Neuron
Clinical Study

Damage to Ventromedial Prefrontal Cortex Impairs Judgment of Harmful Intent

Liane Young,[1,*] Antoine Bechara,[2] Daniel Tranel,[3] Hanna Damasio,[2] Marc Hauser,[4] and Antonio Damasio[2]
[1]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[2]Brain and Creativity Institute and Dornsife Center for Cognitive Neuroimaging, University of Southern California, Los Angeles, CA 90089, USA
[3]Departments of Neurology and Psychology, University of Iowa, Iowa City, IA 52242, USA
[4]Departments of Psychology and Human Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA
*Correspondence: lyoung@mit.edu
DOI 10.1016/j.neuron.2010.03.003

SUMMARY

Moral judgments, whether delivered in ordinary experience or in the courtroom, depend on our ability to infer intentions. We forgive unintentional or accidental harms and condemn failed attempts to harm. Prior work demonstrates that patients with damage to the ventromedial prefrontal cortex (VMPC) deliver abnormal judgments in response to moral dilemmas and that these patients are especially impaired in triggering emotional responses to inferred or abstract events (e.g., intentions), as opposed to real or actual outcomes. We therefore predicted that VMPC patients would deliver abnormal moral judgments of harmful intentions in the absence of harmful outcomes, as in failed attempts to harm. This prediction was confirmed in the current study: VMPC patients judged attempted harms, including attempted murder, as more morally permissible relative to controls. These results highlight the critical role of the VMPC in processing harmful intent for moral judgment.

INTRODUCTION

When we attempt to understand and evaluate other people's actions, we often draw inferences about their beliefs and intentions (Cushman, 2008; Knobe, 2005; Mikhail, 2007; Young et al., 2007). For example, did they believe they would cause harm? Did they intend to cause harm? Typically, these beliefs and intentions match the action's outcomes: when someone thinks she is sweetening her friend's coffee by putting sugar in it, she is usually not mistaken (Young and Saxe, 2009a). Mismatches occur, however, in the case of accidents (e.g., when the "sugar" is in fact poison) and failed attempts to harm (e.g., when the "poison" is in fact sugar). The aim of the current study is to understand the causal role of the ventromedial prefrontal cortex (VMPC) for such moral judgments that rely on assessments of intent (Casebeer and Churchland, 2003; Gazzaniga, 2005; Haidt, 2007; Mikhail, 2007). Using a neuropsychological approach, we show that bilateral damage to the VMPC leads to moral judgments that largely neglect harmful intent, focusing instead on the outcome of the action (e.g., the moral judgment of a failed murder attempt as permissible). Consequently, we suggest that the VMPC plays an integral role in processing negatively valenced intentions for moral judgment.

Prior neuroimaging and neuropsychological evidence has suggested a role for the VMPC in evaluating harmful actions (Schaich Borg et al., 2006; Ciaramelli et al., 2007; Glenn et al., 2009; Greene et al., 2001, 2004; Harenski and Hamann, 2006; Heekeren et al., 2003; Koenigs et al., 2007; Luo et al., 2006; Mendez et al., 2005; Moll et al., 2002; Young and Saxe, 2009b). Specifically, the VMPC was robustly recruited when subjects evaluated emotionally salient harms to an individual that were intended as a means to maximize aggregate welfare, for example, pushing a person into the path of a trolley in order that his body stop the trolley from hitting five other people (Greene et al., 2001, 2004). Furthermore, patients with bilateral damage to the VMPC were more likely to deliver utilitarian moral judgments, that is, to endorse such harmful actions as appropriate, compared to brain damaged or healthy comparison participants (Ciaramelli et al., 2007; Koenigs et al., 2007; Mendez et al., 2005). These results were taken to indicate a causal role for emotional processing, as subserved by the VMPC, in evaluating harmful actions in this context (Young and Koenigs, 2007). This body of work, however, leaves open an important question that we seek to address in the current study: do VMPC patients endorse harmful actions because of a failure to process harmful outcomes or harmful intentions?

Here we probe moral judgment in patients with adult-onset bilateral damage to the VMPC using scenarios that critically disentangle the contributions of intentions and outcomes to moral judgment. By studying patients with damage to this region, we therefore directly investigate the causally necessary role of the VMPC in the processing of intentions and outcomes for moral judgment. We note that the current study also differs from the prior work in several methodological respects: (1) the presentation of more ordinary and perhaps familiar scenario settings (e.g., eating at a restaurant, driving home from work) rather than the somewhat contrived contexts previously tested (e.g., halting runaway trolleys, facing terrorists in a jungle); (2) a focus on third-person moral judgments, as opposed to hypothetical first-person action predictions (e.g., what would *you* do in this situation?); (3) a departure from moral *dilemmas* (i.e., competing norms and no clear socially or legally mandated answers) of stereotypical form (e.g., would you kill one to save many?). These methodological changes allow us to determine whether the role

