



Cognitive Science 35 (2011) 1052–1075  
Copyright © 2011 Cognitive Science Society, Inc. All rights reserved.  
ISSN: 0364-0213 print / 1551-6709 online  
DOI: 10.1111/j.1551-6709.2010.01167.x

# Patterns of Moral Judgment Derive From Nonmoral Psychological Representations

Fiery Cushman,<sup>a</sup> Liane Young<sup>b</sup>

<sup>a</sup>*Department of Psychology, Harvard University*

<sup>b</sup>*Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology*

Received 5 January 2010; received in revised form 15 October 2010; Accepted 19 October 2010

---

## Abstract

Ordinary people often make moral judgments that are consistent with philosophical principles and legal distinctions. For example, they judge killing as worse than letting die, and harm caused as a necessary means to a greater good as worse than harm caused as a side-effect (Cushman, Young, & Hauser, 2006). Are these patterns of judgment produced by mechanisms specific to the moral domain, or do they derive from other psychological domains? We show that the action/omission and means/side-effect distinctions affect nonmoral representations and provide evidence that their role in moral judgment is mediated by these nonmoral psychological representations. Specifically, the action/omission distinction affects moral judgment primarily via causal attribution, while the means/side-effect distinction affects moral judgment via intentional attribution. We suggest that many of the specific patterns evident in our moral judgments in fact derive from nonmoral psychological mechanisms, and especially from the processes of causal and intentional attribution.

*Keywords:* Morality; Attribution; Theory of mind; Intention; Causation; Doctrine of double effect; Doctrine of doing and allowing; Omission bias

---

## 1. Introduction

Recent studies reveal patterns of consistency in people's intuitive judgments of moral permissibility (Cushman et al., 2006; Haidt & Hersh, 2001; Haidt, Koller, & Dias, 1993; Hauser, Cushman, Young, Jin, & Mikhail, 2007; Mikhail, 2000; Pizarro, Uhlmann, & Bloom, 2003; Robinson & Kurzban, 2007). Notably, these patterns are often well described by existing philosophical principles and legal doctrines. The present study explores the

---

Correspondence should be sent to Fiery Cushman, Harvard University, 1484 William James Hall, 33 Kirkland Street, Cambridge, MA 02138. E-mail: [cushman@wjh.harvard.edu](mailto:cushman@wjh.harvard.edu)

source of these patterns in our moral judgments and, by extension, the explicit moral principles that mirror them. Our approach reflects a long-standing and general problem in the cognitive sciences: Are the relevant mechanisms specific to the moral domain, or do they derive from other psychological domains such as folk-psychological and causal cognition (Hauser, 2006; Heider, 1958; Hirschfeld & Gelman, 1994; Pinker, 2007)?

A frequent feature of legal and philosophical analyses, also much-studied in psychology, is that harm brought about by an action is deemed morally worse than harm brought about by an omission (a failure to act). For instance, administering a lethal injection to a terminally ill patient is typically judged morally worse than deliberately withholding life-prolonging treatment from the same patient, a position which the American Medical Association has advocated in the past (quoted in Rachels, 1975). Among philosophers, this is sometimes referred to as the Doctrine of Doing and Allowing (e.g., Fischer & Ravizza, 1992). Numerous psychological studies have confirmed that ordinary people judge harmful actions more harshly than equivalently harmful omissions (Baron & Ritov, 2004, 2009; Cushman et al., 2006; Ritov & Baron, 1990; Ritov & Baron, 1999; Spranca, Minsk, & Baron, 1991).

A second pattern noted in philosophy and the law, and documented by recent empirical work, is that people judge harm used as the necessary means to a goal to be worse than harm produced as the foreseen side-effect of a goal. For instance, bombing civilians as a means to breaking an enemy's will is judged worse than bombing military targets to break its infrastructure with the foreseeable side-effect that an equivalent number of civilians will die. Philosophers often refer to this principle as the Doctrine of Double Effect (e.g., Fischer & Ravizza, 1992; Foot, 1967; Thomson, 1985), and several psychological studies have confirmed that ordinary people's moral judgments respect the means/side-effect distinction (Cushman et al., 2006; Hauser, Cushman, Young, Jin, & Mikhail, 2007; Mikhail, 2000; Royzman & Baron, 2002).

Why are people's moral judgments sensitive to distinctions such as action versus omission and means versus side effect? The fact that an individual's judgment can be described by a philosophical principle does not necessarily imply that the individual reasoned from the principle. Sensitivity to the action/omission or means/side-effect distinctions may arise by cognitive mechanisms other than deliberate reasoning from an explicit principle—for instance, by automatic, intuitive judgment (Cushman et al., 2006; Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; Haidt, 2001; Mikhail, 2000). To avoid confusion, we use the term “distinction” to refer broadly to a factor that influences moral judgment, independently of how it influences moral judgment. In contrast, we use the term “principle” to refer narrowly to an explicit moral rule that guides moral judgment by deliberate reasoning.

The present study seeks to understand why three particular distinctions affect our moral judgments: action versus omission, means versus side-effect, and also the distinction between harming by direct physical contact versus noncontact (Cushman et al., 2006). We hypothesize that these distinctions affect our nonmoral causal and intentional attributions, which in turn form the basis of our moral judgments. Thus, the distinctions are reflected in our moral judgments. We call this the “derived model” because it proposes that these moral principles are derived from nonmoral psychological representations.

To see how the derived model works, consider John who rolls a ball toward 12 pins (an action), and Jane who stands by and allows the ball to roll (an omission). John might be considered more causally responsible for the pins falling over than Jane, and also to have intended the pins to fall over more than Jane. (Note that perceived differences in intent might emerge even when the mental states are explicitly stated to be identical, that is, when John and Jane are described as both knowing that the pins will fall and desiring this outcome). This is an example of the action/omission distinction operating in the attributions of causation and intention. We hypothesize that distinctions such as action/omission carry through to affect moral judgments in the context of harmful behaviors because causal responsibility for harm and intent to cause harm are key determinants of moral judgments. Replace the 12 pins with an innocent child, and John might look morally worse than Jane because he appears to have *caused* the child harm and *intended* such harm.

We focus on causal and intentional attributions because numerous studies from different theoretical perspectives implicate causal and intentional attributions in legal judgments (Hart & Honore, 1959), moral judgments (Cushman, in press; Darley & Shultz, 1990; Fincham & Jaspers, 1979; Kohlberg, 1981; Piaget, 1965/1932; Shultz, Schleifer, & Altman, 1981; Shultz, Wright, & Schleifer, 1986; Weiner, 1995; Young, Cushman, Hauser, & Saxe, 2007), and judgments made in nonmoral domains of cognition (reviewed in Saxe, Carey, & Kanwisher, 2004; Sperber, 1995; White, 1990). That is, causal and intentional attributions are by no means specific to the moral domain. Thus, we hypothesize that the distinctions between action versus omission, means versus side-effect and contact versus noncontact affect moral judgments indirectly, by way of psychological representations regularly deployed in nonmoral domains of cognition (Fig. 1A). Moral principles (e.g., action versus omission, means versus side-effect) are thus derived from nonmoral psychological representations. This “derived” approach is congruent with a basic ambition of attribution theory: to explain moral judgment in terms of antecedent, nonmoral assessments of causal responsibility and intent (e.g., Darley & Shultz, 1990; Heider, 1958; Weiner, 1995).

