A Dissociation Between Moral Judgments and Justifications

MARC HAUSER, FIERY CUSHMAN, LIANE YOUNG, R. KANG-XING JIN AND JOHN MIKHAIL

Abstract: To what extent do moral judgments depend on conscious reasoning from explicitly understood principles? We address this question by investigating one particular moral principle, the principle of the double effect. Using web-based technology, we collected a large data set on individuals' responses to a series of moral dilemmas, asking when harm to innocent others is permissible. Each moral dilemma presented a choice between action and inaction, both resulting in lives saved and lives lost. Results showed that: (1) patterns of moral judgments were consistent with the principle of double effect and showed little variation across differences in gender, age, educational level, ethnicity, religion or national affiliation (within the limited range of our sample population) and (2) a majority of subjects failed to provide justifications that could account for their judgments. These results indicate that the principle of the double effect may be operative in our moral judgments but not open to conscious introspection. We discuss these results in light of current psychological theories of moral cognition, emphasizing the need to consider the unconscious appraisal system that mentally represents the causal and intentional properties of human action.

A dominant perspective in philosophy, psychology, and law centers on the idea that our moral judgments are the product of a conscious decision in which individuals move directly from conscious reasoning to moral verdict (Dworkin, 1998; Kaplow and Shavell, 2002; Kohlberg, 1981; Korsgaard, 1996; Piaget, 1932/1965). For developmental psychologists such as Piaget and Kohlberg who have followed in this tradition, stages of moral development are thought to track the ability to articulate sound justifications for moral judgments. Under Kohlberg's (1981) scheme, for example, individuals reach the final stage of moral maturity when they are able to justify moral decisions on the basis of a central Kantian principle: treat individuals as ends and never merely as means. From this perspective, individuals at Kohlberg's highest stages of development should be able to support their moral judgments with explicit principles. A corollary prediction is that sources of individual variation in consciously held values and beliefs might affect the nature of moral judgments and justifications, creating pockets of homogeneity in the patterns of responses. For instance, individuals exposed to coursework in moral

For discussion of the data, analyses, and earlier drafts of the paper, we thank Jonathan Baron, Alfonso Caramazza, Susan Carey, Joshua Greene, Jonathan Haidt, Richard Nisbett, and Elizabeth Spelke. Funding for this project was provided to MDH by a Guggenheim Award. All of this work was approved by the Harvard IRB.

Address for correspondence: Marc Hauser, Department of Psychology, Harvard University, Cambridge, MA, 02138, USA.
Email: mdh@wjh.harvard.edu

Mind & Language, Vol. 22 No. 1 February 2007, pp. 1–21.

philosophy may reason from a different, or perhaps more extensive, set of moral principles than those with no such exposure; young individuals, as Piaget (1932/1965) suggested, may focus on the consequences of actions as they are not yet sensitive to non-consequential factors such as the agents' intentions. Similar arguments might be invoked with respect to variation in religious or ethnic background. Although such variation is by no means a necessary outcome of the conscious reasoning perspective, evidence of widely shared patterns of moral judgment would be less expected and more difficult to explain.

An alternative theoretical perspective holds that at least some of our moral judgments are the product of unconscious psychological processes, and thus, intuitive. A significant component of the intuitive perspective places a strong emphasis on the role of emotions. For example, Haidt's (2001; 2003) (Wheatley and Haidt, 2005) work shows that when people confront scenarios involving some kind of disgusting action, they engage in a process called 'moral dumbfounding' in which they fail to give sufficient justifications for their confidently delivered moral judgments. Haidt interprets this finding as evidence that emotions are responsible for the judgments, and both neuroimaging and patient studies appear to confirm the importance of emotional areas of the brain in guiding certain aspects of our moral intuitions (Greene *et al.*, 2001; Greene and Haidt, 2002; Berthoz *et al.*, 2002; Moll *et al.*, 2002; Damasio, 1994). The critical prediction of the intuitionist perspective is a dissociation between judgment and justification for certain moral dilemmas (Haidt, 2001; Hauser, 2006; Mikhail, 2000; Mikhail, in press; Hauser *et al.*, in press; Rawls, 1971).