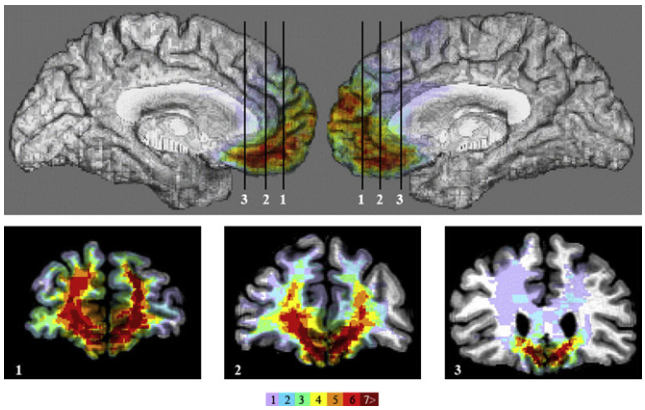

Figure 1. Lesion Overlap of the Nine VMPC Subjects Using the MAP-3 Technique

Top panel shows the left and right mesial views of the template brain. Panels 1–3 show three coronal sections through VMPC at the levels indicated in the top panel. The number of overlaps at each voxel is shown in the color bar.

Table 1. Demographic and Clinical Data

Participant	Age	Educ.	Hand.	Sex	Chronicity	Etiology
0318	69	14	+100	M	34	meningioma resection
0770	67	16	+100	F	24	meningioma resection
1424	73	13	+100	M	24	head trauma
1815	57	20	+100	M	11	meningioma resection
1983	46	13	+100	F	13	SAH; ACoA aneurysm
2352	60	14	+100	F	10	SAH; ACoA aneurysm
2391	63	13	+100	F	9	meningioma resection
2577	69	11	+100	M	10	SAH; ACoA aneurysm
3383	59	12	−100	F	3	SAH; ACoA aneurysm
VMPC						
Mean	62.6	14.0	8 RH	4 M	15.3	
SD	(8.2)	(2.6)	1 LH	5 F	(9.8)	
BDC (n = 7)						
Mean	62.4	16.6	7 RH	4 M	8.9	
SD	(9.5)	(3.0)	0 LH	3 F	(6.8)	
NC (n = 8)						
Mean	64.1	14.1	7 RH	5 M		
SD	(9.7)	(1.7)	1 LH	3 F		

Age, in years. Educ., years of formal schooling. Hand., degree of right- or left-handedness on a scale ranging from full right-handedness (+100) to full left-handedness (−100). Chronicity, years between lesion onset and current experiment. Etiology, cause of brain damage (SAH, subarachnoid hemorrhage; ACoA, anterior communicating artery). The seven brain-damaged comparison patients had brain damage caused by cerebrovascular disease. For Age and Education, there were no significant differences between the three groups, per one-way ANOVA. For Chronicity, the VMPC and BDC groups did not differ, per t test.

of the VMPC in moral judgment extends to (1) more ordinary contexts, (2) judgments as opposed to predictions of behavior, and (3) moral scenarios that feature pure transgressions (e.g., murder attempts) as opposed to moral dilemmas that force a choice between violations of competing moral norms (e.g., "the lesser of two evils").

We tested a sample of nine patients with adult-onset, focal bilateral VMPC lesions (Figure 1) and comparison groups of neurologically normal (NC) and brain-damaged (BDC) participants (Table 1; see Experimental Procedures). Based on prior neuropsychological testing, all of the VMPC patients in the current study exhibited characteristic deficits in social emotional processing (Table 2), while presenting generally intact intellect and cognitive function (Table S1). In general, despite preserved general intelligence, logical reasoning, and declarative knowledge of social and moral norms (Burgess et al., 2006; Saver and Damasio, 1991), patients with VMPC lesions commonly fail to apply such knowledge in daily living and exhibit impairments in processing social emotions such as empathy and embarrassment (Anderson et al., 2006; Barrash et al., 2000; Beer et al., 2003; Camille et al., 2004), as well as counterfactual emotional responses such as guilt and regret (Camille et al., 2004; Krajbich et al., 2009). Other work has demonstrated that VMPC patients are specifically impaired in triggering emotional responses when they must infer an emotional event (Bechara et al., 1997; Camille et al., 2004), as opposed to when they are presented with an actual emotional outcome (e.g., losing money), in which case their emotional responses are relatively spared or even exaggerated (Bechara et al., 2000; Koenigs and Tranel, 2007). This neuropsychological profile is best understood in the context of the functional connectivity of the VMPC. The VMPC projects to the basal forebrain and brainstem regions, which regulate and execute bodily components of emotional responses (Ongür and Price, 2000), while neurons within the VMPC encode the emotional value of stimuli (Rolls, 2000).

Scenarios presented to participants followed a 2 × 2 design (see Figure 2 and Supplemental Information for full text): (1) the protagonist either intended to cause harm to another person (negative intent) or intended to cause no harm (neutral intent), and (2) the protagonist either caused harm to another person (negative outcome) or caused no harm (neutral outcome) (Young et al., 2007). More precisely, the stimuli explicitly specified the agent's *belief* about whether he or she would cause harm, and, on this basis, participants could infer the agent's *intention* to cause harm or not. This design contained two conditions where intentions and outcomes matched and two where they mismatched (i.e., attempted harms and accidental harms). Participants made moral judgments of the protagonist's action on a scale of 1 (morally forbidden) to 7 (morally permissible).