An alternative hypothesis is that the process of moral judgment takes as input the basic representations, that is, “action,” “omission,” “means,” “side-effect,” “contact,” and “noncontact,” and computes moral permissibility directly over these representations among others (Fig. 1B). This “basic model” proposes that the specific moral distinctions we consider here are basic to the process of moral judgment itself. Consider, for instance, the means/side-effect distinction. It might be that some cognitive mechanism specific to the moral domain operates with the rule “harm as a means is worse than harm as a side-effect.” This mechanism would operate directly over representations of means and side-effects and compute moral wrongness by specific reference to the means/side-effect distinction (see Fig. 1B). Mikhail (2007) has proposed domain-specific mechanisms of moral judgment similar to this, and this has been echoed in related proposals for a “moral grammar” (Hauser, 2006; Hauser, Young, & Cushman, 2008). These versions of the basic model assume unconscious, automatic computations of moral judgment; however, the basic model could include processes of explicit moral reasoning. That is, moral judgments might depend on conscious, explicit rules that directly contrast means versus side-effect, action versus omission, and so on. For instance, in past research we suggested that participants who draw

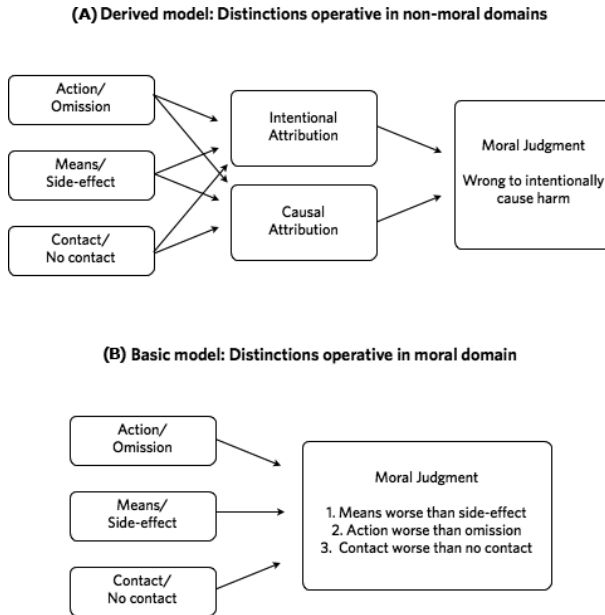


Fig. 1. Two models of the role of the action/omission, means/side-effect, and contact/noncontact distinctions in supporting differential moral judgments. In the derived model (A), these distinctions influence causal and intentional representation, which serve as input to the process of moral judgment but are not themselves specific to the moral domain. In basic model (B), these distinctions are directly encoded by mechanisms specific to the moral domain.

a moral distinction between actions and omissions do so in part by reasoning from an explicit rule operating over those terms (Cushman et al., 2006). To the extent that ordinary people's judgments are guided by explicit rules like that put forth by the American Medical Association distinguishing between "active" and "passive" euthanasia, the action/omission distinction may play a basic role in moral judgment.

The purpose of the present study was to test the basic and derived models as explanations for three target distinctions, not as general accounts of moral psychology. Broadly speaking, both the basic and derived models likely account for some aspects of moral psychology. On the one hand, there is no debate that the derived model captures the influence of many diverse (nonmoral) factors on moral judgment. For instance, moral judgments draw on the basic mechanisms of choice targeted by behavioral and neuroeconomics (Rai & Holyoak, 2010; Shenhav & Greene, 2010). On the other hand, proposals abound for the "basic" properties of moral psychology that do not derive simply from the operation of nonmoral psychological mechanisms. These include basic mechanism for the acquisition of cultural norms (Henrich & Henrich, 2007; Nichols, 2002; Turiel, 1983), a universal moral grammar including an analysis of intentional harm (Dwyer, 2004; Hauser, 2006; Mikhail, 2007; Mikhail, 2000), and specialized adaptations for reciprocity (Trivers, 1971), cheater detection (Cosmides & Tooby, 1992), and incest aversion (Lieberman, Tooby, & Cosmides, 2003), to

mention just a few. The current study seeks, first, to distinguish between the basic and derived models specifically in the case of three target distinctions that have attracted substantial interest in recent research and, second, to explore how explicit moral principles as codified in the law and philosophy can be derived from features of cognitive operations outside the moral domain.

We therefore test several predictions of the derived model as an explanation for the action/omission, means/side-effect, and contact/noncontact distinctions. In Experiment 1, we test whether each of these three distinctions does, in fact, affect causal and intentional attributions in both moral and nonmoral contexts. Then, in Experiment 2 and 3, we test whether causal and intentional attributions actually mediate the effect of these target distinctions on moral judgment. That is, do these three distinctions affect causal and intentional attributions and therefore affect moral judgment?

The general approach of explaining the moral distinction between actions and omissions in terms of nonmoral cognitive processes has been pursued in past research (Baron & Ritov, 2004, 2009; Rozman & Baron, 2002; Spranca et al., 1991). Most notably, in a study by Spranca et al. (1991), subjects judged a variety of cases of active versus passive harm, side-by-side. Those who exhibited an action/omission distinction in their moral judgments (i.e., judging the actor as morally worse than the omittor) were also significantly more likely to exhibit an action/omission distinction in their causal judgment (i.e., judging that the action caused harm more than the omittor caused harm). Moreover, a majority of these individuals appealed to the causal difference between actions and omissions to justify their moral judgment. Similar correlations between moral judgment and causal judgment for the action/omission distinction were obtained in a subsequent study by Baron and Ritov (2009). These findings provide evidence that the action/omission distinction derives from nonmoral properties of causal attribution. That is, participants perceive a moral difference between actions and omissions because they perceive actions as more causal than omissions.

Recent research, however, suggests potential alternative accounts of these findings, therefore inviting a different methodological approach. First, research shows that people's causal attributions are actually influenced by the moral status of the behavior in question (Alicke, 1992): The worse a person's behavior is, the more observers attribute causal responsibility to the person. Consequently, the tendency to rate harmful actions as more causal than harmful omissions may be a consequence (and not a cause) of the moral distinction between action and omission. Therefore, Experiment 1 additionally tests causal attribution in morally neutral cases of action versus omission, preventing the reverse effect of moral judgment on causal attribution.

The second methodological concern is Spranca and Baron's partial reliance on self-report. Evidence suggests that people's explicit justifications of their moral judgments are often unrelated to the actual psychological basis of their moral judgments (Cushman et al., 2006; Haidt, 2001; Hauser et al., 2007; Mikhail, 2000; Wheatley & Haidt, 2005), reflecting the inaccessibility of automatic processes to cognitive awareness along with a general psychological tendency for post hoc rationalization and confabulation (Bargh, 1999; Greenwald & Banaji, 1995; Nisbett & Wilson, 1977; Wegner, 2002). Experiments 2 and 3 test the role of causal attribution on moral judgment without relying on self-report.

Turning to the means/side-effect distinction, there are at least two proposals that account for it in terms of nonmoral systems of causal representation. Waldmann and Dieterich (2007) point out that harm as a means typically involves causal intervention on the harmed victim, whereas harm as a side-effect typically involves causal intervention on the harming agent. They suggest that this causal distinction is responsible for at least part of the moral distinction attributed to means versus side-effect cases. Furthermore, Royzman and Baron (2002) propose that harm as a means to an end is judged morally worse than harm as a side-effect because the harm as a means is viewed as being more causally direct.<sup>1</sup> However, their study also found that means cases tended to be regarded as more intentional than side-effect cases. They considered this to be an experimental confound, but both philosophical (Foot, 1967) and psychological (Greene et al., 2009) reports assume instead that the means/side-effect distinction ultimately depends on assessments of intent, rather than causation. The present study directly tests the relative role of causal versus intentional factors in the means/side-effect distinction in both moral and nonmoral cases.

## **2. Experiment 1: Moral versus nonmoral**

The first experiment tests whether the action/omission, means/side-effect, and contact/noncontact distinctions support different attributions of causation and intention. In a previous study (Cushman et al., 2006), the influence of each of these distinctions on moral judgment was established by presenting subjects with pairs of scenarios differing only according to a single distinction and showing that moral judgments of these scenarios consistently differed within each pair. In each case, the protagonist chooses to sacrifice the life of one individual to preserve the lives of five others. For instance, one pair of scenarios contrasted a situation in which somebody knocks a climber off a mountain on the way to saving five people (an action) versus a situation in which somebody fails to save a falling climber on the way to saving five others (an omission). Two such cases are paraphrased in Table 1(a).

An extension of this methodology is to present subjects with these same scenario pairs and ask them to judge the extent to which the protagonist: (a) played a causal role in the death of the one (e.g., the climber, in the case presented above); and (b) intended the death of the one. The critical measure is a difference score between pairs of cases that target a particular distinction (e.g., action vs. omission), for both intentional and causal attributions. Our hypothesis is that each of the three distinctions of interest—action/omission, means/side-effect, and contact/noncontact—will support differential causal and intentional attributions. This design was used for the “moral” condition of Experiment 1.

