The speed of morality: a high-density electrical neuroimaging study

Jean Decety1,2 and Stephanie Cacioppo1,3
1Department of Psychology, The University of Chicago, Chicago, Illinois; 2Department of Psychiatry and Behavioral Neuroscience, University of Chicago, Chicago, Illinois; and 3Department of Psychology, University of Geneva, Geneva, Switzerland

Decety J, Cacioppo S. The speed of morality: a high-density electrical neuroimaging study. J Neurophysiol 108: 3068–3072, 2012. First published September 5, 2012; doi:10.1152/jn.00473.2012.—Neuroscience research indicates that moral reasoning is underpinned by distinct neural networks including the posterior superior temporal sulcus (pSTS), amygdala, and ventromedial prefrontal cortex, which support communication between computational systems underlying affective states, cognitions, and motivational processes. To characterize real-time neural processing underpinning moral computations, high-density event-related potentials were measured in participants while they viewed short, morally laden visual scenarios depicting intentional and accidental harmful actions. Current source density maxima in the right pSTS as fast as 62 ms poststimulus first distinguished intentional vs. accidental actions. Responses in the amygdala/temporal pole (122 ms) and ventromedial prefrontal cortex (182 ms) were then evoked by the perception of harmful actions, indicative of fast information processing associated with early stages of moral cognition. Our data strongly support the notion that intentionality is the first input to moral computations. They also demonstrate that emotion acts as a gain antecedent to moral judgment by alerting the individual to the moral salience of a situation and provide evidence for the pervasive role of affect in moral sensitivity and reasoning.

MORAL DECISION MAKING is a fundamental aspect of social cognition and is seen as a product of our biological evolutionary and cultural history. In the past decade, research in multiple academic domains including evolutionary biology, developmental science, economics, and cognitive neuroscience has endeavored to more clearly define and investigate this construct. Results from these disciplines suggest that mature moral abilities emerge from a sophisticated integration of cognitive, emotional, and motivational mechanisms (Decety and Howard 2012). For instance, developmental research indicates that very early in life (around 3 mo of age), preverbal babies express what appear to be nascent moral evaluations (Hamlin et al. 2011; Decety and Porges 2011; Greene et al. 2001; Heekeren et al. 2009). Subsequent studies from patients with such lesions support the critical role of this region in moral decision making and moral behavior (e.g., Gleichgerrcht et al. 2011; Koenigs et al. 2007; Moretto et al. 2010). Interestingly, impairments of the ventromedial prefrontal cortex (vmPFC) are not associated deficits in mental states understanding.

In the past decade, functional magnetic resonance imaging (fMRI) has begun to identify a network of brain regions involved in moral cognition (see Moll and Schultkin 2009; Young and Dungan 2012 for recent reviews). These studies indicate that a restricted number of regions are involved in moral reasoning, including the posterior superior temporal sulcus [pSTS, also reported in the literature as the temporoparietal junction (TPJ)], amygdala, and vmPFC (Borg Schaich et al. 2011; Decety and Porges 2011; Greene et al. 2001; Heekeren et al. 2003; Moll et al. 2002, 2007). Moral reasoning seems to be underpinned by specific neural circuitry, but, in fact, these circuits are not unique to morality; rather, they involve regions and systems underlying specific affective states and cognitive and motivational processes.

One serious limitation of a majority of previous neuroimaging studies is that the results rely only on subtraction logic in their designs, which is based on the a priori assumption that one computational process can be added to a preexisting set of processes without affecting them, assuming that there are no interactions among the different components of a given task. Furthermore, characterizing brain activity in terms of functionally segregated regions does not reveal anything about how different brain regions communicate with each other. Connectivity analyses and high-density event-related potentials (ERPs) can identify patterns of communication between regions that contrast analyses may not detect. Such methods are necessary to advance knowledge on the neuroscience of morality. This is also important at a theoretical level. The fact that fMRI studies found activation in emotion-related areas such as the amygdala and vmPFC during morally salient stimuli provides only correlational evidence, showing that emotions are associated with moral cognition, but is insufficient to determine whether affective processing is taking place during moral evaluations or antecedent to them (Huebner et al. 2008). In addition, there is no information on the timing and order of component processes involved in moral judgment.

In the current study, we took advantage of high-density ERPs and analysis of source localizations to examine the spatiotemporal dynamics of the neural processing evoked by...
the perception of visual morally laden scenarios in healthy individuals. We used a well-established paradigm that was validated with fMRI and eye-tracking measures in a large number of participants between ages 4 and 37 yr (Decety et al. 2012). The results of that study showed that the perception of intentional harm vs. accidental harm was associated with increased signal in the amygdala, periaqueductal gray, insula, vmPFC, and right pSTS/TPJ. High-density ERPs can identify patterns of communication between regions that contrast analyses may not detect, and such methods are necessary to advance of knowledge of the neuroscience of morality.

