feel uncomfortable or unable to respond naturally in experimental settings. These dynamics have made studying failures of empathy—and related phenomena such as Schadenfreude—a methodological challenge (Cikara, Bruneau, & Saxe, in press). Using indirect measures such as fMRI to complement self-report helps to circumvent some of the hurdles associated with

measuring socially undesirable emotions and behaviors. In addition to providing converging evidence for the effects of stereotypes on empathic responding, the current fMRI findings add to a growing literature examining disruptions in empathy as it unfolds in the brain.

References

- Avenanti A, Sirigu A, Aglioti SM. Racial bias reduces empathic sensorimotor resonance with other-race pain. *Current Biology*. 2010; 20:1018–1020. [PubMed: 20537539]
- Batson, CD. Empathy-induced altruistic motivation. In: Mickulincer, M.; Shaver, PE., editors. *Prosocial motives, emotions, and behavior: The better angels of our nature*. Washington DC: APA; 2010. p. 15-34.
- Batson CD, Ahmad NY. Using empathy to improve intergroup attitudes and relations. *Social Issues and Policy Review*. 2009; 3:141.
- Batson, CD.; Van Lange, PAM.; Ahmad, N.; Lishner, DA. Altruism and helping behavior. In: Hogg, MA.; Cooper, J., editors. *Sage handbook of social psychology*. London: Sage Publications; 2003. p. 279-295.
- Botvinick MM, Jha P, Bylsma LM, Fabian SA, Solomon PE, Prkachin KM. Viewing facial expressions of pain engages cortical areas involved in the direct experience of pain. *NeuroImage*. 2005; 25:312–319. [PubMed: 15734365]
- Carr L, Iacoboni M, Dubeau MC, Mazziotta JC, Lenzi GL. Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Sciences*. 2003; 100:5497–5502.
- Chiao JY. Neural basis of social status hierarchy across species. *Current Opinion in Neurobiology*. 2010; 20:803–809. [PubMed: 20850964]
- Chiao JY, Harada T, Oby ER, Li Z, Parrish T, Bridge DJ. Neural representations of social status hierarchy in human inferior parietal cortex. *Neuropsychologia*. 2009; 47:354–63. [PubMed: 18992759]
- Chiao JY, Mathur VA. Intergroup empathy: How does race affect empathic neural responses? *Current Biology*. 2010; 20:R478–R480. [PubMed: 20541493]
- Cikara M, Bruneau E, Saxe R. Us and them: Intergroup failures of empathy. *Current Directions in Psychological Science*. 2011; 20:149–153.
- Cikara M, Botvinick MM, Fiske ST. Us v. them: Social identity shapes neural responses to intergroup competition and harm. *Psychological Science*. 2011; 22:306–313. [PubMed: 21270447]
- Cikara M, Farnsworth RA, Harris LT, Fiske ST. On the wrong side of the trolley track: Neural correlates of relative social valuation. *Social Cognitive and Affective Neuroscience*. 2010; 5:404–413. [PubMed: 20150342]
- Cikara M, Fiske ST. Stereotypes and Schadenfreude: Affective and physiological markers of pleasure at outgroups' misfortunes. *Social Psychological and Personality Science*. (in press).
- Cox RW. AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*. 1996; 29:162–173. [PubMed: 8812068]
- Crespo-Facorro B, Kim J, Andreasen NC, O'Leary DS, Bockholt HJ, Magnotta V. Insular cortex abnormalities in schizophrenia: A structural magnetic resonance imaging study of first-episode patients. *Schizophrenia Research*. 2000; 46:35–43. [PubMed: 11099884]
- Cuddy AJC, Fiske ST, Glick P. The BIAS map: Behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology*. 2007; 92:631–648. [PubMed: 17469949]
- Cuddy AJC, Fiske ST, Glick P. Warmth and Competence As Universal Dimensions of Social Perception: The Stereotype Content Model and the BIAS Map. *Advances in Experimental Social Psychology*. 2008; 40:61–149.
- Cuddy AJC, Rock MS, Norton MI. Aid in the aftermath of hurricane katrina: Inferences of secondary emotions and intergroup helping. *Group Processes & Intergroup Relations*. 2007; 10:107–118.
- Decety J. Dissecting the neural mechanisms mediating empathy. *Emotion Review*. 2011; 3:92–108.
- Decety, J.; Ickes, W. *The Social Neuroscience of Empathy*. Cambridge: MIT Press, Cambridge; 2009.

