

# Spatiotemporal neural dynamics of moral judgment: A high-density ERP study



Keith J. Yoder<sup>a</sup>, Jean Decety<sup>a,b,\*</sup>

<sup>a</sup> Department of Psychology, University of Chicago, Chicago, IL, United States

<sup>b</sup> Department of Psychiatry, University of Chicago Medicine, Chicago, IL, United States

## ARTICLE INFO

### Article history:

Received 10 March 2014  
Received in revised form  
25 April 2014  
Accepted 20 May 2014  
Available online 4 June 2014

### Keywords:

Moral judgment  
High-density EEG/ERPs  
Cognitive empathy  
Source localization  
Right temporoparietal junction  
Ventromedial prefrontal cortex

## ABSTRACT

Morality is a pervasive aspect of human nature across all cultures, and neuroscience investigations are necessary for identifying what computational mechanisms underpin moral cognition. The current study used high-density ERPs to examine how moral evaluations are mediated by automatic and controlled processes as well as how quickly information and causal-intentional representations can be extracted when viewing morally laden behavior. The study also explored the extent to which individual dispositions in affective and cognitive empathy as well as justice sensitivity influence the encoding of moral valence when healthy participants make moral judgments about prosocial (interpersonal assistance) and antisocial (interpersonal harm) actions. Moral judgment differences were reflected in differential amplitudes for components associated with cognitive appraisal (LPP) as well as early components associated with emotional salience (N1 and N2). Moreover, source estimation was performed to indicate potential neural generators. A posterior-to-anterior shift was observed, with current density peaks first in right inferior parietal cortex (at the temporoparietal junction), then later in medial prefrontal cortex. Cognitive empathy scores predicted behavioral ratings of blame as well as differential amplitudes in LPP and component activity at posterior sites. Overall, this study offers important insights into the temporal unfolding of moral evaluations, including when in time individual differences in empathy influence neural encoding of moral valence.

© 2014 Elsevier Ltd. All rights reserved.

## 1. Introduction

Humans perceive people, and interactions among them, through a framework that conceptualizes behavior as fundamentally linked with mental states (Malle & Guglielmo, 2012), and intuitively evaluate their behaviors as morally laden (Hamlin, 2014). Previous neuroimaging investigations using functional MRI in healthy individuals converge with lesion studies in neurological patients to identify brain mechanisms implicated in moral thinking. There is now solid evidence for the involvement of a network of regions, which include the ventromedial prefrontal cortex (vmPFC), dorso-lateral prefrontal cortex (dlPFC), medial prefrontal cortex (mPFC), temporoparietal junction (TPJ), amygdala, and insula (e.g., Buckholz & Marois, 2012; Decety, Michalska, & Kinzler, 2012; Fumagalli & Priori, 2012; Mendez, 2009; Moll et al., 2007; Young & Dungan, 2012). Moreover, it is clear that these regions are not specific to the domain of morality. Instead, they comprise several domain-general systems which support intention understanding, decision-making,

affective arousal, and attention (Decety & Cowell, 2014; Decety & Howard, 2013; Moll, Oliveira-Souza, & Eslinger, 2003; Prehn & Heekeren, 2009; Shenhav & Greene, 2010).

More recently, there has been a renewed focus on characterizing when in time aspects of these systems contribute to moral judgment. Since event-related potentials (ERPs) measure neural activity with millisecond resolution, they are uniquely suited to capture the time course of information processing in response to morally laden stimuli, and thus provide valuable insights into the relative contributions of attentional, affective, and cognitive systems (Hajcak, MacNamara, & Olvet, 2010). For instance, one study reported larger N2 amplitudes for moral violations compared to conventional violations (Lahat, Helwig, & Zelazo, 2013). Moral violations, such as hitting and stealing, involve intrinsic negative consequences for others, such as physical harm or issues of fairness. In contrast, social conventions are behavioral uniformities that serve to coordinate individuals' interactions in a social system (Helwig & Turiel, 2011). In another study, peak amplitude and latency of the P3 component predicted participants' moral judgments about others' decisions to provide or withhold help (Chiu Loke, Evans, & Lee, 2011). Recently, high-density ERPs, combined with source localizations analysis, were used to examine the spatiotemporal dynamics of the neural processing evoked

\* Corresponding author at: Department of Psychology, University of Chicago, Chicago, IL, United States.

E-mail address: [decety@uchicago.edu](mailto:decety@uchicago.edu) (J. Decety).

by the perception of visual scenarios depicting intentional versus accidental interpersonal harm. Current source density maxima for intentional harm were detected in the right pSTS/TPJ, as fast as 62 ms post-stimulus, and later responses were found in the amygdala (122 ms) and ventromedial prefrontal cortex (182 ms) (Decety & Cacioppo, 2012).

A primary question in moral cognition that remains to be answered is how moral evaluations are mediated by affective processes as well as how rapidly information and causal-intentional representations can be extracted (Huebner, Dwyer, & Hauser, 2009). While there is some recent ERP evidence that individuals encode the intentionality of harmful actions rapidly (Decety & Cacioppo, 2012), directly investigating the spatiotemporal dynamics of neural encoding during explicit moral judgments is an important next step and can leverage the extensive literature on electrophysiological indices of cognitive control and affective processing (Folstein & Petten, 2008; Hajcak et al., 2010; Olofsson, Nordin, Sequeira, & Polich, 2008). For instance, while valence and arousal can be difficult to tease apart when viewing emotionally laden visual stimuli, very early ERP amplitude modulations (prior to 200 ms after stimulus onset) appear to be influenced by stimulus valence, but not by how arousing individuals find the stimulus (Olofsson et al., 2008). Specifically, there is evidence that emotional valence modulates the amplitude of frontal negativities during the N1 (120–150 ms), with the amplitude for pleasant pictures enhanced as compared to that for unpleasant pictures (Keil et al., 2002). Thus, examining the effect of moral valence on very early ERP components can cast light on the unique contributions of moral valence, without the confounding influence of potential differences in arousal between morally bad and morally good actions.