METHODS

Participants

A total of 10 healthy participants were involved in the study. All provided written informed consent to participate in the experiment, which was approved by the local Committee for Protection of Human Subjects. Three volunteers had to be excluded due to artifacts in their EEG data. Thus brain activity from seven (4 females, 3 males) volunteers was fully analyzed. All participants were right-handed, older than 18 yr (mean age: 21.86 yr; SD: 3.13 yr), native English speakers with normal or corrected-to-normal vision, and were not taking antidepressant medication. None of the participants had prior or current neurological or psychiatric disorders (e.g., traumatic brain injury with loss of consciousness, epilepsy, neurological impairment, or degenerative neurological illness), as ascertained by a detailed anamnesis.

Behavioral Procedure

Participants completed a modified version of a standard Intention Inference Task (IIT), developed by Decety et al. (2012) in studies on empathy and morality, according to a 2 × 2 factorial design with intention type (intentional vs. accidental) and target type (object vs. person) as within-subject factors. While their electrical brain activity was recorded, participants were required to watch the stimuli, to gaze at the center of the screen, and to judge whether the action was performed intentionally or accidentally. During this part of the experiment, no reaction times were collected to avoid any motor artifacts. After the completion of the EEG recordings, one additional behavioral block was run during which accuracy and reaction times were recorded.

Behavioral paradigm

During the IIT, participants watched a series of three-frame video clips that were presented centrally on a monitor screen, as in a previous fMRI study (Decety et al. 2012). The first frame (T1) from the video clip was 500 ms long and displayed an establishing scene; the second frame (T2) was a 700-ms frame displaying either an intentional harm or an accidental harm, followed by a third 1,000-ms frame (T3) confirming the intentional or accidental harm. The three-frame video-clip technique provided a tight control of the kinematics, task duration, and timing. Each trial began with a fixation cross that was presented for 150 ms. A 1,500- to 2,000-ms maximum inter-trial interval was randomly inserted.

Instructions

Participants were instructed to gaze at the center of the screen and to judge whether the action was performed intentionally or accidentally. In addition, participants were asked to refrain from blinking or moving their eyes except during the interval between trials.

Electrophysiological Recordings

Continuous EEG was recorded from 128 AgCl carbon fiber-coated electrodes using an Electrical Geodesic Sensor Net (GSN300; Electrical Geodesic, Eugene, OR; http://www.egi.com/), where EEG electrodes are arrayed in a dense and regular distribution across the head surface with an inter-sensor distance of ~3 cm. The EEG was digitized at 500 Hz (corresponding to a sample bin of 2 ms), bandwidth at 0.01–200 Hz, with the vertex electrode (Cz) serving as an on-line recording reference; impedances were kept below 50 kΩ. Participants were seated in a comfortable chair 150 cm away from a personal computer screen on which stimuli were presented centrally.

Electrophysiological Data Processing

Electrophysiological data were imported and analyzed in Cartool (version 3.32; Denis Brunet, http://brainmapping.unige.ch/Cartool.htm). First, epochs of analysis were visually inspected for oculomotor (saccades and blinks), muscles, and other artifacts in addition to an automated threshold rejection criterion of 100 mV.

After off-line artifact rejections, visual evoked potential (VEPs) were computed from 0 to 400 ms after the onset of T2 (i.e., the second frame displaying either an intentional harmful action or an accidental harmful action). VEP data were then bandpass filtered between 1 and 30 Hz without baseline correction. VEP data were next recalculated off-line against the average reference and normalized to their mean global field power (i.e., GFP) before group-averaging. The GFP is computed as the spatial standard deviation of the scalp electric field, yields larger values for stronger electric fields, and is calculated as the square root of the mean of the squared value recorded at each electrode (vs. the average reference). Channels with corrupted signals and channels showing substantial artifacts and noise throughout the recordings were interpolated to a standard 111-channel electrode array using a three-dimensional spline procedure.