Aligning more closely with the emotion-oriented perspective, but importantly distinct, is a view that focuses on the causes and consequences of an agent's actions. This view builds on some of the insights of moral philosophers (Kamm, 2000), in addition to making an analogy to generative linguistics (Chomsky, 1986; Dwyer, 1999; Dwyer, 2004; Hauser, 2006; Hauser *et al.*, in press; Mikhail, 2000; Mikhail, in press; Rawls, 1971). In particular, it builds on non-consequential moral philosophy by exploring how the psychology of such distinctions as that between killing and letting die and intended harm and foreseen harm bears on the nature of our moral judgments. This view also draws on the analogy to language by making explicit the distinction between operative and expressed principles. Previous work in moral development, especially as championed by Kohlberg and his students, failed to make these distinctions, focusing exclusively on expressed principles. Part of our motivation here is to contrast the unconsciously operative principles guiding people's judgments with their expressed principles. The other part of our motivation is to show how our moral judgments are mediated by an appraisal system that takes into account the causal and intentional properties of human action. Lest it be misinterpreted, the idea that there is an appraisal system as we describe need not be in conflict with the emotional perspective. Both an appraisal system and emotional processes may play a role in generating moral intuitions, though the timing of their effects may differ. We return to some of these issues in the discussion. For now, the primary contrast is between the *conscious reasoning* perspective and a broadly construed *intuitive* perspective.

Several recent empirical studies of moral judgment, most noticeably those by Petrinovich and colleagues (O'Neill and Petrinovich, 1998; Petrinovich *et al.*, 1993), Mikhail and colleagues (1998), and Greene and colleagues (Greene *et al.*, 2001; Greene *et al.*, 2004), have used modified versions of some classic moral dilemmas developed by philosophers to explore our intuitions about the permissibility of harming or helping others. Of these, the best studied are the trolley problems (Fischer and Ravizza, 1992; Foot, 1967; Kamm, 1992; Kamm, 1998b; Thomson, 1970). In the present work, we build on this tradition in three significant ways. First, we use web-based technology to collect data from a much larger and more diverse population than previously tested. Most of the earlier studies used undergraduate subjects, and those that extended the sample to other demographic or cultural categories were limited. Thus, even for the best studied trolley problems (see our scenarios 1 and 2 below), our current understanding of populational variation is poor. Second, with the exception of Mikhail's work, no other empirical study of moral judgments has looked at the relationship between expressed and operative principles. Here, we provide a quantitative analysis of subjects' justifications, with an eye to exploring the fit between what subjects say and how well this accounts for their judgments. Third, we use two carefully controlled dilemmas to systematically test the sufficiency of the emotional account to explain the pattern of moral judgments. This pair of scenarios was designed to target the *principle of the double effect*, which holds that it may be permissible to harm an individual for the greater good if the harm is not the necessary means to the greater good but, rather, merely a foreseen side effect (Fischer and Ravizza, 1992; Kamm, 1998b; Mikhail, 2000; Thomson, 1970). In particular, one of the paired scenarios was constructed in such a way that the only relevant difference was captured by the distinction between means and side effect. The principle of the double effect has received some attention in the psychological literature (Mikhail *et al.*, 1998; Royzman and Baron, 2002), but neither with as diverse a subject population nor with a within-subjects analysis of justifications for judgments provided in controlled scenario pairs. If operative in some way, the principle is directly relevant to the idea that we have an appraisal system that generates moral judgments based on the causal and intentional properties of human action.

Before presenting our results, we make three points of clarification. First, the subject of this research is moral judgment not behavior. The relevant question is 'what do subjects perceive as morally right or wrong?' as opposed to 'how do subjects employ this moral knowledge in their actual day to day conduct?' It is likely that conscious reasoning plays an important role in determining our actions, as in the familiar experience of weighing the pros and cons of a significant moral choice. We do not deny this process, nor do we explore it. Instead, we provide evidence that among the relevant inputs into this process are intuitive judgments of moral right and wrong.

