


The neural mechanisms of threat and reconciliation efforts between Muslims and non-Muslims

Dorottya Lantos, Yong Hui Lau, Winnifred Louis & Pascal Molenberghs


To cite this article: Dorottya Lantos, Yong Hui Lau, Winnifred Louis & Pascal Molenberghs (2020) The neural mechanisms of threat and reconciliation efforts between Muslims and non-Muslims, *Social Neuroscience*, 15:4, 420-434, DOI: [10.1080/17470919.2020.1754287](https://doi.org/10.1080/17470919.2020.1754287)

To link to this article: <https://doi.org/10.1080/17470919.2020.1754287>

 View supplementary material 

 Published online: 04 May 2020.

 Submit your article to this journal 

 Article views: 481

 View related articles 

 View Crossmark data 

 Citing articles: 1 View citing articles 



The neural mechanisms of threat and reconciliation efforts between Muslims and non-Muslims

Dorottya Lantos^{a*}, Yong Hui Lau^b, Winnifred Louis^c and Pascal Molenberghs^d

^aPsychology Department, Goldsmiths, University of London, London, UK; ^bSchool of Psychological Sciences, The University of Melbourne, Australia; ^cSchool of Psychology, The University of Queensland, Australia; ^dISN Psychology, Institute for Social Neuroscience, Melbourne, Australia

ABSTRACT

To reduce the escalation of intergroup conflict, it is important that we understand the processes related to the detection of group-based threat and reconciliation. In the present study, we investigated the neural mechanisms of such processes using functional Magnetic Resonance Imaging (fMRI). Functional neuroimaging techniques may shed light on quick, automatic responses to stimuli that happen outside of conscious awareness and are thus increasingly difficult to quantify relying only on participants' self-reported experiences. They may further provide invaluable insight into physiological processes occurring in situations of sensitive nature, whereby participants—deliberately or not—may withhold their honest responses due to social desirability. Non-Muslim Western Caucasian participants watched short video clips of stereotypical Middle-Eastern Muslim males threatening their ingroup, offering reconciliation to the ingroup, or making a neutral statement. Threatening statements led to increased activation in the amygdala, insula, supramarginal gyrus, and temporal lobe. Reconciliation efforts led to increased activation in the prefrontal cortex, anterior cingulate gyrus, and caudate. The results suggest that threat detection is a relatively automatic process while evaluating and responding to reconciliation offers requires more cognitive efforts. The implications of these findings and future research directions are discussed.

ARTICLE HISTORY

Received 7 October 2019
Revised 6 March 2020
Published online 30 April 2020

KEYWORDS



Intergroup conflict;
intergroup processes; fMRI;
social neuroscience;
terrorism; stereotypes

Attention-grabbing headlines of terrorism threats have become a part of our everyday lives. However, data suggest that this threat is overexaggerated by the media, compared to that warranted by the actual number of deaths caused by terrorism (Ritchie, 2019). In 2016, terrorism as a cause of death was 3906 times overrepresented in the US media. For example, the New York Times covered terrorism as a cause of 35.6% of all their reported deaths occurring in the US in 2016, while in reality it accounted for less than 0.01% of deaths that year. Acts of terrorism are over 350% more likely be covered by the US media if the perpetrator is Muslim (Kearns et al., 2019). Continuous exposure to media representations of certain groups as harboring evil intentions likely leads to the escalation of intergroup conflict and violence (Ahmed & Matthes, 2017; West & Lloyd, 2017), leading to potentially devastating consequences for individuals and societies alike. Prolonged periods of intergroup conflict may lead to skepticism and resistance to reconciliation efforts. However, such efforts are a necessary step toward long-term peace


(Čehajić-Clancy et al., 2016; Tropp et al., 2017). With the stakes so high, it is important to better understand why we are so prone to overreport and attend to news of terrorism threat, as well as both intergroup violence and intergroup reconciliation. To provide some insights into these underlying mechanisms, the present study explored the neural underpinnings of outgroup threats targeting the ingroup, as well as intergroup reconciliation efforts by outgroup members.

Intergroup threat

Humans are inherently social beings. Ingroup members have been critical for our survival throughout our evolutionary history. Being the member of a group meant not only safety from predators, but also support in obtaining food and opportunities for successful reproduction (Baumeister & Leary, 1995; DeWall et al., 2011; Dunbar, 2011; Jetten et al., 2012). From an evolutionary perspective, threat to the ingroup can thus be interpreted as a form of existential threat (Das et al., 2009). Unsurprisingly, humans

CONTACT Pascal Molenberghs  pascalmolenberghs@gmail.com  The Institute for Social Neuroscience, ISN Psychology, R5/443 Upper Heidelberg Rd, Ivanhoe VIC 3079, Australia

*Dorottya Lantos and Yong Hui Lau contributed equally to this paper.

 Supplemental data for this article can be accessed [here](#).

© 2020 Informa UK Limited, trading as Taylor & Francis Group

are highly sensitive to the suffering of ingroup members and threats targeting the ingroup (Boyer et al., 2015; Neuberg & Schaller, 2016). Neuroscientific findings also point toward the ability of the human brain to readily detect outgroup threats and increased responses to ingroup suffering (Hein et al., 2010; Molenberghs et al., 2016; Molenberghs & Louis, 2018; Xu et al., 2009).

Although it is involved in a number of functions related to arousal, activity in the amygdala has been reliably linked to threat detection (e.g., Carlson et al., 2009; Öhman, 2005). For example, individuals with bilateral amygdala lesion detect threat less readily in comparison to those with intact amygdala or unilateral lesion (Anderson & Phelps, 2001). Chekroud et al. (2014) suggest that this threat-detection function of the amygdala is activated to a higher degree when the source of the threat is an outgroup member. Neuroimaging studies have shown increased activity in the amygdala of Caucasian participants while observing Black faces (e.g., Cunningham et al., 2004; Wheeler & Fiske, 2005), a group stereotyped as violent and threatening (Trawalter et al., 2008). Japanese participants' implicit evaluations of South Koreans can be reliably predicted from the activity observed in their amygdala as they observe images related to South Korea (Izuma et al., 2019). This suggests that the amygdala plays an important role in prejudice (Amodio et al., 2003; Phelps et al., 2000; Van Bavel et al., 2008; for reviews, see Amodio, 2014; Cikara & Van Bavel, 2014; Mattan et al., 2018).

A recent fMRI study investigated the neural responses to outgroup members harming ingroup members (Molenberghs et al., 2016). Participants watched short videos of students from their own and another university harming each other or interacting peacefully. Harming others compared to interacting peacefully, led to increased activation in brain areas involved in threat detection, emotional awareness and moral sensitivity such as the amygdala, insula and orbitofrontal cortex (OFC). This effect was especially pronounced in instances when a member of the outgroup harmed a member of the ingroup. However, the harming depicted in this experiment was low on ecological validity. Seeing someone with a shared university membership being harmed by someone holding a different university membership can hardly be thought of as presenting a true existential threat to the survival of the ingroup. In the present experiment we aim to overcome this issue by investigating neural activity in response to a more realistic and potentially dangerous intergroup situation: extremist Muslims threatening to kill Westerners.

Intergroup reconciliation

While there is certainly room to expand our knowledge with regards to the neural responses to intergroup

threat, there are hardly any neuroscientific findings reported on the topic of intergroup reconciliation. Behavioral findings suggest that intergroup reconciliation is a complex and often long process, influenced by multiple factors (Kelman, 2010). For example, Israeli-Jewish participants responded positively to the reconciliation offer of a Palestinian leader showing empathy when their trust toward Palestinians was high, but signs of empathy had negative effects on their evaluation of the offer when trust was low (Nadler & Liviatan, 2006). Reconciliation is predicted by forgiveness, which in turn, is associated with ingroup identity, common ingroup identity, and political ideology, among others (Noor et al., 2008). Individuals consider historical and contextual knowledge when evaluating reconciliation offers, and these are often met with initial skepticism and resistance (Blatz & Philpot, 2010; Van Tongeren et al., 2014; Wohl et al., 2012).

While the initial reaction to a threat is a relatively automatic process centered around the amygdala (Chekroud et al., 2014), we anticipate reconciliation offers to elicit responses in more cognitive regions such as the prefrontal lobe, indicative of higher cognitive functions involved in evaluating the offers. This is in line with the dual-process models of decision-making (e.g., Kahneman, 2003, 2011; for a review on different models see Evans, 2008; Strack & Deutsch, 2015). In System 1, information processing happens in a fast, automatic way, relying on heuristics, and does not exert much cognitive effort. In System 2, information processing takes a longer time, requiring conscious effort to evaluate information. In line with this, social categorization and the activation of relevant stereotypes can happen automatically, outside of our conscious awareness (Banaji et al., 2001; Tajfel, 1969). In contrast, being presented with stereotype incongruent stimuli should motivate careful evaluation of the available information through System 2 processing.

