

The Contribution of Emotion and Cognition to Moral Sensitivity: A Neurodevelopmental Study

Jean Decety^{1,2}, Kalina J. Michalska¹ and Katherine D. Kinzler¹

¹Department of Psychology and Center for Cognitive and Social Neuroscience and ²Department of Psychiatry and Behavioral Neuroscience, The University of Chicago, Chicago, IL 60637

Address correspondence to Dr. Jean Decety. Email: decety@uchicago.edu.

Whether emotion is a source of moral judgments remains controversial. This study combined neurophysiological measures, including functional magnetic resonance imaging, eye-tracking, and pupillary response with behavioral measures assessing affective and moral judgments across age. One hundred and twenty-six participants aged between 4 and 37 years viewed scenarios depicting intentional versus accidental actions that caused harm/damage to people and objects. Morally, salient scenarios evoked stronger empathic sadness in young participants and were associated with enhanced activity in the amygdala, insula, and temporal poles. While intentional harm was evaluated as equally wrong across all participants, ratings of deserved punishments and malevolent intent gradually became more differentiated with age. Furthermore, age-related increase in activity was detected in the ventromedial prefrontal cortex in response to intentional harm to people, as well as increased functional connectivity between this region and the amygdala. Our study provides evidence that moral reasoning involves a complex integration between affective and cognitive processes that gradually changes with age and can be viewed in dynamic transaction across the course of ontogenesis. The findings support the view that negative emotion alerts the individual to the moral salience of a situation by bringing discomfort and thus can serve as an antecedent to moral judgment.

Keywords: amygdala, empathy, insula, moral cognition, neurodevelopment, punishment, ventromedial prefrontal cortex

Introduction

Whether emotion is a source of moral judgments or follows them remains controversial (Huebner et al. 2009). From one perspective, emotion may act as an antecedent to moral judgment (Greene and Haidt 2002), and negative emotion may alert the individual to the moral salience of a situation by bringing discomfort (Swedene 2005). From another, moral judgments may exist even in the absence of an emotional response (Kohlberg 1984). Or mature moral reasoning may rely both on emotion and on cognitive processes responsible for representing and integrating information about intentions, beliefs, and outcomes of action. For instance, a fully developed moral cognition may depend on negative emotion encouraging empathic sensitivity and concern for others, theory of mind providing for the representation of the mental states of others, and also on the ability to integrate information about mental states with information about the consequences and emotion in the context of moral judgment (Killen et al. 2011; Young 2011). The current study was designed to explore the role of emotion and empathy in the development of moral cognition across age by combining functional magnetic resonance

imaging (fMRI), eye-tracking, and behavioral evaluation of morally laden scenarios.

Traditionally, moral reasoning was considered a product of gradually developing cognitive processes and deliberations (Kohlberg 1984). More recent research, however, provides evidence that conscious deliberative processes account for only part of the human moral compass. Studies of both adults and young children suggest that the mechanisms behind moral intuitions are sometimes unknown to the individuals who are experiencing those intuitions (Greene and Haidt 2002) and that even preverbal infants express what appear to be nascent moral evaluations (Hamlin et al. 2007). Behavioral studies provide evidence that within the first year of life, infants engage in rudimentary forms of social, and potentially moral, evaluation. For example, infants as young as 6 months of age preferentially interact with an agent who helped, rather than hindered, the actions of another agent (Hamlin et al. 2007). Prosocial behaviors such as altruistic helping also emerge early in childhood. By 12 months of age, infants begin to comfort victims of distress, and 14- to 18-month-old infants exhibit spontaneous unrewarded instrumental helping behaviors (Zahn-Waxler and Radke-Yarrow 1990; Warneken and Tomasello 2009). Three-year-old children are selectively prosocial toward others as they selectively avoid helping those who cause or even intend to cause others harm (Vaish et al. 2010). During the second year of life, children's empathic responses increase. Prosocial behavior increases over childhood in concert with growing self-other differentiation and social understanding (Svetlova et al. 2010). These natural early-emerging behaviors are thought to be motivated by sympathetic emotion or concern for the well-being of others.

Moreover, empirical evidence suggests that increased experiences of empathy encourage prosocial action and inhibit aggression and other antisocial behavior (Zahn-Waxler et al. 1992; Eisenberg 2000; Eisenberg and Eggum 2009). The contextual cues that link moral emotions to social norms are certainly variable and shaped by culture. Nonetheless, there may be some commonalities across human societies in the role that empathy plays in the formation and cohesion of human groups in which people care for each other and in the observance of a moral code within society (Decety 2011b). The relation between emotion, empathy, and morality is thus important and warrants further investigation.

The importance of empathy for the development of moral action is at the core of the violence inhibition mechanism developed by Blair (1995). According to this model, most humans are predisposed to find the distress of conspecifics aversive, at least outside of competitive contexts. Humans, thus, are in effect punished by signals of others' sadness or fear and generally learn to avoid actions associated with that distress.

Consequently, individuals become less likely to engage in actions that give rise to others' distress and more likely to engage in actions that alleviate others' discomfort. In short, most humans consider the distress of another fellow human aversive, and this has consequences for their social actions (Blair RJR and Blair KS 2009; Decety 2010). Moral action can be elicited by expressive cues that directly reflect the other's feelings or situational cues that convey the impact of external stimuli on the individual.

In the past decade, research in affective and cognitive neuroscience has begun to identify a network of brain regions involved in moral cognition (Young 2011). This diverse literature can be divided into 3 domains: moral emotions, theory of mind, and abstract reasoning. The link between these different domains is becoming increasingly clear. Converging results from lesion (e.g., Bramham et al. 2009; Eslinger et al. 2009; Young et al. 2010) and functional neuroimaging studies (e.g., Greene et al. 2001; Moll et al. 2002, 2007; Harenski and Hamann 2006; Shamay-Tsoory et al. 2006; Young and Saxe 2008) indicate the specific roles of the ventromedial prefrontal cortex (vmPFC), medial prefrontal cortex (mPFC), anterior cingulate cortex (ACC), insula, amygdala, and posterior superior temporal sulcus (pSTS) in moral cognition (Moll et al. 2005 for a review). Notably, an emotional learning system mediated by the amygdala and a system for decision making on the basis of reinforcement expectations mediated by the medial orbitofrontal cortex is crucial for cueing morally appropriate behavior and the acquisition of moral knowledge during childhood (Blair 2009; Blair RJR and Blair KS 2009). Moral reasoning is thus underpinned by specific neural circuitry, but these circuits are not unique to morality; rather, they involve regions and systems underlying specific states of feelings, cognitive and motivational processes. These circuits emerge and are interconnected over the course of development to produce adaptive social behavior (Decety et al. 2011).

Though understanding the ontogenetic trajectory by which neural circuitry is formed has clear implications for how we think about the end state of moral reasoning, as well as the nature of moral development, most past neuroscience research has been conducted only with adult participants. No study yet has tackled developmental changes in the neural circuits subserving the component processes mediating moral sensitivity. Studying subcomponents of more complex behaviors can be particularly useful from a developmental perspective, when it is the case that only some components of or precursors to more complex behaviors are observable. Developmental studies can provide unique opportunities to see how the components of the system interact in ways not possible in adults, where all the components are fully mature and operational (De Haan and Gunnar 2009). A neurodevelopmental approach to morality is especially important as many brain regions that are germane to moral functioning appear to be under maturation until nearly young adulthood, while regions involved in affective arousal and somatosensory resonance develop earlier (Gottay et al. 2004). The construction of the prefrontal cortex—a region that includes the dorsolateral prefrontal cortex (dlPFC), mPFC, and vmPFC—is particularly slow to develop (Paus 2011 for a review) and vital to moral cognition, perspective taking, and empathic concern (Damasio 1994; Moll et al. 2003; Ruby and Decety 2003, 2004).