The second point of clarification is that to probe the nature of people's moral judgments, we have chosen to use artificial dilemmas as opposed to real world cases such as abortion, euthanasia, warfare, or heroic acts of altruism. There are at least three

reasons to employ this methodology. First, by using artificial cases we can guarantee that subjects will have no familiarity with or personal attachment to the particular details of the case. This has several advantages, many of which parallel the arguments made throughout the cognitive sciences. For example, artificial examples eliminate the potential confounding effects of in-group versus out-group biases by simply using anonymous agents. Second, each case can be modified in critical ways in order to isolate salient dimensions. Consequently, the use of artificial moral dilemmas to explore our moral psychology is like the use of theoretical or statistical models with different parameters; parameters can be added or subtracted in order to determine which parameters contribute most significantly to the output. The use of artificial dilemmas also parallels the use of artificial utterances to explore the structure of our linguistic intuitions, or the use of black and white grating patterns and line orientations to explore the psychophysics of vision. Third, philosophers have derived fundamental descriptive and normative principles by considering their own personal intuitions in response to these cases (Fischer and Ravizza, 1992; Kamm, 1998a; Thomson, 1970). By using these moral dilemmas as psychological probes, it is possible to test whether the intuitions of professional philosophers align with those of a larger and more diverse group of people. In addition, although we claim that artificially created moral dilemmas provide a useful method for probing the nature of our intuitions, we also recognize that this approach explores only a fragment of our moral psychology, albeit a potentially significant one.

The third point of clarification concerns the nature of our data set, and the use of web-based technology. The web provides a powerful mechanism for collecting large data sets. There are, however, limitations that we acknowledge up front. In terms of our interest in testing for the effects of demographic and cultural variables, it is clear that our sample is biased. Only English-speaking subjects responded to our dilemmas, even though subjects varied in their nationalities. This bias does not, however, affect the demographic variation we sampled. Our goal is more modest at this point: to explore the extent to which the variation in our sample impacts upon the nature of people's judgments and justifications. Concerning the use of web-based technology more generally, there were early concerns that participants would give more biased or less accurate responses than those obtained by more traditional, pen-and-paper methods. Several recent studies have now shown that the web results replicate those obtained from traditional methods (Baron and Siepmann, 2000; Greenwald *et al.*, 2003; Kraut *et al.*, 2004; Nosek *et al.*, 2002; Schmidt, 1997). In addition, the dilemmas we present on the web have been presented previously using traditional methods (Greene *et al.*, 2001; Mikhail *et al.*, 1998), thus providing an independent standard for comparison with the web-based data.

Methods

1. Subjects

Subjects were voluntary visitors to the Moral Sense Test website (http://www.moral.wjh.harvard.edu) from September 2003 to January 2004. Overall, there were

some 5,000 subjects responding to the dilemmas targeted in this paper, covering 120 countries, but with a strong bias toward English-speaking nationalities. The website was promoted through print and online media coverage, online discussion forums, and word of mouth. All procedures were conducted in accordance with the Institutional Review Board of Harvard University, and followed the testing procedures of other web-based research projects.

2. Testing Procedure

Subjects began by reading a general description of the test and were asked to acknowledge understanding of the test's nature and content. Next, subjects provided personal information. Finally, they received instructions for the test. Subjects were asked to complete the test without interruption, to read through each scenario and associated question once, and to answer the question based solely on the information provided.

The test itself consisted of 19 scenarios divided into four sets. We presented each subject with four scenarios—three moral dilemmas and one control scenario—each drawn randomly from one of four sets. Here, we limit our analysis to four of the 19 scenarios.

Since the scenarios presented in Figure 1 represent abbreviated versions of the actual text, we present the full version in the Appendix, along with the text of the two control scenarios. Order of scenario presentation was randomized, and scenarios were presented on separate pages. Each page contained the scenario on top and the associated permissibility question and response buttons below. For two scenarios—indicated as scenarios 3 and 4 in Figure 1—the text was accompanied by a schematic of the setup. After providing permissibility judgments for all four scenarios, subjects were reminded of their responses to two of the scenarios, presented again with the full text of the scenarios, and asked to provide brief (<150 words) justifications for their judgments. Finally, we gave subjects an opportunity to register their e-mail addresses on the site so they could be contacted for future studies.

3. Data Analysis

We excluded from analysis all subjects who, in one of the two control scenarios, judged that it was permissible to choose a course of action that resulted in a death even though there was a costless alternative. Only 2.7% of subjects failed the control. Subjects were also excluded if they took fewer than four seconds to read and respond to any of the four test scenarios (n = 6); a pilot study revealed that the shortest of these scenarios could not be read and understood in this time. Thus, an overwhelming proportion of subjects were not excluded from the analyses. We adopted these precautions to ensure that subjects included in our analysis appeared to be following the stated instructions and approaching the test in a serious manner.

