

# SCIENTIFIC REPORTS



OPEN

## Neuroanatomical correlates of forgiving unintentional harms

Indrajeet Patil<sup>1,2</sup>, Marta Calò<sup>3,\*</sup>, Federico Fornasier<sup>3,\*</sup>, Liane Young<sup>4</sup> & Giorgia Silani<sup>5</sup>

Received: 07 November 2016

Accepted: 10 March 2017

Published: 06 April 2017

Mature moral judgments rely on the consideration of a perpetrator's mental state as well as harmfulness of the outcomes produced. Prior work has focused primarily on the functional correlates of how intent information is neurally represented for moral judgments, but few studies have investigated whether individual differences in neuroanatomy can also explain variation in moral judgments. In the current study, we conducted voxel-based morphometry analyses to address this question. We found that local grey matter volume in the left anterior superior temporal sulcus, a region in the functionally defined theory of mind or mentalizing network, was associated with the degree to which participants relied on information about innocent intentions to forgive accidental harms. Our findings provide further support for the key role of mentalizing in the forgiveness of accidental harms and contribute preliminary evidence for the neuroanatomical basis of individual differences in moral judgments.

When humans make moral judgments, one critical input is information about intent. Did she spill the hot coffee on her colleague on purpose? Did he step on his competitor's foot by accident? Similarly, the common law tradition relies on presence of culpable mental states (*mens rea*) for criminal conviction. Much recent work in moral psychology and cognitive neuroscience has explored the psychological and neurofunctional basis of intent processing for moral judgment. In the current study, we extend this literature by exploring the neuroanatomical correlates of this process. Behavioral work shows that when intent and outcome information conflict, people primarily rely on information about intent, an effect observed in older children and adults across cultures<sup>1–3</sup>. Representations of others' mental states are constructed by *Theory of Mind* (ToM) or mentalizing, the process of inferring representational content (e.g., beliefs, desires, knowledge, intentions) from observing others' in order to explain and predict their behavior<sup>4,5</sup>. This capacity is neurally implemented in a specific network consisting primarily of the bilateral temporoparietal junction (TPJ), ventromedial prefrontal cortex (vmPFC), dorsomedial prefrontal cortex (dmPFC), temporal poles (TP), superior temporal sulcus (STS), and precuneus/posterior cingulate cortex (PC/PCC)<sup>4</sup>.

There is ample amount of evidence that shows an overlap between the moral reasoning network and ToM network, as highlighted by a recent meta-analysis<sup>6</sup>. Latest work has also begun attributing specific functions that various regions in ToM network may play during moral evaluations. For example, the encoding of mental states occurs very early during information processing in the right TPJ<sup>7</sup>, left TPJ<sup>8</sup>, and amygdala<sup>9</sup>. The rTPJ, dmPFC, and PCC are also involved in integrating belief states<sup>10,11</sup> with other morally relevant pieces of information (e.g., consequences) to construct final moral judgments<sup>12–15</sup>. On the other hand, the dmPFC is involved in encoding the valence (harmful or neutral) of the beliefs<sup>10</sup>. Individual differences in both the overall magnitude of activity and the spatial pattern of activity in rTPJ have been repeatedly associated with the degree to which information about innocent intentions is used to forgive agents who cause accidental harms<sup>16–18</sup>. Additionally, disrupting activity in rTPJ via transcranial magnetic stimulation (TMS) leads to a more lenient assessment of attempted harms<sup>19</sup>, while enhancing this activity via transcranial direct-current stimulation (tDCS) leads to reduced blame for accidents<sup>20</sup>.

In summary, this work shows that neural activity in the ToM network, in general, is involved in encoding and integrating the information about mental states of actors involved in moral situations. This network thereby also underpins cognitive basis of how people forgive<sup>21,22</sup> accidents based on innocent intentions and condemn attempted harms based on malicious intentions.

<sup>1</sup>Scuola Internazionale Superiore di Studi Avanzati, Neuroscience Sector, Trieste, Italy. <sup>2</sup>Department of Psychology, Harvard University, Cambridge, MA, USA. <sup>3</sup>University of Trieste, Trieste, Italy. <sup>4</sup>Department of Psychology, Boston College, Boston, USA. <sup>5</sup>Department of Applied Psychology: Health, Development, Enhancement and Intervention, University of Vienna, Austria. \*These authors contributed equally to this work. Correspondence and requests for materials should be addressed to I.P. (email: patilindrajeet.science@gmail.com) or G.S. (email: giorgia.silani@univie.ac.at)

Although extant work delineates the neurofunctional correlates of intent-based moral judgements, the neurostructural basis of this process remains sparsely studied. Previous morphometry studies relating to moral cognition have examined how regional variation in brain structure relates to individual differences in endorsed moral values<sup>23</sup>, moral development<sup>24</sup>, group-focused moral foundations<sup>25</sup>, injustice sensitivity<sup>26</sup>, indirect reciprocity<sup>27</sup>, prosocial behavior<sup>28–31</sup>, and moral judgments in clinical populations<sup>32,33</sup>. To our knowledge, only one prior voxel-based morphometry (VBM) study has investigated this issue in a small sample ( $n = 19$ ), recruited as controls for comparison with a neurological population<sup>32</sup>. This study found correlation patterns for accidental harm in regions typically associated with ToM, namely PC, vmPFC, and dorsolateral prefrontal cortex (dlPFC), although no claim about their functional properties was made in this work.

Given the consistent prior evidence from fMRI and VBM studies implicating the ToM network in forgiving accidental harms<sup>16–18</sup>, we focused our attention on this context: we predicted that volumetric differences in regions belonging to the ToM network will explain variation in moral judgments of accidental harms.

## Methods and Materials

**Participants.** A total of 50 healthy community members (32 female; 41 right-handed) were recruited to participate and financially compensated for their time and travel. Average age was 23.06 years ( $SD = 3.08$ ), with a range of 18 to 35. All participants gave written informed consent. This study was conducted according to the principles in the Declaration of Helsinki, approved by the Ethics Committee of the Hospital ‘Santa Maria della Misericordia’ (Udine, Italy), and was carried out in accordance with the approved guidelines. All participants provided written informed consent before any study procedure was initiated.

Rule-out criteria for participation included Italian as a secondary language, presence of a diagnosed psychiatric illness and/or history of psychiatric treatment, history of significant neurological illness or brain injury, and current usage of psychoactive drugs. All participants were screened for neurological condition and MRI contraindications in two-step checks, during a pre-scanning telephone interview and before entering the scanner. All data (structural, functional, and behavioral) from one participant were excluded from the final analysis as this participant was consuming clinically prescribed psychoactive drugs and divulged this information only during the post-scan debriefing. Thus, structural and behavioral data were available for 49 participants. Functional data for the moral judgment task from two participants were removed due to excessive head motion, four participants due to high collinearity in regressors, and data from one additional participant could not be collected due to technical error (final  $n = 42$ ).

**Experimental stimuli and procedure.** Participants performed a moral judgment task in the scanner. Additionally, they completed a ToM localizer task. The order of tasks (described in detail below) was counterbalanced across participants. Although the functional data from the moral judgment task are described in detail in a different manuscript, we offer a brief summary of data acquisition and analysis procedures, given our inclusion of exploratory functional data analysis based on our VBM analysis results (see below).