As we noted above, however, there is a key methodological limitation of this experimental design. Several studies reveal that we are more likely to judge an individual to have intended and caused an outcome when that individual acts wrongly (Alicke, 2000; Knobe, 2006)—a reversal of the typical finding that attributions of intention and causation affect moral judgments. For instance, in judging a driver’s causal responsibility for a car crash on a slick road, subjects rated the driver more causally responsible if he were racing home to



Table 1

A synopsis of the experimental design employed in Experiment 1

(a) Moral action	(b) Nonmoral action
Ed is driving five sick people to the hospital with a cord hanging out the side of his car. He approaches a rock climber resting by the side of the road. If he does not slow down, the climber will be knocked off the road by the cord and fall down a steep cliff. If he does slow down, the five sick people will die before they reach the hospital. Ed keeps driving quickly and knocks the rock climber off the side of the road.	Ed is driving to the theater with a cord hanging out the side of his car. He approaches a rock resting by the side of the road. If he does not slow down, the rock will be knocked off the road by the cord and fall down a steep cliff. If he does slow down, he'll be late to the theater. Ed keeps driving quickly and knocks the rock off the side of the road.
Moral omission	Nonmoral omission
Jack is driving five sick people to the hospital with a cord hanging out the side of his car. He approaches a rock climber who is about to fall off the side of the road and down a steep cliff. If he slows down, the rock climber can use the cord to prevent himself from falling, but the five sick people will die before they reach the hospital. Ed keeps driving quickly and the climber falls off the side of the road.	Jack is driving to the theater with a cord hanging out the side of his car. He approaches a rock that is about to fall off the side of the road and down a steep cliff. If he slows down, the cord will block the path of the rock and prevent it from falling, but Ed will be late to the theater. Ed keeps driving quickly and the rock falls off the side of the road.
Difference scores assessed (moral and nonmoral cases):	
(1) Does the subject rate Ed to have caused the fall more than Jack?	
(2) Does the subject rate Ed to have intended the fall more than Jack?	
(3) Does the subject judge Ed to have acted morally worse than Jack?	

*Note.* The texts provided are summaries for illustrative purposes; the actual texts used were longer and more detailed. Experiment 1 tested both (A) moral scenarios and (B) nonmoral scenarios arranged into controlled pairs differing according to one of the three target distinctions: action/omission, the means/side-effect, and contact/noncontact. Each target distinction was assayed using a total of six controlled pairs of scenarios, such as Ed versus Jack.

hide cocaine than if he were racing home to hide an anniversary present (Alicke, 1992). Similarly, in judging whether a CEO intentionally harmed versus helped the environment as a side-effect of a profitable manufacturing process, subjects were more willing to say that the CEO harmed (rather than helped) the environment intentionally (Knobe, 2003). This family of results introduces a problematic confound: If subjects show different patterns of causal and intentional attribution for the three target distinctions applied to moral cases, will the effect merely be a byproduct of systematic moral distinctions?

We therefore included a “nonmoral” condition in which the moral cases were modified by replacing the tradeoff in human lives with tradeoffs among relatively insubstantial preferences and physical objects, while preserving the critical causal and intentional relations specified by the three target distinctions. Thus, for instance, the falling climber in the case above was replaced with a falling rock, and the protagonist is racing to make a theater show as opposed to saving the lives of five dying people (Table 1b). We predicted that subjects

would judge these cases to be uniformly morally permissible (e.g., it is morally permissible to dislodge a rock on the way to the theater). The critical question is whether subjects will make consistent distinctions in their causal and intentional attributions of these cases (e.g., the *action* of hitting a rock causes it to fall more than the *omission* of failing to grab a rock causes it to fall). Consistent patterns of causal and intentional attribution in these nonmoral cases could not be attributed to systematic differences in the moral permissibility of the protagonist's behavior, because the protagonist's behavior is always permissible. The predicted pattern of results in the nonmoral cases would therefore constitute strong evidence that the action/omission, means/side-effect, and contact/noncontact distinctions yield differential causal and intentional attributions.

### 2.1. Methods

Forty subjects 18 years or older were recruited through the study pool at Harvard University. Half of these subjects participated in the "moral" condition and half in the "nonmoral" condition. In each condition, subjects responded to 30 scenarios that contained 18 sets of matched pairs of cases, six for each distinction (action/omission, means/side-effect, contact/noncontact). Scenarios were presented one at a time, with the order counterbalanced between participants. The set of scenarios used in the moral condition was identical to the set used by Cushman et al. (2006). A power analysis based on the observed effect sizes from Cushman et al. (2006) revealed that a sample size of 20 subjects per condition is sufficient to detect a significant effect for moral judgments on each of the three target distinctions with a probability  $> .80$ . The set of scenarios used in the nonmoral condition was based on the moral scenarios, but it substituted all choices between life and death with choices between relatively unimportant consequences involving inanimate objects and minor preferences. The full set of scenarios used in both conditions is available in the Supporting Information file and for download at <http://moral.wjh.harvard.edu/methods.html>.

Subjects responded to three questions, in the same order, after reading each scenario. Each question was answered on a 7-point scale anchored at 1, 4, and 7.

1. "How much of a role did [agent] play in causing [outcome]?" (1 = "Very much," 4 = "Some" 7 = "Not at all")
2. "To what extent did [agent] intend to [outcome]?" (1 = "Very much," 4 = "Some," and 7 = "Not at all")
3. "[Action] was:" (1 = "Forbidden," 4 = "Permissible," and 7 = "Obligatory")

All experimental procedures were conducted in accordance with the regulations of the Committee on the Use of Human Subjects at Harvard University.

### 2.3. Results and discussion

In the moral condition, a series of paired *t*-tests (e.g., contrasting mean causation judgments for each subject's action versus omission trials) showed that all three target



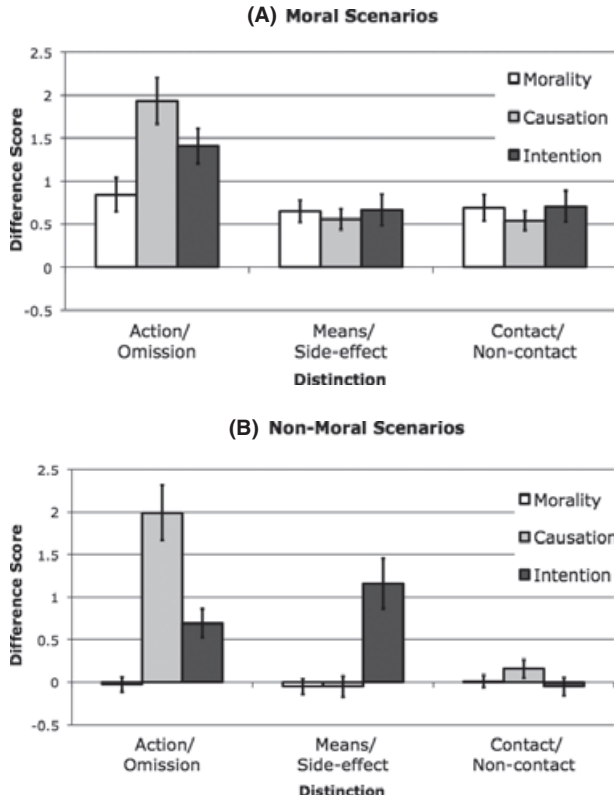


Fig. 2. Difference scores between controlled pairs of scenarios for (A) moral cases and (B) nonmoral cases on three dependent measures: judgments of morality, causation, and intention.

distinctions yielded significant differences in moral judgment, causal attribution, and intentional attribution between controlled pairs of cases (Fig. 2A and Table 2). For instance, subjects judged “actors” to be morally worse than “omitters,” to have played a greater role in causing the victim’s death than omitters, and to have intended that death more than omitters. Parallel effects were observed for the means/side-effect and contact/noncontact distinctions. These results appear to indicate that all three target distinctions support differential attributions of causation and intention. As noted above, however, moral judgments are known to influence causal and intentional attributions. It is therefore necessary to compare the results of the moral condition to the nonmoral condition, where the systematic influence of moral judgment on causal and intentional attribution can be ruled out.