We analyzed the scenarios by comparing responses across pairs. Scenario pairs were designed to be similar along several salient dimensions in order to isolate the potential effect of morally relevant parameters. When a statistically significant number of subjects rated one scenario in a pair more permissible than the other,

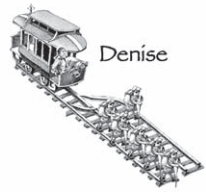
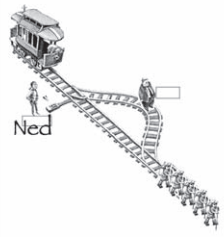
Scenario	Schematic	Description	% 'Yes'
1		Denise is a passenger on a train whose driver has fainted. On the main track ahead are 5 people.The main track has a side track leading off to the left, and Denise can turn the train on to it.There is1 person on the left hand track. Denise can turn the train, killing the 1; or she can refrain from turning the train, letting the 5 die.	Is it morally permissible for Denise to turn the train? 85%
2		Frank is on a footbridge over the train tracks. He sees a train approaching the bridge out of control. There are 5 people on the track. Frank knows that the only way to stop the train is to drop a heavy weight into its path. But the only available, sufficiently heavy weight is1large man, also watching the train from the foot bridge.Frank can shove the 1 man onto the track in the path of the train, killing him; or he can refrain from doing this, letting the 5 die.	Is it morally permissible for Frank to shove the man? 12%
3		Ned is walking near the train tracks when he notices a train approaching out of control. Up ahead on the track are 5 people.Ned is standing next to a switch, which he can throw to turn the train on to aside track. There is a heavy object on the side track. If the train hits the object, the object wills low the train down, giving the men time to escape. The heavy object is 1 man, standing on the side track. Ned can throw the switch, preventing the train from killing the 5 people, but killing the 1 man. Or he can refrain from doing this, letting the 5 die.	Is it morally permissible for Ned to throw the switch? 56%
4		Oscar is walking near the train tracks when he notices a train approaching out of control. Up ahead on the track are 5 people. Oscar is standing next to a switch, which he can throw to turn the train on to aside track. There is a heavy object on the side track. If the train hits the object, the object will slow the train down, giving the 5 people time to escape. There is 1 man standing on the sidetrack infront of the heavy object. Oscar can throw the switch, preventing the train from killing the 5 people, but killing the1man. Or he can refrain from doing this, letting the five die.	Is it morally permissible for Oscar to throw the switch? 72%

Figure 1 *A schematic illustration of the scenario, the description and question provided to each subject, and the proportion of subjects responding affirmatively to the question. The description presented here represents an abbreviated version of the actual text that is described in the supplementary material, along with control scenarios. The schematic illustration is provided here, but was not given to subjects.*

we inferred that this difference was the consequence of the parameter(s) in question. In what follows, we first consider subjects' judgments and then their justifications.

Results

1. Do Subjects Use the Principle of the Double Effect?

Scenarios 1 and 2 (see Figure 1) elicited subjects' judgments for two moral dilemmas that have received significant attention in both philosophy and psychology, but with small sample sizes, limited in both demographic and cultural variation. In scenario 1, Denise is a passenger on an out-of-control train. Denise can let the train hit five individuals on the track ahead, or she can turn the train down a side track towards a single individual. We asked subjects whether it was morally permissible for Denise to turn the train onto the side track. In scenario 2, Frank is standing on a footbridge above the railroad tracks when an out-of-control train approaches. Frank can allow the train to pass and hit five people ahead, or he can shove a heavy man next to him in front of the train, stopping the train in time to save the five. We asked subjects whether it was morally permissible for Frank to shove the man.

In order to eliminate the possibility of order effects, we restricted our analyses to first-trial responses, with comparisons made between-subjects. Analyses focused on the proportion of subjects indicating that an act is permissible or impermissible along with a 95% confidence interval for each mean. All other statistical analyses were conducted as one-tailed chi square tests with significance set at $p < 0.05$. We used one-tailed tests because we had a priori predictions from both philosophical intuition and earlier studies concerning the significance of foreseen over intended consequences.