To characterize the neurodevelopment of moral sensitivity, we collected behavioral evaluations, eye-tracking, and neuro-

hemodynamic response to short animated scenarios from a large sample of individuals ($n = 126$) whose age ranged from 4 to 37 years (Fig. 1A for examples). The scenarios were categorized along a 2×2 factorial design to distinguish participants' responses to viewing people versus object being injured; these injuries occurred via either intentional versus unintentional actions. This latter factor of intentionality is particularly critical to explore because detecting intentionality is the decisive cue in determining whether an action was malicious or not (Cushman 2008; Young et al. 2010; Killen et al. 2011). Further, a sense of agency (i.e., the awareness of both oneself and others as agents who are initiators of actions, desires, thoughts, and feelings) is important in relation to morality in determining that people can be held responsible for their actions (Kahn 1992; Blakemore and Decety 2001; Decety and Porges 2011). This current design thus combines explicit evaluations of morally laden scenarios with implicit neurophysiological reactions (neurohemodynamic and affective arousal) and allows us to distinguish brain regions that process intentional versus accidental harm, in concert with corresponding behavior judgments. Given that the amygdala and insula come online earlier in ontogeny than other structures such as the dorsal and lateral vmPFC—which become progressively specialized for the evaluation of social stimuli and are slower to mature—we expect a differential involvement of these regions across age (Decety and Michalska 2010). Such a pattern of functional changes will support the general notion that the development of affective processing from childhood to adulthood is accompanied by reduced activity within limbic affect processing systems and an increased involvement of other prefrontal systems (Killgore and Yurgelun-Todd 2007) that are crucial for moral reasoning and decision making. Given that toddlers differentially interpret intentional from accidental or nonpurposeful actions (e.g., Carpenter et al. 1998; Woodward 1999) and imitate even uncompleted intentional acts of others (Meltzoff 1995; Woodward and Sommerville 2000; Schilbach et al. 2008), judgments of perceived intentionality should be consistent across age. Whether and how information about intentionality is used for judging the wrongness of an action may be age dependent (Zelazo et al. 1996) and influence recommendations of punishment. Determinations of punishment require a complex integration between the analysis of mental states of the perpetrator and the consequences of his actions. These decisions also involve the assignment of moral blame and should therefore be correlated with activity in the vmPFC (Blair 2009; Blair RJR and Blair KS 2009). Furthermore, based on developmental research that has shown that children who lack false belief knowledge are more likely to attribute negative intentions to an accidental transgressor (Killen et al. 2011), and that young children sometimes use information about the outcome of actions to make judgments about intentions (Yuill and Perner 1988; Leslie et al. 2006; Pettit and Knobe 2009), we anticipate punishments to be less differentiated for intentional versus accidental actions in young children in comparison with older participants.

Articulating affective neuroscience and developmental science approaches provides a powerful way to elucidate the neural and computational mechanisms underpinning moral sensitivity and how the social brain develops (Decety 2010). Such interdisciplinary research also has the potential to contribute to recent theoretical debates on the added value of moral neuroscience, an emerging area of research that has

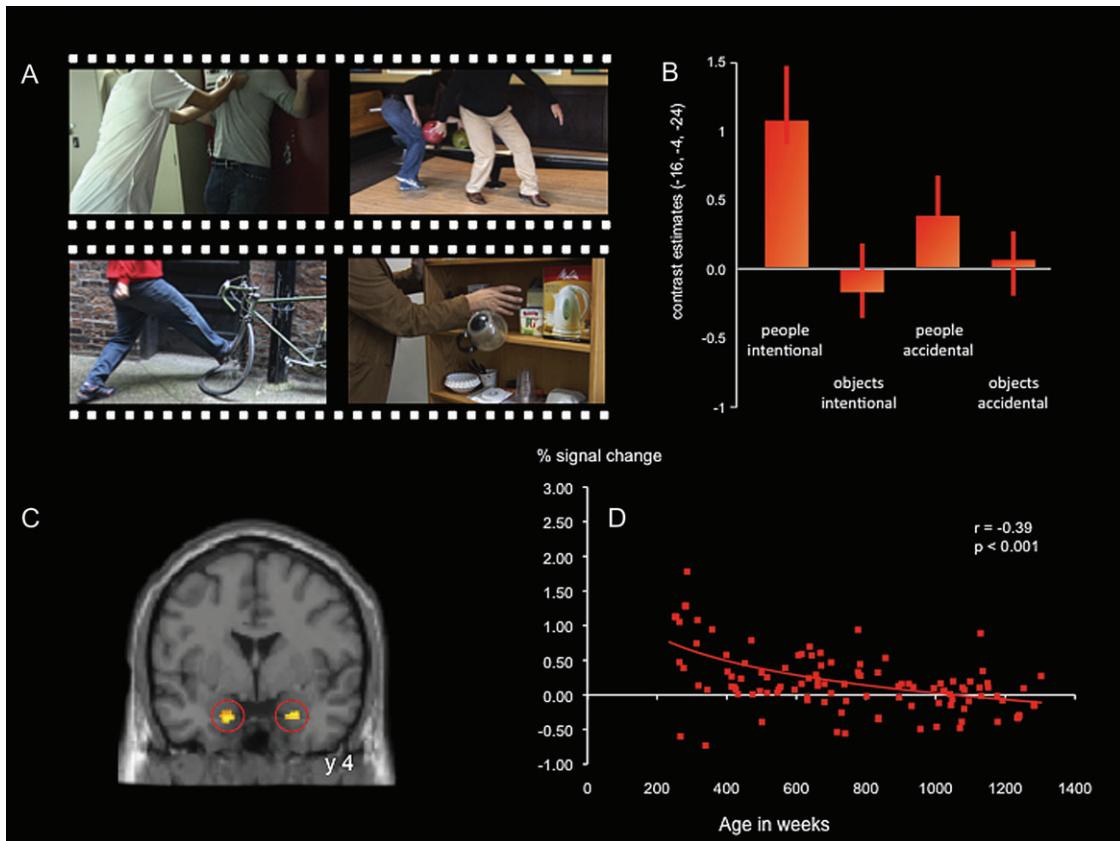


Figure 1. (A) Examples of the visual stimuli used in the study depicting people (top row) or objects (bottom row) being harmed intentionally or by accident. The stimuli were short dynamic visual scenarios (2.2 s). (B) Parameter estimates in the right amygdala for each of the 4 target conditions (ANCOVA with age held constant in the model); (C) ROI in the amygdala; (D) Age-related decrease in the amygdala when the participants viewed scenarios depicting people intentionally being harmed versus people accidentally harmed. A significant negative linear association of age with signal change was observed in the amygdala ($r = -0.39$; $t_{97} = -4.443$; $P < 0.001$). In addition to the linear effect, evidence of a quadratic or curvilinear effect for age was also observed ($t_{97} = 2.283$; $P = 0.025$), indicating that the negative relationship between age and PSC in this cluster of voxels was more extreme at younger ages but grew weaker as age increased.

the potential to improve our understanding of human social and moral behavior (Killen and Smetana 2006; Narvaez 2008; Malti and Latzko 2010).

Experimental Procedures

Participants

One hundred and twenty-six individuals (62 female) were recruited from the local community of the University of Chicago campus and surroundings. Age ranged from 4 to 37 years. The age distribution of participants was as follows: age 4–7 years, mean age 5.73 ± 0.9 ($N = 21$, 13 female); age 8–12, mean age 10.37 ± 1.47 ($N = 30$, 11 female); age 13–17, mean age 14.67 ± 1.59 ($N = 21$, 9 female); age 18–37, mean age 23.74 ± 4.62 ($N = 54$, 29 females). Participants' written consent was obtained. For subjects under age 18, parents' written informed consent was obtained in addition to the minors' verbal assent. All participants were paid for their participation. The study was approved by the University of Chicago Institutional Review Board and conducted in accordance with the Declaration of Helsinki. Of the 126 participants, 102 completed MRI scanning with the following age distribution: $N = 21$ from age 4–7 (13 female), $N = 30$ from age 8–12 (11 female), $N = 21$ from age 13–17 (9 female), and $N = 30$ from age 18–26 (14 female).

Stimuli Preparation

A series of dynamic visual stimuli depicting moral and nonmoral transgressions was created and validated prior the study. Each dynamic stimulus consisted of 3 digital color pictures, which were edited to the same size (600×480 pixels) and presented in a successive manner to imply motion. The durations of the first, second, and third pictures in each animation were 1000, 200, and 1000 ms, respectively. The stimuli belonged to one of 4 categories in a 2 by 2 factorial design ([target (people being hurt/objects being broken; intention (intentional/accidental))] and portrayed the following (see Fig. 1 for examples): 1) A person is shown hurting another person intentionally (person intentional, PI); 2) A person is shown hurting another unintentionally (person unintentional, PU); 3) A person is shown breaking an object intentionally (object intentional, OI); and 4) A person is shown breaking an object unintentionally (object unintentional, OU). One additional baseline stimulus category depicted people in everyday social interactions without any infliction of pain or damage (actions; e.g., a person giving another individual a notebook). The clips showed situations of varying degrees of intensity, portrayed people of multiple races and ethnic groups, as well as various ages. Importantly, the faces of the protagonists were not visible and thus there was no emotional reaction visible to participants. One hundred and fifty dynamic visual stimuli (30 exemplars per

stimulus category, plus 30 in the baseline condition) were created.