**Moral judgment task.** Experimental stimuli were text-based scenarios. Scenarios were largely adapted from previous studies<sup>1,19</sup> and translated into Italian (see Supplementary Text S1 for more details). There were four variations of 36 unique scenarios for a total of 144 stories. The four variations were the result of a  $2 \times 2$  within-subjects design where the factors *belief* (neutral, negative) and *outcome* (neutral, negative) were independently varied. Each participant saw one variation of each scenario, for a total of 36 stories.

Each scenario lasted for 32 s and consisted of four cumulative segments (each lasting for 8 s): (i) *background*: this segment was common to all variations and provided context for the action; (ii) *foreshadow*: this segment foreshadowed the outcome as neutral or harmful; (iii) *mental state information*: this segment provided information about the agent’s neutral or harmful belief; (iv) *consequence*: this final segment described the agent’s action and the outcome. We use the term *mental state information* instead of *belief* and *consequence* instead of *outcome* to avoid confusion as the latter terms represent factors of the experimental design, while the former represent story segments containing information about the agent’s beliefs and the nature of the outcome, respectively. We provide an example of one story below (called “Rabid dog”, presented here in the accidental harm condition):

**Background.** Chiara works at the pound. Several new dogs have just come in. A lady comes in, interested in taking one of the new dogs home with her.

**Foreshadow.** The dogs are very sick with rabies and will make their owners sick too by biting them.

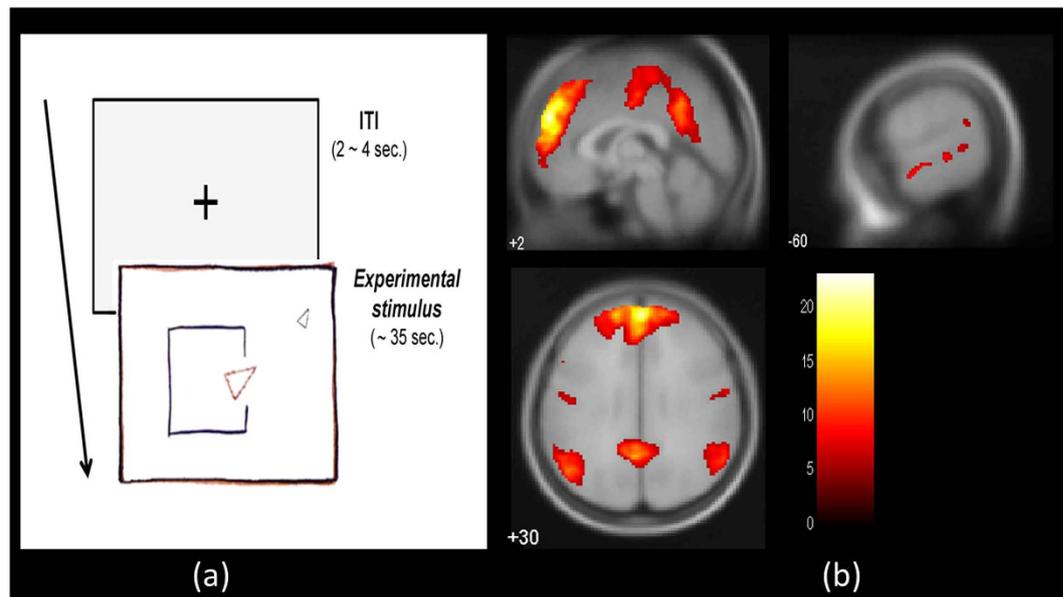
**Mental state information.** Chiara talks with one of the other people at the pound. Chiara thinks that the new dogs have been through a thorough health inspection and will make good pets.

**Consequence.** Chiara gives the lady one of the new dogs. It is infected with rabies and bites the lady on the neck on the very first day.

After reading each story, participants provided two types of moral judgments following prior work<sup>1</sup>, presented in randomized order across trials for each participant:

[1] *acceptability* - “How morally acceptable was [the agent]’s behavior?” (1: *Completely acceptable* to 7: *Not at all acceptable*);

[2] *blame* - “How much blame does [the agent] deserve?” (1: *None at all* to 7: *Very much*).



**Figure 1. Schematics and results for the Theory of Mind (ToM) localizer task.** (a) Participants watched animations involving two triangles interacting with each other in either a meaningful way (experimental) or in a random manner (control). (b) The ToM network localized at the group level using group independent component analysis (gICA) on the ToM functional localizer task. Regions including the bilateral temporoparietal junction (TPJ), sections of medial prefrontal cortex (mPFC), temporal poles (TP), superior temporal sulcus (STS), and precuneus (PC) formed a functionally connected network. Statistical maps represent  $t$ -values thresholded at a voxel-wise threshold of  $p < 0.05$  (FWE-corrected, height threshold:  $t = 5.62$ ) and extent threshold of 10 voxels. The color bar denotes  $t$ -values. A similar network was observed in GLM-based analysis of the data (see Supplementary Text S2). Abbreviation - ITI: intertrial interval.

Each question was presented on the screen for 6 s, and participants could provide judgments using a 7-point Likert scale, on which the cursor could be moved using two fingers. Since these two judgments were highly correlated (neutral:  $r = 0.875$ , accidental:  $r = 0.863$ , attempted:  $r = 0.889$ , intentional:  $r = 0.617$ ), they were averaged to form a single moral judgment per condition, indexing severity of moral condemnation.

**ToM localizer task.** In order to localize functional regions involved in mental state attribution, we used the animated shapes paradigm based on the Frith-Happé animations<sup>34</sup>, a reliable and sensitive method for eliciting spontaneous mentalizing<sup>35</sup>. The stimuli consisted of eight animations (four per condition) featuring two triangles (a big red triangle and a small blue triangle) moving on a white background. Stimuli were matched for overall shape, speed, and orientation changes as closely as possible. ToM animations featured complex scripts in which triangles interacted in a way that gave semblance of bluffing, mocking, persuading, and surprising. By contrast, control animations comprised of random, billiard ball-like motion of the two triangles. The length of videos was matched for the two types of animations (on average 34–35 s), and the order of condition appearance was randomized. While there was no active task, we ensured that participants closely attended to the videos by telling them there would be questions at the end of the task (in reality, there were no questions asked). Intertrial interval (ITI) was randomly jittered with an average of 2 s (jitter range: 0–2 s) (Fig. 1(a)).

**Functional scan acquisition.** All fMRI scans were acquired using a 3 T Philips Achieva scanner equipped with an 8-channel head coil. High-resolution structural images were acquired as 180 T1-weighted transverse images (0.75 mm slice thickness). Functional images were acquired in interleaved manner using a T2\*-weighted echoplanar imaging (EPI) sequence with 33 transverse slices covering the whole brain with the following parameters: slice thickness = 3.2 mm; interslice gap = 0.3 mm; repetition time (TR) = 2000 ms, echo time (TE) = 35 ms; flip angle = 90°, field of view = 230 × 230 mm<sup>2</sup>; matrix size = 128 × 128, SENSE factor 2. The slices were oriented at a 30° oblique angle to the AC-PC.