An evolutionary approach to forgiveness supports the necessity of weighing multiple factors in a complex decision-making process (Billingsley & Losin, 2017; Burnette et al., 2012; McCullough, 2008; McCullough et al., 2013). Notably, individuals facing such a process must consider the value of their relationship to the wrongdoer, assess the extent to which future interactions may be beneficial, and evaluate the likelihood that the wrongdoer will harm them in the future. Similarly, individuals presented with reconciliation offers must draw on their social and cultural knowledge, consider the (often grim) history of the relationship between the ingroup and the outgroup, assess the present situation and message critically, and decide whether or not they are willing to trust the person offering reconciliation. If they decide to trust an

insincere offer, they may expose themselves and the ingroup to potential harm. If they reject a potentially sincere reconciliation offer, they may prolong intergroup hostility and violence, and promote threat to the self and the ingroup in this way. It is thus crucial to make the right decision in this situation, which likely involves a high level of cognitive effort.

A recent review investigating the neural systems related to forgiveness highlights the involvement of two main neural networks (Billingsley & Losin, 2017). First, since the instinctive response to unfair treatment or harm is that of revenge and punishment, an inhibitory network must constrain such a response. Accordingly, regions linked to impulse control, inhibition, and emotion regulation including the dorsolateral prefrontal cortex (dlPFC), the ventrolateral prefrontal cortex (vlPFC), and the dorsal anterior cingulate cortex (dACC) were associated with forgiveness (e.g., Brüne et al., 2013; Ricciardi et al., 2013; Tabibnia et al., 2008; Will et al., 2014). This network may be responsible for downregulating the negative affect, and thus the spontaneous aggressive response, resulting from unfair treatment. Second, understanding others and therefore engaging the theory of mind network is a crucial step toward assessing whether an apology is sincere or not (Billingsley & Losin, 2017; Burnette et al., 2012). Areas corresponding to this network, involving the temporal parietal junction (TPJ), the medial prefrontal cortex (mPFC), and the precuneus have previously shown increased activation in the context of forgiveness (e.g., Ohtsubo et al., 2018; Strang et al., 2014; Will et al., 2014; Yamada et al., 2012; Young & Saxe, 2009).

In addition to the results suggesting the role of an inhibitory network in the process of forgiveness, a recent comprehensive literature review suggests that previous findings connecting emotions and the prefrontal cortex can be integrated into a unified model through appraisal or evaluation processes (Dixon et al., 2017). Through the process of evaluating the reconciliation messages, individuals are confronted with conflict. They see members of an outgroup associated with previous conflict, and yet the outgroup member's message is incongruent with this prior knowledge. Activity in the mPFC and ACC have been associated not only with evaluation, but also specifically with the regulation of emotional conflict (Etkin et al., 2011). These same areas are also implicated in the processing of social information (Amodio & Frith, 2006; Van Overwalle, 2009). Finally, a vital part of the process of evaluating the reconciliation message is deciding how to respond and the mPFC and ACC have also been associated with decision-making (Rushworth

et al., 2012, 2011). Therefore, we anticipate increased activation in the PFC and ACC in response to outgroup reconciliation efforts.

Overview of the present research

In the present research we investigated the neural responses to intergroup threat or reconciliation. Specifically, we recruited non-Muslim Western Caucasian participants, and explored their brain activity using fMRI as they observed short video clips of a stereotypical Muslim person either (1) threatening to kill Westerners, (2) offering reconciliation to Westerners, or (3) making a neutral statement. We hypothesized that participants will exhibit increased activation in the amygdala when watching the threatening videos in comparison to the videos of reconciliation and neutral messages. We further hypothesized increased activation in prefrontal brain regions of participants when presented with reconciliation messages in comparison to threatening or neutral messages, indicative of the cognitive evaluation of the offers, particularly in the PFC and ACC.

Method

Participants

Thirty participants (14 females, 16 males; ages 19–55 ($M = 28.97$, $SD = 10.39$)) took part in both sessions of the present experiment. All participants were non-Muslim Western Caucasian, had normal or corrected-to-normal vision, and had no MRI contraindications. Caucasian adult participants were recruited through flyers and advertisements displayed across Melbourne and posted online. After indicating their interest, participants were interviewed via e-mail or telephone to ensure they were Caucasian, non-Muslim, and MRI safe. Upon arrival to the MRI facilities, they were once again thoroughly interviewed by a qualified radiographer to ensure that they met all the MRI safety criteria before entering the scanner. The participants were unaware of the hypotheses prior to taking part in the experiment and were reimbursed 50AUD for their time. All procedures were approved by the University of Melbourne ethics committee and all participants gave written informed consent.

The sample size was based on the recommendation suggesting that 30 participants is sufficient to detect large ($d = 0.8$) effect sizes in fMRI pairwise comparisons with 80% power and an alpha of .001 (Yarkoni, 2009). We

expected large effects as the experimental stimuli contained statements of extreme negativity (threats) and maximal positivity (reconciliation offers).

Stimuli Design

We created the experimental stimuli using the VSDC Free Video Editor software. Each video clip lasted five seconds and showed a Muslim male with Middle-Eastern appearance (i.e., an outgroup member) making a statement with a neutral facial expression. Participants were informed that the video clips were from non-English speakers, that the audio of each video clip was removed and that the messages were instead presented as subtitles. This ensured that any mismatch between the movement of the lips and the words ostensibly said by the actors did not cause suspicion. In reality, we manipulated the statements to create three conditions: a threat condition, in which the outgroup member threatened the ingroup, e.g., *“Westerners are our enemies. We Muslims have to wipe all of them out without mercy.”*; a reconciliation condition, in which the outgroup member made a reconciliation statement, e.g., *“Muslims love peace. So, we call for greater harmony between us and Westerners.”*; and a neutral control condition, in which the outgroup member made a neutral statement, e.g., *“Most Muslims and Westerners go to school between the age of 5 and 18”*. The complete list of statements can be found in Supplementary Material. All videos used in this study are available on the Open Science Framework as well as the group data: <https://osf.io/e4b6t/>. The video clips were followed by the question *“How do you feel about the*

statement?” and participants were asked to respond within 5 seconds using a 9-point scale (1 = very negative; 9 = very positive; see Figure 1). The nine-digit response options were presented randomly in the form of a 3 × 3 keypad, ensuring similar motor movements across the three experimental conditions.

The number of characters and words in the statements did not significantly differ across the three conditions, $p_{characters} = .87$; $p_{words} = .97$. All video clips were identical in size (1280 × 720 pixels) and speed (30 frames per second). All visual stimuli were presented in colored images on a black background using e-Prime 2 software (Psychology Software Tools Inc, 2013) and a rear projection system connected to the MRI scanner.

Procedure

Participants completed two sessions. In the first session, they provided demographic details and completed forms related to MRI safety on the campus of The University of Melbourne. In order to familiarize participants with the fMRI task, they underwent a 12-minute long practice session, consisting of 60 trials. Following each trial, participants indicated how they felt about the statement presented in the video on a 9-point scale (1 = very negative; 9 = very positive). The stimuli presented during the practice trials were not included in the experimental session.

The second session took place in the Monash Biomedical Imaging Center. Prior to the MRI scanning, participants underwent a final MRI screening with the assistance of a radiographer and were reminded that it

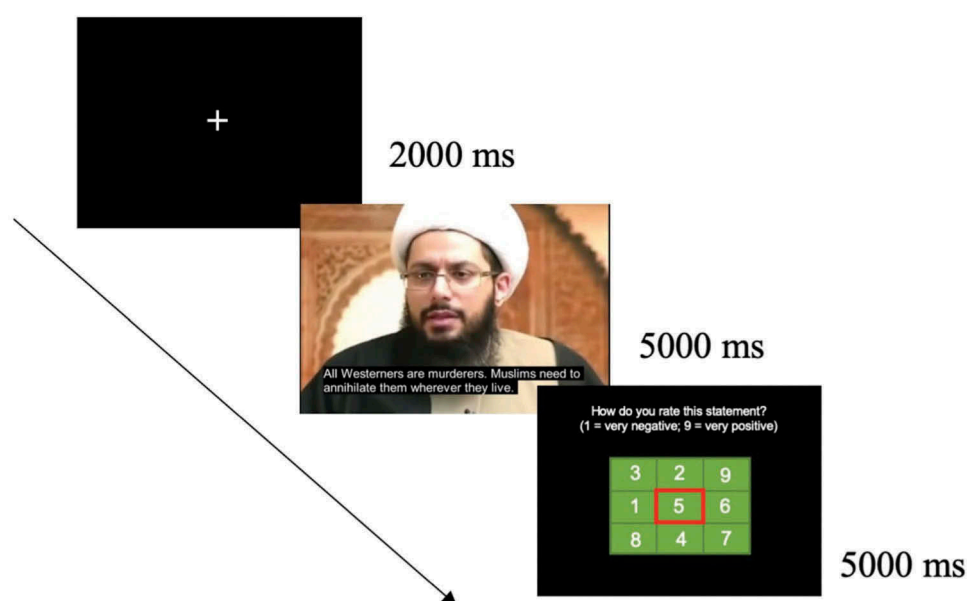


Figure 1. The schematic representation of a sample trial involving a 2 s fixation cross, a 5 s video clip, and a 5 s response.

was critical to respond to every trial and to fully engage with the content by imagining how they would feel in real life when listening to similar statements. Participants were provided with ear plugs in order to minimize the sound disturbance inside the scanner. They were given two response pads with two buttons on each pad, allowing them to navigate among the response options during the task. Foam pads were placed around the head to reduce head movements, and a coil with an attached mirror allowed participants to see the stimuli, presented through a projector. For safety purposes, an emergency response button was placed on participants' chest.