Stimuli were validated prior to the study on computerized visual analog scales for perceived intentionality and empathic concern by a group of 26 participants whose age ranged between 18 and 23 years (Hempel 2009). Eye-tracking and pupillary dilatation data were simultaneously recorded with a Tobii T120 system. Results showed that subjective ratings of empathic concern were higher when participants were watching the stimuli depicting people being hurt intentionally than when watching people whose pain was accidentally caused ($P < 0.001$). Participant's pupil dilations were analyzed using a 2×2 repeated-measures analyses of variance (ANOVA). A main effect of intentionality was found ($F_{1,25} = 30.46$; $P < 0.001$), showing larger pupil dilation in response to clips depicting intentional actions. There was also a main effect of target identity ($F_{1,25} = 31.02$; $P < 0.001$), showing that participants' pupils dilated significantly more if the target of an action depicted on a clip was a person than an object. Planned comparisons revealed that when the target in the clip was a person, participants showed larger pupil dilations when the pain was intentionally inflicted than when it was caused by accident ($P < 0.01$). A similar pattern was found when the target of the clip was an object such that participants had larger pupil dilation in the object intentionally damaged condition than the object unintentionally damaged condition ($P < 0.01$).

Training in a Mock Scanner

Prior to MRI scanning, participants were acclimated to the experimental procedures in a mock scanner. They were asked to lie in the mock scanner while a documentary movie was played. When participants felt comfortable, they were presented with 24 stimuli (6 per condition) depicting situations similar to, but not the same as, those they would watch in the actual scanning sessions. MRI noise was simulated through a recording played during the mock session.

MRI Scanning

Stimuli were presented with E-prime software (Psychology Software Tools, Inc., Pittsburgh, PA) and a back-projection system. A mixed blocked/event-related design paradigm was used with a total of 30 baseline blocks (duration 17.6 s each) during which a fixation cross was presented and 25 active blocks (duration 19.8 s each) during which stimuli from one of the 5 categories were presented. Stimuli were blocked by target (person/object) and intention (intentional/unintentional) was randomized within each block. The presentation order was counterbalanced across runs and across subjects. Each block consisted of 6 stimuli (2200 ms each) with a jittered interstimulus interval (1.69–5.93 s), during which a black fixation cross was presented against a gray background. Participants were shown the stimuli in 5 short sessions (5 active blocks per session, 4.5 min each) to maintain their attention. To avoid confounding motor-related activation in the ACC and pre-SMA/SMA, no overt response was required. Instead, participants were instructed to watch the stimuli carefully.

MRI was performed on a 3 T Philips Achieva Quasar scanner at the University of Chicago Brain Research Imaging Center. The fMRI pulse sequence parameters include time repetition/

time echo (TR/TE) 2200/26, flip angle = 80, contiguous slices with 4 mm thickness, no gap, 230×230 mm field of view (FOV), approximately 76×75 matrix. High resolution structural images were acquired in the sagittal plane using a T_1 -weighted 3D Turbo Field Echo (TFE/MP-RAGE) anatomical scan with the following parameters: TR = 7.4 ms, TE = 3.4 ms, FOV = 25×25 cm, 0.6 mm slice thickness, no gap, 228×228 mm matrix, $1.1 \times 1.1 \times 0.6$ mm voxel size. During the anatomical and diffusion tensor imaging scans, participants watched a movie of their choice (from a selection of various popular documentaries about space, underwater adventures, birds, etc.).

Eye-Tracking Measures and Analysis

For the duration of the scanning, participants' eye gaze fixations and pupil dilation were recorded using ASL 6000 (Applied Science Laboratories, Beford, MA). Eye gaze was tracked at an update rate of 120 Hz. Fixations on the stimuli were measured as constant eye gaze that must be held within 1 degree of the visual angle for at least 100 ms. Data were analyzed using ASL Results v. 1.17.09. Data were successfully collected from 61 participants. Due to technical error, data were analyzed for 57 (22 adults, 29 females). Analysis was constrained to the third frame for 2 reasons: first, subjects cannot differentiate between conditions in the first frame and second, using the third frame ensured that any change in pupillometry could be attributed to emotional arousal instead of differing luminance between the animation and the previous fixation screen.

Moral Evaluations

After the scanning session, participants were presented with the same stimuli (in a randomized order) that they saw in the scanner on a computer desktop and asked to rate whether the action performed by the perpetrator in the video clip was intentional or not (language was adapted to participants' age; e.g., younger participants were asked "was this on purpose?"). They were then asked to respond to a set of 5 questions probing moral judgment using a computer-based visual analogue scale (VAS) ranging from 0 to 100. The questions were designed to assess empathic concern for the victim, personal distress, understanding of the perpetrator's mental state, and moral evaluation, respectively. The following questions were asked: "How sad are you for the person/object that was hurt?" "How upset do you feel about what happened?" "How mean was the person who did this?" "Was it wrong to do this?" and "How much would you punish the person who did this?" A research assistant sat next to the younger children to assist them in these evaluations. One hundred and eight participants (60 female) answered questions related to empathy, personal distress, and mental state understanding.

Data Analysis

Based on exploratory factor analyses in a separate study (de Wied et al. 2007), 6 of the 20 items on the Bryant Empathy Index (1982) were selected for analysis because of their thematic similarity to a subgroup of items on the Interpersonal Reactivity Index (IRI) that measures empathic concern. The scores on these scales were compared to assess the subjects' empathic sadness, which evaluates a more emotional aspect of empathy, namely, responsiveness to another person's sadness.

Since the questions on the Kiddie Mach had been directly adapted from the Mach IV scale and modified for children's use, data from the 2 scales were transformed into *z*-scores in order to look at age-related changes on a continuous scale. All data were analyzed in SPSS.

fMRI data processing was carried out with SPM8 (Wellcome Department of Imaging Neuroscience, London, UK), implemented in MATLAB 7.0 (Mathworks Inc., Sherborn, MA). Preprocessing included correction for head motion, normalization to the echo-planar imaging (EPI) template provided in SPM8, and smoothing using a 6-mm full-width half-maximum isotropic Gaussian kernel. Images were realigned and normalized using standard SPM procedures. A mean T_1 image was created from all participants who completed scanning. Structural T_1 images were first coregistered to the mean EPI image for each participant. The coregistered T_1 images were then spatially normalized and an average of these normalized T_1 images of all the participants was created. All 102 subjects who completed scanning had less than 0.5 voxels of in-plane motion throughout the entire experiment. A 2-level approach for block-design fMRI data was adopted using SPM8. A voxel-by-voxel multiple regression analysis of expected signal changes for each of the 5 block categories, which were constructed using the hemodynamic response function provided by SPM8, was applied to the preprocessed images for each subject. Individual subject data were analyzed using a fixed-effects model. Condition effects at the subject level were modeled by box-car regressors representing the occurrence of each of the 5 block types.

The resulting first-level contrast images were then entered into a 2 (Target: people vs. objects) \times 2 (Intention: intentional vs. unintentional) analysis of covariance (ANCOVA), with age in weeks as covariate. Results for the group analysis were thresholded at $P < 0.005$ corrected for multiple comparisons across the whole volume using the false discovery rate (FDR) approach (Genovese et al. 2002), and only clusters of 10 or more contiguous voxels were reported. The choice of this threshold was determined based on previous studies on empathy for pain and on power considerations for the current paradigm (Jackson et al. 2005; Lamm, Batson, et al. 2007; Lamm, Nusbaum, et al. 2007; Lamm and Decety 2008; Lamm et al. 2010). These included regions associated with theory of mind (pSTS/TPJ, mPFC) and moral sensitivity (vmPFC). Activations were overlaid on the average of the normalized T_1 images of all the participants. For sake of clarity, functional data are presented overlaid on a representative high-resolution structural T_1 -weighted image from a single subject from the SPM8 canonical image set, coregistered to Montreal Neurological Institute (MNI) space. Results were compared to confirm that the ensuing activations were indeed accurate.

In addition, specific regions of interest (ROIs) were computed for activity in the following 4 regions: amygdala, insula, vmPFC, and pSTS/TPJ. Data extraction for the ROI analyses was performed using the rfxplot toolbox (Glässcher et al. 2009) implemented in SPM8. ROIs were defined as a 6-mm spherical region centered on the following MNI coordinates: right amygdala: $x = 16, y = -4, z = -24$, left amygdala: $x = -20, y = -2, z = -24$; insula $x = 44, y = 20, z = -6$; vmPFC $x = 10, y = 28, z = -12$; pSTS $x = 52, y = -42, z = 28$. These coordinates were determined on the basis of neuroanatomical atlases as well as 3 meta-analyses: one on 20 fMRI studies of pain empathy (Jackson et al. 2006), another on 32 studies of empathy for pain

(Lamm et al. 2011), and also one on 70 fMRI studies of theory of mind, perspective taking, empathy, and attention reorientation (Decety and Lamm 2007). The individual mean percent signal change (PSC) values were then subject to ANOVA for repeated measures to test for main effects of target and intent as well as target-by-intent interactions at the group level. For statistical analyses of the ROI data, SPSS was used.