**fMRI data analysis.** Data were preprocessed with SPM12 (Wellcome Department of Imaging Neuroscience, London, UK) running on MATLAB R 2013 a. Each subject's data were motion-corrected and then normalized onto a common stereotactic space (the MNI template). Data were then smoothed by using a Gaussian filter (FWHM = 6 mm at first-level) and high-pass-filtered.

**ToM localizer task.** For technical reasons (see Supplementary Text S2), we preferred model-free group independent component analysis (gICA) over (GLM) to analyze the ToM localizer task data (although same results were obtained with GLM analysis, see below). ICA is a model-free analysis technique that allows one to characterize the spatio-temporal structure of the data. ICA identifies mutually independent sets of regions (component) exhibiting high within-component functional connectivity<sup>36</sup>.

We performed spatial gICA on preprocessed functional datasets ( $n = 49$ ) of equal length using the GIFT toolbox (v4.0, <http://mialab.mrn.org/software/gift/>; Calhoun, Liu, & Adal<sup>37</sup>) to localize the ToM network. The component corresponding to the ToM network was identified (for full details, see Supplementary Text S2) and the coordinates for the network were derived from a one-sample Student's  $t$ -test. Statistical significance was assessed using an FWE value of  $p < 0.05$  corrected for the whole-brain volume at the cluster-level ( $k = 10$ ). The regions thus localized were saved as thresholded binary maps to be used as an inclusive ToM mask for the VBM analysis. We note that the key nodes of the ToM network (dmPFC, vmPFC, PC/PCC, bilateral TPJ, bilateral STS, TP) were observed in both model-free ICA and model-based GLM analyses (See Supplementary Text S4).

**Structural scan acquisition.** High-resolution structural images were acquired as 190 T1-weighted transverse images with a 3D ultrafast gradient echo sequence on a 3T Philips Achieva scanner at the Hospital 'Santa Maria della Misericordia' (Udine, Italy) equipped with an 8-channel SENSE head coil. The following parameters were used: voxel size =  $1 \times 1 \times 1$  mm, slice thickness = 1 mm, TR/TE = 8.2/3.7 ms, matrix size =  $240 \times 240$  mm, field of view = 19 cm, flip angle =  $8^\circ$ , no overcontiguous slices.

**Voxel-based morphometry.** For reporting VBM analyses details, we have followed a set of prior guidelines<sup>38,39</sup>. Additional technical details are provided in Supplementary Text S3. Both preprocessing and statistical analysis of anatomical data were carried out using SPM12 according to procedures outlined in prior work<sup>40–43</sup>.

In order to account for the intensity inhomogeneities present in MR scans at high field strengths ( $\geq 3$  T), bias field correction was applied during the segmentation procedure<sup>44</sup>. Each image was segmented into six different tissues classes (grey matter (GM), white matter (WM), cerebrospinal fluid (CSF), bone, other soft tissues, and air/background) using the modified unified segmentation approach<sup>45,46</sup> implemented in SPM12. The Non-linear deformation field was estimated for each individual image such that tissue probability maps for each tissue class were best aligned. The segmented images were imported (only for GM and WM) both in native space and DARTEL space.

The segmented images (only GM and WM) were then iteratively registered via a fast diffeomorphic registration algorithm<sup>47</sup> (DARTEL) to warp the GM and WM partitions into a study-specific template space representative of the average of all study subjects.

This process created a template image for the group of individuals and also estimated the nonlinear deformation flow fields that best aligned individual images together. This template image was then transformed to MNI stereotactic space ( $2 \times 2 \times 2$  mm) using affine and non-linear spatial transformations to generate normalized, Jacobian-scaled (grey matter amount preserved, i.e.) GM images for each participant. These images were also simultaneously smoothed with an isotropic Gaussian kernel with FWHM of 10 mm. Note that these final smoothed images represent *absolute amount* of local GM at each voxel in the brain<sup>48,49</sup>. These smoothed normalized GM segments were then entered into a statistical model to conduct voxel-wise statistical tests and map significant effects.

The statistical analysis was carried out by fitting a GLM to the data. We included age, age<sup>2</sup> (to model quadratic effects of age), handedness, and gender as nuisance covariates<sup>50,51</sup>. Since total intracranial volume (TIV) was entered as a global for proportional scaling, it was not included in the design matrix as a regressor. Repeating the same analysis by entering TIV values not as globals but as covariates produced similar results. Although no overall grand mean scaling was applied, we used global normalization by entering TIV values as globals (as recommended by Ridgway *et al.*<sup>52</sup>) in proportional scaling, which identifies specific regional changes that are not confounded by global differences.

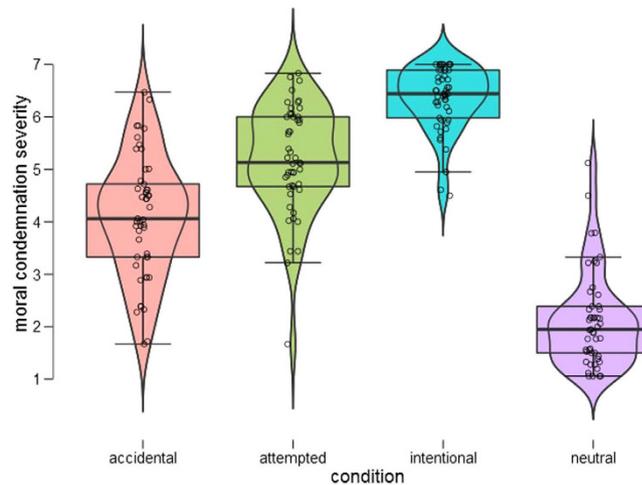
Four separate regression models were created for each condition containing moral judgment score, age, age<sup>2</sup>, handedness, and gender as predictors and GMV as the dependent measure in each model. Two contrasts were created for each model that regressed local GMV on the moral judgment scores, one tracking positive association, while the other negative:

- (a) Positive ( $[0, 1]$ ; greater GMV associated with *increased* moral condemnation) and
- (b) Negative ( $[0, -1]$ ; greater GMV associated with *reduced* moral condemnation).

Importantly, the second-level analyses were restricted only to voxels contained in the inclusive mask derived from the ToM localizer task (see Fig. 1). Note that selection of voxels included in the mask was independent from the data used in the VBM analysis circumventing any circularity<sup>53</sup>. Given recent criticism of parametric cluster-level inference<sup>54,55</sup>, significant clusters were formed by employing the threshold-free cluster enhancement (TFCE) method (as implemented in TFCE toolbox (r95): <http://dbm.neuro.uni-jena.de/tfce/>). The TFCE is a cluster-based thresholding method that circumvents the problem of choosing an arbitrary cluster forming threshold (e.g.,  $p < 0.001$  (uncorrected) and  $k = 10$ ) by taking a raw statistics image and producing an output image in which the voxel-wise values represent the amount of cluster-like local spatial support<sup>56</sup>. This also makes the TFCE inference fairly robust to non-stationarity in the data under varying smoothness levels, degrees of freedom and signal to noise ratios<sup>57,58</sup>. The TFCE image is then turned into voxel-wise  $p$ -values via a permutation-based non-parametric testing (5000 permutations were used). All group comparisons are reported at  $p < 0.05$  after Family-wise Error (FWE) correction and, as recommended<sup>59,60</sup>, we report effect sizes in addition to statistic values.