Prior to the functional scans, participants were given an opportunity to practice the experiment for a couple of minutes inside the scanner. The experiment consisted of three 12-minute-long functional runs. Six minutes of each sequence consisted of the stimuli investigated in the present paper, with the remaining 6-minutes of each collecting data not explored here (see Supplementary Material for details). Each sequence began with reminder instructions, followed by a fixation cross (2 s), 30 video clips (5 s each) investigated in the present paper, 30 video clips not analyzed here (5 s each), with each video clip followed by a feedback opportunity (5 s each). All three runs consisted of the same 60 video clips presented in a pseudorandom fashion, but different combinations of videos and subtitles were presented across participants. The videos were counterbalanced, each one appearing in the threatening, neutral and reconciliation conditions an equal amount of times across participants. This ensured that differences between the conditions cannot be attributed to differences between the video clips. Each participant saw each video clip and actor associated with only one condition. Per run, participants watched 10 different threatening statements, 10 different neutral statements and 10 different reconciliation statements from an out-group member. Two sets of 10 statements were created for each of the three conditions (see Supplementary Material for details) and each list was counterbalanced across participants to ensure that any observed effects are not merely due to the wording of certain stimuli. The same video clips were then repeated across the three runs. Altogether, the entire session took approximately 40 minutes, including a 5-minute structural scan and three 12-minute functional scans. At the end of the second session, participants were debriefed, reimbursed 50AUD and thanked for their participation.

Image Acquisition

A 3-Tesla Siemens Skyra MRI scanner with a 32-channel head volume coil was used to collect the imaging data. A T1-weighted whole-brain structural image was

obtained for anatomical reference with a repetition time (TR) of 2300 ms, echo time (TE) of 2.07 ms, flip angle (FA) of 9°, 192 cube matrix, voxel size of 1 mm³ and slice thickness of 1 mm.

Using the gradient echo planar imaging (EPI) technique, functional images were acquired with following parameters: TR = 2.32 seconds, TE = 30 ms, FA = 80°. Forty transversal slices with 76 × 76 voxels at 3 mm² in-plane resolution covered the whole brain. In total, 319 images were obtained for each functional run (with the first three images from each functional run being removed to allow for steady-state tissue magnetization).

Pre-processing of the fMRI data

We used Statistical Parametric Mapping (SPM12) software run through Matlab software (Mathworks, R2014a) to conduct the pre-processing and statistical analyses of the imaging data. To correct for head movement artifacts, we re-aligned all EPI images to the first scan of each functional run through affine transformation. After saving the re-alignment parameters, we co-registered the T1 anatomical image to the mean functional image. To counter the variations in brain size and anatomy among participants, we normalized the T1 structural scans to the MNI (Montreal Neuropsychological Institute) standard T1 template with a voxel size of 1 × 1 × 1 mm³. We used the same parameters to normalize the functional images with a voxel size of 3 × 3 × 3 mm³. This mathematical transformation process allows selected brain areas to be matched with the MNI template and standardized across all participants. In the final stage of data pre-processing, we smoothed all normalized functional images with an isotropic Gaussian kernel of 9 mm FWHM (full-width at half-maximum). The smoothing process was intended to average the brain signals by enhancing signals with low frequency and removing signals with high frequency.

Statistical analyses

fMRI analyses

We created a general linear model for each participant during the first-level analyses. For each participant, in each condition, an event-related design identified the voxels across the brain which showed significant blood oxygen level dependent (BOLD) changes compared to the baseline. The events were modeled with a canonical hemodynamic response function time-locked to the onset of each video clip and a duration of 5 seconds. We used a high-pass filter of 128 seconds in the first-level models for de-noising.

During the second-level analyses, contrast images at the individual level for all three conditions (versus the baseline)

were included in a one-way ANOVA. The significant main effect was further explored with six T-contrasts: threat > reconciliation, threat > neutral, reconciliation > threat, reconciliation > neutral, neutral > threat, neutral > reconciliation. In line with our a priori hypothesis, we further conducted region of interest (ROI) analyses for the left and right amygdala (as defined by the AAL atlas in the WFU_Pickatlas (https://www.nitrc.org/projects/wfu_pickatlas/) toolbox) for the threat > reconciliation and threat > neutral contrasts. Results were thresholded at an uncorrected voxel-level threshold of $p < .001$. Significance was defined as a family-wise error rate (FWE) voxel-level threshold of $p < .05$ corrected for the whole brain for the whole brain analyses or corrected for the size of the amygdala for the region of interest analyses.

Results

Manipulation check

We used Microsoft Excel and IBM SPSS Statistics software to conduct a manipulation check, analyzing the emotional ratings. First, the rating scores for 30 videos in each condition (i.e., threat, neutral, reconciliation) were averaged. To test whether the emotional ratings of the three conditions differed significantly from each other, we conducted a one-way repeated measures ANOVA. Participants failed to submit a response within the 5-second window in less than 0.01% of the trials. These trials were not included in the analyses.

Mauchly's Test of Sphericity indicated that the assumption of sphericity was violated, $\chi^2(2) = 9.26$,

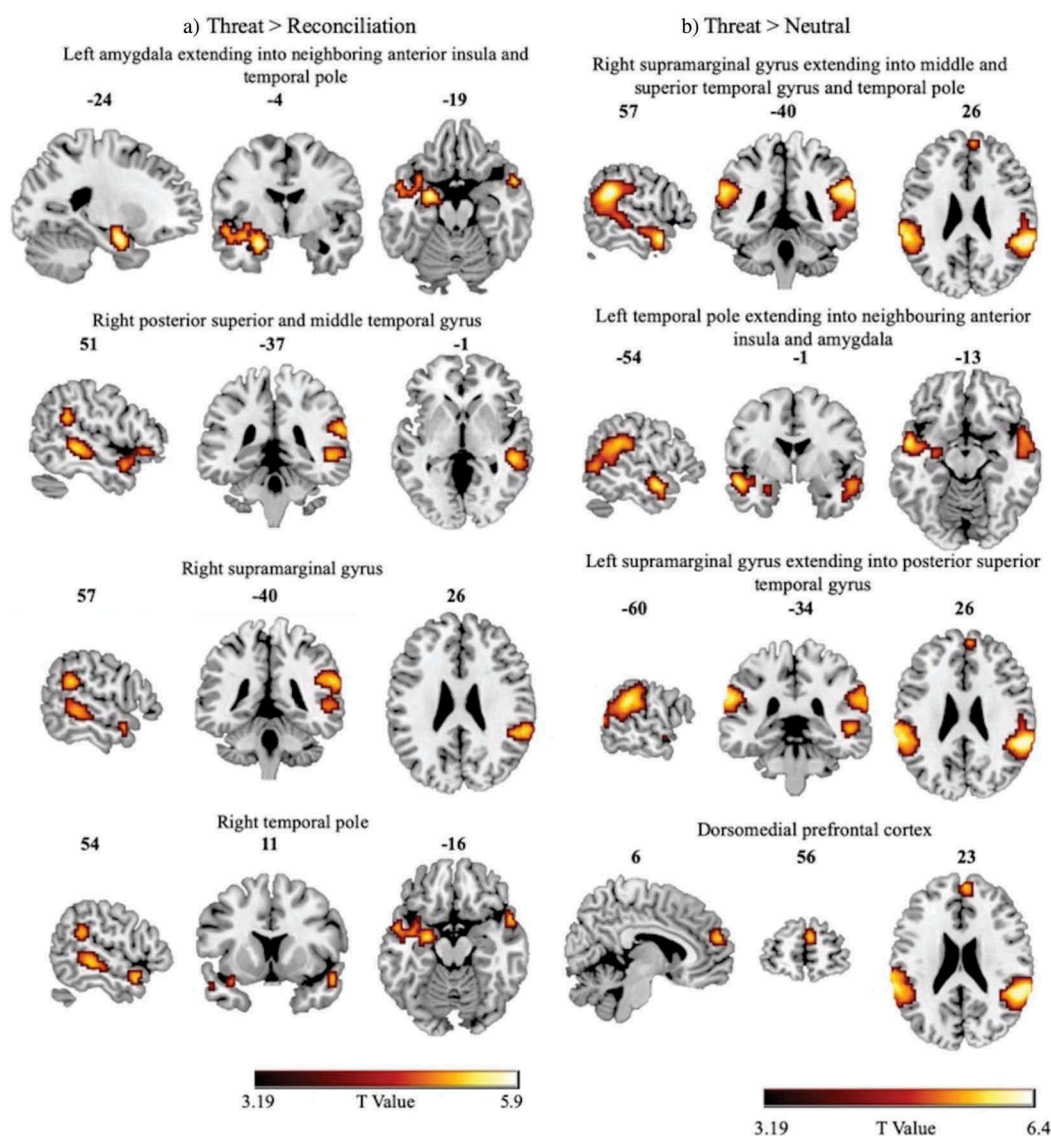


Figure 2. Significant BOLD-signal changes at each significant peak voxel, observed during the A) threat > reconciliation and B) threat > neutral contrasts. The images were created with MRICron using the *ch2better* template.