In order to examine nonlinear effects of age on the neurohemodynamic response, age and age^2 were included as predictors in the ROI-based analyses using linear regression. To improve interpretability, the age variable was initially centered at its mean (calculated as age-mean age) and age^2 was calculated as (age-mean age) 2 . The quadratic effect of age was assessed by the statistical significance of the age^2 term. To characterize this effect further, we examined the intercept and the slope of the curve at different ages, specifically at 1 standard deviation (SD) above and 1 SD below the mean.

For relatively young participants (1 SD below the mean), the intercept was found to be statistically different from zero ($t_{97} = 6.573; P < 0.001$). Similarly, the slope was found to be negative and statistically different from zero ($t_{97} = -3.791; P < 0.001$). For relatively old participants (1 SD above the mean), however, both the intercept and the slope were not statistically different from zero ($t_{97} < 1$). Taken together, data indicate that the PSC in the amygdala appears to follow a curvilinear function, such that the signal is highest at the youngest ages, decreases rapidly through childhood and early adolescence and asymptotes in late adolescence through adulthood.

In order to assess the relationship between behavioral data and brain activity, random effects correlation analyses were performed, and only clusters that survived FDR at $P = 0.05$ were considered. Scores on the Empathic Concern and Personal Distress subscales of the IRI for adults and the Empathic Sadness subscale of the Bryant Empathy Index for children, and the normalized values of the Machiavellian Scale were correlated with parameter estimates of the contrasts People > Objects; Intentional > Unintentional and People intentional > People Unintentional. Correlations with ratings of empathic concern, personal distress, understanding of the perpetrator's mental state and moral evaluation were computed within each of the selected ROIs. Average parameter estimates were computed across all voxels in that region for each condition, for each participant, individually. The correlation between the hemodynamic activity in those regions in response to viewing each of the experimental conditions was then computed with their associated subjective ratings. Evaluations of the morally laden stimuli were computed as difference scores (consisting of the following differences: People – Objects; Intentional – Accidental). Postscan ratings were not included in the generalized linear model and were therefore not used to define the regions to ensure independence of the analyses (Vul et al. 2009).

Functional Connectivity

Psychophysiological interaction (PPI) analysis was used to estimate functional connectivity between a source (vmPFC) and target regions of interest (amygdala, pSTS), during the viewing of moral versus nonmoral actions involving people. PPI analysis assesses the hypothesis that activity in one brain region can be explained by an interaction between the presence of a cognitive process and activity in another part of the brain.

The selection of vmPFC as the PPI source region denoted activity within vmPFC as the physiological regressor in the PPI analysis. Intentionality condition (intentional vs. accidental harm to people, PI vs. PA) was the psychological regressor. A third regressor in the analysis represented the interaction between the first and second regressors.

The psychological variable used was a vector coding for the specific task (1 for PI, -1 for PA) convolved with the hemodynamic response function (HRF). The individual time series for the vmPFC was obtained by extracting the first principal component from all raw voxel time series in a sphere (3 mm radius) centered on the coordinates of the subject-specific vmPFC activations. These time series were mean-corrected and high-pass filtered to remove low-frequency signal drifts. The physiological factor was then multiplied with the psychological factor to constitute the interaction term. PPI analyses were then carried out for each subject involving the creation of a design matrix with the interaction term, the psychological factor, and the physiological factor as regressors. Subject-specific contrast images were then entered into random effects group analyses (thresholded at $P < 0.001$, uncorrected, $k = 10$). PPI analyses were conducted at each of the 4 age groups separately (4–7 years; 8–12 years; 12–17 years; and 18–25 years), in order to identify target brain regions that showed a significant increase in functional coupling with vmMPFC during moral relative to nonmoral harm at each successive age group.

Results

The goal of the study was to assess the contribution of emotion and mental state understanding to the neurodevelopment of moral cognition. Results and their interpretation are organized along the following dimensions: dispositional measures, attention and affective arousal, intention understanding, empathic responses, mental state understanding, and moral evaluation.

Dispositional Measures

IRI scores for empathic concern ($M = 3.67$, $SD = 0.71$) and personal distress ($M = 2.53$, $SD = 0.67$) did not change with age. No gender differences were found. Significant gender differences were found in the empathic sadness subscale of the Bryant empathy index, where boys' scores decreased with age while girls' scores increased (boys: $r_{26} = -0.44$, $P = 0.026$; girls: $r_{27} = 0.36$, $P = 0.048$). Linear regression analyses indicated that this measure predicted how upset the children were when viewing harm in all scenarios ($P < 0.01$). In adults, no significant associations were found between subscales of the IRI and any of the moral evaluations. As age increased, subjects' Mach scores rose significantly, $r_{77} = 0.261$, $P < 0.05$ indicating that participants increasingly acquire knowledge about the ability to manipulate other's emotions. Linear regression analyses showed that participants' scores on the Mach scale predicted rating of wrongness for intentional harm done to people only ($t_{97} = -28$, $P < 0.05$). There were no gender differences in Mach scores, $P > 0.05$. Interestingly, there were no significant correlations between dispositional assessments of empathy and physiological or hemodynamic measures (see Decety, 2011a for a critical appraisal of self-report measures of empathy and neuroimaging research).

Attention and Affective Arousal

Eye-tracking data confirmed that all participants paid attention to the stimuli while being scanned. Eye gaze patterns differed significantly depending on the target of the action, with participants looking at the target of harmful situations (people and objects) significantly longer than at the agent for every condition, $F_{1,53} = 22.77$, $P < 0.001$. When participants viewed people being harmed, there was no main effect of intentionality ($r_{56} = 0.057$, $P = \text{NS}$) or any difference across age.

Pupil dilation showed a main effect of intention, across all ages, such that significantly greater dilation was observed in response to situations depicting intentional over accidental harm, regardless of the target ($F_{1,56} = 12.46$, $P < 0.001$) after controlling for luminance of the different stimuli.

This effect was qualified by a 3-way target \times intention \times age interaction, $F_{1,55} = 7.63$, $P < 0.01$. To decompose this interaction, separate ANCOVAs were conducted examining the effects of age and intention for each type of target (people and objects). These analyses revealed that there was a significant age \times intention interaction for people only, $F_{1,55} = 12.67$, $P < 0.001$, such that less pupil dilation was measured with increasing age for people being accidentally harmed. Interestingly, a statistically significant correlation was found for pupil dilation during the observation of intentionally caused pain to people and activity in the dorsal mid-cingulate cortex, amygdala, and pSTS/TPJ (Supplementary Table 1).

Intention Understanding

All participants, regardless of their age, correctly answered whether an action was intentional or not on 86.7% of the trials. Accuracy was highest for intentional conditions (93.1%) and lowest for unintentional conditions (80.3%). No significant effects of age were found.

A main effect of perceived intentionality, irrespective of the target of the action (PI + OI vs. PU + OU), was associated with significant hemodynamic signal increase in the amygdala (see Fig. 1B,C), temporal pole, periaqueductal gray (PAG), vmPFC, insula, middle frontal gyrus, right pSTS/TPJ, and intraparietal sulcus in both hemispheres (see Supplementary Table 2). ANCOVA indicated a significant interaction between intention and target in the amygdala, insula, vmPFC, and pSTS. Planned comparisons revealed a significant effect of intentionality for the people only trials (PI-PU) indicating a higher hemodynamic response in each of these regions when viewing people being intentionally versus accidentally hurt (see Supplementary Table 3).

Age-related changes were observed in a subset of these regions: the younger the participants, the higher the activity in the PAG, temporal poles, amygdala (see Fig. 2c), and right insula. For relatively young participants (1 SD below the mean), the intercept for the amygdala response was found to be statistically different from zero ($t_{97} = 6.573$; $P < 0.001$). Similarly, the slope was found to be negative and statistically different from zero ($t_{97} = -3.791$; $P < 0.001$). For relatively old participants (1 SD above the mean), however, both the intercept and the slope were not statistically different from zero ($t_{97} < 1$). The older the participants, the higher the activity in the vmPFC (-10, 46, -14) when they viewed intentional versus accidental harm done to people (Fig. 2).