Permissibility judgments for these cases were widely shared, independently of the order in which they were presented in the session: Denise's action was judged permissible by 89% of subjects ($p[.87 < \mu < .91] = 0.05$), while Frank's action was judged permissible by 11% of subjects ($p[.09 < \mu < .13] = 0.05$). These proportions differ significantly ($\chi^2[1, N = 2646] = 1615.96$, $p < 0.001$), with an effect size of $w = 0.78$. This result indicates that, as a group, subjects reconstruct and make use of the information that varies between these two cases in determining their moral judgments. Such information might relate to the principle of the double effect, or might also include the fact that Frank makes physical contact with a person, whereas Denise does not, or that Frank introduces a new threat whereas Denise redirects an existing threat.

Scenarios 3 and 4 tested subjects' judgments of two moral dilemmas tailored to differ along only a single dimension: whether battery to a single individual was an intended means to the saving of five (Ned, scenario 3) or was merely a foreseen side effect (Oscar, scenario 4). For both cases, the act is impersonal (throwing a switch), the potential consequences are the same (harming one or five), and the

train will kill five if unimpeded but only one if it is redirected. If subjects judge the permissibility of these two actions differently, this would implicate the use of the principle of the double effect. Judgments were again compared using first trial data and a between-subjects design.

In Scenario 3, the intended harm case, 56% of subjects judged the action permissible ($p[.53 < \mu < .59] = 0.05$), whereas in scenario 4, the foreseen harm case, 72% of subjects judged the action permissible ($p[.69 < \mu < .74] = 0.05$). These percentages differ significantly ($\chi^2[1, N = 2612] = 72.35$, $p < 0.001$), with an effect size of $w = 0.17$. This result indicates that, as a group, subjects make use of the principle of the double effect.

2. Is the Principle of the Double Effect Observed Across Demographic and Cultural Variation?

To explore the potential influence of demographic and cultural variation in subjects' judgments, we tested whether the pattern of split judgments between scenarios 1 and 2, and also between 3 and 4, was maintained in defined subpopulations of our total subject population. That is, do subjects of different genders, ages, educational backgrounds, ethnicities, religions and national affiliations judge that scenario 1 is more permissible than scenario 2, and that scenario 3 is less permissible than scenario 4? By only testing subpopulations for which we had sufficient statistical power (first at 0.95 and then again at 0.80) to detect a significant discrepancy at $p < .05$, we framed the question of consistency across subpopulations as a falsifiable hypothesis: does any subpopulation exist for which we have the power to find a significant effect, but for which no significant effect exists? The threshold for statistical power was determined using the effect sizes obtained for the entire subject population.

Based on the effect size for the entire subject population judging scenarios 1 and 2 ($w = 0.74$), we calculated the minimum sample size required to detect a significant difference between these cases with an equivalent effect size at $p < .05$ with a probability $\beta = 0.95$; this sample size was 22. What this means is that for every subset of 22 subjects, we should have a 95% probability of detecting a difference between scenarios 1 and 2 equivalent to the difference observed for the total population. A total of 33 subpopulations were defined and in every case we found a significant difference in subjects' judgments for scenarios 1 and 2 (see Table 1) at $p < .001$.

The procedure was repeated with subsets of 13 or more subjects, lowering the probability of detecting the effect to $\beta = .80$ and presenting a more stringent test of cross-cultural consistency. An additional 3 subsets were included. For all three of these subsets the difference between scenarios was detected at $p < 0.05$.

Based on the effect size of the contrast between scenarios 3 and 4 for the whole population ($w = 0.16$) we calculated that a minimum of 391 subjects would be needed to detect a significant difference between these cases of equivalent effect size ($p < 0.05$) with a probability $\beta = 0.95$. Because of the larger number of subjects needed, many fewer demographic subsets were available for scenarios

Table 1. *Statistical analyses testing for a difference in permissibility between scenarios 1 and 2 across subpopulations defined by likely sources of variation in judgments. Plain text rows refer to analyses where the power required to detect a statistical difference was 0.95. Italicized text rows refer to analyses where the power required to detect a statistical difference was 0.80. Values under each scenario refer to the mean proportion of subjects judging each case as morally permissible.*