Given the critical role of the VMPC in triggering emotional responses to inferred or abstract events (Bechara et al., 1997; Damasio et al., 1990), we predicted that patients with VMPC damage would fail to perceive the emotional significance of harmful intentions (e.g., unobservable mental states) and therefore deliver abnormal moral judgments in the case that judgments depend on emotional responses to such abstract representational content. We predicted that, as a direct result, VMPC patients would instead judge actions primarily on the basis of the actions' outcomes, which are represented concretely in the world. In particular, we predicted that patients with VMPC damage would judge attempted harms as more

Table 2. Emotional and Social Functioning Data for VMPC Patients

Patient	SCRs	Social Emotions	Acquired Personality Changes
0318	lower SCR	diminished (3)	yes (3)
0770	lower SCR	diminished (3)	yes (3)
1424	lower SCR	diminished (2)	yes (2)
1815	lower SCR	diminished (2)	yes (2)
1983	lower SCR	diminished (3)	yes (3)
2352	lower SCR	diminished (2)	yes (3)
2391	lower SCR	diminished (3)	yes (2)
2577	lower SCR	diminished (3)	yes (3)
3383	lower SCR	diminished (3)	yes (3)

Skin conductance responses (SCRs) to emotionally charged social stimuli (e.g., pictures of social disasters, mutilations, and nudes, using methods described previously [Damasio et al., 1990]). None of the seven brain-damaged comparison patients had SCR impairments to emotionally charged stimuli. Social Emotions, the patient's demonstrated capacity for empathy, embarrassment, and guilt, as determined from reports from a collateral source (spouse or family member) provided on the Iowa Scales of Personality Change (Barrash et al., 2000) and from data from clinical interviews. Acquired Personality Changes, postlesion changes in personality (e.g., irritability, emotional dysregulation, and impulsivity), as determined from data from the Iowa Scales of Personality Change. For Social Emotions and Acquired Personality Changes, the degree of severity is designated in parentheses (1, mild; 2, moderate; 3, severe). None of the seven brain-damaged comparison patients had defective social emotions or postmorbid personality changes.

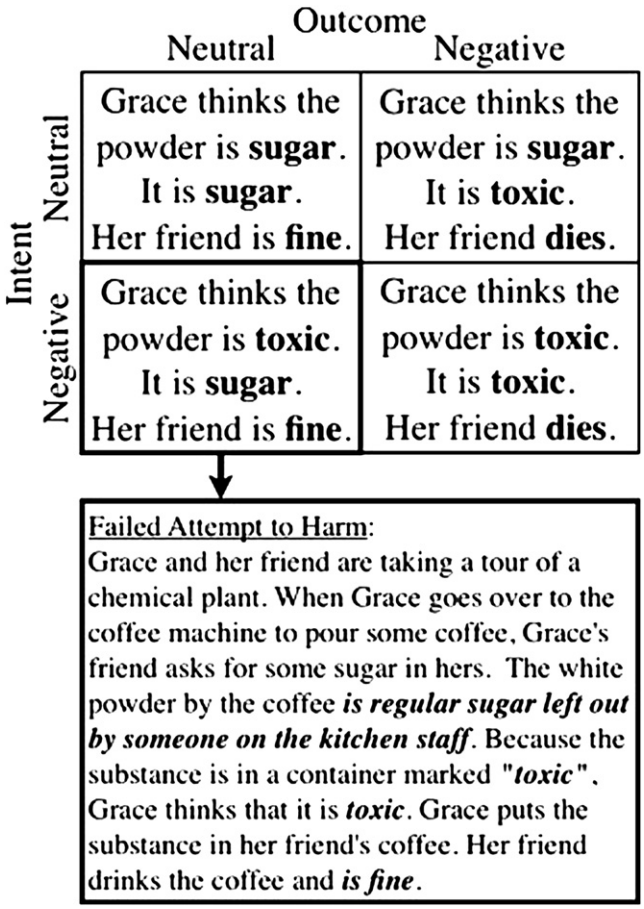

Figure 2. Experimental Design and Stimuli

(Top) The combination of intent (neutral versus negative) and outcome (neutral versus negative) factors yielded a 2 × 2 design with four conditions. (Bottom) Full text of an example "failed attempt to harm" scenario. Bold sections indicate words that differed across conditions.

morally permissible than control participants and, consequently, use the neutral outcome as the relevant moral metric. Notably, moral judgment of accidental harms (neutral intent, negative outcome) also requires the processing of an unobservable mental state; however, in this case, the mental state is a neutral intent, which does not necessarily elicit an emotional response that is critical for moral judgment. We therefore predicted that VMPC patients would show a selective deficit only when moral judgment requires an emotional response to mental state content. In other words, we predicted a deficit for attempted harms, not accidental harms. This pattern of results would indicate that, in the absence of a normally functioning VMPC and normal emotional responses subserved by the VMPC that are typically associated with perceiving harmful intentions, individuals will deliver abnormal moral judgments.