$p = .010$, thus its degree of freedom was subsequently corrected using the Huynh-Feldt estimates of sphericity ($\epsilon = .82$). A one-way repeated measures ANOVA revealed that there was a significant difference between negative statements ($M = 1.11$, $SD = .28$), neutral statements ($M = 5.86$, $SD = 1.19$) and positive statements ($M = 8.47$, $SD = .63$), $F(1.64, 49.27) = 668.51$, $p < .001$. Bonferroni corrected pairwise comparisons indicated that the negative statements were rated more negatively than the neutral, $t(29) = -21.15$, $p < .001$ and positive statements, $t(29) = -52.04$, $p < .001$, while the positive statements were rated significantly more positively than the neutral statements, $t(29) = 11.18$, $p < .001$. These results show that our stimuli significantly differed in their valence, and thus were suitable for the manipulations we aimed for.

fMRI Results

The F-contrast investigating the differences between the three conditions produced significant main effects,

illustrated in Supplementary Table 1. To further investigate which conditions produced the observed effects, we next conducted pairwise comparisons. The threat > reconciliation and threat > neutral contrasts revealed increased activation in left amygdala, left insula, right posterior superior and middle temporal gyrus, right supramarginal gyrus and bilateral temporal pole (Figure 2 and Table 1). The threat > neutral contrast also revealed increased activation in the left supramarginal gyrus, left middle and superior temporal gyrus and the dorsomedial prefrontal cortex. In line with our a priori hypothesis, ROI analyses revealed increased activation in the bilateral amygdala for both the threat > reconciliation (left amygdala: $-24, -4, -19$, $Z = 5.07$, extent 40, $p_{FWE} < .001$; right amygdala: $30, -4, -16$, $Z = 3.60$, extent 17, $p_{FWE} = .008$) and threat > neutral contrasts (left amygdala: $-24, -4, -19$, $Z = 3.52$, extent 14, $p_{FWE} = .01$; right amygdala: $27, -7, -13$, $Z = 3.90$, extent = 13, $p_{FWE} = .003$). The reconciliation > threat and reconciliation > neutral conditions revealed increased activation in the prefrontal cortex, anterior cingulate

Table 1. The results of the T-contrasts investigating the hypotheses.

Threat > Reconciliation						
Anatomical Regions	Cluster size	Peak Z-value	Peak $p_{FWE-corr}$	MNI Coordinates		
				x	y	z
Left amygdala extending into neighboring anterior insula and temporal pole	269	5.07	.005	-24	-4	-19
Right posterior superior and middle temporal gyrus	151	4.67	.028	51	-37	-1
Right supramarginal gyrus	131	4.66	.030	57	-40	26
Right temporal pole	59	4.58	.040	54	11	-16
Threat > Neutral						
Anatomical Regions	Cluster size	Peak Z-value	Peak $p_{FWE-corr}$	MNI Coordinates		
				x	y	z
Right supramarginal gyrus extending into middle and superior temporal gyrus and temporal pole	945	5.78	<.001	57	-40	26
Left temporal pole extending into neighboring anterior insula and amygdala	188	5.20	.003	-54	-1	-13
Left supramarginal gyrus extending into posterior superior temporal gyrus	518	5.17	.003	-60	-34	26
Dorsomedial prefrontal cortex	46	4.77	.018	6	56	23
Reconciliation > Threat						
Anatomical Regions	Cluster size	Peak Z-value	Peak $p_{FWE-corr}$	MNI Coordinates		
				x	y	z
Bilateral middle and superior frontal gyrus extending into the medial prefrontal cortex and anterior cingulate gyrus	1466	5.56	<.001	-39	47	-7
Bilateral caudate	193	4.58	.040	-12	20	5
Reconciliation > Neutral						
Anatomical Regions	Cluster size	Peak Z-value	Peak $p_{FWE-corr}$	MNI Coordinates		
				x	y	z
Medial prefrontal cortex extending into anterior cingulate gyrus and bilateral caudate	819	5.66	<.001	6	56	20

gyrus, and bilateral caudate (Figure 3 and Table 1). The activations of the neutral > threat (Supplementary Table 2) and neutral > reconciliation (Supplementary Table 3) contrasts are presented in Supplementary Material.

Discussion

The human brain evolved under circumstances where survival depended on being the member of a group (Baumeister & Leary, 1995; DeWall et al., 2011; Dunbar, 2011; Jetten et al., 2012). Threats targeting the ingroup can pose an existential threat to the survival of the ingroup (Das et al., 2009), and individuals readily detect them (Molenberghs et al., 2016; Molenberghs & Louis, 2018). The media is biased to

focus on sensational reports of terrorism, especially when the perpetrators are from a stereotypically violent ethnic or religious backgrounds (Ahmed & Matthes, 2017; Kearns et al., 2019; Powell, 2011). Frequent exposure to such framing of certain groups promotes the endorsement of negative stereotypes and hostility, and the arising emotional conflict may lead to skepticism when presented with offers of reconciliation (Saleem et al., 2017). To prevent the escalation of intergroup conflict, we must gain a better understanding of the way individuals respond to threats and reconciliation offers from perceived dangerous outgroups. The present study was the first to investigate the neural mechanisms of intergroup threat and reconciliation using an ecologically valid, realistic paradigm.

Frequent exposure to stereotypical news about outgroup members increases negative attitudes toward them (Arendt & Northup, 2015). This phenomenon is amplified through the use and accessibility of traditional and social media. Internalizing threatening stereotypes (consciously or not) may more easily than ever before in human history translate to not only prejudice, but potentially even hostility, in simple daily interactions. The prevention of stereotypical thinking and negative bias stemming from the reinforcement of linking certain outgroups to violent behavior is a timely question requiring investigation.

We explored the neural responses of non-Muslim Caucasian Western participants to short video clips showing Muslim males with a stereotypical Middle-Eastern appearance threatening their group. Results revealed increased activation in the amygdala and insula while observing threatening messages in comparison to reconciliation messages or neutral messages. The amygdala is associated with responding to high-arousal situations, representing fear-related stimuli, encoding threat and distress, and contributes to the recognition of emotions (Adolphs, 2008; Blair, 2003, 2008; Cikara & Van Bavel, 2014; Fusar-Poli et al., 2009). Similarly, activation in the insula is tied to introspective awareness, the recognition of emotions, and the processing of threatening stimuli, and is associated with greater activation while observing intergroup harm (Klumpp et al., 2012; Molenberghs et al., 2016; Stein et al., 2007; Veit et al., 2011).

Observing threatening statements further led to increased activation in the left middle and superior temporal gyrus and bilateral temporal pole. These areas are associated with semantic processing (Davey et al., 2016; Visser et al., 2012), and more specifically the representation and retrieval of social knowledge and stereotypes (Contreras et al., 2012; Olson et al., 2013; Skipper et al.,

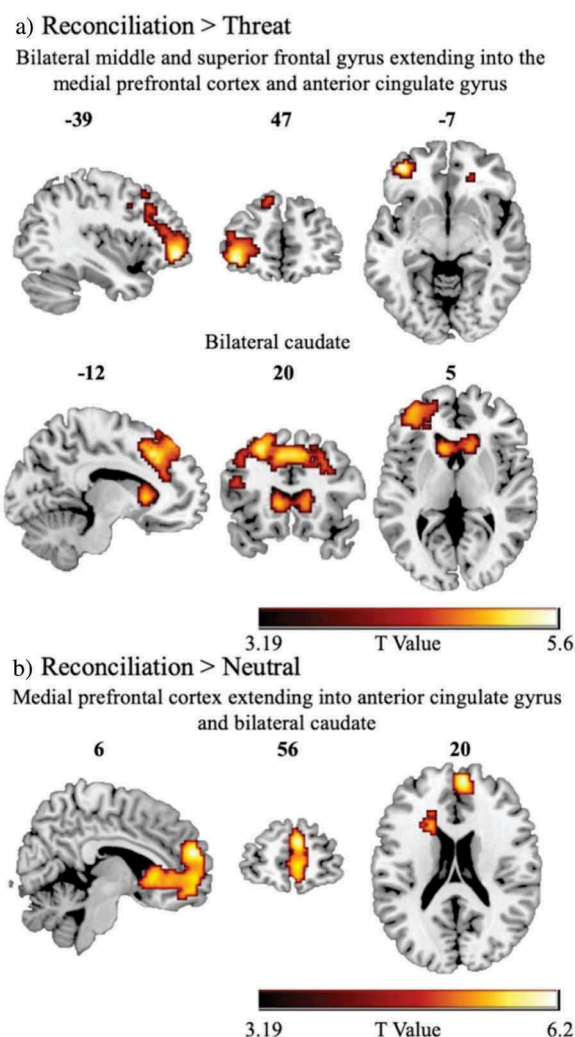


Figure 3. Significant BOLD-signal changes at each significant peak voxel, observed during the A) reconciliation > threat and B) reconciliation > neutral contrasts. The images were created with MRICron using the *ch2better* template.