Second-Level Electrical Data Analysis

Topographical analyses. VEP data were analyzed with space-oriented brain electric field analysis using Cartool. Because this electrophysiological method has been extensively detailed previously (see Brunet et al. 2011), here we provide only the essential details. This space-oriented brain electric field approach is based on the empirical observation that the brain electric field configuration changes stepwise over time. Epochs of quasi-stable field configurations (“microstates”) are concatenated by abrupt transitions in the brain electric field configurations. Microstates thus are assumed to implement specific brain functions. To identify start and end of each optimal microstate, a standard cluster analysis previously described was employed using the grand-mean ERPs of each condition. This cluster analysis uses a hierarchical agglomerative cluster algorithm to identify the predominant topographies (i.e., maps) and their sequence within a data set (these methods are implemented in Cartool). The optimal number of maps (i.e., the minimal number of maps that accounts for the greatest variance of the data set) is determined on the basis of a modified Krzanowski-Lai criterion. Importantly, this pattern analysis is reference free and insensitive to amplitude modulation of the same scalp potential field across conditions, since normalized maps are compared. We also applied the constraint that a given scalp topography must be observed for at least five consecutive data points (i.e., 10 ms at a 500-Hz digitization rate) in the group-averaged data. This criterion is effectively similar to that frequently applied in the analysis of VEP waveform modulations. This pattern analysis was performed across time and experimental conditions to determine whether and when different conditions recruited different configurations of intracranial generators. This conservative approach allowed us to analyze the VEP’s responses for intentional vs. accidental harmful actions. Next, the pattern of maps observed in the group-
averaged data was statistically tested by comparing each of these maps with the moment-by-moment scalp topography of individual subjects’ VEP from each condition. To do so, each time point of each VEP from each subject was labeled according to the map with which it best correlated spatially. In other words, the optimal number of maps was fitted into the original data for each individual subject, using a competitive fitting procedure. This “fitting” procedure determines whether a given experimental condition is more often described by one map versus another. From this “fitting” procedure, a large amount of information can be extracted and analyzed for a given condition across subjects, such as the total amount of time for a given stable configuration. This latter value represents the frequency with which a given microstate was observed within a given time period for each experimental condition. This is the information used here. The extracted values of interest were then subjected to a repeated-measure ANOVA. Results were accepted as significant at \( P < 0.05 \).

The hierarchical cluster analysis of the VEP topographies in this “action type” condition revealed a total of six time periods of stability that significantly differed in both timing and location for intentional harm and accidental harm in the 400-ms poststimulus period that explained 90.14% of variance in the collective data set (Fig. 1). The windows of occurrence for these stable time periods corresponded to the following time intervals: microstate 1 (M1): 0–20 ms; microstate 2 (M2): 22–60 ms; microstate 3 (M3): 62–120 ms for intentional actions and 62–140 ms for accidental actions; microstate 4 (M4): 122–180 ms for intentional actions and 142–180 ms for accidental actions; microstate 5 (M5): 182–274 ms for intentional actions and 182–304 ms for accidental actions; and microstate 6 (M6): 276–400 ms for intentional actions and 306–400 ms for accidental actions (see Table 1).

Distributed intracranial source estimations. The intracranial generators for each condition were estimated using a distributed linear inverse solution based on a low-resolution brain electromagnetic tomography (LORETA) model of the unknown current density in the brain. Here, LORETA was used with a lead field (solution space) that was calculated on a realistic head model that included 3,005 solution points, selected from a 6 × 6 × 6 mm grid equally distributed within the gray matter. Source estimations were rendered on the MNI/McGill average standard brain as supplied by Cartool. Accuracy of anatomical labeling was ascertained with a visual inspection of the Duvernoy (1991) brain atlas. The maximum current density for each microstate was the following: M1: 344, −7 (xyz Talairach coordinates); M2: 52, −60, 23; M3: 52, −60, 21; M4: 32, 0, −30; M5: 3, 38, −7; and M6: 3, 38, −7.

### RESULTS AND DISCUSSION

To determine the timing and order of component processes implicated in moral cognition and whether affective processing occurs during moral evaluations or antecedent to them, we used high-density ERPs to examine the spatiotemporal dynamics of the neural processing evoked by the perception of visual morally laden scenarios. The perception of intentional harm was associated with better (90% vs. 71%) and faster reaction times for intentional harm compared with accidental harm (\( P < 0.05 \)), as well as the specific involvement of the right pSTS, amygdala/temporal pole, and vmPFC. These regions were found activated in an fMRI study using the same stimuli and same contrast (Decety et al. 2012). Interestingly, intentional harmful actions were significantly distinguished from the accidental harmful actions in three main time periods (i.e., from 62 to 140 ms and from 122 to 180 ms poststimulus, see Fig. 1). More precisely, a specific scalp potential field with a current source density maximum in the right pSTS (52, 60, 21; xyz Talairach coordinates) characterized accidental harm during the first time period, whereas a different scalp topography with a current source density in the right amygdala (32, 0, −30; xyz Talairach coordinates from 122 to 180 ms) and vmPFC (3, 38, −7; xyz Talairach coordinates from 182 to 304 ms) characterized the perception of intentional harm (Table 1).