As expected, and based on previous developmental research (Meltzoff 1995; Woodward 1998, 1999; Woodward and

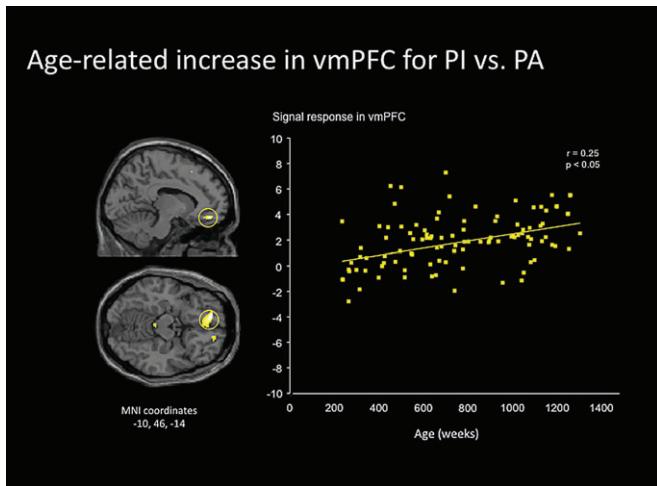


Figure 2. Age-related hemodynamic increase in the vmPFC when participants perceive visual scenarios depicting people being harmed intentionally versus accidentally. A positive linear association of age with signal change in the vmPFC was determined ($r = 0.25$; $t_{95} = 2.723$; $P < 0.05$). The quadratic effect in this cluster of voxels was nonsignificant ($P = 0.08$).

Sommerville 2000), these results demonstrate that perception of intentionality does not vary with age (at least in the range tested here), and as predicted in our hypothesis, the neural response to morally laden scenario in the insula and amygdala in young children is not due to demands of processing of intentional information itself but rather is likely indicative of response to the moral content.

Empathic Responses

The ANOVA revealed a main effect both of intentionality and of target for participant's empathic responses. All participants reported feeling more sad and more upset when watching intentional harm, as opposed to accidental harm (sad: $F_{1,100} = 187.45$, $P < 0.001$); upset: ($F_{1,100} = 300.59$, $P < 0.001$). They similarly reported being more sad and more upset when seeing a person being hurt than when seeing an object being damaged (sad: $F_{1,100} = 185.41$, $P < 0.001$; upset $F_{1,100} = 117.70$, $P < 0.001$). For empathic sadness ratings, there was no interaction with age for either intentionality or target. Ratings of empathic sadness for the victim correlated with activity in the left insula, thalamus, and subgenual prefrontal cortex (Supplementary Table 4). The main effect of intention on ratings of personal distress (i.e., "how upset do you feel?") was qualified by a significant age \times intention interaction ($F_{1,100} = 6.87$, $P < 0.01$). This interaction was driven by the fact that with increasing age, subjects reported feeling less upset when viewing objects being damaged ($r_{101} = -0.32$, $P < 0.001$). Reports of being upset when viewing harm to people did not change with age ($r_{101} = -0.12$, $P = \text{NS}$), however. The more upset the participants were the stronger the neural response in the right amygdala and medial orbitofrontal cortex (OFC) when seeing people being harmed as compared with objects being damaged (see Supplementary Table 4).

Mental State Understanding and Moral Evaluations

Across all ages, participants reported intentional harm as significantly more wrong than accidental harm, regardless of target ($F_{1,85} = 392.51$, $P < 0.001$). They also rated harm done to people as significantly more wrong than damage done to

objects regardless of intent ($F_{1,85} = 198.53$, $P < 0.001$). Critically, judgments about how wrong an action was did not differ significantly by age ($F_{1,85} = 1.41$, $P = \text{NS}$ for target \times intentionality \times age interaction). Linear regression analyses indicate that ratings of wrongness of an action predicted ratings of deserved punishment for all scenarios ($P < 0.001$). The more an action was evaluated as wrong, the stronger the activity in the right insula, temporal poles, inferior frontal gyrus, and ACC (see Supplementary Table 4).

There was a main effect of both intentionality and target for ratings of the perpetrator's mental state. Participants rated agents as more mean if they intentionally harmed a person or damaged an object rather than accidentally doing so ($F_{1,100} = 578.76$, $P < 0.001$) and if they harmed a person rather than an object ($F_{1,100} = 257.66$, $P < 0.001$). While the 2-way interaction of intentionality \times target was not significant ($P = 0.06$), the ANCOVA revealed a significant target \times intention \times age interaction ($F_{1,100} = 8.845$, $P < 0.005$). With age, agents were rated as even more mean if they had harmed a person intentionally than if they had harmed a person accidentally ($r_{93} = 0.259$, $P < 0.01$). Also, the difference between intentionally and accidentally harming objects decreased with age ($r_{93} = 0.346$, $P < 0.001$). Linear regression analyses indicate that ratings of meanness of an agent predicted ratings of wrongness for all scenarios ($P < 0.001$). Ratings of meanness correlated with increased activity in the dorsolateral PFC and vmPFC.

Across all ages, participants reported that they would punish agents of intentional harm more than agents of accidental harm ($F_{1,86} = 379.17$, $P < 0.001$) and would punish individuals who harmed people more than those who damaged objects ($F_{1,86} = 102.79$, $P < 0.001$). In contrast to wrongness ratings, there was a significant target \times intention \times age interaction ($F_{1,86} = 6.09$, $P < 0.01$), which was driven by the fact that with increasing age, participants were less willing to punish agents who intentionally damaged objects than those who intentionally harmed people ($P < 0.01$, see Fig. 2). The more participants would punish someone for an action, the greater the activity in the orbitofrontal cortex and right hippocampus (see Supplementary Table 4).

Functional connectivity analysis showed a significant interaction between intentionality (PI $>$ PA) and vmPFC activation, expressed in the PAG in the youngest age group (aged 4 to 7). In early childhood, the PAG showed greater functional connectivity with vmPFC during moral than nonmoral actions. Within the middle childhood group (aged 8 to 12), PPI analysis revealed a significant interaction between intentionality (PI $>$ PA) and vmPFC activation, expressed in left pSTS, left amygdala, left inferior temporal gyrus and left fusiform gyrus. Within the adolescent group (aged 13 to 17), PPI analysis revealed a significant interaction between intentionality (PI $>$ PA) and vmPFC activation, expressed in bilateral amygdala and mid-cingulate cortex (aMCC). Within the adult group (aged 18 to 25), PPI analysis revealed a significant interaction between intentionality (PI $>$ PA) and vmMPFC activation, expressed in the amygdala bilaterally, in right pSTS/TPJ, and superior parietal lobe. Comparison across the three groups showed a significant difference ($P < 0.05$) between the early childhood group and the adult group in the connectivity between the vmPFC and amygdala reflecting increased functional integration between these 2 regions in adults.

Discussion

It is commonly agreed that moral reasoning necessitates attention, emotion, and cognition. However, the respective contribution of these processes and their causal structure remains unclear (Huebner et al. 2009). This study adopted a neurodevelopmental approach to elucidate the computational mechanisms that underlie moral sensitivity and reasoning by combining implicit physiological measures of attention, arousal, and neural activity with explicit evaluations of morally laden scenarios. This is the first study to examine brain and behavior relationships in response to moral and nonmoral situations from a developmental perspective. Such a perspective is particularly valuable because it provides unique opportunities to chart out how the various components interact across age in ways that are not possible in a purely adult sample (as it is often the case in neuroimaging research), where all the components are mature and operational. A schematic representation of the neural components involved in moral sensitivity and patterns of age-related changes in these regions is illustrated in Supplementary Figure 1.

Attention and Affective Arousal

Measures of attention allocation and emotional arousal recorded simultaneously with fMRI scanning, using eye-tracking, indicate 2 important findings. First, all participants across age attended equally to the scenarios and to the actions depicted in them. They paid significantly more attention (looking time) to people being harmed and objects being damaged than to the agents responsible for these actions. This suggests that action analysis was consistent across both age and target conditions and that more time was spent looking at the “victim” of harmful actions. Second, pupil dilation was significantly greater for intentional actions than accidental actions, and this difference was constant across age. Pupillary movements are determined by the state of the iridic musculature under the direct control of both the sympathetic and parasympathetic branches of the autonomic nervous system. Thus, pupillometry provides a reliable index of autonomic arousal. Increase in pupil size in our study was associated with a specific enhancement in hemodynamic activity in the right amygdala, pSTS, and anterior aMCC, a region that plays a key role in a general salience system concerned with bodily monitoring, emotional arousal, response selection, and skeletomotor body orientation. The aMCC has also been activated during anxiety in response to aversive conditional stimuli, pain anticipation and pain imagination (e.g., Porro et al. 2003; Jackson et al. 2006; Lamm, Batson, et al. 2007), and it appears that noxious activation in aMCC is associated with fear and anxiety that are critical to avoidance behaviors (Vogt et al. 2003). In general, activity in this region, functionally connected with brainstem nuclei and PAG (Kong et al. 2010), correlates and predicts cardiovascular and electrodermal arousal evoked by a range of emotional and motivational tasks (Critchley et al. 2002). Overall, eye-tracking data demonstrate that all participants, irrespective of their age, attended equally to the stimuli in each category and that intentional actions evoked higher affective arousal. These findings are of particular importance since no explicit task was required of the observers in the scanner, and yet these implicit measures provide assurance that the scenarios were well attended to and discriminated based on the intention of the agent. Consequently, age-related

differences in hemodynamic and behavioral measures are unlikely due to differences in attention or affective arousal.