## Results

**Behavioral data.** A 2-by-2 repeated measures ANOVA on moral judgment data revealed the expected main effects of intent ( $F(1, 48) = 217.778$ ,  $p < 0.001$ ,  $\eta^2 = 0.819$ ), outcome ( $F(1, 48) = 122.012$ ,  $p < 0.001$ ,  $\eta^2 = 0.718$ ), and their interaction ( $F(1, 48) = 30.393$ ,  $p < 0.001$ ,  $\eta^2 = 0.388$ ). In other words, agents who acted with harmful



**Figure 2. Moral condemnation ratings across conditions.** Full distribution of moral condemnation ratings across conditions is shown using combination of box and violin plots<sup>96</sup>. Box plot within the violin plot contains thick black line for the median and the box indicates the interquartile range, while the added rotated kernel density plot shows the probability density of the data at different values. As can be seen, there was more variation in accidental and attempted harm cases, where intent and outcome was misaligned, as compared to neutral and intentional cases.

intent or who produced a harmful outcome were condemned more severely than those acting with innocent intention or who produced a neutral outcome, respectively (Fig. 2; for descriptive statistics, see Supplementary Text S5). Additionally, the intent and outcome information interacted such that the degree to which the presence or absence of harmful consequence affected severity of moral condemnation depended on whether the intent was neutral or negative (greater difference in severity of moral judgment in accidental versus neutral comparison than intentional versus attempted comparison).

**Functional localizer results.** The gICA on the ToM localizer task revealed a component consisting of the regions involved in mentalizing (see Fig. 1(b)), including bilateral TPJ, PC/PCC, dmPFC, TP, posterior STS, anterior STS, etc. (for more details on the component characteristic, see Supplementary Text S2) All VBM analyses were restricted to the binary mask comprising of voxels belonging to this network.

**Anatomical data.** *Neutral, attempted, and intentional harm conditions.* No supra-threshold voxels were found for positive (greater GMV associated with *increased* moral condemnation) or negative (greater GMV associated with *reduced* moral condemnation) contrasts.

*Accidental harm condition.* Regression analyses revealed that more severe moral condemnation for accidental harm was associated with reduced GMV in left ( $x = -62, y = -12, z = -12; \beta = -0.0252, \text{TFCE} = 13.26; k = 202; p(\text{FWE-corrected}) = 0.002$ ) anterior STS (aSTS) (see Fig. 3). Note that although we refer to this region here as aSTS<sup>61–63</sup>, other studies have also referred to the same region with the anatomical label “middle temporal gyrus (MTG)”<sup>6,35,64–67</sup>.

No supra-threshold voxels were found for positive contrast. Additionally, no suprathreshold voxels outside the localized ToM network were found for any contrast or for any condition in the whole-brain analysis.

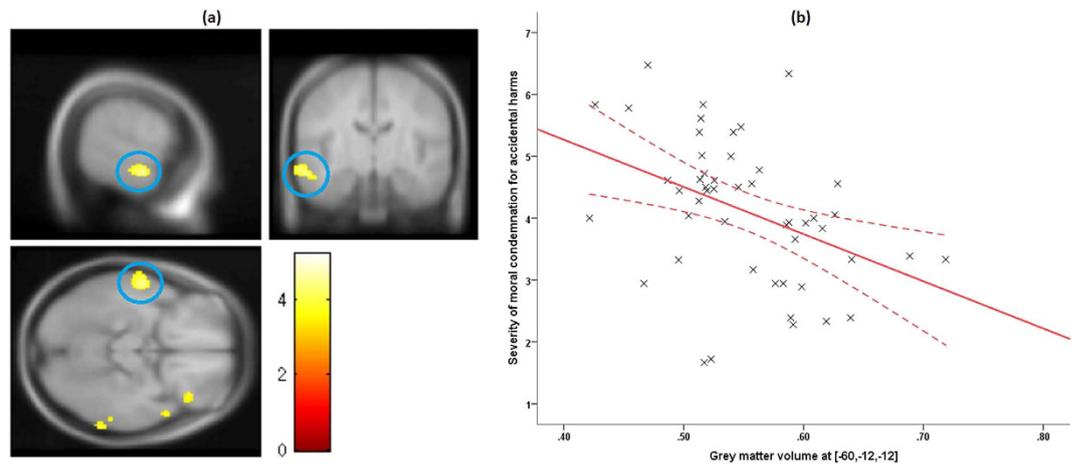
**Data availability statement.** Unthresholded VBM statistical maps of reported contrasts are available on Neurovault<sup>68</sup> at the following address: <http://neurovault.org/collections/1689/>.

All the behavioral data are available at: <https://osf.io/akx6a/>.

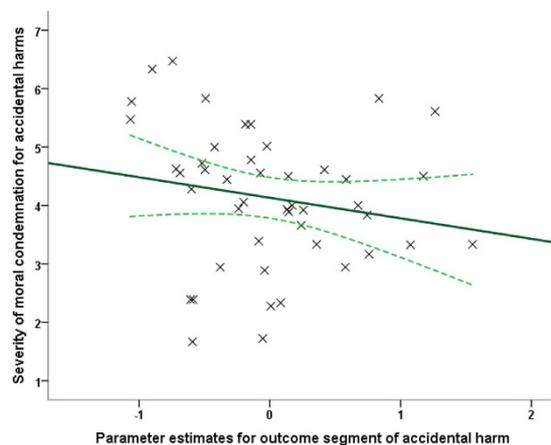
**Post hoc exploratory functional data analysis.** Given that functional data were also available for each participant, we decided to explore *post hoc* whether the extent to which innocent intentions are taken to mitigate condemnation for accidental harms is correlated with the functional activity at this region. Note that this region has not received much attention in prior work on this topic and thus we wanted to ascertain that functional activity in this region was predictive of moral judgments.

For ROI analysis, the data from spherical ROIs with a radius of 8 mm was extracted from l-aSTS at coordinates observed in the VBM analysis and was analyzed using the MarsBar toolbox for SPM (v0.44, <http://marsbar.sourceforge.net/>)<sup>69</sup>. Within the ROI, we extracted parameter estimates ( $\beta$ s) from all segments of interest (*mental state information, consequence, acceptability, blame*) for the accidental harm condition and correlated these with behavioral ratings, i.e., moral condemnation. Data defining ROIs were independent from data used in the repeated measures statistics<sup>53,70</sup>, and restricting analysis to a few ROIs reduced Type-I error by drastically limiting the number of statistical tests performed<sup>71</sup>.