Another important component relevant to the study of morality is the frontal N2 (negative deflections over frontocentral sites between 200 and 350 ms after stimulus onset or feedback). This component has reliably been used to index cognitive control, novelty, and expectancy violation (for a review, see Folstein & Petten, 2008). N2 is consistently localized to the anterior cingulate cortex (ACC) (Hajcak, Moser, Yeung, & Simons, 2005), a region implicated in conflict monitoring (Botvinick, Cohen, & Carter, 2004) and saliency processing (Harsay, Spaan, Wijnen, & Ridderinkhof, 2012). Importantly, N2 also reflects the affective evaluation and motivational significance of negative outcomes (Hajcak et al., 2005). Moreover, a number of ERP studies have shown that N2 is elicited by observation of others in physical pain, independent of the task demands, suggesting that N2 may index an early automatic component of sensitivity to pain (Chen, Yang, & Cheng, 2012; Cheng, Hung, & Decety, 2012; Decety, Yang, & Cheng, 2010; Fan & Han, 2008; Han, Fan, & Mao, 2008; Perry, Bentin, Bartal, Lamm, & Decety, 2010). This component is particularly interesting in the context of moral judgments because in ultimatum games, fair and unfair offers produce differential N2 amplitudes (Boksem & De Cremer, 2010), and this effect is influenced by the social relationships between the participants (Wu, Leliveld, & Zhou, 2011).

Importantly, while early (< 300 ms) ERPs are thought to index relatively obligatory processes such as attention capture (Macnamara, Foti, & Hajcak, 2009; Weinberg & Hajcak, 2010), later components reflect elaborative processes such as cognitive appraisal (Olofsson et al., 2008; Weinberg & Hajcak, 2010). One such component is the late positive potential (LPP), which is thought to be a marker of attentional mechanisms that occur after initial stimulus processing (Cacioppo, Crites, & Gardner, 1996). In the context of emotion-processing, LPP is sensitive to fine-grained aspects of stimuli, rather than a broad pleasantness dimension (Macnamara et al., 2009). Moreover, recent work has shown that such late positivities not only distinguish between prosocial and antisocial actions, but that they are also sensitive to personality differences (Chiu Loke et al., 2011) and social context (Wu et al., 2011).

Finally individual dispositions in empathy and justice sensitivity may play a role at different stages of the information processing

when individuals evaluate morally-laden scenarios. Empathy is a multidimensional construct comprising dissociable components that interact and operate in parallel fashion, including affective, motivational and cognitive components (Decety & Jackson, 2004; Decety & Svetlova, 2012). Justice sensitivity reflects the individual's concern for justice and is an important predictor of justice-related emotion and behavior (Baumert, Rothmund, Thomas, Gollwitzer, & Schmitt, 2013).

To specifically assess the rapid encoding of moral valence, the present study used dynamic visual scenarios that depict everyday interactions of helping and hurting, thus varying along one dimension of well-established moral valence (Sousa, Holbrook, & Piazza, 2009; Turiel, 2008). In a previous fMRI study (Yoder & Decety, 2014), these scenarios reliably elicited neuro-hemodynamic activity in regions involved in moral cognition, including the right pSTS/TPJ and dlPFC, which suggests that morally bad actions could be evaluated with less reliance on cognitive representations implemented in the prefrontal cortex. Here, a similar pattern of response was expected. It was hypothesized that scenarios depicting prosocial and antisocial actions would be automatically distinguished as indexed by differential N1 and N2 amplitudes. Moreover, amplitude differences were expected to persist into later components, specifically, LPP. Individual dispositions in empathy were expected to specifically impact the ERPs associated with later cognitive reappraisal in moral evaluation. Finally, if justice sensitivity is more related to cognitive elaboration than to emotional arousal (see Yoder and Decety, 2014), then such a dimension may not impact neural responses on the electrophysiological time scale to which ERPs are sensitive.

## 2. Materials and methods

### 2.1. Participants

Thirty-eight healthy adults participated in the study. Participants were compensated with either cash or course credit. Due to experimenter error, data from two subjects was not usable. Additionally, one subject was removed for less than 10 artifact-free trials per condition. Thus, the final sample consisted of 35 healthy adults (23 female, 12 male; age = 20.81 ± 2.10 years). This study was approved by University of Chicago Institutional Review Board.