Perception of Intention

We hypothesized that while people’s intentions are readily perceived at very young ages (Woodward 1999), age-related differences would be observed in regions coding for affective saliency and emotion processing when participants attend to morally laden scenarios. This was indeed the case. Postscan evaluations showed that all participants, irrespective of their age, correctly differentiated intentional actions from accidental actions. At the group level, perceiving intentional harm to people versus accidental harm was associated with increase activation in regions sensitive to the perception, prediction, and interpretation of others’ actions and intentions such as the right pSTS/TPJ (Blakemore et al. 2003; Pelphrey and Carter 2008), as well as regions processing the affective consequences of these actions, namely the temporal poles, the insula, vmPFC, and amygdala. Of singular importance, the activity in the amygdala showed, at the group level, a distinct pattern such that the evoked response was stronger when participants viewed intentional actions harming people than for accidental harm to people (Fig. 1B). The former situation is certainly more arousing as confirmed by the pupillary data, and this is coherent with a general role of the amygdala in processing relevant and salient stimuli (Norris et al. 2004). Furthermore, the more participants reported being upset about people being harmed versus objects being damaged, the higher the activity in the amygdala. Interestingly, while participants’ behavioral ratings of being upset while watching people being harmed did not change with age, a significant hemodynamic decrease was detected in the amygdala, temporal poles, insula, and PAG (regions involved in emotion processing). Regression analyses indicate that the signal change in the amygdala appears to follow a curvilinear function, such that the signal is highest at the youngest ages, decreases rapidly through childhood and early adolescence and asymptotes in late adolescence through adulthood. The age-dependent signal change in the amygdala and correlation with ratings of empathic distress supports the role of this region in the normal development of empathic understanding (Decety and Michalska 2010). It has been suggested that the amygdala factors essentially into the development of social behavior, specifically in the generation of emotion to a social stimulus that, depending on the context, may either inhibit or facilitate social interaction (Bauman et al. 2011). Importantly, lesion studies in humans have shown that individuals with early damage to the amygdala are impaired relative to all other groups on more advanced tests of theory of mind reasoning (Stone et al. 2003; Shaw et al. 2004). In contrast, patients who sustained damage to the amygdala in adulthood are not impaired in mental state reasoning relative to both clinical and healthy controls. This developmental neuro-psychological research, in conjunction with the results presented here, provides evidence that ostensibly similar signatures of behavior may be subserved by distinct neural pathways in young children and adults.

The neurohemodynamic signal in older participants increased in the dlPFC and mPFC in the comparison between intentional versus accidental actions (Fig. 2). These regions are associated with meta-cognitive representations that enable us to reflect on the values linked to outcomes and actions

(Amodio and Frith 2006). In addition, the ratings of sadness for the victim, which were strongest in young children, decreased gradually with age and correlated with activity in the insula and subgenual prefrontal cortex. This latter area has extensive connections with circuits implicated in emotional behavior and autonomic/neuroendocrine response to stressors, including the amygdala, lateral hypothalamus, and brainstem serotonergic, noradrenergic, and dopaminergic nuclei (Drevets et al. 1997). Its lesion is associated with abnormal autonomic responses to emotional experiences, inability to experience emotion related to concepts that ordinarily evoke emotion, and impaired comprehension of the adverse consequences of pernicious social behaviors (Bechara et al. 1996). In addition, patterns of functional connectivity during the perception of morally laden scenarios (intentional harm) relative to accidental harm provide complementary evidence for increased integration between prefrontal cortex and amygdala. Changes in functional integration between vmPFC and amygdala were observed, such that the older participants showed significant coactivation in these regions during the PI relative to PA condition, whereas the youngest children only exhibited a significant covariation between the vmPFC and PAG. Furthermore, adult participants showed the strongest connectivity between vmPFC and pSTS/TPJ while viewing moral relative to nonmoral actions than the younger participants, suggestive of developmental changes in functional integration within the mentalizing system.

In addition, the fact that morally laden stimuli evoked stronger empathic sadness in younger participants combined with an enhanced response in neural networks coding affective saliency provides strong support to the notion that emotion plays a critical role in guiding the development trajectory of our moral capacities (Zahn-Waxler and Radke-Yarrow 1990; Blair 1997; Anderson et al. 1999; Decety et al. 2008; Decety et al. 2011).

Moral Evaluations

Nonetheless, determining whether an action is moral or not does not rely only on affective sensitivity. Determinations of moral relevance also require the capacity to integrate a representation of the mental states of others together with the consequences of their actions (Leslie et al. 2006; Young et al. 2007; Killen et al. 2011). Moral judgments stem from a complex interaction between the assessment of causal responsibility for harm and an assessment of an intent to harm (Cushman 2008). Interestingly, judgments of wrongness did not change across age; all participants rated intentional harm as more wrong than accidental harm. While this finding accords with evidence suggesting that children use intentionality to determine moral relevance from an early age in development (Smetana and Killen 2008), it contrasts with studies showing that judgments of wrongness are considerably influenced by perceived intentionality in an age-dependent manner (e.g., Yuill and Perner 1988; Zelazo et al. 1996). This discrepancy is likely due to the fact that our youngest participants were around 5 years old, and older than in previous studies. Furthermore, when participants were asked about the malevolence of the agent (which taps into mental state understanding), their behavioral evaluations indicated a more differentiated appraisal with age. Whereas young children considered all agents malicious, irrespective of intention and targets (i.e., people and objects), older participants perceived the perpetrator as

clearly less mean when carrying out an accidental action, and even more so when the target was an object. Ratings of deserved punishment changed similarly with age (Fig. 3). As age increased, participants punished an agent who damaged an object less severely than an agent who harmed a person. Though even young children attend to both intentionality and target in guiding their own empathic responses and judgment of wrongness, an increased discrimination of intentionality and target in determining moral culpability with age is seemingly consistent with the developmental shift in moral judgment dominated by a focus on outcomes to the integration of both intent to harm and consequences (Cushman 2008; Killen et al. 2011). It should further be noted that the importance of objects to children and adults may differ: the objects we used here were everyday objects. It seems possible that damaging objects with a particularly important history or value might be seen as immoral among adults.

Finally, an age-related hemodynamic increase was found in the vmPFC in response to scenarios depicting intentional harm versus accidental harm to people (Fig. 2). This result is consistent with lesion studies that documented a specific impairment of judgment of intent in patients with bilateral damage of this region (Anderson et al. 1999; Bramham et al. 2009; Young et al. 2010). In addition, ratings of meanness correlated with increased activity in the dlPFC, and ratings of punishments with activity on the OFC. Together, both implicit brain measures and explicit moral evaluations converge on this developmental transition.

How much would you punish?

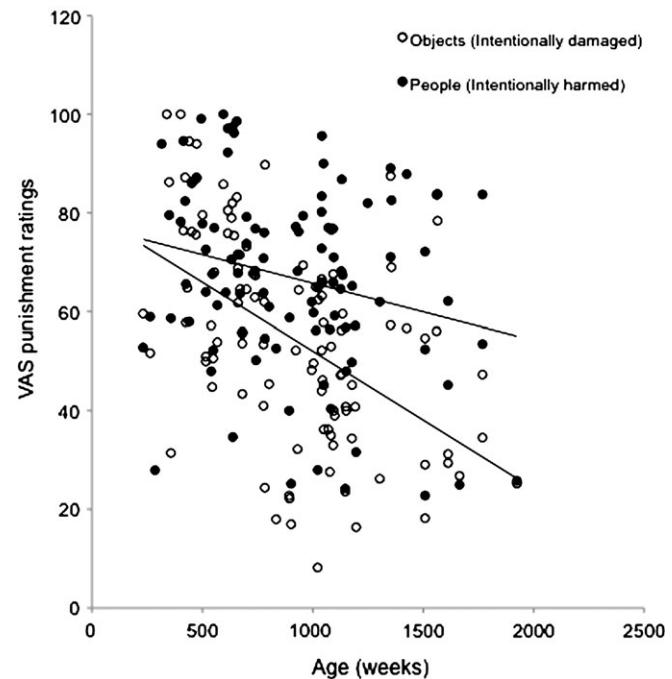


Figure 3. Across all ages, participants will punish people more for an intentionally versus accidentally harmful action, regardless of whether the target is a person or an object ($P < 0.001$). They will punish people more if they harm a person than if they damage an object, regardless of the intention of the action ($P < 0.01$). The graph shows age-related change in ratings of deserved punishment. With age, participants' severity of punishment decreases significantly more for intentional damage done to objects than intentional harm done to people ($P < 0.01$).

Conclusion

Our study provides strong evidence that moral reasoning involves a complex integration between emotion and cognition that gradually changes with age, which can be viewed in dynamic transaction across the course of ontogenesis (see Textbox 1 for a summary of major findings).