2011). Activation in these areas is furthermore related to thoughts regarding the characteristics of people rather than objects (Amodio, 2014; Zahn et al., 2007), and increased activity in the right temporal pole has previously been associated with observing fearful vs. neutral faces (Preibisch et al., 2009). Increased activation in these areas may thus contribute to threat detection through the monitoring of social cues. Finally, the supramarginal gyrus also showed increased activation in response to threatening messages compared to the reconciliation and neutral conditions. The supramarginal gyrus has been associated with increased attention to specific semantic information (Molenberghs et al., 2017). This might suggest that the threatening statements captured the attention of participants to a higher degree than the non-threatening statements. Recent research points toward the greater attention-grabbing properties of emotional and moral words in the context of social media posts (Brady et al., 2020), which likely recruits similar neural regions.

To our knowledge, no previous neuroimaging studies have examined reactions to outgroup reconciliation statements. Reconciliation messages increased activation in the medial prefrontal cortex (mPFC), the anterior cingulate cortex (ACC) and the caudate. The prefrontal cortex is often involved in appraisal and evaluation (Dixon et al., 2017). The mPFC and ACC are often associated with conflict evaluation, regulation, and monitoring (Egner et al., 2007; Etkin et al., 2011, 2006). The mPFC and ACC are further linked to both the conditioning, and the extinction of fear (Etkin et al., 2011; Giustino & Maren, 2015; Mechias et al., 2010), decision-making (Rushworth et al., 2011), and the processing of social information and stereotypical thinking (Amodio & Frith, 2006; Van Overwalle, 2009). These areas overlap with some of the most often reported neural correlates of forgiveness, a precursor to reconciliation (Billingsley & Losin, 2017; Noor et al., 2008). Specifically, the prefrontal cortex has been associated with forgiveness as well as impulse control, inhibition, and emotion regulation (e.g., Brüne et al., 2013; Ricciardi et al., 2013; Tabibnia et al., 2008; Will et al., 2014). These regions likely contribute to the downregulation of negative affect and to the inhibition of spontaneous aggressive responses to unfair treatment and harm. In the context of the present procedure, they may play a role in inhibiting negative stereotypical attributions to the actor in the short video clips upon hearing their harmless, friendly intention. The mPFC has been implicated in the forgiveness literature as a key component of the theory of mind network (Billingsley & Losin, 2017; Ohtsubo et al., 2018; Will et al., 2014; Yamada et al., 2012). It is thus possible that, alternatively to the above interpretation of the activation in the mPFC

being associated with (uniquely) conflict evaluation, regulation, and monitoring, it may (additionally) play a key role in engaging theory of mind. Taking the perspective of the actor can allow participants to make better informed decisions about whether or not their intentions are genuine.

We further observed an increased activation in the caudate in response to outgroup reconciliation offers. The caudate is consistently linked with reward processing (Haruno & Kawato, 2006; Knutson et al., 2001; O'Doherty et al., 2001; Rolls, 2000). Activity in the striatum (the region encompassing the caudate) is further related to social rewards, positive social interactions, and cooperation with others (Bhanji & Delgado, 2014; Decety et al., 2004; Fareri et al., 2012; Izuma et al., 2008). The increased activity in this area is likely associated with the highly positive wording of the reconciliation messages, supported by participants' high positivity-ratings for these messages.

The brain regions and corresponding processes activated by the threatening messages—the amygdala, insula, temporal lobes, and supramarginal gyri, which are linked to threat detection, stereotypical thinking, increased attention—suggest quick, automatic neural responses. This is evolutionarily adaptive, since these processes allow humans to identify threat and danger in a quick and efficient way, enabling them to act on it quickly. This is also in line with the dual-process models of decision-making, suggesting that strong associations recruit System 1 processing based on heuristics and intuition (e.g., Kahneman, 2003, 2011).

These results indicate that there may be a network of brain regions responding to terrorism threats in a relatively automatic way, setting off an alarm signal in the brain to which we respond instantly in order to avoid potential danger. This might explain why people are so fascinated with news stories relating to terrorism and why the media is biased to overreport such stories.

On the other hand, reconciliation messages activated a vastly different neural network—the mPFC and ACC—often associated with higher level cognitive processing. Participants are likely exerting effort into the evaluation of the reconciliation messages, which may be incongruent with prior experiences with the behavior of stereotypical outgroup members. This is especially likely since the reconciliation messages are presented in a randomized order along the threatening (and neutral) messages. The results of behavioral studies suggest that reconciliation is a complex process affected by multiple factors, including the perceived personality and behavior of the outgroup leader, historical and contextual knowledge, as well as the identity and ideology of the individual (Blatz & Philpot, 2010; Kelman, 2010; Nadler &

Liviatan, 2006; Noor et al., 2008). Individuals need to consider multiple questions during the decision-making process before making up their mind about the reconciliation offer. Can I trust this person? Can I trust this message? What are the consequences of trusting this message? What are the consequences of *not* trusting this message? In the present experiment, this likely entails both the activation of fear through the threatening stereotypical looks of those in the videos and potentially fear extinction through the positive message, as well as conflict evaluation and conflict regulation, it draws on social knowledge and involves decision-making processes, all processes previously associated with the mPFC and ACC (Amodio & Frith, 2006; Egner et al., 2007; Etkin et al., 2011, 2006; Rushworth et al., 2011; Van Overwalle, 2009).

These findings further support dual-process models of decision-making, which suggest that effortful System 2 processing takes place when presented with incongruent information or with information that one needs to consciously evaluate (e.g., Kahneman, 2003, 2011). The dual-process model is not without criticisms (Keren, 2013; Mugg, 2015), and a competing unimodel of judgment has been introduced, suggesting that both deliberative and intuitive cognitions rely on the same processes (Erb et al., 2003; Kruglanski & Gigerenzer, 2011). However, the present results indicate two distinct neural networks are engaged during automatic and deliberative thinking, providing support for the dual-process model, rather than the unimodel of judgment.

This knowledge may inform policy makers and social scientist about the ways in which targeting reconciliation efforts may be the most beneficial. Understanding the cognitive load that the processing of a reconciliation situation puts on individuals may motivate policy makers to deliberately emphasize certain pieces of information. For example, emphasizing positive intergroup events, the currently safe social environment, or the benefits which the reconciliation holds for the ingroup are all ways to reinforce the individual's positive evaluation of the situation. The salience of such information makes decision-making somewhat easier by stressing the positive aspects of intergroup history and the positive potential outcomes of the reconciliation. In the absence of such positive information, the automatic response might be to recall the negative prior events with the outgroup and to focus on the potential detrimental outcomes of future interactions which undermines reconciliation.

However, it is important to consider whether it is the incongruency of the reconciliation offers coming from individuals with negative stereotypes that drives the observed effects (at least in the ACC), rather than the

content of the messages. The ACC is associated with the mediation of conflict in the case of violated expectations (Egner et al., 2007; Etkin et al., 2011), and the present experiment cannot rule out such processes underlying our findings. Future experiments should follow up on this limitation.

While the present study yields important information about the neural mechanisms involved in the processing of intergroup threat and reconciliation, it is important to keep in mind that we only explored the perception of one ingroup (non-Muslims Western Caucasian participants) toward only one outgroup (stereotypical Middle-Eastern Muslims). The same processes between different ingroups and outgroups could result in different patterns of neural activity, depending on the groups and associated stereotypes. Future studies should aim to replicate the present experiment using a different ingroup and a different outgroup to explore the generalizability of the present results. It would further support the generalizability of the findings to compare participants' neural responses to threat and reconciliation statements between groups that differ along race, ethnicity, culture, gender, religion, or nationality.

Additionally, the effects observed here of both the threatening and the reconciliation messages should be compared to the effects of similar messages coming from groups associated with less negative stereotypes. This would give more insight into the observed mechanisms, and whether they are produced purely by the intergroup processes or the types of messages used here. To understand whether the present results are unique to outgroup threats or threats in general, future studies should also include a condition in which an ingroup member threatens to attack other ingroup members. Finally, it is not only important to understand how participants respond to different outgroups for the sake of generalizability, but to investigate characteristic of the studied sample as well. For example, future experiments should test whether participants with different socio-political views exhibit similar neural responses to intergroup threats and reconciliation efforts or not.

Future experiments should further investigate whether the neural processes observed in response to the threatening messages are indeed unique to threatening stimuli or perhaps whether similar patterns of activations may be present in response to negative stimuli in general. In order to clarify this, the threatening messages should be compared to other types of negative messages. Similarly, the neural responses elicited by the reconciliation messages should be investigated in relation to comparable

positive messages in order to uncover whether or not the observed effects are unique to messages involving reconciliation efforts. The neutral statements used in the present experiment aimed to represent both the ingroup and the outgroup. Omitting the groups from these statements could lead to vastly different neural responses, such as a lack of activity in areas related to social processing, resulting in an imprecise control condition. These statements highlight some similarities between the ingroup and the outgroup, and may have unintentionally led to the blurring of intergroup boundaries, related to a reduction of implicit prejudice and stereotype threat and increased empathy (Beckes et al., 2012; Hall et al., 2009; Rosenthal & Crisp, 2006). A replication of the present experiment using neutral statements phrased without putting an emphasis on the similarities between the groups could clarify whether such a confounding effect influenced the extensive overlap between the threat > reconciliation and threat > neutral contrasts. Despite this possibility, a manipulation check suggests that participants rated the reconciliation statements in the present experiment significantly more positively than the neutral statements. It is thus reasonable to assume that there was a difference in the intended direction between the perceived valence of the neutral and the reconciliation statements. That said, studies using similar stimuli in the future are encouraged to conduct a norming study before data collection to determine “how positive” or “how negative” the stimuli are perceived by participants.