- Dunning D, Hayes AF. Evidence of egocentric comparison in social judgment. *Journal of Personality and Social Psychology*. 1996; 71:213–229.
- Fiske, ST. *Envy up, scorn down: How status divides us*. New York: Russell Sage; 2011.
- Fiske ST, Cuddy AJC, Glick P, Xu J. A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*. 2002; 82:878–902. [PubMed: 12051578]
- Fiske ST, Cuddy AJC, Glick P. First judge warmth, then competence: Fundamental social dimensions. *Trends in Cognitive Sciences*. 2007; 11:77–83. [PubMed: 17188552]
- Gray K, Wegner DM. Moral typecasting: Divergent perceptions of moral agents and moral patients. *Journal of Personality and Social Psychology*. 2009; 96:505–520. [PubMed: 19254100]
- Harris, LT.; Cikara, M.; Fiske, ST. Envy as predicted by the stereotype content model: A volatile ambivalence. In: Smith, R., editor. *Envy: Theory and Research*. New York: Oxford University Press; 2008.
- Harris LT, Fiske ST. Dehumanizing the lowest of the low: Neuro-imaging responses to extreme outgroups. *Psychological Science*. 2006; 17:847–853. [PubMed: 17100784]
- Harris LT, Fiske ST. Social neuroscience evidence for dehumanised perception. *European Review of Social Psychology*. 2009; 20:192–231.
- Hein G, Silani G, Preuschoff K, Batson CD, Singer T. Neural responses to ingroup and outgroup members' suffering predict individual differences in costly helping. *Neuron*. 2010; 68:149–160. [PubMed: 20920798]
- Jackson PL, Meltzoff AN, Decety J. How do we perceive the pain of others? A window into the neural processes involved in empathy. *NeuroImage*. 2005; 24(3):771–779. [PubMed: 15652312]
- Jackson PL, Rainville P, Decety J. To what extent do we share the pain of others? Insight from the neural bases of pain empathy. *Pain*. 2006; 125:5–9. [PubMed: 16997470]
- Jacob P. What do mirror neurons contribute to human social cognition? *Mind and Language*. 2008; 23:190–223.
- Johnson JD, Simmons CH, Jordan A, MacLean L, Taddei J, Thomas D. Rodney King and O. J. revisited: The impact of race and defendant empathy induction on judicial decisions. *Journal of Applied Social Psychology*. 2002; 32:1208–1223.
- Krendl AC, Macrae CN, Kelley WM, Heatherton TF. The good, the bad, and the ugly: An fMRI investigation of the functional anatomic correlates of stigma. *Social Neuroscience*. 2006; 1:5–15. [PubMed: 18633772]
- Lamm C, Batson CD, Decety J. The neural substrate of human empathy: effects of perspective-taking and cognitive appraisal. *Journal of Cognitive Neuroscience*. 2007; 19:42–58. [PubMed: 17214562]
- Leach CW, Spears R. Dejection at in-group defeat and Schadenfreude toward second- and third-party out-groups. *Emotion*. 2009; 9:659–665. [PubMed: 19803588]
- Mathur VA, Harada T, Lipke T, Chiao J. Neural basis of extraordinary empathy and altruistic motivation. *NeuroImage*. 2010; 51:1468–1475. [PubMed: 20302945]
- Paus T, Tomaiuolo F, Otaky N, MacDonald D, Petrides M, Atlas J, Morris R, Evans AC. Human cingulate and paracingulate sulci: Pattern, variability, asymmetry, and probabilistic map. *Cerebral Cortex*. 1996; 6:207–214. [PubMed: 8670651]
- Poldrack RA. Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*. 2006; 10:59–63. [PubMed: 16406760]
- Preston SD, de Waal FBM. Empathy: its ultimate and proximate bases. *Behavioral and Brain Sciences*. 2002; 25:1–72. [PubMed: 12625087]
- Saarela MV, Hlushchuk Y, Williams AC, Schurmann M, Kalso E, Hari R. The compassionate brain: humans detect intensity of pain from another's face. *Cerebral Cortex*. 2007; 17:230–237. [PubMed: 16495434]
- Sanfey AG, Rilling JK, Aronson JA, Nystrom LE, Cohen JD. The neural basis of economic decision-making in the Ultimatum Game. *Science*. 2003; 300:1755–1758. [PubMed: 12805551]
- Singer T, Seymour B, O'Doherty J, Kaube H, Dolan RJ, Frith CD. Empathy for pain involves the affective but not sensory components of pain. *Science*. 2004; 303:1157–1162. [PubMed: 14976305]