Neurodevelopmental changes are clearly seen in structures that are implicated in emotion saliency (amygdala and insula), with a gradual decrease with age. Conversely, activity in regions of the medial and ventral prefrontal cortex that are reciprocally connected with the amygdala and that are involved in decision making and evaluation increases with age and these regions become more functionally coupled. Overall, these results highlight the importance of affect in the development of morality (Blair 1997; Greene and Haidt 2002; Narvaez 2008; Warneken and Tomasello 2009; Malti and Latzko 2010; Decety et al. 2011), which has consequences for theories of moral cognition. They are also consistent with the view that morality is instantiated by functionally integrating several distributed areas/networks. While neuroimaging studies alone are insufficient to determine what computations are involved in moral reasoning, combining fMRI data, physiological recordings of the autonomic nervous system with explicit evaluations of

mental states, moral sensitivity, and empathy in a developmental perspective provides a more complete account of the component processes that mediate moral reasoning, from which we hope that future research will build. Finally, our study indicates that neurodevelopmental approaches provide added value to developmental psychology, even in places in which behavioral responses are remarkably similar across age.

Supplementary Material

Supplementary material can be found at: <http://www.cercor.oxfordjournals.org/>

Funding

National Science Foundation award (BCS-0718480 to J.D.).

Notes

We are grateful to Alexa Tompary for assistance in scanning children and eye-tracking data analysis. We thank all the families and their children who participated in the study. *Conflict of Interest:* None declared.

References

Text box 1: A summary of the major findings.

- Greater looking time (eye-tracking) at the victim of harmful actions than the perpetrator and increased pupil size for situations depicting intentional harm. Pupil size correlated with activation in pSTS and amygdala. No age-related changes in looking time or pupil size.
- Perception of intentional harm associated with increase signal in the amygdala, PAG, insula, vmPFC, and right pSTS/TPJ. The younger the participants the higher the activity in amygdala, PAG and insula when they viewed intentional harm versus accidental harm, whereas the older the participants the higher the activity in the vmPFC.
- Increased functional connectivity with age between the vmPFC and the amygdala when participants viewed scenarios depicting intentional harm.
- Ratings of empathic sadness for the victim of intentional harm correlated with the neurohemodynamic activity in the insula, thalamus, and subgenual prefrontal cortex.
- Judgments of wrongness did not change across age; all participants rated intentional harm as more wrong than accidental harm.
- Whereas young children considered all agents malicious, irrespective of intention and targets (people and objects), older participants viewed perpetrators as clearly less mean when carrying out accidental actions, particularly when the target was an object.
- With increasing age, participants were less willing to punish agents who intentionally damaged objects than those who intentionally harmed people. The more participants punished someone for an action, the greater the activity in the orbitofrontal cortex and right hippocampus.

- Amadio DM, Frith CD. 2006. Meeting of minds: the medial frontal cortex and social cognition. *Nat Rev Neurosci.* 7:268–277.
- Anderson SW, Bechara A, Damasio H, Tranel D, Damasio AR. 1999. Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nat Neurosci.* 2:1032–1037.
- Bauman MD, Bliss-Moreau E, Machado CJ, Amaral DG. 2011. The neurobiology of primate social behavior. In: Decety J, Cacioppo JT, editors. *The oxford handbook of social neuroscience.* New York: Oxford University Press. p. 683–701.
- Bechara A, Tranel D, Damasio H, Damasio AR. 1996. Failure to respond autonomically to anticipated future outcomes following damage to the prefrontal cortex. *Cereb Cortex.* 6:215–225.
- Blakemore SJ, Boyer P, Pachot-Clouard M, Meltzoff A, Segebarth C, Decety J. 2003. The detection of contingency and animacy from simple animations in the human brain. *Cereb Cortex.* 13:837–844.
- Blakemore SJ, Decety J. 2001. From the perception of action to the understanding of intention. *Nat Rev Neurosci.* 2:561–567.
- Blair RJR. 1995. A cognitive developmental approach to morality: investigating the psychopath. *Cognition.* 57:1–29.
- Blair RJR. 1997. Moral reasoning and the child with psychopathic tendencies. *Pers Individ Dif.* 22:731–739.
- Blair RJR. 2009. Neuro-cognitive systems involved in moral reasoning. In: Verplaetse J, De Schrijver J, Vanneste S, Braeckman J, editors. *The moral brain.* New York: Springer. p. 87–105.
- Blair RJR, Blair KS. 2009. Empathy, morality, and social convention: evidence from the study of psychopathy and other psychiatric disorders. In: Decety J, Ickes W, editors. *The social neuroscience of empathy.* Cambridge (MA): MIT press. p. 139–152.
- Bramham J, Norris RG, Hornak J, Bullock P, Polkey CE. 2009. Social and emotional functioning following bilateral and unilateral neurosurgical prefrontal cortex lesions. *J Neuropsychol.* 3:125–143.
- Carpenter M, Akhtar M, Tomasello M. 1998. Fourteen- through 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behav Dev.* 21:315–330.
- Critchley HD, Melmed RN, Featherstone E, Mathias CJ, Dolan RJ. 2002. Volitional control of autonomic arousal: a functional magnetic resonance study. *Neuroimage.* 16:909–919.
- Cushman F. 2008. Crime and punishment: distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition.* 108:353–380.
- Damasio AR. 1994. *Descartes' error.* New York: Putnam Publishing.
- Decety J. 2010. The neurodevelopment of empathy in humans. *Dev Neurosci.* 32:257–267.