Conclusions

The aim of the present experiment was to uncover the neural responses associated with intergroup threat and reconciliation. Our results give new insights into the complex psychological and neural mechanisms underlying intergroup processes. They extend the growing body of literature exploring the physiological basis of ingroup bias and prejudice (Amodio, 2014; Cikara & Van Bavel, 2014; Han, 2018; Kubota et al., 2012; Molenberghs, 2013). These results suggest that the neural responses to threat involve more implicit processes, whereas reconciliation efforts involve more cognitive effort. These findings contribute to our understanding of the mechanisms behind the inflated attention given to terrorism threats in the media and the obstacles undermining intergroup reconciliation. A better understanding of the neural mechanisms behind intergroup reconciliation may equip researchers and policy makers with new insights to improve their

strategies to reduce intergroup conflict and improve intergroup reconciliation.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Australian Bicentennial Scholarship awarded to Dorottya Lantos by the Menzies Centre for Australian Studies; European Association of Social Psychology Travel Grant awarded to Dorottya Lantos; Heart Foundation Future Leader Fellowship [1000458] awarded by the Australian Heart Foundation to Pascal Molenberghs; ARC Discovery Grant [DP130100559] awarded by the Australian Research Council to Pascal Molenberghs.

ORCID

Dorottya Lantos  <http://orcid.org/0000-0001-8306-7101>
 Winnifred Louis  <http://orcid.org/0000-0003-2996-982X>
 Pascal Molenberghs  <http://orcid.org/0000-0003-0350-6659>

References

- Adolphs, R. (2008). Fear, faces, and the human amygdala. *Current Opinion in Neurobiology*, 18(2), 166–172. <https://doi.org/10.1016/j.conb.2008.06.006>
- Ahmed, S., & Matthes, J. (2017). Media representation of Muslims and Islam from 2000 to 2015: A meta-analysis. *International Communication Gazette*, 79(3), 219–244. <https://doi.org/10.1177/1748048516656305>
- Amodio, D. M. (2014). The neuroscience of prejudice and stereotyping. *Nature Reviews. Neuroscience*, 15(10), 670–682. <https://doi.org/10.1038/nrn3800>
- Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews. Neuroscience*, 7(4), 268–277. <https://doi.org/10.1038/nrn1884>
- Amodio, D. M., Harmon-Jones, E., & Devine, P. G. (2003). Individual differences in the activation and control of affective race bias as assessed by startle eyeblink responses and self-report. *Journal of Personality and Social Psychology*, 84(4), 738–753. <https://doi.org/10.1037/0022-3514.84.4.738>
- Anderson, A. K., & Phelps, E. A. (2001). Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature*, 411(6835), 305–309. <https://doi.org/10.1038/35077083>
- Arendt, F., & Northup, T. (2015). Effects of long-term exposure to news stereotypes on implicit and explicit attitudes. *International Journal of Communication*, 9, 2370–2390.
- Banaji, M. R., Lemm, K. M., & Carpenter, S. J. (2001). The social unconscious. In A. Tesser & N. Schwarz (Eds.), *Blackwell Handbook of Social Psychology: Intraindividual Processes* (pp. 134–157). Blackwell Publishers Ltd.
- Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental

- human motivation. *Psychological Bulletin*, 117(3), 497–529. <https://doi.org/10.1037/0033-2909.117.3.497>
- Beckes, L., Coan, J. A., & Hasselmo, K. (2012). Familiarity promotes the blurring of self and other in the neural representation of threat. *Social Cognitive and Affective Neuroscience*, 8(6), 670–677. <https://doi.org/10.1093/scan/nss046>
- Bhanji, J. P., & Delgado, M. R. (2014). The social brain and reward: Social information processing in the human striatum. *Wiley Interdisciplinary Reviews. Cognitive Science*, 5(1), 61–73. <https://doi.org/10.1002/wcs.1266>
- Billingsley, J., & Losin, E. A. (2017). The neural systems of forgiveness: An evolutionary psychological perspective. *Frontiers in Psychology*, 8, 737. <https://doi.org/10.3389/fpsyg.2017.00737>
- Blair, R. J. R. (2003). Facial expressions, their communicatory functions and neuro-cognitive substrates. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 358(1431), 561–572. <https://doi.org/10.1098/rstb.2002.1220>
- Blair, R. J. R. (2008). The amygdala and ventromedial prefrontal cortex: Functional contributions and dysfunction in psychopathy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1503), 2557–2565. <https://doi.org/10.1098/rstb.2008.0027>
- Blatz, C. W., & Philpot, C. (2010). On the outcomes of intergroup apologies: A review. *Social and Personality Psychology Compass*, 4(11), 995–1007. <https://doi.org/10.1111/j.1751-9004.2010.00318.x>
- Boyer, P., Firat, R., & van Leeuwen, F. (2015). Safety, threat, and stress in intergroup relations: A coalitional index model. *Perspectives on Psychological Science*, 10(4), 434–450. <https://doi.org/10.1177/1745691615583133>
- Brady, W. J., Gantman, A. P., & Van Bavel, J. J. (2020). Attentional capture helps explain why moral and emotional content go viral. *Journal of Experimental Psychology. General*, 149(4), 746–756. <https://doi.org/10.1037/xge0000673>
- Brüne, M., Juckel, G., & Enzi, B. (2013). “An eye for an eye”? Neural correlates of retribution and forgiveness. *PloS One*, 8(8), e73519. <https://doi.org/10.1371/journal.pone.0073519>
- Burnette, J. L., McCullough, M. E., Van Tongeren, D. R., & Davis, D. E. (2012). Forgiveness results from integrating information about relationship value and exploitation risk. *Personality & Social Psychology Bulletin*, 38(3), 345–356. <https://doi.org/10.1177/0146167211424582>
- Carlson, J. M., Reinke, K. S., & Habib, R. (2009). A left amygdala mediated network for rapid orienting to masked fearful faces. *Neuropsychologia*, 47(5), 1386–1389. <https://doi.org/10.1016/j.neuropsychologia.2009.01.026>
- Čehajić-Clancy, S., Goldenberg, A., Gross, J. J., & Halperin, E. (2016). Social-psychological interventions for intergroup reconciliation: An emotion regulation perspective. *Psychological Inquiry*, 27(2), 73–88. <https://doi.org/10.1080/1047840X.2016.1153945>
- Chekroud, A. M., Everett, J. A., Bridge, H., & Hewstone, M. (2014). A review of neuroimaging studies of race-related prejudice: Does amygdala response reflect threat? *Frontiers in Human Neuroscience*, 8, 179. <https://doi.org/10.3389/fnhum.2014.00179>
- Cikara, M., & Van Bavel, J. J. (2014). The neuroscience of intergroup relations: An integrative review. *Perspectives on Psychological Science*, 9(3), 245–274. <https://doi.org/10.1177/1745691614527464>
- Contreras, J. M., Banaji, M. R., & Mitchell, J. P. (2012). Dissociable neural correlates of stereotypes and other forms of semantic knowledge. *Social Cognitive and Affective Neuroscience*, 7((7)), 764–770. <https://doi.org/10.1093/scan/nsr053>
- Cunningham, W. A., Johnson, M. K., Raye, C. L., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2004). Separable neural components in the processing of black and white faces. *Psychological Science*, 15(12), 806–813. <https://doi.org/10.1111/j.0956-7976.2004.00760.x>
- Das, E., Bushman, B. J., Bezemer, M. D., Kerkhof, P., & Vermeulen, I. E. (2009). How terrorism news reports increase prejudice against outgroups: A terror management account. *Journal of Experimental Social Psychology*, 45(3), 453–459. <https://doi.org/10.1016/j.jesp.2008.12.001>
- Davey, J., Thompson, H. E., Hallam, G., Karapanagiotidis, T., Murphy, C., De Caso, I., Krieger-Redwood, K., Bernhardt, B. C., Smallwood, J., & Jefferies, E. (2016). Exploring the role of the posterior middle temporal gyrus in semantic cognition: Integration of anterior temporal lobe with executive processes. *Neuroimage*, 137, 165–177. <https://doi.org/10.1016/j.neuroimage.2016.05.051>
- Decety, J., Jackson, P. L., Sommerville, J. A., Chaminade, T., & Meltzoff, A. N. (2004). The neural bases of cooperation and competition: An fMRI investigation. *Neuroimage*, 23(2), 744–751. <https://doi.org/10.1016/j.neuroimage.2004.05.025>
- DeWall, C. N., Deckman, T., Pond, J. R., & Bonser, I. (2011). Belongingness as a core personality trait: How social exclusion influences social functioning and personality expression. *Journal of Personality*, 79(6), 1281–1314. <https://doi.org/10.1111/j.1467-6494.2010.00695.x>
- Dixon, M. L., Thiruchselvam, R., Todd, R., & Christoff, K. (2017). Emotion and the prefrontal cortex: An integrative review. *Psychological Bulletin*, 143(10), 1033–1081. <https://doi.org/10.1037/bul0000096>
- Dunbar, R. (2011). Evolutionary basis of the social brain. In J. Decety & J. T. Cacioppo (Eds.), *The Oxford Handbook of Social Neuroscience*. Oxford University Press.
- Egner, T., Etkin, A., Gale, S., & Hirsch, J. (2007). Dissociable neural systems resolve conflict from emotional versus non-emotional distracters. *Cerebral Cortex*, 18(6), 1475–1484. <https://doi.org/10.1093/cercor/bhm179>
- Erb, H. P., Kruglanski, A., Chun, W. Y., Pierro, A., Mannetti, L., & Spiegel, S. (2003). Searching for commonalities in human judgement: The parametric unimodel and its dual mode alternatives. *European Review of Social Psychology*, 14(1), 1–47. <https://doi.org/10.1080/10463280340000009>
- Etkin, A., Egner, T., & Kalisch, R. (2011). Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends in Cognitive Sciences*, 15(2), 85–93. <https://doi.org/10.1016/j.tics.2010.11.004>
- Etkin, A., Egner, T., Peraza, D. M., Kandel, E. R., & Hirsch, J. (2006). Resolving emotional conflict: A role for the rostral anterior cingulate cortex in modulating activity in the amygdala. *Neuron*, 51(6), 871–882. <https://doi.org/10.1016/j.neuron.2006.07.029>
- Evans, J. S. B. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, 59(1), 255–278. <https://doi.org/10.1146/annurev.psych.59.103006.093629>
- Fareri, D. S., Chang, L. J., & Delgado, M. R. (2012). Effects of direct social experience on trust decisions and neural reward