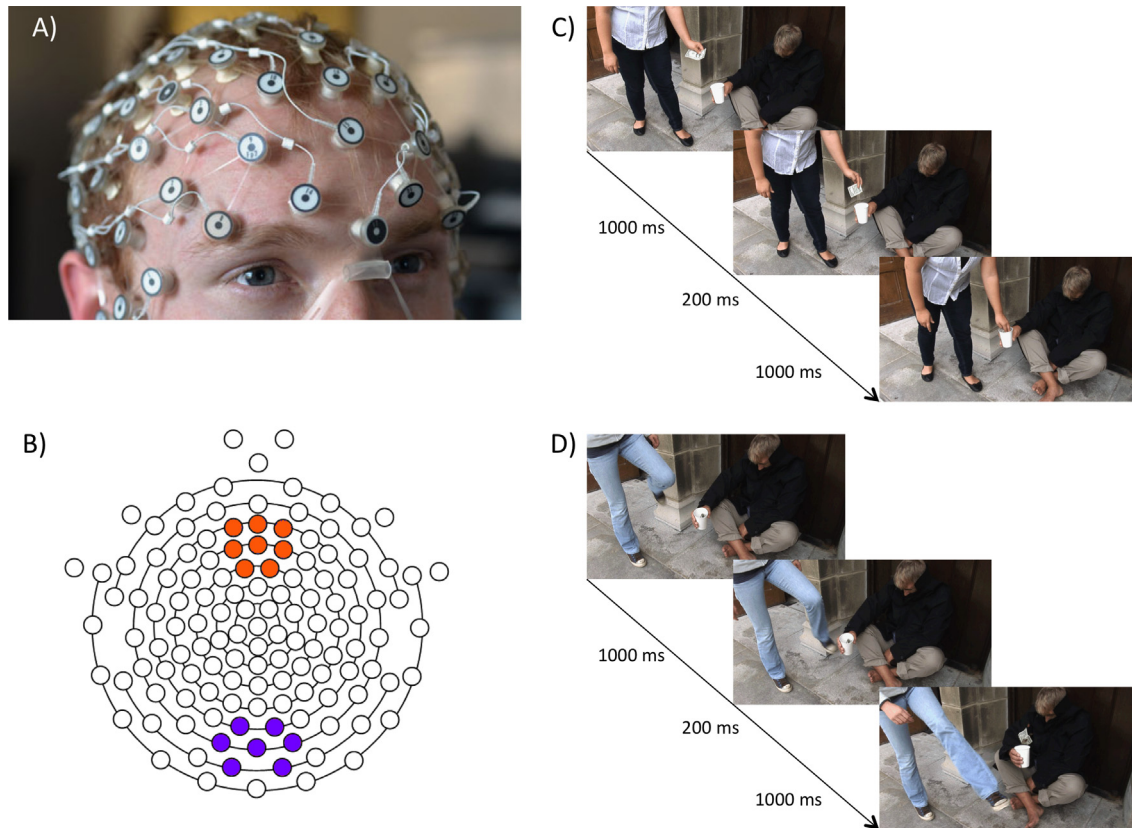
### 2.2. Dispositional measures

At least one week before coming to lab, participants completed an online form that assessed basic demographics such as age, gender, and education level. Additionally, trait cognitive and affective empathy were assessed using the Questionnaire for Cognitive and Affective Empathy (QCAE; Reniers, Corcoran, Drake, Shryane, & Völlm, 2011), and trait justice sensitivity was measured using the Justice Sensitivity Inventory (Schmitt, Baumert, Gollwitzer, & Maes, 2010). This inventory is a self-report psychometric measure that assesses four perspectives of justice sensitivity and produces four scores between 0 and 6, which index an individual's disposition to react to unfair situations.

### 2.3. Stimuli and task

A validated stimulus set was taken from a previous fMRI study of moral judgment (Yoder & Decety, 2014). Briefly, actors were videotaped carrying out intentional actions directed towards another person. Ninety scenarios depicted a range of ecologically valid dyadic interactions of intentional assistance and harm (e.g., helping someone up off of the floor, pulling hair). Importantly, the individuals' faces were not visible so no emotional reaction was visible to study participants. Three still frames were extracted from each clip, cropped to 720 × 420 pixels, and presented serially (1000 ms, 200 ms, 1000 ms; Fig. 1 for examples and the EEG net).

Participants classified each scene six times in a mixed block/event-related design. Prior to recording, a brief training session familiarized participants with the experiment structure. At the start of each block, a cue word appeared to indicate whether they should "focus on the motive or intention of the person performing the action" ("Motive") or "focus on the outcome of the action" ("Outcome"). Following the 1-s cue, six scenes were presented. Each scene was followed by a fixation cross (500 ms) to prevent motor potentials from contaminating the ERP. Finally, a response screen (1000 ms) appeared and participants indicated whether the preceding action had been "Good," "Bad," or "Neither" by pressing the 1, 2, or



**Fig. 1.** Electrode placement and task structure. The sensor net (A) and electrode diagram (B) with frontal (orange) and posterior (violet) clusters shaded. Examples of the three picture sequence with their duration for (C) a good prosocial action, a person putting a dollar bill in a beggar's cup, and (D) a bad antisocial action, a person kicking a beggar's cup away. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3 key on a keypad. Trials without a response during the 1000 ms response screen were scored as misses and not included in any analysis. A jittered fixation cross followed each response screen ( $1000 \pm 500$  ms). Motive and Outcome conditions were collapsed into a single "moral" condition. Blocks, scene order within blocks, and response-key mapping were pseudo-randomized for each participant. Participants viewed each scene one during the course of run (6.5 min). Rest periods were included between each runs, and impedances were checked during this time. Six runs resulted in a total of 540 trials. The experiment was programmed in MATLAB 7.12 (Mathworks Inc., Sherborn, MA) using the Psychophysics Toolbox (Brainard, 1997), and participants viewed stimuli while seated in a chair ~75 cm away from a 21 in. CRT monitor set to  $1024 \times 768$  resolution at 100 Hz refresh rate. After the EEG/ERP recording session, participants were asked to evaluate each of the scenarios using a computerized visual analog scale to indicate how much they would blame or praise the individual who performed the actions.

#### 2.4. EEG/ERP acquisition and analysis

Data were collected using a Geodesic Hydrocel system with 128 channels (Electrical Geodesics, Eugene, OR). Impedances were kept below 50 k $\Omega$ . Samples were digitized at 250 Hz referenced to the vertex using EGI software. Data were re-referenced offline to the average of all channels and IIR filtered from 0.1 to 30 Hz, with a 60 Hz notch filter, in Brainvision Analyzer 2 (Brain Products, Gilching, Germany). For the purpose of frequency analyses, continuous EEG was segmented into good and bad scenes based on subjects' classifications beginning 1000 ms before the onset of the first picture and continuing until the offset of the third picture (total 3200 ms epochs).

After segmentation, data were imported into the ERP PCA Toolkit (Dien, 2010). Bad channel interpolation was conducted trial-wise, and EEGLAB's runICA algorithm (Delorme & Makeig, 2004) was utilized to remove blinks and movement artifacts. ERPs were baseline-corrected using the 100 ms prior to onset of the first picture. In a separate validation study, 40 healthy adults who did not participate in the EEG study were randomly presented with either the first picture or second picture from each of the scenarios and were asked to guess whether the outcome of the depicted action was harmful or helpful. By the first picture, accuracy ratings were well above chance ( $t(39)=26.87, p < 0.001$ ). Thus, the onset of the first picture was used to define all ERPs. Mean amplitudes for N1 (100–150 ms), N2 (175–300 ms), LPP/early slow wave (300–600 ms) and late slow wave (600–1000 ms) were extracted for bad and good actions within frontal (EGI electrodes

4, 5, 10, 11, 12, 16, 18, and 19) and posterior (EGI electrodes 70, 71, 74, 75, 76, 82, and 83) clusters (see Fig. 1B). The effect of moral valence was assessed within each time window using paired *t*-tests and corrected for multiple comparisons using the conservative Bonferroni method. In order to visualize the likely neural generators of scalp voltages within these temporal windows, the grand average was imported into Brainvision Analyzer 2 and four (each time window) separate LORETA source estimates were generated.

In order to investigate the component activity underlying the measured scalp voltages, data were subjected to a two-step principal component analysis (PCA). First, a temporal PCA was conducted with a Promax rotation to decompose microvolts at each electrode across subjects. This temporal decomposition was followed by a spatial decomposition with an Infomax (ICA) rotation. This procedure is especially well-suited to extract component activity that overlaps in both time and space (Dien, Khoe, & Mangun, 2007). Based on the Scree plots, 21 components were retained from the temporal PCA, and five factors were retained from the subsequent spatial PCA. Components were then transformed back into  $\mu$ V. Finally, log mean power density in the alpha frequency range (8–12 Hz) at the frontal and posterior clusters, as well as C3, C4, and the average across all recording sites, was extracted for 1000 ms preceding stimulus onset and subtracted from 1000 ms following stimulus onset.

### 3. Results

#### 3.1. Subjective moral evaluations

Participant's classified interpersonal harm as morally wrong and deserving of blame and classified interpersonal assistance as morally good and deserving of praise ( $p < 0.001$ ).