Results revealed a negative correlation between parameter estimates extracted from the *consequence/outcome* segment and condemnation for accidents for l-aSTS ( $\rho(40) = -0.294, p = 0.029, n = 42$ , one-tailed) (Fig. 4). Note



**Figure 3. VBM results.** (a) VBM result for accidental harm condition at group random effects analysis. Maps are thresholded at  $p(\text{uncorrected}) < 0.001$ ,  $k > 10$ , for illustrative purposes. The left anterior superior temporal sulcus (l-aSTS) is highlighted in the blue circle. The color bar denotes  $t$ -values. (b) A scatter plot illustrating the negative linear association between the grey matter volume (GMV) in l-aSTS ( $\rho(47) = -0.547$ , 95% CI  $[-0.726, -0.296]$ ,  $p < 0.001$ ,  $n = 49$ , two-tailed) and the severity of moral condemnation of accidental harm, accounting for nuisance variables. The solid lines indicate a linear fit to the data, while the curved lines represent mean 95% confidence intervals for these lines. Extracted grey matter volume data presented in figures are non-independent of the statistical test used to find effect at this region and thus should *not* be used for effect-size estimates<sup>97</sup>. They are included here only as a visual aid for interpretation of results. Abbreviation - VBM: voxel-based morphometry.



**Figure 4. Brain-behavior correlation in fMRI data.** A scatter plot illustrating the negative linear association ( $\rho(40) = -0.294$ ,  $p = 0.029$ ,  $n = 42$ , one-tailed) between the parameter estimates ( $\beta$ s) extracted from l-aSTS  $[-60, -12, -12]$  during the *consequence/outcome* segment (when outcome information was revealed) of the moral judgment task and the severity of moral condemnation of accidental harms. The solid lines indicate a linear fit to the data, while the curved lines represent mean 95% confidence intervals for this line.

that we have used Spearman's  $\rho$  as our correlation measure, as it is more robust to univariate outliers<sup>72</sup> and one-tailed correlation tests given our strong directional hypotheses<sup>73</sup>. None of the other correlations were significant ( $p$ s  $> 0.05$ , one-tailed).

## Discussion

In the present study, VBM was used to investigate whether inter-individual variation in intent-based moral judgments could be predicted from variation in the local GMV from regions belonging to the ToM network. We found that only variation in GMV in the l-aSTS, which was localized using an independent functional localizer task, could explain variance in moral condemnation of accidental harms: higher GMV was associated with increased tendency to forgive accidents (i.e., unintended harm to others). A similar profile was also observed in an exploratory analysis of functional data. These results raise several questions: What is the function of aSTS in the ToM network and its involvement in moral reasoning? What does the observed VBM effect signify? We discuss these issues below.

**The functional role of the l-aSTS.** Over the last two decades, over 400 studies have investigated the neural basis of the ToM network, and this work has revealed a remarkably consistent set of brain regions, including bilateral TPJ, PC, sections of mPFC, bilateral STS, and TP. In the current study, we were able to delineate a network anatomically consistent with the meta-analytic definitions of the ToM network using a well-validated functional localizer task, featuring social animations, and VBM effect was found at left aSTS, which was part of the ToM network. Here, we speculate what psychological process this region may support during moral evaluations.

Quantitative meta-analyses and large-scale studies focusing on ToM tasks have consistently found aSTS/MTG<sup>64,65,74–77</sup>. Asking participants to attend to *why* an agent is performing a certain action as compared to *how* the action is being performed elicits increased activity in l-aSTS<sup>63</sup>. The l-aSTS is also involved in representing the valence dimension of mental state representations, which captures the difference between positive and negative mental states<sup>78</sup>. While reasoning about agents acting with false (versus true) beliefs, l-aSTS shows significant signal change from baseline<sup>62,79</sup> and also comes online as early as 200–300 ms<sup>80</sup>. This work indicates the involvement of l-aSTS in ToM<sup>4</sup>. But exactly which aspect of ToM does this region support? The umbrella term ToM in its broadest sense represents the capacity to process the representational mental states of other agents, but this process relies on other basic component processes as well (e.g., causal reasoning, self/other distinction, face recognition, gaze processing, etc.)<sup>81</sup>. Therefore, it has been argued that there is not one homogeneous ToM network or mentalizing system but rather different neural regions for distinct aspects of ToM<sup>82</sup>. Indeed, recent meta-analyses<sup>64,65</sup> suggest heterogeneous functional profiles associated with different types of ToM tasks (e.g. false belief, social animations, strategic games, etc.). What specific aspect of ToM is the l-aSTS responsible for? Some prior work sheds light on this issue. A systematic study contrasting social animations versus false belief tasks revealed that the TPJ primarily supports predicting behavior based on covert mental states, while both aSTS and pSTS support decoding intentions or mental states from animate motion or perceived actions<sup>61</sup>, consistent as well with more exhaustive meta-analytic investigations<sup>64</sup>. According to these authors, the broader term “intention” belies the graded distinction between two types of mental states<sup>83</sup>: (i) “intention-in-action”: perceiving intentions and goals in the actions of agents (e.g., social animation tasks) (subserved by l-aSTS); (ii) “prior intention”: representing intentions based on covert mental states (e.g., false belief tasks) that may or may not lead to immediate action (subserved by TPJ). Indeed, some developmental evidence also supports this distinction: the representation of others’ intentions decoded from actions develops before the representation of covert mental states<sup>84</sup>.

According to this view, the role of l-aSTS during the moral judgment task is to infer the nature of the agent’s intent (malicious or innocent) based on the action the agent performed. Notably, in real life, we rarely have access to agents’ internal mental states; agents’ external actions are what we tend to rely on when considering their goals and intentions<sup>85,86</sup>. Thus, we speculate that the observed structure-behavior correlation represents participants’ dispositional tendency in daily life to infer the intentions underlying actions based on observation of actions rather than by explicit representation of mental state information. The greater this tendency, the more adept the participant will be in attributing reduced harmful intent to accidental harm-doers. The current result is also consistent with prior work showing activity in l-aSTS during forgivability judgments<sup>67</sup>.

To some, it may be surprising that we did not find the VBM effect at rTPJ, given the amount of research that places rTPJ at the center of morally relevant mental state reasoning<sup>16–18</sup>. We discuss this null effect at length in Supplementary Text S6.

**What do the current VBM results signify?** In the VBM analysis, whole-brain voxel-wise regression analyses were restricted to the functionally defined ToM network to investigate the link between individual differences in moral condemnation and variation in local GMV in the network. This analysis revealed one robust effect: more lenient moral judgments of accidental harms were correlated with greater GMV at l-aSTS. What does this result signify at the mechanistic level?

It is still unclear how and why individual differences in brain morphometry are found to be correlated with personality traits and task performance, but it is often assumed that greater GMV is associated with better computational efficiency of that region<sup>48</sup>, which in turn leads to enhanced task performance. Indeed, grey matter reduction in l-aSTS is associated with ToM deficits in schizophrenia patients<sup>66</sup>. Therefore, the current findings can be interpreted in the following way: individuals with greater GMV in l-aSTS tend to exculpate an agent who causes harm accidentally because they are better at generating a robust representation of an agent’s innocent intentions needed to compete with prepotent negative arousal elicited by harmful consequences, which would lead to condemnation<sup>87</sup>. In other words, greater GMV in l-aSTS enhances computational efficiency of generating and processing mental state representation, which in turn leads to greater reliance on this information for moral judgment.