Category	N	Scenario 1	Scenario 2	p value (one-tailed)
Exposure to Moral Philosophy				
No	2098	0.89	0.11	<.001
Yes	548	0.91	0.11	<.001
National Affiliation				
Australia	49	0.90	0.15	<.001
Brazil	31	0.82	0.29	0.002
Canada	93	0.84	0.12	<.001
India	24	0.88	0.00	<.001
United States	2218	0.90	0.10	<.001
United Kingdom	58	0.83	0.15	<.001
Ethnicity				
American Indian/ Alaskan Native	*18*	*0.40*	*0.07*	*0.048*
Asian/Pacific Islander	85	0.80	0.12	<.001
Black Non-Hispanic	27	0.90	0.26	<.001
Hispanic	293	0.80	0.16	<.001
White Non-Hispanic	9183	0.90	0.10	<.001
Current Religion				
Buddhist	71	0.88	0.16	<.001
Catholic	375	0.88	0.06	<.001
Orthodox Christian	36	0.73	0.14	<.001
Protestant	778	0.93	0.08	<.001
Christian (Other)	347	0.85	0.12	<.001
Hindu	17	0.81	0.00	<.001
Jewish	72	0.93	0.14	<.001
Muslim	*20*	*0.78*	*0.90*	*0.001*
None	784	0.89	0.16	<.001
Highest Educational Level Attained				
Middle School	*15*	*0.78*	*0.33*	*0.043*
Some High School	96	0.83	0.14	<.001
High School	362	0.83	0.08	<.001
Some College	938	0.90	0.10	<.001
BA	709	0.92	0.11	<.001
Masters	308	0.91	0.15	<.001
PhD	207	0.90	0.13	<.001
Age				
<20	117	0.80	0.21	<.001
20s	398	0.87	0.18	<.001
30s	458	0.89	0.15	<.001
40s	601	0.93	0.09	<.001
50s	606	0.89	0.06	<.001

(Continued)

Table 1. *Continued.*

Category	N	Scenario 1	Scenario 2	p value (one-tailed)
60s	337	0.89	0.08	<.001
70s	116	0.88	0.10	<.001
Gender				
Male	1490	0.88	0.90	<.001
Female	1156	0.90	0.13	<.001

3 and 4 than for scenarios 1 and 2. Subsets were designated according to formal exposure to moral philosophy, religion, highest educational level attained, age, or gender. For ethnicity and nationality, there were insufficient subjects beyond the predominant categories of US and white non-Hispanic.

A total of 11 subsets were defined and in every case the difference between scenarios 3 and 4 was detected at $p < 0.05$ (see Table 2). As with scenarios 1 and 2, the procedure was repeated with subsets of 224 or more subjects, lowering the probability of detecting the effect to $\beta = 0.80$. An additional 5 subsets were included, and for all five a significant difference was observed.

3. Are there Differences in the Extent to which Subjects Employ the Principle of the Double Effect across Different Sub-populations?

The approach employed in the previous section confirms that the principle of the double effect is operative across a wide range of sub-populations. Here we ask whether there are significant differences in the extent to which it is used across subpopulations of our subjects. The appropriate method in this case is to look at whether there are significant differences between subpopulations in the proportion of subjects who gave the opposite judgments to cases 1 and 2, or to cases 3 and 4. Since this necessitates a comparison across two scenarios for each subject, however, we can no longer limit our analyses to first-trial responses. A within-subjects design introduces the influence of order effects and interference between scenarios. Particularly in the case of scenarios 3 and 4, which were phrased identically except for the critical difference between foreseen and intended harm, subjects were unlikely to judge both cases differently in a single test session, even though, as discussed above, between-subjects analysis of first-trial responses indicates a significant difference in permissibility ratings between the two scenarios.

Where subpopulations were categorically defined (gender, exposure to moral coursework, ethnicity, religion and national affiliation) analyses were conducted by chi square tests. All subpopulations tested in analysis 2 were included in the chi square tests unless their inclusion resulted in an expected value of less than 5 in the chi square procedure. Where subpopulations were continuously defined (age and educational level) we used a linear regression.

Table 2. *Statistical analyses testing for the use of the principle of the double effect in scenarios 3 and 4 across subpopulations defined by likely sources of variation in judgments. Plain text rows refer to analyses where the power required to detect a statistical difference was 0.95. Italicized text rows refer to analyses where the power required to detect a statistical difference was 0.80. Values under each scenario refer to the mean proportion of subjects judging each case as morally permissible.*