#### 3.2. Dispositional measures

Individuals who scored high on cognitive empathy ( $r=0.69, p < 0.001$ ) or high on justice sensitivity ( $r=0.34, p=0.0467$ ) assigned greater blame for bad actions. Interestingly, affective empathy was not significantly related to blame ratings ( $p > 0.7$ ).

Justice sensitivity, affective empathy, and cognitive empathy did not significantly predict praise ratings of good action ( $p > 0.1$ ). Justice sensitivity was significantly positively correlated with both cognitive empathy ( $r=0.48$ ,  $p=0.003$ ) and affective empathy ( $r=0.39$ ,  $p=0.021$ ).

### 3.3. Moral valence

Following the onset of the first picture (Fig. 2A and C), both morally good and morally bad actions were associated with a biphasic deflection that was positive at posterior sites, and negative at frontal sites. N1 mean amplitude was significantly greater for good actions at the posterior cluster ( $t(34)=-3.66$ ,  $p=0.007$ ). There was no N1 amplitude difference at the frontal cluster after correction for multiple comparisons ( $t(34)=2.35$ ,  $p > 0.2$ ). The frontal N2 was significantly more negative for morally good actions than morally bad actions ( $t(34)=3.31$ ,  $p=0.018$ ). Mean frontal LPP was significantly greater for morally good actions ( $t(34)=2.96$ ,  $p=0.044$ ), but there was no significant difference at the posterior cluster ( $p > 0.16$ ). There was no significant difference in mean amplitude during late slow wave at either site (both  $p > 0.4$ ).

For the PCA analysis, 16 factor combinations accounted for at least one percent of the overall variance. Visual inspection revealed a single factor whose activity and spatial distribution corresponded to established ERP components and occurred during the presentation of the first picture (Dien, Beal, & Berg, 2005). This factor, Temporal Factor 1, Spatial Factor 1 (TF1/SF1) was maximal at electrode 90 and occurred 368–372 ms after the onset of the

first picture. This factor was significantly greater for good scenes ( $t(34)=-2.92$ ,  $p=0.012$ ), and the magnitude of this difference (bad–good) was significantly correlated with cognitive empathy ( $r=0.37$ ,  $p=0.027$ ; Fig. 3C) and blame ratings ( $r=0.43$ ,  $p=0.002$ ; Fig. 3B). Justice sensitivity was not significantly related to TF1/SF1 amplitude differences.

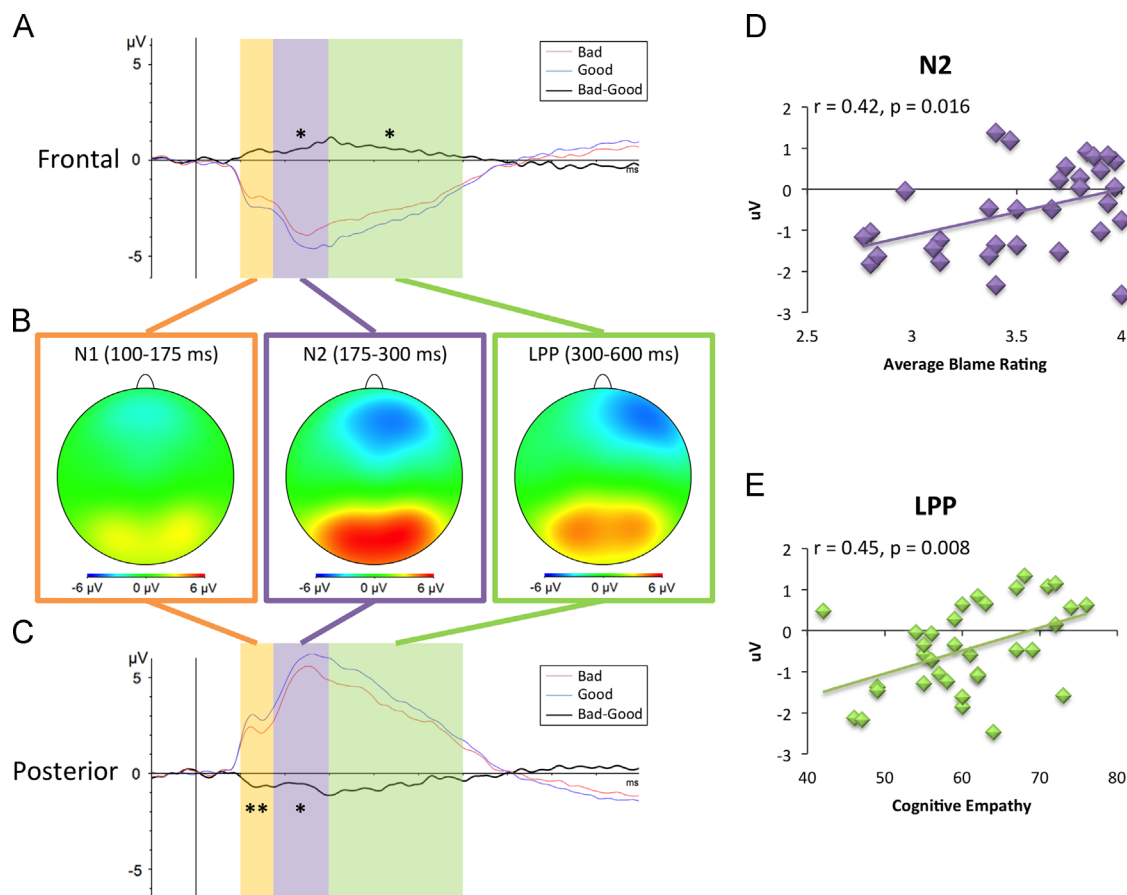
Subjective moral evaluations of the scenarios indicated that blame ratings were significantly correlated with the N2 amplitude of the bad–good difference wave ( $r=0.42$ ,  $p=0.016$ ; Fig. 2D). Cognitive empathy scores, but not affective empathy ( $p > 0.8$ ) were significantly correlated with differential LPP amplitudes ( $r=0.45$ ,  $p=0.008$ ; Fig. 2E). Justice sensitivity and subjective praise ratings were not significantly correlated with any amplitude differences (all  $p > 0.2$ ).