RESULTS

A 2 (intent: neutral versus negative) × 2 (outcome: neutral versus negative) × 3 (group: VMPC versus BDC versus NC) mixed-effects ANOVA of participants' moral judgments yielded main effects of intent ($F(1,21) = 136.0$ $p = 1.2 \times 10^{-10}$), outcome ($F(1,21) = 94.4$ $p = 3.2 \times 10^{-9}$), and an interaction between intent and outcome ($F(1,21) = 7.0$ $p = 0.015$) (Figure 3). Importantly, these effects were observed in the context of interaction effects involving the participant group variable, specifically, a two-way interaction between intent and participant group ($F(1,21) = 9.7$ $p = 0.001$) and a three-way interaction between intent, outcome, and participant group ($F(1,21) = 3.9$ $p = 0.036$). There were no statistically significant interaction effects involving the participant group variable for reaction time (intent × participant group, $F(1,21) = 1.4$ $p = 0.27$; belief × outcome by participant group, $F(1,21) = 0.50$ $p = 0.61$; see also Supplemental Analyses).

To interpret these interaction effects, planned comparisons were conducted, yielding significant differences between participant groups only for attempted harms. VMPC participants judged attempted harms as more permissible than BDC participants ($t(14) = 4.0$, $p = 0.001$) and NC participants ($t(15) = 4.6$, $p = 3.3 \times 10^{-4}$). There was no difference between BDC and NC participants in their moral judgments of attempted harms ($t(13) = 0.73$, $p = 0.48$) or any other condition. Moreover, there were no other significant differences for any pair of participant groups (VMPC, BDC, NC) on any of the other conditions: non-harm, accidental harm, or successful attempt to harm. Importantly, there were no differences between VMPC participants and either comparison group on non-harms (BDC: $t(14) = -0.40$, $p = 0.70$; NC: $t(15) = -0.21$, $p = 0.84$), accidental harms (BDC: $t(14) = -0.16$, $p = 0.89$; NC: $t(15) = -0.71$, $p = 0.49$), or successful

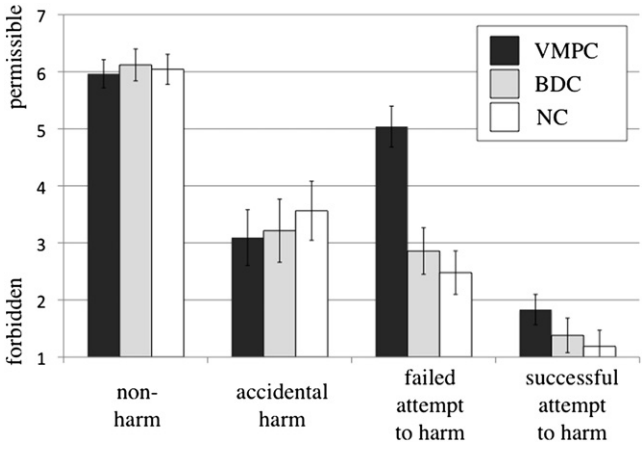

Figure 3. Moral Judgments for All Four Conditions

Judgments are shown for each participant group, on a seven-point scale. Error bars represent standard error of the mean. VMPC participants judged failed attempts to harm as significantly more permissible than the brain-damaged comparison (BDC) participants and the normal comparison (NC) participants (p values < 0.001).

attempts to harm (BDC: t(14) = 0.94, p = 0.37; NC: t(15) = 1.6, p = 0.13).

VMPC participants' judgments did reflect a difference between attempted harms and non-harms (t(8) = 2.97, p = 0.018) and a difference between accidental harms and successful attempts to harm (t(8) = 6.2, $p = 2.5 \times 10^{-4}$). Thus, VMPC participants were able to distinguish between these conditions by representing the content of negative beliefs and intentions. The difference between attempted harms and non-harms also emerged in the NC group (t(6) = 7.3, $p = 3.5 \times 10^{-4}$) and the BDC group (t(7) = 12.7, $p = 4.5 \times 10^{-6}$), as did the difference between accidental harms and successful attempts to harm (NC: t(6) = 2.7, p = 0.038, BDC: (t(7) = 4.9, p = 0.002).

Notably, VMPC participants also judged attempted harms as significantly more permissible than accidental harms (t(8) = 3.7, p = 0.006), a pattern that was significantly different from the pattern observed in the BDC participant group ($F(1,14) = 5.3$ p = 0.037) and the NC participant group ($F(5,10) = 12.0$ p = 0.003). Moral judgments of accidental and attempted harms in the BDC and NC groups reflected a difference in the opposite direction, though this difference did not reach significance (combined analysis for BDC and NC groups: t(14) = 1.3, p = 0.2). Strikingly, all nine VMPC participants showed the same reversal of judgments of attempted and accidental harms; this pattern was significantly different from the pattern of judgments in the BDC and NC participant groups (Kruskal-Wallis test, H = 8.3, 2 d.f., p = 0.016). Furthermore, this difference was significant for the comparison between both VMPC and BDC participants (Mann-Whitney U test, U = 13.5, p = 0.01), and between VMPC and NC participants (U = 13.5, p = 0.006).

DISCUSSION

The primary aim of this study was to examine the causal role of the VMPC in specific aspects of moral judgment: processing intentions versus outcomes. Because the emotional valence of an intention (e.g., negative versus neutral) greatly influences normal moral judgments (e.g., negative intentions are judged as immoral) and because neuropsychological studies of these VMPC patients reveal deficits in emotional processing, we predicted that VMPC patients would show a selective neglect of negative intentions in moral judgment. The current results are consistent with this prediction: VMPC participants judged attempted harms as more morally permissible than both comparison groups. VMPC participants even judged attempted harms (e.g., attempting, but failing to poison someone) as more permissible than accidental harms (e.g., accidentally poisoning someone).