- circuitry. *Frontiers in Neuroscience*, 6, 148. <https://doi.org/10.3389/fnins.2012.00148>
- Fusar-Poli, P., Placentino, A., Carletti, F., Landi, P., Allen, P., Surguladze, S., ... Perez, J. (2009). Functional atlas of emotional faces processing: A voxel-based meta-analysis of 105 functional magnetic resonance imaging studies. *Journal of Psychiatry & Neuroscience*, 34(6), 418–432.
- Giustino, T. F., & Maren, S. (2015). The role of the medial prefrontal cortex in the conditioning and extinction of fear. *Frontiers in Behavioral Neuroscience*, 9, 298. <https://doi.org/10.3389/fnbeh.2015.00298>
- Hall, N. R., Crisp, R. J., & Suen, M. W. (2009). Reducing implicit prejudice by blurring intergroup boundaries. *Basic and Applied Social Psychology*, 31(3), 244–254. <https://doi.org/10.1080/01973530903058474>
- Han, S. (2018). Neurocognitive basis of racial ingroup bias in empathy. *Trends in Cognitive Sciences*, 22(5), 400–421.
- Haruno, M., & Kawato, M. (2006). Different neural correlates of reward expectation and reward expectation error in the putamen and caudate nucleus during stimulus-action-reward association learning. *Journal of Neurophysiology*, 95(2), 948–959. <https://doi.org/10.1152/jn.00382.2005>
- Hein, G., Silani, G., Preuschoff, K., Batson, C. D., & Singer, T. (2010). Neural responses to ingroup and outgroup members' suffering predict individual differences in costly helping. *Neuron*, 68(1), 149–160. <https://doi.org/10.1016/j.neuron.2010.09.003>
- Izuma, K., Aoki, R., Shibata, K., & Nakahara, K. (2019). Neural signals in amygdala predict implicit prejudice toward an ethnic outgroup. *NeuroImage*, 189, 341–352. <https://doi.org/10.1016/j.neuroimage.2019.01.019>
- Izuma, K., Saito, D. N., & Sadato, N. (2008). Processing of social and monetary rewards in the human striatum. *Neuron*, 58(2), 284–294. <https://doi.org/10.1016/j.neuron.2008.03.020>
- Jetten, J., Haslam, S. A., & Haslam, C. (2012). The case for a social identity analysis of health and well-being. In C. Haslam, J. Jetten, & A. S. Haslam (Eds.), *The social cure: Identity, health and well-being*. Psychology Press.
- Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, 58(9), 697–720. <https://doi.org/10.1037/0003-066X.58.9.697>
- Kahneman, D. (2011). *Thinking fast and slow*. Straus and Giroux.
- Kearns, E. M., Betus, A. E., & Lemieux, A. F. (2019). Why do some terrorist attacks receive more media attention than others? *Justice Quarterly*, 1–24.
- Kelman, H. C. (2010). Conflict resolution and reconciliation: A social-psychological perspective on ending violent conflict between identity groups. *Landscapes of Violence*, 1(1), 5.
- Keren, G. (2013). A tale of two systems: A scientific advance or a theoretical stone soup? Commentary on Evans & Stanovich (2013). *Perspectives on Psychological Science*, 8(3), 257–262. <https://doi.org/10.1177/1745691613483474>
- Klumpp, H., Angstadt, M., & Phan, K. L. (2012). Insula reactivity and connectivity to anterior cingulate cortex when processing threat in generalized social anxiety disorder. *Biological Psychology*, 89(1), 273–276. <https://doi.org/10.1016/j.biopsycho.2011.10.010>
- Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *Journal of Neuroscience*, 21(16), RC159–RC159. <https://doi.org/10.1523/JNEUROSCI.21-16-j0002.2001>
- Kruglanski, A. W., & Gigerenzer, G. (2011). Intuitive and deliberate judgments are based on common principles. *Psychological Review*, 118(1), 97–109. <https://doi.org/10.1037/a0020762>
- Kubota, J. T., Banaji, M. R., & Phelps, E. A. (2012). The neuroscience of race. *Nature Neuroscience*, 15(7), 940–948. <https://doi.org/10.1038/nn.3136>
- Mattan, B. D., Wei, K. Y., Cloutier, J., & Kubota, J. T. (2018). The social neuroscience of race-based and status-based prejudice. *Current Opinion in Psychology*, 24, 27–34. <https://doi.org/10.1016/j.copsyc.2018.04.010>
- McCullough, M. E. (2008). *Beyond Revenge: The Evolution of the Forgiveness Instinct*. Jossey-Bass.
- McCullough, M. E., Kurzban, R., & Tabak, B. A. (2013). Cognitive systems for revenge and forgiveness. *Behavioral and Brain Sciences*, 36(1), 1–15. <https://doi.org/10.1017/S0140525X11002160>
- Mechias, M. L., Etkin, A., & Kalisch, R. (2010). A meta-analysis of instructed fear studies: Implications for conscious appraisal of threat. *NeuroImage*, 49(2), 1760–1768. <https://doi.org/10.1016/j.neuroimage.2009.09.040>
- Molenberghs, P. (2013). The neuroscience of in-group bias. *Neuroscience and Biobehavioral Reviews*, 37(8), 1530–1536. <https://doi.org/10.1016/j.neubiorev.2013.06.002>
- Molenberghs, P., Gapp, J., Wang, B., Louis, W. R., & Decety, J. (2016). Increased moral sensitivity for outgroup perpetrators harming ingroup members. *Cerebral Cortex*, 26(1), 225–233. <https://doi.org/10.1093/cercor/bhu195>
- Molenberghs, P., & Louis, W. R. (2018). Insights from fMRI studies into ingroup bias. *Frontiers in Psychology*, 9, 1868. <https://doi.org/10.3389/fpsyg.2018.01868>
- Molenberghs, P., Prochilo, G., Steffens, N. K., Zacher, H., & Haslam, S. A. (2017). The neuroscience of inspirational leadership: The importance of collective-oriented language and shared group membership. *Journal of Management*, 43(7), 2168–2194. <https://doi.org/10.1177/0149206314565242>
- Mugg, J. (2015). Two minded creatures and dual-process Theory. *Journal of Cognition and Neuroethics*, 3(3), 87–112.
- Nadler, A., & Liviatan, I. (2006). Intergroup reconciliation: Effects of adversary's expressions of empathy, responsibility, and recipients' trust. *Personality & Social Psychology Bulletin*, 32(4), 459–470. <https://doi.org/10.1177/0146167205276431>
- Neuberg, S. L., & Schaller, M. (2016). An evolutionary threat-management approach to prejudices. *Current Opinion in Psychology*, 7, 1–5. <https://doi.org/10.1016/j.copsyc.2015.06.004>
- Noor, M., Brown, R., Gonzalez, R., Manzi, J., & Lewis, C. A. (2008). On positive psychological outcomes: What helps groups with a history of conflict to forgive and reconcile with each other? *Personality & Social Psychology Bulletin*, 34(6), 819–832. <https://doi.org/10.1177/0146167208315555>
- O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience*, 4(1), 95–102. <https://doi.org/10.1038/82959>
- Öhman, A. (2005). The role of the amygdala in human fear: Automatic detection of threat. *Psychoneuroendocrinology*, 30(10), 953–958. <https://doi.org/10.1016/j.psyneuen.2005.03.019>
- Ohtsubo, Y., Matsunaga, M., Tanaka, H., Suzuki, K., Kobayashi, F., Shibata, E., Hori, R., Umemura, T., & Ohira, H. (2018). Costly apologies communicate conciliatory intention: An fMRI