### 3.4. Alpha suppression

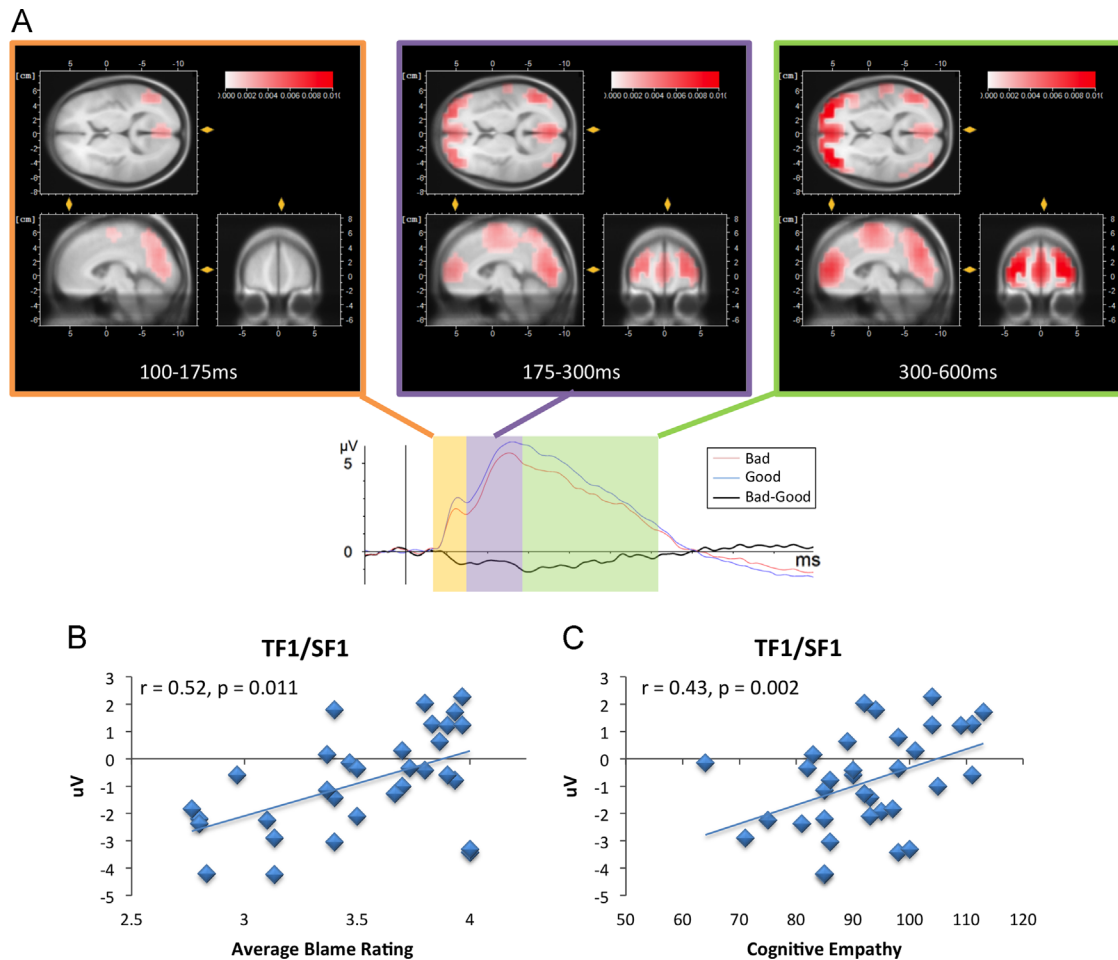
A significant decrease in log alpha power was observed at all sites (all  $p < 0.05$ ). The reduction at central sites was not significantly different from the reduction observed at the frontal and posterior clusters ( $p > 0.6$ ), indicating a global reduction in alpha, rather than centrally-localized reduction as seen in mu-suppression. Alpha suppression did not differ between morally good and morally bad actions ( $p > 0.6$ ).

### 3.5. LORETA

Source estimation was conducted within each time window (Fig. 3A). During the first two windows, the maximal source was



**Fig. 2.** (A, C) ERP traces for bad (antisocial) (red), good (prosocial) (blue), and bad–good difference waves (black). \*  $p < 0.05$ ; \*\*  $p < 0.01$  (Bonferroni corrected). (B) Scalp plots for the grand average within specific time windows. (D) Mean amplitude of the frontal N2 bad–good difference was significantly correlated with average blame ratings. (E) Mean amplitude of the bad–good difference wave for frontal LPP was significantly correlated with cognitive empathy scores. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 3.** (A) LORETA source estimation for the grand average ERP during the three time windows that showed significant differences between bad and good morally laden actions. The ERP from the posterior cluster is reproduced for reference. The peak amplitude of the bad–good difference wave for Temporal Factor 1, Spatial Factor 1 (TF1/SF1) was significantly correlated with average blame ratings (B) and cognitive empathy scores (C).

located in right parietal cortex. During the early and later slow wave window, the maximal source occurred in vmPFC.

#### 4. Discussion

This study was designed to examine the spatiotemporal neural dynamics of moral judgments of everyday actions, an important aspect to illuminate theories of moral cognition. Our results provide compelling evidence that electrophysiological indices of both automatic and controlled processes are influenced by moral valence. Moreover, dispositional differences in cognitive empathy predict neural differentiation during moral evaluations, but only after 300 ms. Consistent with a growing body of recent literature (Lahat et al., 2013; Sarlo et al., 2012), these findings support a complex conception of morality involving early differentiation between prosocial and antisocial actions, followed by a secondary reappraisal of these actions. It is this later reappraisal that is heavily impacted by individual dispositions.

Soon after the appearance of morally laden scenes, participants showed a rapid distinction between bad and good social interactions as evidenced by differential N1 amplitudes (Fig. 2). Consistent with previous literature, the N1 amplitude was greater for good actions than bad (Keil et al., 2002), perhaps because participants found the prosocial scenarios more pleasant and rewarding than the antisocial scenarios (Decety & Porges, 2011; Knutson, Taylor, Kaufman, Peterson, & Glover, 2005; Yoder &

Decety, 2014). Similarly, N2 amplitudes were greater for prosocial scenes. Moreover, this N2 differentiation was significantly negatively correlated with subjective blame ratings, indicating that those individuals who assigned more blame showed more similar N2 amplitudes for morally good and morally bad actions (Fig. 2). Further, the temporospatial PCA revealed an occipitally maximal factor whose sensitivity to moral valence was also related to blame severity and dispositional cognitive empathy (Fig. 3B and C). From a conflict-monitoring standpoint, this result would suggest that participants expected to see antisocial actions and so found the good actions novel (Botvinick et al., 2004). A less grim interpretation of these findings is that participants in our study found prosocial actions more personally relevant. Interestingly, previous studies have shown that N2 peaks arise in anticipation of others' in pain (Chen et al., 2012; Cheng et al., 2012). Thus, N2 in this study may also reflect additional attentional resources devoted to processing the anticipation of prosocial outcomes as compared to antisocial outcomes.