This raises an even more interesting question as to why some people have greater GMV at aSTS than others to begin with. While a cross-sectional study like ours can’t arbitrate on this issue (or determine the causal direction of the relationship between brain structure and moral judgment), we offer some speculation here. On the one hand, it is known that individual differences in GMV at focal brain areas are highly heritable<sup>88,89</sup>. It is therefore possible that genes contribute to variation in GMV at l-aSTS and thus to variation in moral judgments (gene → structure → judgment). On the other hand, the alternative causal pathway is equally valid (environment → judgment → structure). Different environments (cultures, societies, etc.) differ in the *degree* to which they place emphasis on intent versus outcome for moral judgment<sup>2,90,91</sup>. These environmental influences are associated with variation in brain regions due to use-dependent brain plasticity such that the size of a brain region is influenced by its use (neurons that regularly “fire together, wire together”<sup>92</sup>). In addition to these two possibilities, much more complicated interplay can also be expected between genes and environment via either interaction and/or correlation<sup>88</sup>.

**Limitations.** As a limitation, we note that although, traditionally, behavioral data for VBM studies are collected out of the scanner, this was not the case for the current study. Some recent work shows that participants are slower and exhibit poor focus on the task while in scanner as opposed to the lab environment, though this difference was observed for a perceptual decision-making task<sup>93</sup> and its relevance for social decision-making remains unclear. Future studies can also explore how variety of other demographic details<sup>94</sup> (Big Five personality traits, education, ethnicity, etc.) and more realistic contexts<sup>95</sup> affect moral judgments rather than relying primarily on a uniform group university students and hypothetical text vignettes as sampled in the current study.

## Conclusion

In conclusion, we have shown here that the interindividual differences in the severity of moral judgments about unintentional harmful behaviors are associated with volumetric differences in the left aSTS, a region implicated in reasoning about others' mental states, such that the greater the grey matter volume, the less accidental harm-doers are condemned.

## References

- Cushman, F. Crime and punishment: distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition* **108**, 353–80 (2008).
- Barrett, H. C. *et al.* Small-scale societies exhibit fundamental variation in the role of intentions in moral judgment. *Proc. Natl. Acad. Sci.* **113**, 4688–4693 (2016).
- Nobes, G., Panagiotaki, G. & Bartholomew, K. J. The influence of intention, outcome and question-wording on children's and adults' moral judgments. *Cognition* **157**, 190–204 (2016).
- Koster-hale, J. & Saxe, R. In *Understanding Other Minds* (eds Baron-Cohen, S., Lombardo, M. & Tager-Flusberg, H.) 132–163 (Oxford University Press, 2013).
- Gershman, S. J., Gerstenberg, T., Baker, C. L. & Cushman, F. Plans, Habits, and Theory of Mind. *PLoS One* **11**, e0162246 (2016).
- Bzdok, D. *et al.* Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. *Brain Struct. Funct.* **217**, 783–96 (2012).
- Decety, J. & Cacioppo, S. The speed of morality: a high-density electrical neuroimaging study. *J. Neurophysiol.* **108**, 3068–3072 (2012).
- Gan, T. *et al.* Temporal Dynamics of the Integration of Intention and Outcome in Harmful and Helpful Moral Judgment. *Front. Psychol.* **6**, 2022 (2016).
- Hesse, E. *et al.* Early detection of intentional harm in the human amygdala. *Brain* **139**, 54–61 (2016).
- Young, L. & Saxe, R. The neural basis of belief encoding and integration in moral judgment. *Neuroimage* **40**, 1912–20 (2008).
- Ginther, M. R. *et al.* Parsing the Behavioral and Brain Mechanisms of Third-Party Punishment. *J. Neurosci.* **36**, 9420–9434 (2016).
- Treadway, M. T. *et al.* Corticolimbic gating of emotion-driven punishment. *Nat. Neurosci.* **17**, 1270–5 (2014).
- Yu, H., Li, J. & Zhou, X. Neural Substrates of Intention–Consequence Integration and Its Impact on Reactive Punishment in Interpersonal Transgression. *J. Neurosci.* **35**, 4917–4925 (2015).
- Bellucci, G. *et al.* Effective connectivity of brain regions underlying third-party punishment: functional MRI and Granger causality evidence. *Soc. Neurosci.* **12**, 124–134 (2017).
- Patil, I., Young, L., Sinay, V. & Gleichgerricht, E. Elevated moral condemnation of third-party violations in multiple sclerosis patients. *Soc. Neurosci.*, doi: 10.1080/17470919.2016.1175380 (2016).
- Young, L. & Saxe, R. Innocent intentions: A correlation between forgiveness for accidental harm and neural activity. *Neuropsychologia* **47**, 2065–2072 (2009).
- Koster-Hale, J., Saxe, R., Dungan, J. & Young, L. Decoding moral judgments from neural representations of intentions. *Proc. Natl. Acad. Sci.* **110**, 5648–53 (2013).
- Chakroff, A. *et al.* When minds matter for moral judgment: intent information is neurally encoded for harmful but not impure acts. *Soc. Cogn. Affect. Neurosci.* **11**, 476–484 (2016).
- Young, L., Camprodon, J. A., Hauser, M., Pascual-Leone, A. & Saxe, R. Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. *Proc. Natl. Acad. Sci.* **107**, 6753–6758 (2010).
- Sellaro, R. *et al.* Increasing the role of belief information in moral judgments by stimulating the right temporoparietal junction. *Neuropsychologia* **77**, 400–408 (2015).
- Peets, K., Hodges, E. V. E. & Salmivalli, C. Forgiveness and its determinants depending on the interpersonal context of hurt. *J. Exp. Child Psychol.* **114**, 131–145 (2013).
- Ricciardi, E. *et al.* How the brain heals emotional wounds: the functional neuroanatomy of forgiveness. *Front. Hum. Neurosci.* **7**, 839 (2013).
- Lewis, G. J., Kanai, R., Bates, T. C. & Rees, G. Moral values are associated with individual differences in regional brain volume. *J. Cogn. Neurosci.* **24**, 1657–63 (2012).
- Prehn, K. *et al.* Neural Correlates of Post-Conventional Moral Reasoning: A Voxel-Based Morphometry Study. *PLoS One* **10**, e0122914 (2015).
- Nash, K., Baumgartner, T. & Knoch, D. Group-focused morality is associated with limited conflict detection and resolution capacity: Neuroanatomical evidence. *Biol. Psychol.* **123**, 235–240 (2017).
- Baumgartner, T., Saulin, A., Hein, G. & Knoch, D. Structural Differences in Insular Cortex Reflect Vicarious Injustice Sensitivity. *PLoS One* **11**, e0167538 (2016).
- Watanabe, T. *et al.* Two distinct neural mechanisms underlying indirect reciprocity. *Proc. Natl. Acad. Sci.* **111**, 3990–5 (2014).
- Thijssen, S. *et al.* Cortical thickness and prosocial behavior in school-age children: A population-based MRI study. *Soc. Neurosci.* **10**, 571–582 (2015).
- Marsh, A. *et al.* Neural and cognitive characteristics of extraordinary altruists. *Proc. Natl. Acad. Sci.* **111**, 15036–41 (2014).
- Yamagishi, T. *et al.* Cortical thickness of the dorsolateral prefrontal cortex predicts strategic choices in economic games. *Proc. Natl. Acad. Sci.* **113**, 5582–5587 (2016).
- Patil, I., Zanon, M., Novembre, G., Zangrando, N., Chittaro, L. & Silani, G. Neuroanatomical basis of concern-based altruism in virtual environment. *Neuropsychologia*. doi:10.1016/j.neuropsychologia.2017.02.015 (2017).
- Baez, S. *et al.* Integration of Intention and Outcome for Moral Judgment in Frontotemporal Dementia: Brain Structural Signatures. *Neurodegener. Dis.* **16**, 206–17 (2016).
- Baez, S. *et al.* Orbitofrontal and limbic signatures of empathic concern and intentional harm in the behavioral variant frontotemporal dementia. *Cortex* **75**, 20–32 (2015).
- Castelli, F., Happé, F., Frith, U. & Frith, C. Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *Neuroimage* **12**, 314–25 (2000).
- Moessnang, C. *et al.* Specificity, reliability and sensitivity of social brain responses during spontaneous mentalizing. *Soc. Cogn. Affect. Neurosci.* **11**, 1687–1697 (2016).