- study on forgiveness in response to costly apologies. *Evolution and Human Behavior*, 39(2), 249–256. <https://doi.org/10.1016/j.evolhumbehav.2018.01.004>
- Olson, I. R., McCoy, D., Klobusicky, E., & Ross, L. A. (2013). Social cognition and the anterior temporal lobes: A review and theoretical framework. *Social Cognitive and Affective Neuroscience*, 8(2), 123–133. <https://doi.org/10.1093/scan/nss119>
- Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience*, 12(5), 729–738. <https://doi.org/10.1162/089892900562552>
- Powell, K. A. (2011). Framing Islam: An analysis of US media coverage of terrorism since 9/11. *Communication Studies*, 62(1), 90–112. <https://doi.org/10.1080/10510974.2011.533599>
- Preibisch, C., Lanfermann, H., Wallenhorst, T., Walter, H., & Erk, S. (2009). Neuroanatomical correlates of visual field bias: A sensitive system for detecting potential threats? *Brain Research*, 1263, 69–77. <https://doi.org/10.1016/j.brainres.2009.01.051>
- Psychology Software Tools, Inc. [E-Prime 2.0]. (2013). Retrieved from <https://www.pstnet.com>
- Ricciardi, E., Rota, G., Sani, L., Gentili, C., Gaglianese, A., Guazzelli, M., & Pietrini, P. (2013). How the brain heals emotional wounds: The functional neuroanatomy of forgiveness. *Frontiers in Human Neuroscience*, 7, 839. <https://doi.org/10.3389/fnhum.2013.00839>
- Ritchie, H. (2019). *Does the news reflect what we die from?* Retrieved from <https://ourworldindata.org/does-the-news-reflect-what-we-die-from>
- Rolls, E. T. (2000). The orbitofrontal cortex and reward. *Cerebral Cortex*, 10(3), 284–294. <https://doi.org/10.1093/cercor/10.3.284>
- Rosenthal, H. E., & Crisp, R. J. (2006). Reducing stereotype threat by blurring intergroup boundaries. *Personality & Social Psychology Bulletin*, 32(4), 501–511. <https://doi.org/10.1177/0146167205281009>
- Rushworth, M. F., Kolling, N., Sallet, J., & Mars, R. B. (2012). Valuation and decision-making in frontal cortex: One or many serial or parallel systems? *Current Opinions in Neurobiology*, 22(6), 946–955. <https://doi.org/10.1016/j.conb.2012.04.011>
- Rushworth, M. F., Noonan, M. P., Boorman, E. D., Walton, M. E., & Behrens, T. E. (2011). Frontal cortex and reward-guided learning and decision-making. *Neuron*, 70(6), 1054–1069. <https://doi.org/10.1016/j.neuron.2011.05.014>
- Saleem, M., Prot, S., Anderson, C. A., & Lemieux, A. F. (2017). Exposure to Muslims in media and support for public policies harming Muslims. *Communication Research*, 44(6), 841–869. <https://doi.org/10.1177/0093650215619214>
- Skipper, L. M., Ross, L. A., & Olson, I. R. (2011). Sensory and semantic category subdivisions within the anterior temporal lobes. *Neuropsychologia*, 49(12), 3419–3429. <https://doi.org/10.1016/j.neuropsychologia.2011.07.033>
- Stein, M. B., Simmons, A. N., Feinstein, J. S., & Paulus, M. P. (2007). Increased amygdala and insula activation during emotion processing in anxiety-prone subjects. *American Journal of Psychiatry*, 164(2), 318–327. <https://doi.org/10.1176/ajp.2007.164.2.318>
- Strack, F., & Deutsch, R. (2015). The duality of everyday life: Dual-process and dual system models in social psychology. In M. Mikulincer, P. R. Shaver, E. Borgida, & J. A. Bargh (Eds.), *APA handbook of personality and social psychology, Volume 1: Attitudes and social cognition* (pp. 891–927). American Psychological Association.
- Strang, S., Utikal, V., Fischbacher, U., Weber, B., & Falk, A. (2014). Neural correlates of receiving an apology and active forgiveness: An fMRI study. *PloS One*, 9(2), e87654. <https://doi.org/10.1371/journal.pone.0087654>
- Tabibnia, G., Satpute, A. B., & Lieberman, M. D. (2008). The sunny side of fairness: Preference for fairness activates reward circuitry (and disregarding unfairness activates self-control circuitry). *Psychological Science*, 19(4), 339–347. <https://doi.org/10.1111/j.1467-9280.2008.02091.x>
- Tajfel, H. (1969). Cognitive aspects of prejudice. *Journal of Biosocial Science*, 1(S1), 173–191. <https://doi.org/10.1017/S0021932000023336>
- Trawalter, S., Todd, A. R., Baird, A. A., & Richeson, J. A. (2008). Attending to threat: Race-based patterns of selective attention. *Journal of Experimental Social Psychology*, 44(5), 1322–1327. <https://doi.org/10.1016/j.jesp.2008.03.006>
- Tropp, L. R., Hawi, D. R., O'Brien, T. C., Gheorghiu, M., Zetes, A., & Butz, D. A. (2017). Intergroup contact and the potential for post-conflict reconciliation: Studies in Northern Ireland and South Africa. *Peace and Conflict: Journal of Peace Psychology*, 23(3), 239–249. <https://doi.org/10.1037/pac0000236>
- Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2008). The neural substrates of in-group bias: A functional magnetic resonance imaging investigation. *Psychological Science*, 19(11), 1131–1139. <https://doi.org/10.1111/j.1467-9280.2008.02214.x>
- Van Overwalle, F. (2009). Social cognition and the brain: A meta-analysis. *Human Brain Mapping*, 30(3), 829–858. <https://doi.org/10.1002/hbm.20547>
- Van Tongeren, D. R., Burnette, J. L., O'Boyle, E., Worthington, J. E. L., & Forsyth, D. R. (2014). A meta-analysis of intergroup forgiveness. *The Journal of Positive Psychology*, 9(1), 81–95. <https://doi.org/10.1080/17439760.2013.844268>
- Veit, R., Singh, V., Sitaram, R., Caria, A., Rauss, K., & Birbaumer, N. (2011). Using real-time fMRI to learn voluntary regulation of the anterior insula in the presence of threat-related stimuli. *Social Cognitive and Affective Neuroscience*, 7(6), 623–634. <https://doi.org/10.1093/scan/nsr061>
- Visser, M., Jefferies, E., Embleton, K. V., & Ralph, M. A. (2012). Both the Middle Temporal Gyrus and the Ventral Anterior Temporal Area Are Crucial for Multimodal Semantic Processing: Distortion-corrected fMRI Evidence for a Double Gradient of Information Convergence in the Temporal Lobes. *Journal of Cognitive Neuroscience*, 24(8), 1766–1778. https://doi.org/10.1162/jocn_a_00244
- West, K., & Lloyd, J. (2017). The role of labelling bias in the portrayals of acts of 'terrorism': Representations of Muslims versus non-Muslims. *Journal of Muslim Minority Affairs*, 37(2), 211–222. <https://doi.org/10.1080/13602004.2017.1345103>
- Wheeler, M. E., & Fiske, S. T. (2005). Controlling racial prejudice: Social-cognitive goals affect amygdala and stereotype activation. *Psychological Science*, 16(1), 56–63. <https://doi.org/10.1111/j.0956-7976.2005.00780.x>
- Will, G.-J., Crone, E. A., & Guroglu, B. (2014). Acting on social exclusion: Neural correlates of punishment and forgiveness

- of excluders. *Social Cognitive and Affective Neuroscience*, 10 (2), 209–218. <https://doi.org/10.1093/scan/nsu045>
- Wohl, M. J. A., Hornsey, M. J., & Bennett, S. H. (2012). Why group apologies succeed and fail: Intergroup forgiveness and the role of primary and secondary emotions. *Journal of Personality and Social Psychology*, 102(2), 306–322. <https://doi.org/10.1037/a0024838>
- Xu, X., Zuo, X., Wang, X., & Han, S. (2009). Do you feel my pain? Racial group membership modulates empathic neural responses. *Journal of Neuroscience*, 29(26), 8525–8529. <https://doi.org/10.1523/JNEUROSCI.2418-09.2009>
- Yamada, M., Camerer, C. F., Fujie, S., Kato, M., Matsuda, T., Takano, H., Ito, H., Suhara, T., & Takahashi, H. (2012). Neural circuits in the brain that are activated when mitigating criminal sentences. *Nature Communications*, 3(1), 759. <https://doi.org/10.1038/ncomms1757>
- Yarkoni, T. (2009). Big correlations in little studies: Inflated fMRI correlations reflect low statistical power—Commentary on Vul et al. (2009). *Perspectives on Psychological Science*, 4(3), 294–298. <https://doi.org/10.1111/j.1745-6924.2009.01127.x>
- Young, L., & Saxe, R. (2009). Innocent intentions: A correlation between forgiveness for accidental harm and neural activity. *Neuropsychologia*, 47(10), 2065–2072. <https://doi.org/10.1016/j.neuropsychologia.2009.03.020>
- Zahn, R., Moll, J., Krueger, F., Huey, E. D., Garrido, G., & Grafman, J. (2007). Social concepts are represented in the superior anterior temporal cortex. *Proceedings of the National Academy of Sciences*, 104(15), 6430–6435.