Prosocial (good) actions were also associated with greater LPP amplitude than antisocial (bad) actions (Fig. 2). This finding is similar to previous work showing enhanced positivity for rewards in a gambling task (Hajcak, Holroyd, Moser, & Simons, 2005), and supports the hypothesis that actions with positive outcomes are more personally meaningful. Further, this later differentiation occurred during the time when fine-grained distinctions and deliberative processes begin to influence ERP amplitudes (Olofsson et al., 2008; Weinberg & Hajcak, 2010). Here, participants with higher cognitive empathy scores

showed reduced LPP amplitude in the bad–good difference wave (Fig. 2E). This suggests that participants who have high dispositional cognitive empathy, which are the same individuals that assign more blame to harmful perpetrators, may spend more time dwelling on the antisocial actions. Conversely, this could suggest that individuals low on trait empathy might perform this task by relying on more basic discrimination processes, and so engage in less reflection about the antisocial actions.

Source estimation was used to localize possible neural generators for the observed ERPs. LORETA results are estimates of the neural generators rather than veridical voxel coordinates, so they are not ideal for making strong claims about which exact cortical regions are involved (Polich, 2007). However, when combined with other studies using more spatially precise neuroimaging methods, source estimation can be useful in integrating spatial and temporal neuroimaging findings. Here, LORETA results conform with previous functional MRI data using the same scenarios, and show a general posterior to frontal shift (Yoder & Decety, 2014). This shift was characterized by an initial recruitment of superior parietal regions, followed by widespread recruitment of prefrontal regions such as mPFC, ACC, and dlPFC (Fig. 3A). Specifically, early (< 300 ms) voltages were localized to pSTS/rTPJ which fits with the established role of rTPJ in moral judgments associated with mental state understanding (Decety & Cacioppo, 2012; Saxe, Xiao, Kovacs, Perrett, & Kanwisher, 2004; Yoder & Decety, 2014; Young & Saxe, 2009).

The LORETA analysis suggested that the peak current for LPP was driven by frontal regions (corresponding to vmPFC, ACC and dlPFC). This result matches, on a different time scale, the effects seen in several fMRI studies, where superior parietal regions encode and integrate moral information before passing this information to prefrontal regions for response selection (Buckholz & Marois, 2012). These findings, along with the N2 results, fit well with the proposed role of the ACC in coordinating the relative recruitment of default and cognitive control networks in service of moral judgment (Chiong et al., 2013), and provide further support for the hypothesis that prosocial actions may be more personally salient than antisocial actions.

Our results demonstrate that the encoding of morally laden stimuli taps into relatively early and automatic processing systems as well as later controlled elaborative processes. Interestingly, while dispositions in cognitive empathy predicted LPP amplitudes, there was no evidence for an influence of individual differences in affective or cognitive empathy on earlier processing. Moreover, dispositional justice sensitivity, while predictive of subjective blame ratings, was not significantly related to any ERP measure. Given that justice sensitivity has previously been linked to neurohemodynamic activity within pSTS/rTPJ and prefrontal regions involved in mental state understanding and maintaining goal representations (Yoder & Decety, 2014), these results suggest that the early processes that give rise to moral judgments may be the locus at which dispositions in empathy influence moral evaluations, rather than at later processing stages where justice sensitivity is more important. Moreover, this pattern of results reflects the importance of considering that different neuroimaging methodologies may not be equally sensitive to individual differences on personality traits. For some dispositions, fMRI may be better suited because the influences of individual differences require more time to unfold, making EEG not as precise to measure such influences.

Finally, there was no evidence for mu-suppression playing a role in the neural encoding of moral valence (see Cheng et al., 2012). Instead, the reduction in global alpha power suggests that individuals were equally engaged in the task. Moreover, morally bad and morally good actions were associated with similar decreases in alpha power, suggesting that the observed ERP differences are not due to general differences in attention between them.

Taken together, this study provides important and new insights into the spatiotemporal neural dynamics of moral judgment. Interpersonal harm and assistance, which serve as fundamental prototypes of morally bad and good actions, are rapidly encoded. Moreover, prosocial actions, as compared to antisocial actions, are perceived as more personally salient, as evidenced by greater amplitudes for prosocial actions in both relatively automatic components, such as N1 and N2, and later components, such as LPP, which are thought to reflect cognitive re-appraisal. Finally, individual differences in cognitive empathy did not influence moral valence encoding until after 300 ms, suggesting that the influence of empathy on moral judgment comes from differences in deliberative reasoning, rather than enhanced early sensitivity to harm.

## Acknowledgments

This study was supported by National Institute of Mental Health R01 Grant (1R01-MH084934) and grants from the John Templeton Foundation (The Science of Philanthropy Initiative and Wisdom Research) to JD. We thank Jason M. Cowell for his advice on data analysis and interpretation.