### Table 1. Local maxima of current source density obtained from LORETA brain source estimations of EEG data for intentional harmful actions and accidental harmful actions

<table>
<thead>
<tr>
<th>Microstate Time Periods for Intentional Harmful Actions</th>
<th>Microstate Time Periods for Accidental Harmful Actions</th>
<th>Brain Microstate</th>
<th>Talairach Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1 0–20 ms</td>
<td>M2 22–60 ms</td>
<td>M3 62–120 ms</td>
<td>M4 122–180 ms</td>
</tr>
<tr>
<td>3</td>
<td>44</td>
<td>7</td>
<td>3</td>
</tr>
</tbody>
</table>

LORETA, low-resolution brain electromagnetic tomography analysis.
As expected from previous work on the perception of intentions (Grafton 2009; Ortigue et al. 2009), the first significant brain microstate emerged around 60 ms over the pSTS. Interestingly, this microstate remained for a longer period for scenarios depicting accidental actions (from 62 to 140 ms; Fig. 2), a finding consistent with the involvement of this region in the visual analysis of other people’s actions and intentions. Activity in the pSTS is greater for incongruent than for congruent actions, demonstrating a need for different levels of processing for observed goal-directed and non-goal-directed observed actions (Pelphrey and Carter 2008). The STS region, which receives input from both the ventral and dorsal visual pathways and which projects to the amygdala and vmPFC, plays a critical role in the visual analysis of perceptual information about bodily motion, action prediction, and evaluation of the intentions behind other people’s behaviors (Allison et al. 2000). Previous studies have shown the presence of recruitment in STS and intraparietal sulcus early in visual processing in line with recent electrophysiological evidence from both animal and human studies arguing for a bidirectional mechanism where these associative areas can be recruited in very early stages of information processing (~60 ms) (Ortigue et al. 2009). The fast involvement of the right pSTS in differentiating intentional from accidental harm demonstrates that automatic perception of intentionality is a critical input to the perception of moral valence of an action. This component, in both timing and localization, is different from emotional negativity bias and stimulus evaluation that are associated with later responses (Huang and Luo 2006). Previous research using ERPs and magnetoencephalography has reported a frontotemporal negativity component, 110–140 ms after visual stimulation, whose source density originates in the dorsal anterior cingulate cortex, is interpreted as an automatic attentional response, whereas the P300 reflects stimulus evaluation and classification.

The early engagement of the right amygdala/temporal pole and vmPFC evoked by the perception of intentional harmful actions suggests that affective processes precede cognitive evaluative processes. The amygdala, through reciprocal connection with the pSTS, underlies rapid and prioritized processing of emotion signals (Sander et al. 2003). The vmPFC projects to the basal forebrain (the major cholinergic output) and brainstem regions, which contains all afferent and efferent systems necessary for survival, including basic affective responses (Ongür and Price 2000), and neurons within the vmPFC encode the emotional value of stimuli (Elliott et al. 2010). Furthermore, the vmPFC, reciprocally connected with the amygdala and hypothalamus, is involved in the autonomic component of emotion and seems essential in the evaluation of harmful intent. Dysfunction within these regions or their functional connectivity has reliably been associated with impaired moral decision making in psychopathy (Anderson and Kiehl 2011; Blair 2007). Although previous neuroimaging research indicates that emotional processing is an integral aspect of moral cognition, the exact point at which this effect occurs has not yet been examined. Here, taking advantage of the high-density ERPs, coupled with brain source analysis, and building on a paradigm validated with fMRI, we demonstrate for the first time how intention understanding in the right pSTS (62 ms after stimulus onset) and then affective processing occurs in very early stages of moral cognition processing (i.e., 122–180 ms after stimulus onset over the amygdala/temporal pole and 182 ms in the vmPFC). These results support the view that intentionality judgments both precede and guide moral cognition (Malle and Guglielmlo 2011).

The timing of early engagement of the amygdala/temporal pole and vmPFC during the perception of intentional harm is
consistent with the notion that emotion acts as a gain antecedent to moral judgment by alerting the individual to the moral salience of a situation and provides evidence for the pervasive role of affect in moral cognition (Blair and Fowler 2008). Finally, our results are also important for neurodevelopmental research with typical and atypical populations, such as children with callous-unemotional traits, which indicates that affective arousal is a necessary (although not sufficient) precursor to mature empathic understanding and moral decision making (Blair 1995; Cheng et al. 2012; Decety et al. 2011).

A limitation of the present study is that it included only a small number of participants. Replication with a greater number of subjects will increase our confidence in the generalizability of these findings.

ACKNOWLEDGMENTS

We thank Tee Zhou for technical help.

GRANTS

J. Decety was supported by NSF Grant BCS-0718480. S. Cacioppo was supported by Swiss National Science Foundation Grant FNS_PPO01_126599/1.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

J.D. and S.C. conception and design of research; J.D. and S.C. performed experiments; J.D. and S.C. analyzed data; J.D. and S.C. interpreted results of experiments; J.D. and S.C. prepared figures; J.D. and S.C. drafted manuscript; J.D. and S.C. edited and revised manuscript; J.D. and S.C. approved final version of manuscript.

REFERENCES