- Decety J. 2011a. Dissecting the neural mechanisms mediating empathy. *Emotion Rev.* 3:92-108.
- Decety J. Forthcoming 2011b. The neuroevolution of empathy. *Ann N Y Acad Sci*.
- Decety J, Lamm C. 2007. The role of the right temporoparietal junction in social interaction: how low-level computational processes contribute to meta-cognition. *Neuroscientist*. 13:580-593.
- Decety J, Michalska KJ. 2010. Neurodevelopmental changes in the circuits underlying empathy and sympathy from childhood to adulthood. *Dev Sci.* 13:886-899.
- Decety J, Michalska KJ, Akitsuki Y. 2008. Who caused the pain? A functional MRI investigation of empathy and intentionality in children. *Neuropsychologia*. 46:2607-2614.
- Decety J, Michalska KJ, Kinzler KD. Forthcoming 2011. The developmental neuroscience of moral sensitivity. *Emot Rev*.
- Decety J, Porges EC. Forthcoming 2011. Imagining being the agent of actions that carry different moral consequences: an fMRI study. *Neuropsychol*.
- De Haan M, Gunnar MR. 2009. The brain in a social environment. Why study development? In: De Haan M, Gunnar MR, editors. *Handbook of developmental social neuroscience*. New York: The Guilford Press. p. 3-10.
- de Wied M, Matthys W, Goudena P. 2007. Bryant's empathy index: a closer examination of its internal structure. *Eur J Psychol Assess*. 23:99-104.
- Drevets WC, Price JL, Simpson JR, Todd RD, Reich T, Vannier M, Raichle ME. 1997. Subgenual prefrontal cortex abnormalities in mood disorders. *Nature*. 386:824-827.
- Eisenberg N. 2000. Emotion, regulation, and moral development. *Ann Rev Psychol*. 51:665-697.
- Eisenberg N, Eggum ND. 2009. Empathic responding: sympathy and personal distress. In: Decety J, Ickes W, editors. *The social neuroscience of empathy*. Cambridge (MA): MIT Press. p. 71-83.
- Eslinger PJ, Robinson-Long M, Realmuto J, Moll J, de Oliveira-Souza R, Tovar-Moll F, Wang J, Yang QX. 2009. Developmental frontal lobe imaging in moral judgment: Arthur Benton's enduring influence 60 years later. *J Clin Exp Neuropsychol*. 31:158-169.
- Genovese CR, Lazar NA, Nichols T. 2002. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage*. 15:870-878.
- Glässcher J. 2009. Visualization of group inference data in functional neuroimaging. *Neuroinformatics*. 7:73-82.
- Gogtay N, Giedd JN, Lusk L, Hayashi KM, Greenstein D, Vaituzis AC, Nugent TM, Herman DH, Clasen LS, Toga AW, et al. 2004. Dynamic mapping of human cortical development during childhood through early adulthood. *Proc Natl Acad Sci U S A*. 101:8174-8179.
- Greene J, Haidt J. 2002. How (and where) does moral judgment work? *Trends Cogn Sci*. 12:517-523.
- Greene J, Sommerville RB, Nystrom LE, Daley JM, Cohen JD. 2001. An fMRI investigation of emotional engagement in moral judgment. *Science*. 293:2105-2108.
- Hamlin JK, Wynn K, Bloom P. 2007. Social evaluation by preverbal infants. *Nature*. 450:557-559.
- Harenski CL, Hamann S. 2006. Neural correlates of regulating negative emotions related to moral violations. *Neuroimage*. 30:313-324.
- Hempel J. 2009. Eye-tracking as a method to investigate empathy and sympathy. Honors Thesis under the direction of Dr. J. Decety. University of Chicago, Chicago, IL, USA.
- Huebner B, Dwyer S, Hauser M. 2009. The role of emotion in moral psychology. *Trends Cogn Sci*. 13:1-6.
- Jackson PL, Meltzoff AN, Decety J. 2005. How do we perceive the pain of others: a window into the neural processes involved in empathy. *Neuroimage*. 24:771-779.
- Jackson PL, Rainville P, Decety J. 2006. To what extent do we share the pain of others? Insight from the neural bases of pain empathy. *Pain*. 125:5-9.
- Kahn PH. 1992. Children's Obligatory and discretionary moral judgments. *Child Development*. 63:416-430.
- Killen M, Mulvey KL, Richardson C, Jampol N, Woodward A. Forthcoming 2011. The accidental transgressor: morally relevant theory of mind. *Cognition*.
- Killen M, Smetana JG. 2006. *Handbook of moral development*. Mahwah (NJ): Lawrence Erlbaum Associates, publishers.
- Killgore WDS, Yurgelun-Todd DA. 2007. Unconscious processing of facial affect in children and adolescents. *Soc Neurosci*. 2:28-47.
- Kohlberg L. 1984. Essays in moral development, volume I: The psychology of moral development. New York: Harper and Row.
- Kong J, Tu P, Zyloney C, Su T. 2010. Intrinsic functional connectivity of the periaqueductal gray, a resting fMRI study. *Behav Brain Res*. 211:215-219.
- Lamm C, Batson CD, Decety J. 2007. The neural substrate of human empathy: effects of perspective-taking and cognitive appraisal. *J Cogn Neurosci*. 19:42-58.
- Lamm C, Decety J. 2008. Is the extrastriate body area (EBA) sensitive to the perception of pain in others? *Cereb Cortex*. 18:2369-2373.
- Lamm C, Decety J, Singer T. 2011. Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *Neuroimage*. 54:2492-2502.
- Lamm C, Meltzoff AN, Decety J. 2010. How do we empathize with someone who is not like us? *J Cogn Neurosci*. 2:362-376.
- Lamm C, Nusbaum H, Meltzoff AN, Decety J. 2007. What are you feeling? Using functional magnetic resonance imaging to assess the modulation of sensory and affective responses during empathy for pain. *PLoS ONE*. 12:e1292.
- Leslie A, Knobe J, Cohen A. 2006. Acting intentionally and the side-effect: theory of mind and moral judgment. *Psychol Sci*. 17:421-427.
- Malti T, Latzko B. 2010. Children's moral emotions and moral cognition: towards an integrative perspective. *New Dir Child Adolesc Dev*. 129:1-10.
- Meltzoff AN. 1995. Understanding the intentions of others: re-enactment of intended acts by 18-month-old children. *Dev Psychol*. 31:838-850.
- Moll J, de Oliveira-Souza R, Bramati IE, Grafman J. 2002. Functional networks in emotional moral and non-moral social judgments. *Neuroimage*. 16:696-703.
- Moll J, de Oliveira-Souza R, Eslinger P. 2003. Morals and the human brain. *Neuroreport*. 14:299-305.
- Moll J, de Oliveira-Souza R, Garrido GJ, Bramati IE, Caparelli-Daquer EMA, Paiva ML, Zahn R, Grafman J. 2007. The self as a moral agent: linking the neural bases of social agency and moral sensitivity. *Soc Neurosci*. 2(3-4):336-352.
- Moll J, Zahn R, de Oliveira-Souza R, Krueger F, Grafman J. 2005. The neural basis of human moral cognition. *Nat Rev Neurosci*. 4:799-809.
- Narvaez D. 2008. Triune ethics: the neurobiological roots of our multiple moralities. *New Ideas Psychol*. 26:95-119.
- Norris CJ, Chen EE, Zhu DC, Small S, Cacioppo JT. 2004. The interaction of social and emotional processes in the brain. *J Cogn Neurosci*. 16:1818-1829.
- Paus T. 2011. Brain development during childhood and adolescence. In: Decety J, Cacioppo JT, editors. *The oxford handbook of social neurosci*. New York: Oxford University Press. p. 293-1213.
- Pelphrey KE, Carter EJ. 2008. Charting the typical and atypical development of the social brain. *Dev Psychopathol*. 20:1081-1102.
- Pettit D, Knobe J. 2009. The pervasive impact of moral judgment. *Mind Lang*. 24:586-604.
- Porro CA, Cettolo V, Francescato MP, Baraldi P. 2003. Functional activity mapping of the mesial hemispheric wall during anticipation of pain. *Neuroimage*. 19:1738-1747.
- Ruby P, Decety J. 2003. What you believe versus what you think they believe? A neuroimaging study of conceptual perspective taking. *Eur J Neurosci*. 17:2475-2480.
- Ruby P, Decety J. 2004. How would you feel versus how do you think she would feel? A neuroimaging study of perspective taking with social emotions. *J Cogn Neurosci*. 16:988-999.
- Schilbach L, Eickhoff SB, Mojzisch A, Vogeley K. 2008. What's in a smile? Neural correlates of facial embodiment during social interaction. *Soc Neurosci*. 3:37-50.
- Shamay-Tsoory SG, Tibi-Elhanamy Y, Aharon-Peretz J. 2006. Ventromedial prefrontal cortex is involved in understanding affective but not cognitive theory of mind stories. *Soc Neurosci*. 1(3-4):149-166.
- Shaw P, Lawrence EJ, Radbourne C, Bramham J, Polkey CE, David AS. 2004. The impact of early and late damage to the human amygdala on theory of mind reasoning. *Brain*. 127:1535-1548.

- Smetana JG, Killen M. 2008. Moral cognition, emotions, and neuroscience: an integrative developmental view. *Eur J Dev Sci.* 2:324-339.
- Stone VE, Baron-Cohen S, Calder A, Keane J, Young A. 2003. Acquired theory of mind impairments in individuals with bilateral amygdala lesions. *Neuropsychologia.* 41:209-220.
- Svetlova M, Nichols SR, Brownell CA. 2010. Toddlers' prosocial behavior: from instrumental to empathic to altruistic helping. *Child Dev.* 81:1814-1827.
- Swedene JK. 2005. Feeling better about moral dilemmas. *J Moral Edu.* 34:43-55.
- Vaish A, Carpenter M, Tomasello M. 2010. Children selectively avoid helping people with harmful intentions. *Child Dev.* 81:1661-1665.
- Vogt BA, Berger GR, Derbyshire SW. 2003. Structural and functional dichotomy of human midcingulate cortex. *Eur J Neurosci.* 18:3134-3144.
- Vul E, Harris C, Winkielman P, Pashler H. 2009. Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspect Psychol Sci.* 4:274-290.
- Warneken F, Tomasello M. 2009. The roots of human altruism. *Br J Psychol.* 100:455-471.
- Woodward AL. 1998. Infants selectively encode the goal object of an actor's reach. *Cognition.* 69:1-34.
- Woodward AL. 1999. Infants' ability to distinguish between purposeful and non-purposeful behaviors. *Infant Behav Dev.* 22:145-160.
- Woodward AL, Sommerville JA. 2000. Twelve-month-old infants interpret action in context. *Psychol Sci.* 11:73-77.
- Young L, Dungan J. Forthcoming 2011. Where in the brain is morality? Everywhere and maybe nowhere. *Soc Neurosci.*
- Young L, Bechara A, Tranel D, Damasio H, Hauser M, Damasio A. 2010. Damage to ventromedial prefrontal cortex impairs judgment of harmful intent. *Neuron.* 65:1-7.
- Young L, Cushman F, Hauser M, Saxe R. 2007. The neural basis of the interaction between theory of mind and moral judgment. *Proc Natl Acad Sci U S A.* 104:8235-8240.
- Young L, Saxe R. 2008. The neural basis of belief encoding and integration in moral judgment. *Neuroimage.* 40:1912-1920.
- Yuill N, Perner J. 1988. Intentionality and knowledge in children's judgments of actor's responsibility and recipient's emotion reaction. *Dev Psychol.* 24:358-365.
- Zahn-Waxler C, Radke-Yarrow M. 1990. The origins of empathic concern. *Motiv Emot.* 14:107-130.
- Zahn-Waxler C, Radke-Yarrow M, Wagner E, Chapman M. 1992. Development of concern for others. *Dev Psychol.* 28:126-136.
- Zelazo PD, Helwig CC, Lau A. 1996. Intention, act and outcome in behavioral prediction and moral judgement. *Child Dev.* 67: 2478-2492.