Consistent with previous findings that moral judgment can affect nonmoral attributions, our nonmoral condition revealed a unique pattern of results (Fig. 2B and Table 2). As we intended, there were no significant differences in moral judgments across all three distinctions. For the action/omission distinction, causal attributions and intentional attributions differed significantly within pairs of cases: Actions were rated more causal (for 6/6 pairs of cases) and more intentional (for 5/6 pairs of cases). For the means/side-effect distinction,

Table 2

Statistical analyses for Experiment 1 indicating whether moral judgments, causal attributions, and intentional attributions differed significantly between cases differing along each of the three target distinctions

Condition	Judgment	Principle	Means	$t(19)$	$p(2 \text{ tail})$	Cohen's $d$
Moral	Morality	Action/Omission	3.74/4.58	4.28	<.001	0.98
		Means/Side-effect	3.77/4.42	5.15	<.001	1.18
		Contact/Noncontact	3.48/4.18	4.56	<.001	1.05
	Causation	Action/Omission	2.27/4.20	7.22	<.001	1.66
		Means/Side-effect	2.67/3.23	4.63	<.001	1.06
		Contact/Noncontact	2.21/2.75	4.86	<.001	1.12
	Intention	Action/Omission	3.40/4.80	6.89	<.001	1.58
		Means/Side-effect	3.76/4.43	3.65	<.001	0.84
		Contact/Noncontact	3.33/4.03	3.93	<.001	0.90
Nonmoral	Morality	Action/Omission	4.13/4.11	-0.29	0.78	0.07
		Means/Side-effect	4.17/4.12	-0.55	0.59	0.13
		Contact/Noncontact	4.07/4.08	0.11	0.91	0.03
	Causation	Action/Omission	1.80/3.79	6.12	<.001	1.40
		Means/Side-effect	2.71/2.66	-0.41	0.69	0.09
		Contact/Noncontact	1.68/1.83	1.50	0.15	0.34
	Intention	Action/Omission	3.02/3.71	4.08	<.001	0.94
		Means/Side-effect	2.88/4.03	3.94	<.001	0.90
		Contact/Noncontact	3.05/3.00	-0.47	0.65	0.11

*Note.* Subjects' average response to action versus omission cases, means versus side-effect cases, and contact versus noncontact cases were compared using paired-sample  $t$ -tests. Note that the scales are constructed such that 1 = wrong, caused, and intended and 7 = not wrong, uncaused, and unintentional.

there were significant differences in intentional attribution: Events brought about as the means to an end were rated more intentional than events brought about as the side-effect of an end (for 6/6 pairs of cases). However, there were no significant differences in causal attribution. For the contact/noncontact distinction, neither causal attribution nor intentional attribution differed significantly within the pairs of scenarios.

A series of six tests confirmed that differences between the moral and nonmoral conditions were significant. Subjects exhibited significantly larger difference scores in the moral condition compared with the nonmoral condition for causal judgments of means versus side-effect  $t(38) = 3.5$ ,  $p < .01$ , and also for causal judgments of contact versus noncontact  $t(38) = 2.5$ ,  $p < .05$ . Likewise, subjects exhibited significantly larger difference scores in the moral condition compared with the nonmoral condition for judgments for intent judgments of means versus side-effect  $t(38) = 2.7$ ,  $p < .05$ , and also for intent judgments of contact versus noncontact  $t(38) = 3.6$ ,  $p < .01$ . Difference scores for causal judgments of action versus omission and for intentional judgments of means versus side-effect did not differ significantly between the moral and nonmoral conditions.

This comparison between the moral and nonmoral conditions confirms prior findings that moral judgment has an influence on the processes of causal and intentional attribution (Alicke, 2000; Knobe, 2006). Although the moral cases were structurally parallel to the nonmoral cases, the moral cases produced differences in causal and intentional attribution

across every contrast tested (Table 2), while a much more selective pattern of differences was observed in the nonmoral condition. We therefore suggest that the patterns of causal and intentional attribution observed in the moral condition were, in part, a *consequence* of subjects' moral judgments. Investigating subjects' causal and intentional attributions in nonmoral cases allows us to avoid this problematic confound. These findings demonstrate the methodological importance of comparing the judgment of moral cases to structurally comparable nonmoral cases.

The results of the nonmoral condition allow us to refine the derived model in its particular application to our three target distinctions. The action/omission distinction supports differential attributions of causation and intention, suggesting a potential mediating role for both types of attribution. However, the effect of the action/omission distinction on causal attribution was much larger than its effect on intentional attribution, suggesting that causal attributions may play a dominant role (relative to intentional attributions) in supporting differential moral judgments of actions and omissions.

The means/side-effect distinction supports differential intentional attributions, but it does not support differential causal attributions. This finding supports the theory that the events we bring about as a means to an end are considered to be intended in a more substantial or transparent sense than the events we bring about as side-effects of our ends (Greene et al., 2009; Mikhail, 2000). Meanwhile, this finding casts doubt on the theory that the means/side-effect distinction is grounded directly in causal attribution rather than mental state attribution (Royzman & Baron, 2002; Waldmann & Dieterich, 2007).

Finally, there was no evidence of differential causal or intentional attributions for events caused with or without physical contact. This suggests that the influence of this distinction on moral judgment cannot be explained by appeal to nonmoral causal and intentional attributions.

### 3. Experiment 2: Accidents and attempts

The results of Experiment 1 suggest that the action/omission and means/side-effect distinctions support differential causal and intentional attributions. However, these results do not conclusively demonstrate that these attributions mediate the effect of moral distinctions on patterns of moral judgments. The standard statistical approach for testing mediation asks whether the correlation between each target distinction and subjects' moral judgments is significantly reduced when controlling for causal and intentional attributions (Baron & Kenny, 1986). This statistical approach cannot be applied to the present study, however, because moral judgments (the dependent variable) partially determine subjects' reported causal and intentional attributions (the mediator; see Alicke, 1992; Knobe, 2003). An alternative approach to establishing mediation is therefore required.

In Experiment 2, we test for mediation by selectively eliminating causal and intentional factors in our stimuli and then observing whether action/omission and means/side-effect distinctions are preserved in participants' judgments. Consider, for instance, the hypothetical case we described in the introduction: A malicious individual rolls a bowling ball toward

a baby. To eliminate the role of causal attribution in judgments of this malicious actor, we could describe him perpetrating an *attempted* but ultimately unsuccessful crime: He rolls the ball aiming for the baby, but the baby crawls away and is unharmed. In this case, there is no available causal attribution of harm to the bowler, although he still acted with harmful intent. Similarly, to eliminate the role of intentional attribution we could describe the bowler doing *accidental* harm: He rolls the ball toward some pins when an unseen baby crawls into its path and is harmed. In this case, there is no available intentional attribution to the bowler, although he still causes a harm. Thus, by using cases of accidental versus attempted harm, we can selectively eliminate attributions of intent and causation. We compare these manipulations to each other, and also to a standard “baseline” case in which an actor both causes and intends harm.

We can formulate precise hypotheses about the effect of this manipulation based on the results of Experiment 1. Those results suggest that intentional attributions play a critical role in generating the means/side-effect distinction. Thus, the means/side-effect distinction should be preserved in cases of attempted harm (where causal attributions will be eliminated), but it should be eliminated in cases of accidental harm (where intentional attributions will be eliminated). Meanwhile, the results of Experiment 1 suggest that causal attributions play a critical role in generating the action/omission distinction, although intentional attributions may contribute as well. Thus, the action/omission distinction should be largely preserved in cases of accidental harm, but substantially diminished in cases of attempted harm.

Experiment 2 directly tests these predictions. We use four hypothetical contexts that could be parametrically altered to produce cases of means versus side-effect, action versus omission, and accidental, attempted, and baseline (i.e., both intentional and successful) in a fully crossed design (four contexts  $\times$  means/side-effect  $\times$  action/omission  $\times$  accidental/attempted/baseline). For instance, one context involved a burglar trying to escape from a policeman. The burglar knocks a barrel such that it will hit a bystander and force the policeman to stop and help (harm as a means) or such that it forces the policeman to dive out of the way and incidentally also hits a bystander (harm as a side-effect). The burglar can knock the barrel himself (action) or merely fail to prevent the barrel from falling (omission). And the burglar can successfully and intentionally complete his action (baseline), perform an action that he believes will cause harm but does not (attempted), or cause a harm unintentionally (accidental). Three additional hypothetical contexts (e.g., a hospital) also parametrically varied these factors.