36. Eickhoff, S & Müller, V. I. In *Brain Mapping: An Encyclopedic Reference* (ed. Toga, A. W.) 187–201, doi: 10.1016/B978-0-12-397025-1.00212-8 (Academic Press, 2015).
37. Calhoun, V. D., Liu, J. & Adalı, T. A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data. *Neuroimage* **45**, S163–S172 (2009).
38. Ridgway, G. R. *et al.* Ten simple rules for reporting voxel-based morphometry studies. *Neuroimage* **40**, 1429–1435 (2008).
39. Kurth, F., Gaser, C. & Luders, E. A 12-step user guide for analyzing voxel-wise gray matter asymmetries in statistical parametric mapping (SPM). *Nat. Protoc.* **10**, 293–304 (2015).
40. Ashburner, J. & Friston, K. Why voxel-based morphometry should be used. *Neuroimage* **14**, 1238–1243 (2001).
41. Ashburner, J. Computational anatomy with the SPM software. *Magn. Reson. Imaging* **27**, 1163–1174 (2009).
42. Kurth, F., Luders, E. & Gaser, C. In *Brain Mapping: An Encyclopedic Reference* (ed. Toga, A. W.) 345–349, doi: 10.1016/B978-0-12-397025-1.00304-3 (Academic Press, Elsevier, 2015).
43. Mietchen, D. & Gaser, C. Computational morphometry for detecting changes in brain structure due to development, aging, learning, disease and evolution. *Front. Neuroinform.* **3**, 25 (2009).
44. Poldrack, R., Mumford, J. & Nichols, T. E. *Handbook of functional MRI data analysis* (Cambridge University Press, 2011).
45. Ashburner, J. & Friston, K. J. Unified segmentation. *Neuroimage* **26**, 839–51 (2005).
46. Malone, I. B. *et al.* Accurate automatic estimation of total intracranial volume: A nuisance variable with less nuisance. *Neuroimage* **104**, 366–372 (2015).
47. Ashburner, J. A fast diffeomorphic image registration algorithm. *Neuroimage* **38**, 95–113 (2007).
48. Kanai, R. & Rees, G. The structural basis of inter-individual differences in human behaviour and cognition. *Nat. Rev. Neurosci.* **12**, 231–242 (2011).
49. Ashburner, J. & Friston, K. In *Human brain function* (eds Frackowiak, R. S. J. *et al.*) 707–722, doi: 10.1016/B978-012264841-0 (Academic Press, 2004).
50. Barnes, J. *et al.* Head size, age and gender adjustment in MRI studies: A necessary nuisance? *Neuroimage* **53**, 1244–1255 (2010).
51. Nostro, A. D., Müller, V. I., Reid, A. T. & Eickhoff, S. B. Correlations Between Personality and Brain Structure: A Crucial Role of Gender. *Cereb. Cortex*, doi: 10.1093/cercor/bhw191 (2016).
52. Ridgway, G. R. *et al.* Issues with threshold masking in voxel-based morphometry of atrophied brains. *Neuroimage* **44**, 99–111 (2009).
53. Poldrack, R. Region of interest analysis for fMRI. *Soc. Cogn. Affect. Neurosci.* **2**, 67–70 (2007).
54. Eklund, A., Nichols, T. E. & Knutsson, H. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proc. Natl. Acad. Sci.* **113**, 7900–7905 (2016).
55. Woo, C. W., Krishnan, A. & Wager, T. D. Cluster-extent based thresholding in fMRI analyses: Pitfalls and recommendations. *Neuroimage* **91**, 412–419 (2014).
56. Smith, S. M. & Nichols, T. E. Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* **44**, 83–98 (2009).
57. Salimi-Khorshidi, G., Smith, S. M. & Nichols, T. E. Adjusting the effect of nonstationarity in cluster-based and TFCE inference. *Neuroimage* **54**, 2006–2019 (2011).
58. Li, H., Nickerson, L. D., Nichols, T. E. & Gao, J.-H. Comparison of a non-stationary voxelation-corrected cluster-size test with TFCE for group-Level MRI inference. *Hum. Brain Mapp.* **38**(3), 1269–1280, doi: 10.1002/hbm.23453 (2017).
59. Chen, G., Taylor, P. A. & Cox, R. W. Is the Statistic Value All We Should Care about in Neuroimaging? *Neuroimage* **147**, 952–959, doi: 10.1016/j.neuroimage.2016.09.066 (2017).
60. Reddan, M. C., Lindquist, M. A. & Wager, T. D. Effect Size Estimation in Neuroimaging. *JAMA Psychiatry* **4**, 863 (2017).
61. Gobbini, M. I., Koralek, A. C., Bryan, R. E., Montgomery, K. J. & Haxby, J. V. Two takes on the social brain: a comparison of theory of mind tasks. *J. Cogn. Neurosci.* **19**, 1803–14 (2007).
62. Deen, B., Koldewyn, K., Kanwisher, N. & Saxe, R. Functional Organization of Social Perception and Cognition in the Superior Temporal Sulcus. *Cereb. Cortex* **25**, 4596–4609 (2015).
63. Spunt, R. P. & Adolphs, R. Validating the Why/How contrast for functional MRI studies of Theory of Mind. *Neuroimage* **99**, 301–311 (2014).
64. Schurz, M., Radua, J., Aichhorn, M., Richlan, F. & Perner, J. Fractionating theory of mind: a meta-analysis of functional brain imaging studies. *Neurosci. Biobehav. Rev.* **42**, 9–34 (2014).
65. Molenberghs, P., Johnson, H., Henry, J. D. & Mattingley, J. B. Understanding the minds of others: A neuroimaging meta-analysis. *Neurosci. Biobehav. Rev.* **65**, 276–291 (2016).
66. Hooker, C. I., Bruce, L., Lincoln, S. H., Fisher, M. & Vinogradov, S. Theory of mind skills are related to gray matter volume in the ventromedial prefrontal cortex in schizophrenia. *Biol. Psychiatry* **70**, 1169–78 (2011).
67. Farrow, T. F. D. *et al.* Quantifiable change in functional brain response to empathic and forgiveness judgments with resolution of posttraumatic stress disorder. *Psychiatry Res. Neuroimaging* **140**, 45–53 (2005).
68. Gorgolewski, K. J. *et al.* NeuroVault.org: a web-based repository for collecting and sharing unthresholded statistical maps of the human brain. *Front. Neuroinform.* **9**, 8 (2015).
69. Brett, M., Anton, J., Valabregue, R. & Poline, J. Region of interest analysis using the MarsBar toolbox for SPM 99. *Neuroimage* **16**, S497 (2002).
70. Vul, E. & Kanwisher, N. In *Foundational Issues in Human Brain Mapping* (eds Hanson, S. J. & Bunzl, M.) 71–92 (MIT Press, 2010).
71. Saxe, R., Brett, M. & Kanwisher, N. Divide and conquer: a defense of functional localizers. *Neuroimage* **30**, 1088–96 (2006).
72. Pernet, C., Wilcox, R. & Rousselet, G. Robust correlation analyses: False positive and power validation using a new open source matlab toolbox. *Front. Psychol.* **3**, 606 (2013).
73. Cho, H.-C. & Abe, S. Is two-tailed testing for directional research hypotheses tests legitimate? *J. Bus. Res.* **66**, 1261–1266 (2013).
74. Mar, R. The neural bases of social cognition and story comprehension. *Annu. Rev. Psychol.* **62**, 103–134 (2011).
75. Spreng, R. N., Mar, R. & Kim, A. S. N. The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. *J. Cogn. Neurosci.* **21**, 489–510 (2009).
76. Krall, S. C. *et al.* The role of the right temporoparietal junction in attention and social interaction as revealed by ALE meta-analysis. *Brain Struct. Funct.* **220**, 587–604 (2014).
77. Kanske, P., Böckler, A., Trautwein, F.-M. & Singer, T. Dissecting the social brain: Introducing the EmpaToM to reveal distinct neural networks and brain-behavior relations for empathy and Theory of Mind. *Neuroimage* **122**, 6–19 (2015).
78. Tamir, D. I., Thornton, M. A., Contreras, J. M. & Mitchell, J. P. Neural evidence that three dimensions organize mental state representation: Rationality, social impact, and valence. *Proc. Natl. Acad. Sci.* **113**, 194–199 (2016).
79. Döhl, K. *et al.* Functional activity of the right temporo-parietal junction and of the medial prefrontal cortex associated with true and false belief reasoning. *Neuroimage* **60**, 1652–61 (2012).
80. Mossad, S. I. *et al.* Thinking about the thoughts of others; temporal and spatial neural activation during false belief reasoning. *Neuroimage* **134**, 320–327 (2016).
81. Happé, F. G., Cook, J. & Bird, G. The Structure of Social Cognition: In(ter) dependence of Sociocognitive Processes. *Annu. Rev. Psychol.* **68**, 243–267 (2017).
82. Schaafsma, S. M., Pfaff, D. W., Spunt, R. P. & Adolphs, R. Deconstructing and reconstructing theory of mind. *Trends Cogn. Sci.* **19**, 65–72 (2015).