Notably, the pattern of moral judgments delivered by the VMPC patients represents not just a departure from but also a reversal of the normal pattern of moral judgments. Among healthy adults and even young children, attempted harms are generally judged quite harshly and usually more harshly than accidental harms (Cushman, 2008; Piaget, 1965). In contrast, when VMPC patients confront the same cases, they neglect the protagonist's negative intention, focusing instead on the action's neutral outcome. This results in unusually lenient moral judgments of failed attempts to harm. Importantly, VMPC participants did not exhibit a global deficit in moral judgment in judging all actions as either more permissible or more forbidden. Instead, their deficit was highly selective, restricted to the context of attempted harms.

In conjunction with prior evidence (Bechara et al., 1997; Beer et al., 2003; Damasio et al., 1990), we suggest that the current pattern of results may be due to impaired emotional processing, subserved by the VMPC. That is, the results are consistent with the possibility that when VMPC participants encounter a failed attempt to harm, they may not experience the aversive emotions that normally arise from perceiving that one person intends to harm another. More specifically, due to a deficit in triggering emotions in response to inferred, abstract, imagined, or recalled events, previously termed as secondary emotion induction (Bechara and Damasio, 2005), VMPC patients may fail to respond appropriately to an agent's intention to cause harm. We note that in the current study this information is both abstract, insofar as mental state representations are abstract representations, and inferred, insofar as the agent's intention is inferred from the agent's belief that he or she would cause harm. Indeed, future research ought to characterize in further cognitive detail the dimensions of representational content that fails to elicit appropriate emotional responding in VMPC patients, in both moral and nonmoral contexts. Engaging an emotional response to harmful intent may normally lead to judging attempted harms as morally forbidden (Valdesolo and DeSteno, 2006; Wheatley and Haidt, 2005). We suggest that VMPC patients may lack this guiding emotional response (Koenigs et al., 2007; Saver and Damasio, 1991). VMPC patients may therefore rely instead on explicit outcome information to formulate their moral judgments. Because failed attempts to harm result in neutral outcomes (e.g., no harm), VMPC patients judge failed attempts as more permissible. By the same logic, VMPC patients judge successful attempts to harm as forbidden, on the basis of negative outcomes. This pattern is therefore consistent with intact

processing of outcome information in VMPC patients, but impaired processing of emotional aspects of intention for moral judgment.

In the current study, we did not measure emotional responding during the moral judgment task itself. However, the characteristics of the emotional deficit exhibited in these patients, i.e., impaired emotional responding to inferred events (''secondary induction''), but not actual outcomes (''primary induction''), have been studied and documented over almost two decades of research (for a review, see Bechara and Damasio, 2005). Although a specific impairment in triggering emotions from inferred or abstract events is the most parsimonious explanation for the observed results given the available evidence, here we consider two alternative hypotheses for the pattern of judgments provided by the VMPC participants.

First, VMPC participants may have produced an abnormal pattern of moral judgments because of deficits in domain-general cognitive abilities, rather than social-emotional deficits. This alternative hypothesis appears unlikely for a number of reasons. VMPC patients, including the ones we tested, showed preserved general intelligence, logical reasoning, and declarative knowledge of social and moral norms (Burgess et al., 2006; Saver and Damasio, 1991). Consistent with this neuropsychological profile, these patients also provided normal moral judgments on all but one condition (i.e., attempted harms) in the current study and showed no reaction time differences as compared to either control group on any condition. Furthermore, the attempted harm condition was not more difficult for any group, as indicated by reaction time. Given the VMPC participants' cognitive profile, as well as their performance on the current task, it is unlikely that their selective deficit on attempted harms is due to generic cognitive deficits.

A second alternative hypothesis is that VMPC participants' performance on the moral judgment task may be attributed to a deficit in basic theory of mind or false belief understanding. In other words, damage to the VMPC in the current participants could have resulted in a deficit in attributing intentions across all conditions. This alternative hypothesis also appears unlikely because VMPC participants did not make abnormal moral judgments across all conditions. Instead, VMPC participants showed a selective deficit for attempted harms. Importantly, VMPC participants exhibited normal performance on accidental harms and distinguished accidental harms from successful attempts to harm. Moral judgments of successful attempts to harm could be made on the basis of outcome information alone. However, moral judgments of accidental harms require attributing beliefs and intentions. VMPC participants' moral judgments of accidental harms may therefore reflect intact processing of neutral intentions as well as negative outcomes. Indeed, VMPC participants were even able to discriminate between attempted harms and non-harms, suggesting an intact capacity to represent the specific content of negative beliefs and intentions. VMPC participants' selective failure on attempted harms cannot therefore be due to a deficit in representing the content of either a negative or a neutral mental state, a belief or an intention. In light of the current pattern of results, as well as prior work on the role of the VMPC in emotional processing, we suggest instead that VMPC participants' abnormal responding to attempted harms may be mediated by a specific deficit in triggering a sufficiently robust emotional response to these representations, in this case, an aversive response to harmful intent (Bechara et al., 1997; Beer et al., 2003; Damasio et al., 1990). While we did not measure VMPC participants' theory of mind or false belief understanding outside the moral judgment task, nor did we measure explicit intention understanding during the task, the full pattern of results suggests that VMPC patients are not impaired in basic theory of mind.