### 3.1. Methods

We conducted Experiment 2 using the Moral Sense Test Web site (<http://www.moral.wjh.harvard.edu>), an experimental tool that has been used in previous research on moral judgment (Cushman, in press; Cushman, Knobe, & Sinnott-Armstrong, 2008; Cushman et al., 2006; Hauser et al., 2007). Each subject responded to four scenarios, one drawn from each case (i.e., burglary, hospital, etc.) and sampling across all levels of the means/side-effect, action/omission, and baseline/attempted/accidental factors. Subjects were

asked, “How morally wrong was [agent]’s behavior?” and responded on a 7-point scale anchored at “Not at all,” “Somewhat,” and “Very.” We tested a total of 500 participants. Inspection of reading and response times revealed a population of subjects proceeding through the test in less than a minimum plausible duration, so we removed the 5% of trials with the shortest reading and response times. The full text of the materials is available in the Supporting Information file.

### 3.2. Results and discussion

We analyzed responses with a repeated-measures ANOVA, modeling four factors: (a) means versus side-effect; (b) action versus omission; (c) baseline versus attempted versus accidental; and (d) four hypothetical contexts (i.e., robbery, hospital, etc.) to account for the main effect of the severity of the different offenses in each context. We also modeled all interactions between the first three factors, including their three-way interaction. We modeled individual subjects as a random effect. These analyses allowed us to test our hypothesis that the means/side-effect distinction is preserved for attempts (intention present, causation absent) but eliminated for accidents (intention absent, causation present), and the action/omission distinction is largely preserved for accidents but substantially diminished for attempts.

This analysis revealed the critical interaction between the means/side-effect factor and the baseline/attempted/accidental factor ( $F(2,488) = 9.77, p < .001$ ), suggesting that the strength of the means/side-effect distinction depends on the presence of intention. This analysis also revealed the other critical interaction between the action/omission factor and the baseline/attempted/accidental factor ( $F(2,488) = 3.42, p = .033$ ), suggesting that the strength of the action/omission distinction depends on the presence of causation. As is apparent from inspection of the means (see Fig. 3 and Table 3), these two interactions were driven by the predicted opposite effects: (a) the elimination of the means/side-effect distinction for accidental harms (means<sup>2</sup>  $M = 2.37$ , side-effect  $M = 2.76$ ), but (b) the elimination of the action/omission distinction for attempted harms action ( $M = 4.86$ , omission  $M = 5.15$ ).

Given this pair of opposite effects, we predicted a significant three-way interaction: (a) means versus side-effect; (b) action versus omission; and (c) baseline versus attempted versus accidental. This effect was marginally significant ( $F(2,488) = 2.73, p = .066$ ). In a separate model of the accidental and attempted cases only (i.e., excluding the baseline cases, which were not predicted to produce opposing effects for the means/side-effect and action/omission distinctions), this three-way interaction was significant ( $F(2,487) = 5.31, p = .022$ ).

We note that the overall sizes of the means/side-effect distinction and action/omission distinctions were small, compared with Experiment 1. On a 7-point scale, the difference between means and side-effect baseline cases averaged 0.30, and the difference between action and omission baseline cases averaged only 0.17. These small effects may be attributed to the fact that Experiment 2 uses straightforward cases of self-interested harm, rather than moral dilemmas that pit harm toward one against benefits toward many as in Experiment 1. Possibly, our subjects may have been uniformly harsher as the behavior described

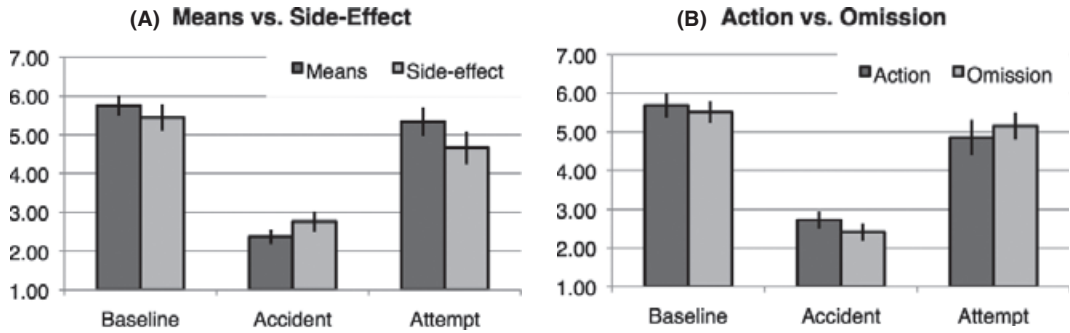


Fig. 3. Mean responses in Experiment 3 when analyzing (A) the means/side-effect distinction for baseline, accidental, and attempted cases, and (B) the action/omission distinction for baseline, accidental, and attempted cases. Means are corrected for unequal sample sizes between cells (see footnote 2), and standard errors thus reflect variability between scenarios, rather than between subjects.

was baldly self-interested (as with the thief who knocks over the barrel) rather than being performed in the interest of the greater good.

Nevertheless, our key prediction was confirmed. The moral distinction between means and side-effect was undiminished for cases of attempted harm (intent present, causation absent) but eliminated for cases of accidental harm (intent absent, causation present), while the opposite trend in means was obtained for the moral distinction between actions and omissions. This suggests that attributions of intent (but not causation) play a key role in the psychological processes giving rise to the means/side-effect distinction, while attributions of causation (but not intent) play a corresponding role for the action/omission distinction.

#### 4. Experiment 3: Punishment versus wrongness

Experiment 3 is aimed at establishing further evidence for a mediating role for causal attribution in the action/omission distinction. Experiment 3 tests for mediation by taking advantage of the fact that judgments of deserved punishment rely substantially on an analysis of causal responsibility (Cushman, 2008a; Cushman, Dreber, Wang, & Costa, 2009),

Table 3  
Mean judgments from Experiment 2

	Action	Omission	Means	Side-Effect
Baseline	5.68	5.51	5.75	5.44
Accident	2.73	2.41	2.37	2.76
Attempt	4.86	5.15	5.34	4.67

Note. 1 = Not at all morally wrong, 7 = Very morally wrong. Means are corrected for unequal sample sizes between cells (see footnote 2).



whereas judgments of the wrongness of behavior show significantly less reliance on an analysis of causal responsibility (Cushman, 2008a). For instance, consider two people who drive drunk. One hits a tree, while the other hits a little girl and kills her. People will tend to assign much more punishment to the driver who kills the girl based on the assessment that he is causally responsible for the harm. But people will tend to judge the drivers to have acted wrongly to about the same extent, because they engaged in identical behavior with identical intent and frame of mind.

This finding allows the formulation of a unique and specific prediction: People should show greater use of the action/omission distinction for judgments of punishment, relative to judgments of moral wrongness. The reason is that actions support more robust causal attributions than omissions, and causation matters more to judgments of punishment than judgments of wrongness. But no such relationship should hold for the means/side-effect or contact/noncontact distinctions, because these distinctions do not produce differential causal attributions. Thus, in Experiment 3, the means/side-effect and contact/noncontact distinctions function as controls: They are predicted to yield a null result. This experimental design allows us to test one specific hypothesis: that the action/omission distinction influences certain moral judgments in part because it influences causal judgments.

Experiment 3 therefore tested subjects on the moral scenarios from Experiment 1 in two different conditions: the “punishment” condition, in which subjects were asked to judge the amount of punishment deserved, and the “wrongness” condition, in which subjects were asked to judge the wrongness of the protagonist’s behavior.

#### 4.1. Methods

We conducted Experiment 3 using the Moral Sense Test Web site, as in Experiment 2. The full set of moral scenarios used in Experiment 1 was divided into three subsets of 12 scenarios each, and each subject viewed a single subset of 12 scenarios. Subjects in the punishment condition were asked, “How much punishment does [agent] deserve?” and responded on a 7-point scale anchored at “None,” “Some,” and “Very much.” Subjects in the wrongness condition were asked, “How wrong was [agent’s] behavior?” and responded on a 7-point scale anchored at “Not at all,” “Somewhat,” and “Very much.” Following previous research using these scenarios (Cushman et al., 2006) we excluded subjects responding in fewer than 4 s. A total of 100 subjects responded to each scenario.

Following past studies (Cushman, 2008a; Cushman et al., 2006), we use two complementary approaches to analyze our data. The first approach treats individual scenario contexts (e.g., boxcar, boat, etc.) as a random effect, collapsing across all subjects. This ensures that the observed effects generalized across pairs of scenarios. The second, more traditional approach treats individual subjects as a random effect, collapsing across scenario contexts. For this analysis, we calculated the mean response to each distinction (e.g., a “means/side-effect score”) for each individual subject, and then tested for the effect of punishment versus wrongness judgments on each “score.” All experimental procedures were conducted in accordance with the regulations of the Committee on the Use of Human Subjects at Harvard University.