## References

- Baumert, A., Rothmund, T., Thomas, N., Gollwitzer, M., & Schmitt, M. (2013). Justice as a moral motive: belief in a just world and justice sensitivity as potential indicator of the justice motive. In: K. Heinrichs, F. Oser, & T. Lovat (Eds.), *Handbook of moral motivation* (pp. 159–180). Rotterdam: Sense Publishers.
- Boksem, M. A. S., & De Cremer, D. (2010). Fairness concerns predict medial frontal negativity amplitude in ultimatum bargaining. *Social Neuroscience*, 5(1), 118–128. <http://dx.doi.org/10.1080/17470910903202666>.
- Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Sciences*, 8(12), 539–546. <http://dx.doi.org/10.1016/j.tics.2004.10.003>.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10(4), 433–436.
- Buckholz, J. W., & Marois, R. (2012). The roots of modern justice: cognitive and neural foundations of social norms and their enforcement. *Nature Neuroscience*, 15(5), 655–661. <http://dx.doi.org/10.1038/nn.3087>.
- Cacioppo, J. T., Crites, S. L., & Gardner, W. L. (1996). Attitudes to the right: evaluative processing is associated with lateralized late positive event-related brain potentials. *Personality and Social Psychology Bulletin*, 22(12), 1205–1219. <http://dx.doi.org/10.1177/01461672962212002>.
- Chen, C., Yang, C.-Y., & Cheng, Y. (2012). Sensorimotor resonance is an outcome but not a platform to anticipating harm to others. *Social Neuroscience*, 7(6), 578–590. <http://dx.doi.org/10.1080/17470919.2012.686924>.
- Cheng, Y., Hung, A.-Y., & Decety, J. (2012). Dissociation between affective sharing and emotion understanding in juvenile psychopaths. *Development and Psychopathology*, 24(2), 623–636. <http://dx.doi.org/10.1017/S095457941200020X>.
- Chiong, W., Wilson, S. M., D'Esposito, M., Kayser, A. S., Grossman, S. N., Poorzand, P., et al. (2013). The salience network causally influences default mode network activity during moral reasoning. *Brain: A Journal of Neurology*, 136(6), 1929–1941. <http://dx.doi.org/10.1093/brain/awt066>.
- Chiu Loke, I., Evans, A. D., & Lee, K. (2011). The neural correlates of reasoning about prosocial-helping decisions: an event-related brain potentials study. *Brain Research*, 1369, 140–148. <http://dx.doi.org/10.1016/j.brainres.2010.10.109>.
- Decety, J., & Cacioppo, S. (2012). The speed of morality: a high-density electrical neuroimaging study. *Journal of Neurophysiology*, 108(11), 3068–3072. <http://dx.doi.org/10.1152/jn.00473.2012>.
- Decety, J., & Cowell, J. M. (2014). The complex relationship between morality and empathy. *Trends in Cognitive Sciences* (In press).
- Decety, J., & Howard, L. H. (2013). The role of affect in the neurodevelopment of morality. *Child Development Perspectives*, 7(1), 49–54. <http://dx.doi.org/10.1111/cdep.12020>.
- Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews*, 3, 71–100. <http://dx.doi.org/10.1177/1534582304267187>.
- Decety, J., Michalska, K. J., & Kinzler, K. D. (2012). The contribution of emotion and cognition to moral sensitivity: a neurodevelopmental study. *Cerebral Cortex*, 22(1), 209–220. <http://dx.doi.org/10.1093/cercor/bhr111>.
- Decety, J., & Porges, E. C. (2011). Imagining being the agent of actions that carry different moral consequences: an fMRI study. *Neuropsychologia*, 49(11), 2994–3001. <http://dx.doi.org/10.1016/j.neuropsychologia.2011.06.024>.
- Decety, J., & Svetlova, M. (2012). Putting together phylogenetic and ontogenetic perspectives on empathy. *Developmental Cognitive Neuroscience*, 2(1), 1–24. <http://dx.doi.org/10.1016/j.dcn.2011.05.003>.