Prior evidence has suggested a specific role for the VMPC in processing affective aspects of another person's mental states (Jenkins and Mitchell, 2009; Mitchell et al., 2006; Shamay-Tsoory and Aharon-Peretz, 2007; Völlm et al., 2006). The current finding that the VMPC is associated with processing intentions with high emotional content, i.e., negative intentions, for moral judgment, is consistent with the role of the VMPC in ''affective'' or ''hot'' theory of mind (Jenkins and Mitchell, 2009; Mitchell et al., 2006; Shamay-Tsoory and Aharon-Peretz, 2007; Völlm et al., 2006). The current results are also consistent with a recent fMRI finding of a selective positive correlation between the average response in the VMPC and moral judgments of attempted harms (Young and Saxe, 2009b). Healthy adult participants with a high VMPC response assigned more moral blame to agents for harmful intentions, in the absence of any actual harmful outcome. Together, these results support the significance of the VMPC in moral judgments of harmful intentions and therefore attempted harms. We note, though, that the VMPC targeted in neuroimaging and neuropsychological work spans a large cortical region; future work is therefore needed in order to further elucidate the functional organization of the VMPC and its contribution to different aspects of the decision-making process.

A fundamental component of normal moral judgment is the ability to blame those who intend harm, even when they fail to cause harm. We recognize failed attempts to harm as deserving of moral blame; failed attempts represent instances in which we might even be motivated to punish at a cost to ourselves (Cushman et al., 2009; de Quervain et al., 2004; Moll et al., 2006). In fact, the ability to blame for failed attempts not only features prominently in mature moral judgments but emerges quite early in development: typically developing children use mental state information (i.e., harmful intent) to assign blame for attempted harms, well before they are able to use mental state information (i.e., neutral intent) to mitigate blame for accidental harms (Baird and Astington, 2004; Piaget, 1965). Thus, while the standard challenge for healthy children and adults lies in forming exculpatory moral judgments or forgiveness (e.g., judging accidental harms as morally permissible on the basis of agents' neutral intentions), the opposite seems to hold true in the case of VMPC damage. Attributing moral blame even for failed murder attempts therefore poses a unique challenge for VMPC patients.

The current results reveal an important aspect of VMPC function for moral judgment, specifically, its role in evaluating harmful intent. In conjunction with prior work on the role of the VMPC in emotional processing, these results further suggest that an emotional response to harmful intent is crucial for condemning failed attempts. Given the critical role of intent in moral judgment,

and social cognition more generally, understanding the neural basis of how intent is processed will be essential in helping us understand human moral judgment.

EXPERIMENTAL PROCEDURES

Subjects

Nine patients with bilateral, adult-onset damage to the VMPC and seven brain-damaged comparison patients who had lesions that excluded structures thought to be important for emotions (VMPC, amygdala, insula, right somatosensory cortices) were recruited from the Patient Registry of the Division of Cognitive Neuroscience at the University of Iowa. Five of these nine VMPC participants were previously tested on a moral dilemmas task, described above (Koenigs et al., 2007); see Supplemental Analyses. Eight healthy comparison subjects with no brain damage were recruited from the Iowa community. Groups were age, gender, and ethnicity matched. All participants gave written informed consent.

Neuroanatomical Analysis

All subjects had MR scans (3) or CT scans (6) obtained in the chronic epoch of their lesions, and all scans were reconstructed in three dimensions using Brainvox (Damasio and Frank, 1992; Frank et al., 1997). The lesions were analyzed on each individual scan. Subsequently, the contours of the nine target lesions were mapped onto a nonlesioned standard brain, using the MAP-3 technique (Damasio, 2005), to visualize the region of maximal overlap (Figure 1).

Stimuli and Task

We presented participants with 24 scenarios, selected from a previously published set (Young et al., 2007; Young and Saxe, 2008). There were four variations (conditions) of each scenario, following a 2 × 2 design: (1) protagonists either harmed another person (negative outcome) or did no harm (neutral outcome); (2) protagonists either believed they would cause harm (negative intent) or believed they would cause no harm (neutral intent). Each possible belief was true for one outcome and false for the other outcome; the agent held true beliefs in the all-neutral and all-negative conditions and false beliefs in the accidental harm and attempted harm conditions. Subjects saw one version of each scenario. Stimuli were presented in a pseudorandom order; conditions were counterbalanced across subjects. Subjects read six stimuli per each of the four conditions. Across subjects, every scenario occurred in each of the four conditions.

Word count was matched across conditions (mean ± SD for the all-neutral condition: 103 ± 10; accidental harm: 101 ± 9; attempted harm: 103 ± 10; intentional harm: 103 ± 9). On average, scenarios featuring negative beliefs contained the same number of words as scenarios featuring neutral beliefs ($F(1, 23) = 0.15$ $p = 0.70$, partial $h^2 = 0.006$); scenarios featuring negative outcomes contained the same number of words as scenarios featuring neutral outcomes ($F(1, 23) = 0.17$ $p = 0.68$, partial $h^2 = 0.007$).