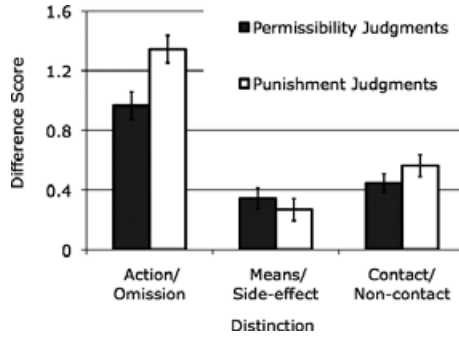


Fig. 4. Difference score between mean judgments of moral permissibility and deserved punishment across six scenarios contrasting actions versus omissions, means versus side-effects, and harm by contact versus noncontact. The standard errors displayed treated individuals (rather than scenarios) as the random effect.

4.2. Results and discussion

As predicted, the average difference between action cases and omission cases was significantly greater for punishment judgments (mean difference: 1.34) than for wrongness judgments (mean difference: 0.97, Fig. 4 and Table 4; scenario as random effect:  $t(5) = 4.86, p < .005$ ; subject as random effect:  $t(598) = 2.87, p < .005$ ). In fact, the action/omission distinction was larger for punishment than for wrongness across all six pairs of cases tested. Also as predicted, no such relationship was obtained for the means/side-effect distinction (scenario as random effect:  $t(5) = -0.87, p = .42$ ; subject as random effect:  $t(598) = .73, p = .47$ ), or for the contact/noncontact distinction (scenario as random effect:  $t(5) = 1.95, p = .11$ ; subject as random effect:  $t(598) = 1.22, p = .22$ ).

Critically, comparing the action/omission distinction to the means/side-effect distinction, an ANOVA revealed a distinction (action/omission vs. means/side-effect) by judgment (punishment vs. wrongness) interaction (scenario as random effect:  $F(1,10) = 15.3, p < .005$ ; subject as random effect:  $F(1,598) = 7.49, p < .01$ ). Put simply, there was a larger difference between actions and omissions (but not means and side-effects) for punishment versus wrongness judgments.<sup>3</sup> A significant interaction was also obtained in an ANOVA comparing the action/omission distinction to the contact/noncontact distinction (scenario as random effect:  $F(1,10) = 7.09, p < .05$ ; subject as random effect:  $F(1,598) = 1.20, p < .05$ ).

In summary, our subjects exhibited a greater difference in their judgments of actions and omissions for punishment judgments than wrongness judgments. This was predicted by the

Table 4  
Mean judgments of deserved punishment and moral wrongness for each of the three distinctions tested in Experiment 3

	Action	Omission	Means	Side-effect	Contact	Noncontact
Punishment	3.5	2.15	2.63	2.36	3.42	2.86
Wrongness	3.75	2.78	3.27	2.93	3.68	3.24

finding that actions generate stronger causal attributions than omissions (Experiment 1), and that punishment judgments depend more on the analysis of causal responsibility than do wrongness judgments (Cushman, 2008a). No such effect was predicted for the means/side-effect or contact/noncontact distinction, which do not generate differential causal attributions (Experiment 1); and no such effect was found for those distinctions. Consequently, Experiment 3 provides evidence that the action/omission distinction influences moral judgments—especially punitive judgments—in part because actions generate more robust causal representations than omissions.

Notably, in U.S. law harmful actions are indeed punished more often than harmful omissions (Loewy, 2003, p. 157). There are circumscribed instances where failing to aid can be punished; for instance, parents and employers have specialized obligations toward their children and employees. Broadly speaking, however, the law has general provisions prohibiting acts that cause harm, but no comparably general provisions prohibiting inactions that allow harm. On the one hand, this introduces a potential confound in our experimental design. Possibly the enhancement of the action/omission distinction in subjects' punitive judgments was not driven by the enhanced role of causal attribution for punitive judgment, but rather by a direct mirroring of legal codes. On the other hand, the central role of the action/omission distinction in legal codes may in fact reflect a more fundamental feature of the psychology of punishment: It depends largely on the assignment of casual responsibility, which in turn depends on a natural distinction between active and passive causation. The present study does not decide between these hypotheses and so the data must be interpreted cautiously.

## **5. General discussion**

People's moral judgments are characterized by consistent patterns that are often complex. In the present study, we sampled a small corner of this complexity, focusing on three distinctions that have been demonstrated to play a role in shaping our moral judgments: action versus omission, means versus side-effect, and contact versus noncontact. To account for these patterns, and the distinctions that appear to underlie them, we explored two potential models.

According to the basic model, mechanisms of moral judgment operate directly over representations such as action, omission, means, side-effect, and so forth. In other words, the computations that yield moral judgment are stated in terms of these representations (e.g., ‘harm caused by an action is morally bad;’ ‘harm caused by an omission is not morally bad’).

According to the derived model, by contrast, distinctions such as means versus side-effect and action versus omission shape our causal and intentional attributions, which subsequently form the basis of our moral judgments. On this model some patterns observed in our moral judgments are derived from the architecture of cognition in non-moral domains such as folk-psychological and causal cognition, rather than from a domain-specific moral faculty. The results of Experiment 1 contribute important detail to the derived model, demonstrating that in nonmoral situations there is (a) an effect of the action/omission distinction on causal attribution and, to a lesser extent, on intentional

attribution; (b) an effect of the means/side-effect distinction on intentional attribution but not causal attribution; and (c) no effect of the contact/noncontact distinction on either type of attribution.

Drawing on these results, we proposed that the action/omission distinction affects moral judgment principally via causal attribution, while the means/side-effect distinction affects moral judgment principally—perhaps exclusively—via intentional attribution. Experiment 2 provided support for this specific hypothesis, demonstrating that the means/side-effect distinction is selectively disrupted in judgments of accidental crimes (where there is a clear lack of intent), while the action/omission distinction is selectively disrupted in judgments of attempted crimes (where no harm is caused). Experiment 3 provided further support for the derived model, demonstrating that the magnitude of the moral distinction between actions and omissions, especially, is predicted by the differential influence of causal attributions on judgments of punishment versus wrongness.

The derived model is broadly consistent with several past proposals that suggest the reliance of moral principles on psychological factors such as cause and intent (Baron & Ritov, 2004, 2009; Cushman et al., 2006; Greene et al., 2009; Mikahil, 2007; Rozman & Baron, 2002; Spranca et al., 1991; Waldmann & Dieterich, 2007). For example, Mikahil (2007) accounts for the means/side-effect distinction in terms of an analysis of “action plans” specifying generic and not necessarily morally specific means/end reasoning. Greene et al. (2009) extend this “action plan” account by proposing a system that monitors one’s own means-end action plans. Waldmann and Dieterich (2007) account for the effect of the “locus of intervention” on moral judgment in terms of underlying causal structure. Specifically, they demonstrate that intervening directly on a victim (e.g., muffling a grenade by throwing a person onto the grenade) is considered worse than intervening directly on a threatening object (e.g., muffling a grenade by throwing the grenade under a person). Baron and colleagues have argued that both the action/omission distinction (Baron & Ritov, 2004, 2009; Spranca et al., 1991) and the means/side-effect distinction (Rozman & Baron, 2002) depend in part on a general property of causal attribution: the notion of “direct” versus “indirect” causation.

In their reliance on causal and intentional analyses, these theories are highly compatible with the derived model. To varying degrees, each of them also suggests that certain moral distinctions may derive from the properties of computations external to the moral domain. But there have been few attempts to experimentally distinguish between the hypothesis that the relevant causal and intentional distinctions are unique to the process of moral judgment and the alternative hypothesis that the relevant distinctions derive from general attributional processes that operate outside the moral domain (Baron & Ritov, 2009; Spranca et al., 1991). Our use of nonmoral scenarios in Experiment 1 provides unique support for this latter hypothesis, while Experiments 2 and 3 provide direct evidence that intentional and causal attributions mediate the effects of the means/side-effect and action/omission distinctions on ordinary moral judgments.