- Smith RH, Powell CAJ, Combs DJY, Schurtz RD. Exploring the when and why of *Schadenfreude*. *Social and Personality Psychology Compass*. 2009; 3:530–546.
- Smith RH, Turner TJ, Garonzik R, Leach CW, Urch V, Weston C. Envy and *Schadenfreude*. *Personality and Social Psychology Bulletin*. 1996; 22:158–168.
- Sommerville JA, Decety J. Weaving the fabric of social interaction: Articulating developmental psychology and cognitive neuroscience in the domain of motor cognition. *Psychonomic Bulletin & Review*. 2006; 13:179–200. [PubMed: 16892982]
- Stephan WG, Finlay K. The role of empathy in improving intergroup relations. *Journal of Social Issues*. 1999; 4:729–743.
- Takahashi H, Kato M, Matsuura M, Mobbs D, Suhara T, Okubo Y. When your gain is my pain and your pain is my gain: Neural correlates of envy and *Schadenfreude*. *Science*. 2009; 323:937–939. [PubMed: 19213918]
- Van Dijk WW, Ouwerkerk JW, Goslinga S, Nieweg M, Gallucci M. When people fall from grace: Reconsidering the role of envy in *Schadenfreude*. *Emotion*. 2006; 6:156–160. [PubMed: 16637759]
- Vul E, Harris C, Winkielman P, Pashler H. Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspectives in Psychological Science*. 2009; 4:274–290.
- Wedell DH. Contextual contrast in evaluative judgments: A test of pre- versus postintegration models of contrast. *Journal of Personality and Social Psychology*. 1994; 66:1007–1019.
- Wicker B, Keysers C, Plailly J, Royet JP, Gallese V, Rizzolatti G. Both of us disgusted in my insula: The common neural basis of seeing and feeling disgust. *Neuron*. 2003; 40:655–664. [PubMed: 14642287]
- Wills TA. Downward comparison principles in social psychology. *Psychological Bulletin*. 1981; 90:245–271.
- Xu X, Zuo X, Wang X, Han S. Do you feel my pain? racial group membership modulates empathic neural responses. *The Journal of Neuroscience*. 2009; 29:8525–8529. [PubMed: 19571143]

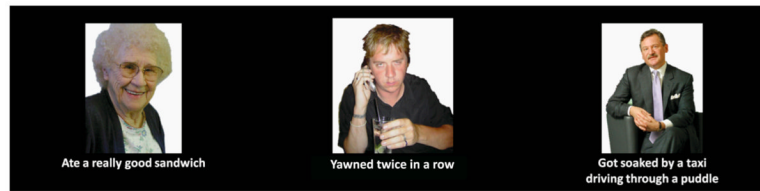


Figure 1.
Stimulus examples (left to right): pity target/positive event, disgust target/neutral event, envy target/negative event.

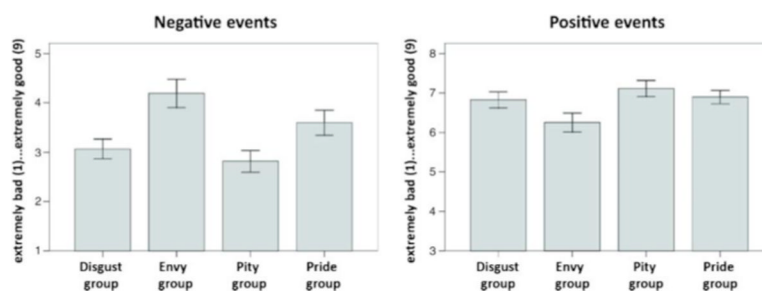


Figure 2.
Ratings of how participants would feel in response to negative events and positive events (note that the y-axis is different for the two graphs). Bars represent standard error.