We presented each story in four cumulative segments (previous segments remained on the screen when later segments were added): (1) background information to set the scene, (2) facts foreshadowing the eventual outcome, (3) the protagonist's belief (from which intent could be inferred), (4) the protagonist's action and its outcome. The question and response scale were then added to the screen. Participants made moral judgments of the protagonist's action on a scale of 1 (forbidden) to 7 (permissible), using a computer keyboard. Participants read and responded at their own pace, pressing the spacebar to add the next segment of the story and finally the question. There was no time limit for reading or responding.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, Results, Discussion, and one table and can be found with this article online at doi:10.1016/j.neuron.2010.03.003.

ACKNOWLEDGMENTS

We thank Rebecca Saxe and Josh Greene for their helpful comments on an earlier draft of this manuscript. This research was supported in part by a grant from the NSF-HSD (A.B., A.D., M.H.), NINDS (P01 NS19632) (A.D., H.D., D.T.), gift funds from J. Epstein and S. Shuman (M.H.), by NIDA (DA022549) (D.T.), and the NSF (L.Y.).

Accepted: February 23, 2010
Published: March 24, 2010

REFERENCES

Anderson, S.W., Barrash, J., Bechara, A., and Tranel, D. (2006). Impairments of emotion and real-world complex behavior following childhood- or adult-onset damage to ventromedial prefrontal cortex. J. Int. Neuropsychol. Soc. *12*, 224–235.

Baird, J.A., and Astington, J.W. (2004). The role of mental state understanding in the development of moral cognition and moral action. New Dir. Child Adolesc. Dev. *103*, 37–49.

Barrash, J., Tranel, D., and Anderson, S.W. (2000). Acquired personality disturbances associated with bilateral damage to the ventromedial prefrontal region. Dev. Neuropsychol. *18*, 355–381.

Bechara, A., and Damasio, A.R. (2005). The somatic marker hypothesis: A neural theory of economic decision. Games Econ. Behav. *52*, 336–372.

Bechara, A., Damasio, H., Tranel, D., and Damasio, A.R. (1997). Deciding advantageously before knowing the advantageous strategy. Science *275*, 1293–1295.

Bechara, A., Damasio, H., and Damasio, A.R. (2000). Emotion, decision making and the orbitofrontal cortex. Cereb. Cortex *10*, 295–307.

Beer, J.S., Heerey, E.A., Keltner, D., Scabini, D., and Knight, R.T. (2003). The regulatory function of self-conscious emotion: insights from patients with orbitofrontal damage. J. Pers. Soc. Psychol. *85*, 594–604.

Burgess, P.W., Alderman, N., Forbes, C., Costello, A., Coates, L.M., Dawson, D.R., Anderson, N.D., Gilbert, S.J., Dumontheil, I., and Channon, S. (2006). The case for the development and use of "ecologically valid" measures of executive function in experimental and clinical neuropsychology. J. Int. Neuropsychol. Soc. *12*, 194–209.

Camille, N., Coricelli, G., Sallet, J., Pradat-Diehl, P., Duhamel, J.R., and Sirigu, A. (2004). The involvement of the orbitofrontal cortex in the experience of regret. Science *304*, 1167–1170.

Casebeer, W., and Churchland, P. (2003). The neural mechanisms of moral cognition: A multiple-aspect approach to moral judgment and decision-making. Biol. Philos. *18*, 169–194.

Ciaramelli, E., Muccioli, M., Làdavas, E., and di Pellegrino, G. (2007). Selective deficit in personal moral judgment following damage to ventromedial prefrontal cortex. Soc. Cogn. Affect. Neurosci. *2*, 84–92.

Cushman, F. (2008). Crime and punishment: distinguishing the roles of causal and intentional analyses in moral judgment. Cognition *108*, 353–380.

Cushman, F., Dreber, A., Wang, Y., and Costa, J. (2009). Accidental outcomes guide punishment in a "trembling hand" game. PLoS ONE *4*, e6699.

Damasio, H. (2005). Human Brain Anatomy in Computerized Images, Second Edition (New York: Oxford University Press).

Damasio, H., and Frank, R. (1992). Three-dimensional in vivo mapping of brain lesions in humans. Arch. Neurol. *49*, 137–143.

Damasio, A.R., Tranel, D., and Damasio, H. (1990). Individuals with sociopathic behavior caused by frontal damage fail to respond autonomically to social stimuli. Behav. Brain Res. *41*, 81–94.

de Quervain, D.J., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., and Fehr, E. (2004). The neural basis of altruistic punishment. Science *305*, 1254–1258.

Frank, R.J., Damasio, H., and Grabowski, T.J. (1997). Brainvox: an interactive, multimodal visualization and analysis system for neuroanatomical imaging. Neuroimage *5*, 13–30.

Gazzaniga, M.S. (2005). The Ethical Brain (New York: Dana Press).

Glenn, A.L., Raine, A., and Schug, R.A. (2009). The neural correlates of moral decision-making in psychopathy. Mol. Psychiatry *14*, 5–6.

Greene, J.D., Sommerville, R.B., Nystrom, L.E., Darley, J.M., and Cohen, J.D. (2001). An fMRI investigation of emotional engagement in moral judgment. Science *293*, 2105–2108.