We suggest that this case-study of action/omission and means/side-effect distinctions points toward an important, broader implication: Moral judgments computed over nonmoral representations inherit the unique and sometimes complex properties of those

nonmoral domains. In particular, the computation of a moral judgment employing the simple rule such as “it is wrong to intentionally cause harm” inherits the distinctions that arise during the attribution of intent and the attribution of causal responsibility. This point is perhaps best appreciated by first-year law students. The Model Penal Code defines murder in a single sentence: “A person commits criminal homicide if he purposely, knowing, recklessly or negligently causes the death of another human being.” But as students of the law quickly learn, the brevity of this rule is belied by intricate legal doctrines that attempt to define the relevant standards of causation and culpability. Research in moral psychology faces the parallel challenge of distinguishing domain-specific moral computations from the effects of other domains on moral judgment (Hauser, 2006). This is a challenge that echoes throughout the cognitive sciences as we discover how many “higher” mental functions depend upon a core set of conceptual primitives, and thereby reflect their idiosyncratic structure (Pinker, 2007).

The present study did not reveal any effect of the contact/noncontact distinction on attributions of intention or causation. At a minimum, this finding suggests that some alternative theory is necessary to account for the role of this distinction in shaping moral judgment. Recent research suggests that “personal” moral dilemmas involving immediate physical harm to specific individuals engage an affective response that influences moral judgment (Greene, Nystrom, Engell, Darley, & Cohen, 2004; Greene et al., 2001; Koenigs et al., 2007). It is possible that the emotional salience of harm by direct physical contact contributes to its role in moral judgment.<sup>4</sup>

As noted above, some philosophers endorse the moral relevance of the distinction between actions and omissions (Quinn, 1989) and between necessary means and foreseen side-effect (Foot, 1967), and these moral distinctions also arise in policy and the law. We can now begin to understand their cognitive origins. A common method in moral philosophy is to derive explicit principles from consistent patterns of intuition (e.g., Fischer & Ravizza, 1992). However, the process of abstracting explicit moral principles from consistent patterns of intuitive moral judgment can take distinctions operative in nonmoral domains and restate them as specifically moral; for example, it is possible to restate the action/omission distinction operative in causal reasoning as a moral rule concerning “killing” versus “letting die.” Haidt and colleagues have investigated the tendency of people to adduce apparently rational bases for their moral judgments, and typically emphasize the insufficiency of these post hoc rationalizations as descriptive accounts of the criteria for moral judgment (reviewed in Haidt, 2001). But when subjects were asked to justify their moral judgments for our target principles, a large majority cited the distinction between actions and omissions as a criterion for moral judgment, and a minority also cited the distinction between necessary means and foreseen side-effect (Cushman et al., 2006). In these instances, subjects may have used their capacity for post hoc rationalization to successfully pick out the relevant distinctions between cases. Like philosophers, however, subjects sometimes state their explicit moral rules directly over terms such as action, omission, necessary means, and side-effect, thereby “moralizing” distinctions that emerge in nonmoral contexts as well and affect moral judgment only indirectly via nonmoral representations such as cause and intent. For instance, in our previous research one subject

explained his judgments by writing, “All things being equal, inaction is more just than action.” In these instances, specifically moral principles reflect the structure of nonmoral psychological representations of causation and intent (Cushman, 2008b; Cushman & Greene, in press).

At first blush, it may be surprising that moral distinctions rooted in causal or intentional attribution are often described at the level of explicit principles in terms of action versus omission, or means versus side-effect. Why distinguish “active euthanasia” and “passive euthanasia,” rather than just stating the difference as “a death you cause, versus a death you don’t?” And why construct the elaborate doctrine of double effect to account for the difference between pushing a man in front of a train versus diverting a train onto a man, rather than just saying “a death you intend, versus a death you don’t?”

Our study cannot definitively answer this question, but a hint comes from the ratings of causation and intention in the nonmoral condition of Experiment 1. The action/omission and means/side-effect distinctions caused statistically significant but modest shifts in causal and intentional attributions. The modest pattern, we suggest, may be attributed to an explicit sense that actions and omissions can in fact both be causal, and means and side-effects can be brought about with equal intent. In other words, ordinary people may hold at an explicit level that there is no actual difference along the dimensions of causation and intent between actions and omissions and means and side effects. Thus, they may resist stating moral principles directly in terms of causal and intentional differences (see also Cushman, 2008a; Cushman & Greene, in press).

In summary, the results of the present study support the hypothesis that prominent distinctions evident in moral judgment are derivative of psychological mechanisms not specific to the moral domain. The action/omission distinction supports different moral judgments principally because of its effect on representations of an agent’s causal role in bringing about a harm, while the distinction between necessary means and foreseen side-effect supports different moral judgments because of its effect on representations of an agent’s intention to bring about harm. We suggest that these results reflect a general property of our moral psychology: Complex patterns of moral judgment may be grounded in complex processes of causal and intentional attribution filtered through relatively simple moral computations. Finally, we suggest that these distinctions become codified as explicit moral rules by a process of post hoc reasoning that is common to philosophers, lawmakers, and ordinary people alike.

## Notes

1. In Experiment 3, Royzman and Baron rely on the “causal directness” interpretation of the means/side-effect distinction as evidence that causal directness is also responsible for the action/omission distinction. Specifically, they show that the magnitude of the action/omission distinction is reduced in cases of side-effects and argue that this is because the action is no longer perceived to be causally direct. However, our own results fail to support the “causal directness” interpretation of means/side-effect



distinction, as we discuss further below. The interaction between the action/omission and means/side-effect distinctions and its psychological basis are important topics for further research.

2. These mean values are corrected for unequal numbers of participants in each cell of our design by calculating a mean value for all participants within a cell, and then calculating the mean value of each effect of interest across cells.
3. Some caution is warranted in interpreting these ANOVA results specifically because the size of the moral difference between actions and omissions is much larger than for the other two moral distinctions. However, there is some reason to think that this cannot account for our finding—after all, the trend we observe for the action/omission cases is not just smaller for the means/side-effect cases; rather, it is fully reversed. This argument can be also be made quantitatively. The size of the action/omission effect was 0.38 points larger for punishment judgments than wrongness judgments. And the means/side-effect effect for wrongness judgments was 36% as large as the action/omission effect for wrongness judgments. Thus, we expect means/side-effect effect for punishment judgments to be about  $0.38 \text{ points} \cdot 36\% = 0.14 \text{ points}$  larger than means/side-effect effect for wrongness judgments. However, the observed difference between punishment and wrongness judgments was  $-0.08$  points, and the upper bound of the 95% confidence interval for this difference is 0.13 points. This scaling procedure is imperfect—in particular, it ignores the difficulty of scaling variance estimates for the action/omission distinction. But it does suggest that the predicted interaction between dependent measure and moral distinction is unlikely to be a pure artifact of effect size differences between the moral distinctions.
4. Notably, in past research (Cushman et al., 2006) we found that a disproportionately large number of people disavow the contact/noncontact distinction, compared with the means/side-effect and action/omission distinctions. Could this disavowal relate to the fact that the contact/noncontact distinction is unique in its nonreliance on causal or intentional attribution? This question deserves further research.

## Acknowledgments

We thank Susan Carey and Jonathan Baron for their comments on this manuscript, and especially Marc Hauser for guidance during the project as well as detailed and insightful revisions. Funding for this research was provided by a grant from the Department of Psychology at Harvard University. We are also grateful for support from the Harvard Mind/Brain/Behavior Initiative.

## References

Alicke, M. (1992). Culpable causation. *Journal of Personality and Social Psychology*, 63(3), 368–378.