- Decety, J., Yang, C.-Y., & Cheng, Y. (2010). Physicians down-regulate their pain empathy response: an event-related brain potential study. *NeuroImage*, 50(4), 1676–1682, <http://dx.doi.org/10.1016/j.neuroimage.2010.01.025>.
- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9–21, <http://dx.doi.org/10.1016/j.jneumeth.2003.10.009>.
- Dien, J. (2010). The ERP PCA Toolkit: an open source program for advanced statistical analysis of event-related potential data. *Journal of Neuroscience Methods*, 187(1), 138–145, <http://dx.doi.org/10.1016/j.jneumeth.2009.12.009>.
- Dien, J., Beal, D. J., & Berg, P. (2005). Optimizing principal components analysis of event-related potentials: matrix type, factor loading weighting, extraction, and rotations. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 116(8), 1808–1825, <http://dx.doi.org/10.1016/j.clinph.2004.11.025>.
- Dien, J., Khoe, W., & Mangun, G. R. (2007). Evaluation of PCA and ICA of simulated ERPs: promax vs. infomax rotations. *Human Brain Mapping*, 28(8), 742–763, <http://dx.doi.org/10.1002/hbm.20304>.
- Fan, Y., & Han, S. (2008). Temporal dynamic of neural mechanisms involved in empathy for pain: an event-related brain potential study. *Neuropsychologia*, 46(1), 160–173, <http://dx.doi.org/10.1016/j.neuropsychologia.2007.07.023>.
- Folstein, J., & Petten, C. Van. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology*, 45(1), 152–170.
- Fumagalli, M., & Priori, A. (2012). Functional and clinical neuroanatomy of morality. *Brain: A Journal of Neurology*, 135(7), 2006–2021, <http://dx.doi.org/10.1093/brain/awr334>.
- Hajcak, G., Holroyd, C. B., Moser, J. S., & Simons, R. F. (2005). Brain potentials associated with expected and unexpected good and bad outcomes. *Psychophysiology*, 42(2), 161–170, <http://dx.doi.org/10.1111/j.1469-8986.2005.00278.x>.
- Hajcak, G., MacNamara, A., & Olivet, D. M. (2010). Event-related potentials, emotion, and emotion regulation: an integrative review. *Developmental Neuropsychology*, 35(2), 129–155, <http://dx.doi.org/10.1080/87565640903526504>.
- Hajcak, G., Moser, J. S., Yeung, N., & Simons, R. F. (2005). On the ERN and the significance of errors. *Psychophysiology*, 42(2), 151–160, <http://dx.doi.org/10.1111/j.1469-8986.2005.00270.x>.
- Hamlin, J. K. (2014). The origins of human morality: complex socio-moral evaluations by preverbal infants. In: J. Decety, & Y. Christen (Eds.), *New frontiers in social neuroscience* (pp. 165–188). New York, NY: Springer.
- Han, S., Fan, Y., & Mao, L. (2008). Gender difference in empathy for pain: an electrophysiological investigation. *Brain Research*, 1196, 85–93, <http://dx.doi.org/10.1016/j.brainres.2007.12.062>.
- Harsay, H. A., Spaan, M., Wijnen, J. G., & Ridderinkhof, K. R. (2012). Error awareness and salience processing in the oddball task: shared neural mechanisms. *Frontiers in Human Neuroscience*, 6, 246, <http://dx.doi.org/10.3389/fnhum.2012.00246>.
- Helwig, C. C., & Turiel, E. (2011). Children's social and moral reasoning. In: C. Hart, & P. Smith (Eds.), *Wiley-Blackwell handbook of childhood social development, Revised* (2nd ed.). Hoboken, NJ: Wiley-Blackwell.
- Huebner, B., Dwyer, S., & Hauser, M. (2009). The role of emotion in moral psychology. *Trends in Cognitive Sciences*, 13(1), 1–6, <http://dx.doi.org/10.1016/j.tics.2008.09.006>.
- Keil, A., Bradley, M. M., Hauk, O., Rockstroh, B., Elbert, T., & Lang, P. J. (2002). Large-scale neural correlates of affective picture processing. *Psychophysiology*, 39(5), 641–649 (doi:10.1017/S0048577202394162).
- Knutson, B., Taylor, J., Kaufman, M., Peterson, R., & Glover, G. (2005). Distributed neural representation of expected value. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 25(19), 4806–4812, <http://dx.doi.org/10.1523/JNEUROSCI.0642-05.2005>.
- Lahat, A., Helwig, C. C., & Zelazo, P. D. (2013). An event-related potential study of adolescents' and young adults' judgments of moral and social conventional violations. *Child Development*, 84(3), 955–969, <http://dx.doi.org/10.1111/cdev.12001>.
- Macnamara, A., Foti, D., & Hajcak, G. (2009). Tell me about it: neural activity elicited by emotional pictures and preceding descriptions. *Emotion (Washington, D.C.)*, 9(4), 531–543, <http://dx.doi.org/10.1037/a0016251>.
- Malle, B. F., & Guglielmo, S. (2012). Are intentionality judgments fundamentally moral? In: R. Langdon, & C. Mackenzie (Eds.), *Emotions, imagination, and moral reasoning* (pp. 275–293). New York, NY: Psychology Press.
- Mendez, M. (2009). The neurobiology of moral behavior: review and neuropsychiatric implications. *CNS Spectrums*, 14(11), 608–620.
- Moll, J., de Oliveira-Souza, R., & Eslinger, P. (2003). Morals and the human brain: a working model. *Neuroreport*, 14(3), 299–305, <http://dx.doi.org/10.1097/01.wnr.0000057866.05120.28>.
- Moll, J., de Oliveira-Souza, R., Garrido, G. J., Bramati, I. E., Caparelli-Daquer, E. M. a., Paiva, M. L. M. F., et al. (2007). The self as a moral agent: linking the neural bases of social agency and moral sensitivity. *Social Neuroscience*, 2(3–4), 336–352, <http://dx.doi.org/10.1080/17470910701392024>.
- Olofsson, J. K., Nordin, S., Sequeira, H., & Polich, J. (2008). Affective picture processing: an integrative review of ERP findings. *Biological Psychology*, 77(3), 247–265, <http://dx.doi.org/10.1016/j.biopsycho.2007.11.006>.
- Perry, A., Bentin, S., Bartal, I. B.-A., Lamm, C., & Decety, J. (2010). "Feeling" the pain of those who are different from us: modulation of EEG in the mu/alpha range. *Cognitive, Affective & Behavioral Neuroscience*, 10(4), 493–504, <http://dx.doi.org/10.3758/CABN.10.4.493>.
- Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 118(10), 2128–2148, <http://dx.doi.org/10.1016/j.clinph.2007.04.019>.
- Prehn, K., & Heekeren, H. R. (2009). Moral judgment and the brain: a functional approach to the question of emotion and cognition in moral judgment integrating psychology, neuroscience and evolutionary biology. In: J. Versplaatse, J. De Schrijver, S. Vanneste, & J. Braeckman (Eds.), *The moral brain* (pp. 129–154). New York, NY: Springer.
- Reniers, R. L. E. P., Corcoran, R., Drake, R., Shryane, N. M., & Völlm, B. a. (2011). The QCAE: a questionnaire of cognitive and affective empathy. *Journal of Personality Assessment*, 93(1), 84–95, <http://dx.doi.org/10.1080/00223891.2010.528484>.
- Sarlo, M., Lotto, L., Manfrinati, A., Rumiati, R., Gallicchio, G., & Palomba, D. (2012). Temporal dynamics of cognitive-emotional interplay in moral decision-making. *Journal of Cognitive Neuroscience*, 24(4), 1018–1029, [http://dx.doi.org/10.1162/jocn\\_a.00146](http://dx.doi.org/10.1162/jocn_a.00146).
- Saxe, R., Xiao, D.-K., Kovacs, G., Perrett, D. I., & Kanwisher, N. (2004). A region of right posterior superior temporal sulcus responds to observed intentional actions. *Neuropsychologia*, 42(11), 1435–1446, <http://dx.doi.org/10.1016/j.neuropsychologia.2004.04.015>.
- Schmitt, M., Baumert, A., Gollwitzer, M., & Maes, J. (2010). The justice sensitivity inventory: factorial validity, location in the personality facet space, demographic pattern, and normative data. *Social Justice Research*, 23(2–3), 211–238, <http://dx.doi.org/10.1007/s11211-010-0115-2>.
- Shenhav, A., & Greene, J. (2010). Moral judgments recruit domain-general valuation mechanisms to integrate representations of probability and magnitude. *Neuron*, 67(4), 667–677, <http://dx.doi.org/10.1016/j.neuron.2010.07.020>.
- Sousa, P., Holbrook, C., & Piazza, J. (2009). The morality of harm. *Cognition*, 113(1), 80–92, <http://dx.doi.org/10.1016/j.cognition.2009.06.015>.
- Turiel, E. (2008). Thought about actions in social domains: morality, social conventions, and social interactions. *Cognitive Development*, 23(1), 136–154, <http://dx.doi.org/10.1016/j.cogdev.2007.04.001>.
- Weinberg, A., & Hajcak, G. (2010). Beyond good and evil: the time-course of neural activity elicited by specific picture content. *Emotion (Washington, D.C.)*, 10(6), 767–782, <http://dx.doi.org/10.1037/a0020242>.
- Wu, Y., Leiveld, M. C., & Zhou, X. (2011). Social distance modulates recipient's fairness consideration in the dictator game: an ERP study. *Biological Psychology*, 88(2–3), 253–262, <http://dx.doi.org/10.1016/j.biopsycho.2011.08.009>.
- Yoder, K. J., & Decety, J. (2014). *The good, the bad, and the just: justice sensitivity predicts neural response during moral evaluation of actions performed by others*, 34(12), 4161–4166, <http://dx.doi.org/10.1523/JNEUROSCI.4648-13.2014>.
- Young, L., & Dungan, J. (2012). Where in the brain is morality? Everywhere and maybe nowhere. *Social Neuroscience*, 7(1), 1–10.
- Young, L., & Saxe, R. (2009). An fMRI investigation of spontaneous mental state inference for moral judgment. *Journal of Cognitive Neuroscience*, 21(7), 1396–1405, <http://dx.doi.org/10.1162/jocn.2009.21137>.