Greene, J.D., Nystrom, L.E., Engell, A.D., Darley, J.M., and Cohen, J.D. (2004). The neural bases of cognitive conflict and control in moral judgment. Neuron *44*, 389–400.

Haidt, J. (2007). The new synthesis in moral psychology. Science *316*, 998–1002.

Harenski, C.L., and Hamann, S. (2006). Neural correlates of regulating negative emotions related to moral violations. Neuroimage *30*, 313–324.

Heekeren, H.R., Wartenburger, I., Schmidt, H., Schwintowski, H.P., and Villringer, A. (2003). An fMRI study of simple ethical decision-making. Neuroreport *14*, 1215–1219.

Jenkins, A.C., and Mitchell, J.P. (2009). Mentalizing under uncertainty: Dissociated neural responses to ambiguous and unambiguous mental state inferences. Cereb. Cortex *20*, 404–410.

Knobe, J. (2005). Theory of mind and moral cognition: exploring the connections. Trends Cogn. Sci. *9*, 357–359.

Koenigs, M., and Tranel, D. (2007). Irrational economic decision-making after ventromedial prefrontal damage: evidence from the Ultimatum Game. J. Neurosci. *27*, 951–956.

Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., and Damasio, A. (2007). Damage to the prefrontal cortex increases utilitarian moral judgements. Nature *446*, 908–911.

Krajbich, I., Adolphs, R., Tranel, D., Denburg, N.L., and Camerer, C.F. (2009). Economic games quantify diminished sense of guilt in patients with damage to the prefrontal cortex. J. Neurosci. *29*, 2188–2192.

Luo, Q., Nakic, M., Wheatley, T., Richell, R., Martin, A., and Blair, R.J. (2006). The neural basis of implicit moral attitude—an IAT study using event-related fMRI. Neuroimage *30*, 1449–1457.

Mendez, M.F., Anderson, E., and Shapira, J.S. (2005). An investigation of moral judgement in frontotemporal dementia. Cogn. Behav. Neurol. *18*, 193–197.

Mikhail, J.M. (2007). Universal moral grammar: theory, evidence and the future. Trends Cogn. Sci. *11*, 143–152.

Mitchell, J.P., Macrae, C.N., and Banaji, M.R. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. Neuron *50*, 655–663.

Moll, J., de Oliveira-Souza, R., Eslinger, P.J., Bramati, I.E., Mourão-Miranda, J., Andreiuolo, P.A., and Pessoa, L. (2002). The neural correlates of moral sensitivity: a functional magnetic resonance imaging investigation of basic and moral emotions. J. Neurosci. *22*, 2730–2736.

Moll, J., Krueger, F., Zahn, R., Pardini, M., de Oliveira-Souza, R., and Grafman, J. (2006). Human fronto-mesolimbic networks guide decisions about charitable donation. Proc. Natl. Acad. Sci. USA *103*, 15623–15628.

Ongür, D., and Price, J.L. (2000). The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. Cereb. Cortex *10*, 206–219.

Piaget, J. (1965). The Moral Judgment of the Child (New York: Free Press).

Rolls, E. (2000). The orbitofrontal cortex and reward. Cereb. Cortex *3*, 284–294.

Saver, J.L., and Damasio, A.R. (1991). Preserved access and processing of social knowledge in a patient with acquired sociopathy due to ventromedial frontal damage. Neuropsychologia *29*, 1241–1249.

Schaich Borg, J., Hynes, C., Van Horn, J., Grafton, S., and Sinnott-Armstrong, W. (2006). Consequences, action, and intention as factors in moral judgments: an FMRI investigation. J. Cogn. Neurosci. *18*, 803–817.

Shamay-Tsoory, S.G., and Aharon-Peretz, J. (2007). Dissociable prefrontal networks for cognitive and affective theory of mind: a lesion study. Neuropsychologia *45*, 3054–3067.

Valdesolo, P., and DeSteno, D. (2006). Manipulations of emotional context shape moral judgment. Psychol. Sci. *17*, 476–477.

Völlm, B.A., Taylor, A.N., Richardson, P., Corcoran, R., Stirling, J., McKie, S., Deakin, J.F., and Elliott, R. (2006). Neuronal correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. Neuroimage *29*, 90–98.

Wheatley, T., and Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. Psychol. Sci. *16*, 780–784.

Young, L., and Koenigs, M. (2007). Investigating emotion in moral cognition: a review of evidence from functional neuroimaging and neuropsychology. Br. Med. Bull. *84*, 69–79.

Young, L., and Saxe, R. (2008). The neural basis of belief encoding and integration in moral judgment. Neuroimage *40*, 1912–1920.

Young, L., and Saxe, R. (2009a). An FMRI investigation of spontaneous mental state inference for moral judgment. J. Cogn. Neurosci. *21*, 1396–1405.

Young, L., and Saxe, R. (2009b). Innocent intentions: a correlation between forgiveness for accidental harm and neural activity. Neuropsychologia *47*, 2065–2072.

Young, L., Cushman, F., Hauser, M., and Saxe, R. (2007). The neural basis of the interaction between theory of mind and moral judgment. Proc. Natl. Acad. Sci. USA *104*, 8235–8240.