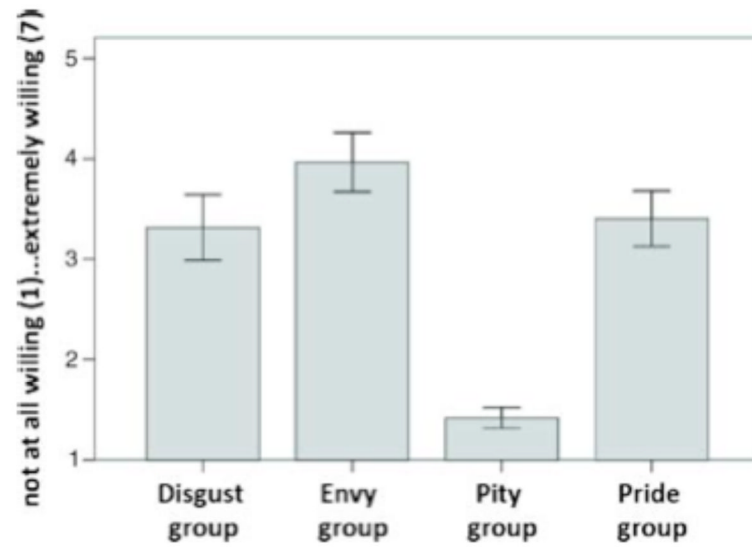
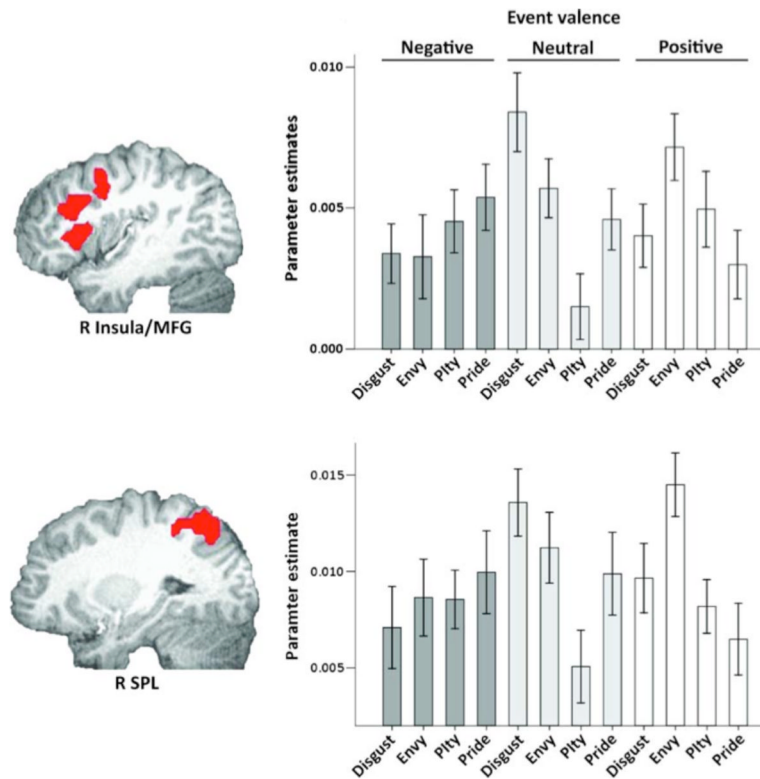


Figure 3. Ratings of how willing participants would be to volunteer targets to receive electric shocks. Bars represent standard error.

**Figure 4.**

Top: Right insula/MFG cluster identified by the event by target interaction. Bottom: Right SPL cluster identified by the event by target interaction. Mean parameter estimates are plotted for each type of target, in response to negative (dark gray), neutral (light gray), and positive events (white). Bars represent standard error.

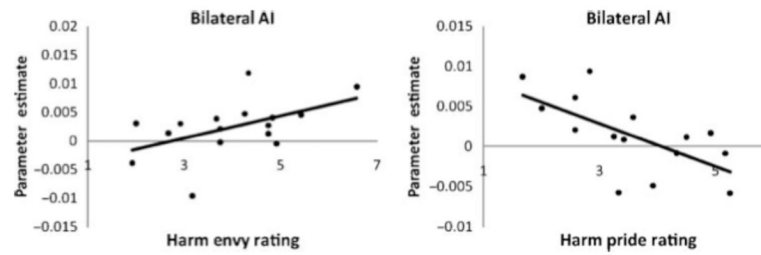


Figure 5.

Left: Average parameter estimates in anatomically-defined bilateral anterior insula in response to viewing envy targets paired with positive events plotted against willingness to harm envy targets. Right: Average parameter estimates in anatomically-defined bilateral anterior insula in response to viewing pride targets paired with positive events plotted against willingness to harm pride targets.

Table 1

Whole-brain analyses: Event Type X Target Group Level ANOVA

Regions	<i>x</i>	<i>y</i>	<i>z</i>	Cluster Size (Voxels)
Event				

Target				

Event X Target				
R insula/middle frontal gyrus	45	18	21	432
R superior parietal Lobule	-28	-62	48	184

Note. Peak voxel and cluster size (1 voxel = 3m³). Voxelwise significance threshold, $p < .005$. Coordinates refer to the Montreal Neurological Institute stereotaxic space. The insula/MFG cluster includes anterior insula and regions of BA 9/6; the SPL cluster is centered on BA 7, extending rostrally to include a dorsal region of BA 31.

Table 2

Regressions Predicting Willingness to Harm

Parameter	<u>Envy Target</u>		<u>Pride/Ingroup Target</u>	
	b	SE	b	SE
Constant	6.23	3.25	-6.24	6.06
Negative Event Affect	-0.51	0.27	0.45	0.43
Positive Event Affect	-0.08	0.37	1.16	0.67
Neutral Event AI	-3.28	57.77	93.58	78.86
Negative Event AI	3.16	38.18	13.15	49.06
Positive Event AI	150.93 *	61.44	-135.31 *	54.66

Note.

* $p < .05$.

B = unstandardized coefficient