83. Searle, J. R. The Intentionality of Intention and Action. *Cogn. Sci.* **4**, 47–70 (1980).
84. Meltzoff, A. N., Gopnik, A. & Repacholi, B. M. In *Developing Theories of Intention: Social Understanding and Self-control* (eds Zelazo, P. D., Astington, J. W. & Olson, D. R.) 17–42 (Erlbaum, 1999).
85. Dungan, J. A., Stepanovic, M. & Young, L. Theory of mind for processing unexpected events across contexts. *Soc. Cogn. Affect. Neurosci.* **11**, 1183–92 (2016).
86. Koster-Hale, J. & Saxe, R. Theory of Mind: A Neural Prediction Problem. *Neuron* **79**, 836–848 (2013).
87. Patil, I. & Silani, G. Alexithymia increases moral acceptability of accidental harms. *J. Cogn. Psychol.* **26**, 597–614 (2014).
88. Gu, J. & Kanai, R. What contributes to individual differences in brain structure? *Front. Hum. Neurosci.* **8**, 262 (2014).
89. Winkler, A. M. *et al.* Cortical thickness or grey matter volume? The importance of selecting the phenotype for imaging genetics studies. *Neuroimage* **53**, 1135–1146 (2010).
90. Cohen, A. B. & Rozin, P. Religion and the morality of mentality. *J. Pers. Soc. Psychol.* **81**, 697–710 (2001).
91. Clark, C. J., Bauman, C. W., Kamble, S. V. & Knowles, E. D. Intentional Sin and Accidental Virtue? Cultural Differences in Moral Systems Influence Perceived Intentionality. *Soc. Psychol. Personal. Sci.* **8**, 74–82 (2016).
92. Hebb, D. O. *The Organization of Behavior*, doi: 10.2307/1418888 (Wiley & Sons, 1949).
93. van Maanen, L., Forstmann, B. U., Keuken, M. C., Wagenmakers, E.-J. & Heathcote, A. The impact of MRI scanner environment on perceptual decision-making. *Behav. Res. Methods* **48**, 184–200 (2016).
94. Banerjee, K., Huebner, B. & Hauser, M. Intuitive Moral Judgments are Robust across Variation in Gender, Education, Politics and Religion: A Large-Scale Web-Based Study. *J. Cogn. Cult.* **10**, 253–281 (2010).
95. Patil, I., Cogoni, C., Zangrando, N., Chittaro, L. & Silani, G. Affective basis of judgment-behavior discrepancy in virtual experiences of moral dilemmas. *Soc. Neurosci.* **9**, 94–107 (2014).
96. Allen, E. A., Erhardt, E. B. & Calhoun, V. D. Data visualization in the neurosciences: overcoming the curse of dimensionality. *Neuron* **74**, 603–8 (2012).
97. Vul, E. & Pashler, H. In *Psychological science under scrutiny: Recent challenges and proposed solutions* (eds Lilienfeld, S. O. & Waldman, I. D.) 196–220 (Wiley, 2017).

## Acknowledgements

The authors gratefully acknowledge Henryk Bukowski, Mallory Feldman, Alessandra Nostro, Tobias Schuwerk, Emily Wasserman, and Marco Zanon for their helpful comments on the earlier version of the manuscript. This work was supported by the Vienna Science and Technology Fund (WWTF; Project CS15-003) to G.S. This article was supported by the Open Access Publishing Fund of the University of Vienna.

## Author Contributions

I.P. and G.S. designed fMRI research; I.P., M.C., and F.F. acquired data; I.P. analyzed the data; I.P. wrote the paper; G.S. and L.Y. provided critical revisions. All authors approved the submitted version of the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing Interests:** The authors declare no competing financial interests.

**How to cite this article:** Patil, I. *et al.* Neuroanatomical correlates of forgiving unintentional harms. *Sci. Rep.* **7**, 45967; doi: 10.1038/srep45967 (2017).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017