- Alicke, M. (2000). Culpable control and the psychology of blame. *Psychological Bulletin*, 126(4), 556–574.
- Bargh, J. A. (1999). The unbearable automaticity of being. *American Psychologist*, 54(7), 462–479.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173–1182.
- Baron, J., & Ritov, I. (2004). Omission bias, individual differences, and normality. *Organizational Behavior and Human Decision Processes*, 94, 74–85.
- Baron, J., & Ritov, I. (2009). Protected values and omission bias as deontological judgments. In D. M. Bartels, C. W. Bauman, L. J. Skitka, & D. Medin (Eds.), *Moral judgment and decision making* (Vol. 50, pp. 133–167). San Diego, CA: Academic Press.
- Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind* (pp. 163–228). New York: Oxford University Press.
- Cushman, F. A. (2008a). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, 108(2), 353–380.
- Cushman, F. A. (2008b). *The origins of moral principles*. Unpublished Doctoral Dissertation, Harvard University.
- Cushman, F. A. (in press). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, 108, 353–380.
- Cushman, F. A., Dreber, A., Wang, Y., & Costa, J. (2009). Accidental outcomes guide punishment in a ‘trembling hand’ game. *PLoS ONE*, 4(8), e6699. doi: 6610.1371/journal.pone.0006699.
- Cushman, F. A., & Greene, J. D. (in press). The philosopher in the theater. In M. Mikulincer & P. R. Shaver (Eds.), *Social psychology of morality: The origins of good and evil*. Washington, DC: APA Press.
- Cushman, F. A., Knobe, J., & Sinnott-Armstrong, W. (2008). Moral appraisals affect doing/allowing judgments. *Cognition*, 108(1), 281–289.
- Cushman, F. A., Young, L., & Hauser, M. D. (2006). The role of conscious reasoning and intuitions in moral judgment: Testing three principles of harm. *Psychological Science*, 17(12), 1082–1089.
- Darley, J. M., & Shultz, T. R. (1990). Moral rules – their content and acquisition. *Annual Review of Psychology*, 41, 525–556.
- Dwyer, S. (2004). How good is the linguistic analogy Available at: <http://www.umbc.edu/philosophy/dwyer>. Accessed February 25, 2004.
- Fincham, F. D., & Jaspers, J. (1979). Attribution of responsibility to the self and other in children and adults. *Journal of Personality and Social Psychology*, 37(9), 1589–1602.
- Fischer, J. M., & Ravizza, M. (1992). *Ethics: Problems and principles*. New York: Holt, Rinehart & Winston.
- Foot, P. (1967). The problem of abortion and the doctrine of double effect. *Oxford Review*, 5, 5–15.
- Greene, J. D., Cushman, F. A., Stewart, L. E., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2009). Pushing moral buttons: The interaction between personal force and intention in moral judgment. *Cognition*, 111(3), 364–371.
- Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron*, 44, 389–400.
- Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, 293, 2105–2108.
- Greenwald, A., & Banaji, M. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102(1), 4.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108, 814–834.
- Haidt, J., & Hersh, M. A. (2001). Sexual morality: The cultures and emotions of conservatives and liberals. *Journal of Applied Social Psychology*, 31(1), 191–221.
- Haidt, J., Koller, S. H., & Dias, M. G. (1993). Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology*, 65(4), 613–628.
- Hart, H. L. A., & Honore, T. (1959). *Causation in the law*. Oxford, England: Clarendon Press.

- Hauser, M. D. (2006). *Moral minds: How nature designed a universal sense right and wrong*. New York: Harper Collins.
- Hauser, M. D., Cushman, F. A., Young, L., Jin, R., & Mikhail, J. M. (2007). A dissociation between moral judgment and justification. *Mind and Language*, 22(1), 1–21.
- Hauser, M. D., Young, L., & Cushman, F. A. (2008). Reviving Rawls' linguistic analogy. In W. Sinnott-Armstrong (Ed.), *Moral psychology* (Vol. 2, pp. 107–143). Cambridge, MA: MIT Press.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.
- Henrich, N., & Henrich, J. (2007). *Why humans cooperate: A cultural and evolutionary explanation*. New York: Oxford University Press.
- Hirschfeld, L. A., & Gelman, S. A. (1994). *Mapping the mind. Domain specificity in cognition and culture*. Cambridge: Cambridge University Press.
- Knobe, J. (2003). Intentional action in folk psychology: An experimental investigation. *Philosophical Psychology*, 16, 309–324.
- Knobe, J. (2006). The concept of intentional action: A case study in the uses of folk psychology. *Philosophical Studies*, 130, 203–231.
- Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F. A., Hauser, M. D., & Damasio, A. (2007). Damage to the prefrontal cortex increases utilitarian moral judgments. *Nature*, 446, 908–911.
- Kohlberg, L. (1981). *Essays on moral development, Volume 1: The philosophy of moral development*. New York: Harper Row.
- Lieberman, D., Tooby, J., & Cosmides, L. (2003). Does morality have a biological basis? An empirical test of the factors governing moral sentiments relating to incest. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(1517), 819.
- Loewy, A. H. (2003). *Criminal law*. St. Paul, MN: Thompson West.
- Mikhail, J. (2007). Universal moral grammar: Theory, evidence, and the future. *Trends in Cognitive Science*, 11(4), 143–152.
- Mikhail, J. M. (2000). *Rawls' linguistic analogy: A study of the 'generative grammar' model of moral theory described by John Rawls in 'A theory of justice'*. Unpublished Doctoral Dissertation, Cornell University.
- Nichols, S. (2002). Norms with feeling: Toward a psychological account of moral judgment. *Cognition*, 84, 221–236.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3), 231–259.
- Piaget, J. (1965/1932). *The moral judgment of the child*. New York: Free Press.
- Pinker, S. (2007). *The stuff of thought*. New York: Viking.
- Pizarro, D. A., Uhlmann, E., & Bloom, P. (2003). Causal deviance and the attribution of moral responsibility. *Journal of Experimental Social Psychology*, 39, 653–660.
- Quinn, W. S. (1989). Actions, intentions, and consequences: the doctrine of doing and allowing. *The Philosophical Review*, 145, 287–312.
- Rachels, J. (1975). Active and passive euthanasia. *New England Journal of Medicine*, 292, 78–80.
- Rai, T., & Holyoak, K. (2010). Moral principles or consumer preferences? Alternative framings of the trolley problem. *Cognitive Science*, 34(2), 311–321.
- Ritov, I., & Baron, J. (1990). Reluctance to vaccinate: Omission bias and ambiguity. *Journal of Behavioral Decision Making*, 3, 263–277.
- Ritov, I. I., & Baron, J. (1999). Protected values and omission bias. *Organizational Behavior and Human Decision Processes*, 79(2), 79–94.
- Robinson, P., & Kurzban, R. (2007). Concordance & conflict in intuitions of justice. *Minnesota Law Review*, 91(6), 1829–1907.
- Royzman, E., & Baron, J. (2002). The preference for indirect harm. *Social Justice Research*, 15(2), 165–184.
- Saxe, R., Carey, S., & Kanwisher, N. (2004). Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annual Review of Psychology*, 55, 1–27.

- Shenhav, A., & Greene, J. (2010). Moral judgments recruit domain-general valuation mechanisms to integrate representations of probability and magnitude. *Neuron*, 67(4), 667–677.
- Shultz, T. R., Schleifer, M., & Altman, I. (1981). Judgments of causation, responsibility, and punishment in cases of harm-doing. *Canadian Journal of Behavioral Science*, 13(3), 238–253.
- Shultz, T. R., Wright, K., & Schleifer, M. (1986). Assignment of moral responsibility and punishment. *Child Development*, 57(1), 177–184.
- Sperber, D. (1995). Introduction. In D. Sperber, D. Premack, & A. Premack (Eds.), *Causal cognition* (pp. xv–xx). Oxford, England: Clarendon Press.
- Spranca, M., Minsk, E., & Baron, J. (1991). Omission and commission in judgment and choice. *Journal of Experimental Social Psychology*, 27(1), 76–105.
- Thomson, J. J. (1985). The trolley problem. *The Yale Law Journal*, 94(6), 1395–1415.
- Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, 46, 35–57.
- Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge: Cambridge University Press.
- Waldmann, M. R., & Dieterich, J. H. (2007). Throwing a bomb on a person versus throwing a person on a bomb: Intervention myopia in moral intuitions. *Psychological Science*, 18(3), 247–253.
- Wegner, D. (2002). *The illusion of conscious will*. Cambridge, MA: MIT Press.
- Weiner, B. (1995). *Judgments of responsibility: A foundation for a theory of social conduct*. New York: Guilford Press.
- Wheatley, T., & Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. *Psychological Science*, 16(10), 780–784.
- White, P. A. (1990). Ideas about causation in philosophy and psychology. *Psychological Bulletin*, 108(1), 3–18.
- Young, L., Cushman, F. A., Hauser, M. D., & Saxe, R. (2007). The neural basis of the interaction between theory of mind and moral judgment. *Proceedings of the National Academy of Sciences*, 104(20), 8235–8240.

### Supporting Information

Additional Supporting Information may be found in the online version of this article on Wiley Online Library:

Full Set of Stimuli Used in Experiments 1, 2, and 3.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.