Category	N	Scenario 3	Scenario 4	p value (one-tailed)
Exposure to Moral Philosophy				
No	2030	0.57	0.70	<.001
Yes	582	0.53	0.78	<.001
Current Religion				
Catholic	*365*	*0.56*	*0.75*	*<.001*
Protestant	799	0.58	0.74	<.001
None	765	0.53	0.72	<.001
Highest Educational Level Attained				
High School	*336*	*0.51*	*0.66*	*0.002*
Some College	930	0.56	0.71	<.001
BA	686	0.60	0.76	<.001
Masters	*327*	*0.57*	*0.75*	*0.001*
Age				
20s	*406*	*0.48*	*0.73*	*<.001*
30s	461	0.53	0.71	<.001
40s	575	0.57	0.66	0.014
50s	591	0.56	0.75	<.001
60s	*327*	*0.61*	*0.75*	*0.003*
Gender				
Male	1492	0.49	0.68	<.001
Female	1120	0.61	0.75	<.001

For scenarios 1 and 2, none of the 5 chi-square tests yielded significant differences between subpopulations and neither of the regressions turned up significant linear relationships. The results of these tests are summarized in Table 3.

For scenarios 3 and 4, chi square analysis could not be run for subpopulations defined by ethnicity or national affiliation because in each case only one subpopulation had an expected value exceeding 5. Of the remaining three chi-square tests, one achieved significance. There was a significant difference in the proportions of Catholics, Protestants and Atheists who judged scenarios 3 and 4 differently ($\chi^2[1, N = 623] = 7.56$, $p = 0.023$). However, the differences between groups were small: 5.6% of Catholics, 2.0% of Protestants and 7.2% of Atheists judged scenarios 3 and 4 differently. There were no significant differences between subjects who had and had not taken formal coursework on moral philosophy, or between men and women.

Separate linear regressions were performed for both age and educational level. Linear regression revealed that age significantly predicted the proportion of subjects who judged scenarios 3 and 4 differently ($F[1,841] = 11.96$, $p = .001$). However, age accounted for only 1.4% of the variance. Educational level did not

Table 3. *Statistical analyses testing for the influence of likely sources of variation on the probability of judging scenarios 1 and 2 differently.*

Subpopulations	Test Statistic
Exposure to Moral Philosophy	
No	$\chi^2(1) = 0.239$, $p = 0.62$
Yes	
National Affiliation	
Australia	$\chi^2(3) = 1.027$, $p = 0.80$
Canada	
United States	
United Kingdom	
Ethnicity	
American Indian/Alaskan Native	$\chi^2(2) = 1.340$, $p = 0.51$
Asian/Pacific Islander	
White non-Hispanic	
Current Religion	
Buddhist	$\chi^2(4) = 3.062$, $p = 0.55$
Catholic	
Protestant	
Jewish	
None	
Highest Educational Level Attained	
Middle School	$F(1, 847) = .005$, $p = .944$, $r = .002$
Some High School	
High School	
Some College	
BA	
Masters	
PhD	
Age	
All Ages	$F(1, 847) = 2.633$, $p = .105$, $r = .056$
Gender	
Male	$\chi^2(1) = 2.914$, $p = 0.088$
Female	

significantly predict the proportion of subjects who judged scenarios 3 and 4 differently ($F[1,841] = 1.24$, $p = .27$). The results of these tests are summarized in Table 4.

4. Are Subjects Able to Provide Sufficient Justifications for their Judgments?

To explore subjects' ability to explicitly articulate the principle(s) responsible for their pattern of judgments, we targeted those subjects who provided different judgments to scenarios 1 and 2, or to scenarios 3 and 4, and asked them to justify their contrasting judgments. Subjects' justifications were coded into three categories by the experimenters: (1) sufficient justification, (2) insufficient justification, and

Table 4. *Statistical analyses testing for the influence of likely sources of variation on the probability of judging scenarios 3 and 4 differently.*

Subpopulations	Test Statistic
Exposure to Moral Philosophy	
No	χ^2 (1) = 0.916, p = 0.339
Yes	
Current Religion	
Catholic	χ^2 (2) = 7.555, p = 0.023
Protestant	
None	
Highest Educational Level Attained	
Middle School	F (1,842) = 1.236, p = .266, r = .038
Some High School	
High School	
Some College	
BA	
Masters	
PhD	
Age	
All Ages	F (1,841) = 11.956, p = .001, r = .118
Gender	
Male	χ^2 (1) = 3.625, p = 0.057
Female	

(3) discountable justification. These categories were derived after reading through several hundred justifications, extracting common patterns, and then achieving high inter-observer reliabilities in coding:

Category 1: A sufficient justification was one that correctly identified any factual difference between the two scenarios and claimed the difference to be the basis of moral judgment. We adopted this extremely liberal criterion so as not to prejudge what, for any given individual, counts as a morally relevant distinction; needless to say, in evaluating the merits of some justifications, it is clear that some distinctions (e.g. the agent's gender) will not carry any explanatory weight. The differences that subjects typically identified included: (1) in scenario 1, the death of the one man on the side track is not a necessary means to saving the five, while in scenario 2, the death of the one man on the bridge is a necessary means to saving the five; (2) in scenarios 1, 3 and 4, an existing threat (of the train) is redirected, while in scenario 2, a new threat (of being pushed off the bridge) is introduced; (3) in scenarios 1, 3 and 4, the action (turning the train) is impersonal, while in scenario 2, the action (pushing the man) is personal or emotionally salient.

Category 2: An insufficient justification was one that failed to identify a factual difference between the two scenarios. Insufficient justifications typically fell into one of three subcategories. First, subjects explicitly expressed an inability to account for their contrasting judgments by offering statements such as 'I don't know how to explain it', 'It just seemed reasonable', 'It struck me that way', and 'It was a gut

feeling'. Second, subjects explained that death or killing is 'inevitable' in one case but not in the other without offering any further explanation of how they reasoned this to be the case. Third, subjects explained their judgment of one case using utilitarian reasoning (maximizing the greater good) and their judgment of the other using deontological reasoning (acts can be objectively identified as good or bad) without resolving their conflicting responses. Subjects using utilitarian reasoning referred to numbers (e.g., save 5 versus 1 or choose 'the lesser of two evils'). Subjects using deontological reasoning referred to principles, or moral absolutes, such as (1) killing is wrong, (2) playing God, or deciding who lives and who dies, is wrong, and (3) the moral significance of not harming trumps the moral significance of providing aid.

Category 3: Responses that either were blank or in any way included added assumptions were discounted. Examples of assumptions include: (1) men walking along the tracks are reckless, while men working on the track are responsible, (2) the conductor is responsible for acting, while the bystander is not, (3) a man's body cannot stop a train, (4) the five men will be able to hear the train approaching and escape in time, and (5) a third option for action such as self-sacrifice exists and should be considered. Though we do not discuss the cause of these added assumptions further, other research suggests that they arise because subjects are incapable of accounting for the pattern of their judgments (Cushman *et al.*, in press).

For a subset of subjects that judged scenario 1 as permissible and scenario 2 as impermissible, we analyzed their justifications of these two cases with respect to each other. That is, we analyzed the nature of subjects' explanations for why scenario 1 represents a permissible action whereas scenario 2 represents an impermissible action. Two of the authors (LY and FC) coded a subset of cases. Though coding open-ended text like this is difficult, LY and FC achieved a high inter-observer reliability (86%; n = 29). LY then coded the complete set of 597 subjects. Of these, 267 were coded into category 3 and therefore excluded from analysis. Of the remainder, 70% provided insufficient responses (category 2; $p[.65 < \mu < .75] = 0.05$), and 30% provided sufficient responses (category 1; $p[.25 < \mu < .35] = .05$). The sufficiency of justifications was not predicted by age, gender or religion. However, a significantly greater proportion of subjects who had been exposed to readings in moral philosophy were able to provide a sufficient justification (41%) compared to those who had not (27%; $\chi^2[1, N = 330] = 4.650$, $p = 0.031$, two-tailed).

As discussed in section 2 of our results, the proportion of subjects who judged scenarios 3 and 4 differently within a single session was quite small (5.8%), presumably because viewing such superficially similar scenarios one after the other induced subjects to provide identical responses. In order to generate a new, larger sample of subjects with potentially conflicting judgments in scenarios 3 and 4, we re-contacted subjects who had been presented with only one of the scenarios and who had judged scenario 3 as impermissible or scenario 4 as permissible, and asked them to make a judgment on the corresponding case; the interval of time from the first to the second scenario was, on average, 20 weeks. Of the 207 subjects who

responded, 33% judged the foreseen case (scenario 4, Oscar) permissible and the intended case (scenario 3, Ned) impermissible. These subjects were then asked to justify their conflicting judgments vis-à-vis each other. Coding of these justifications was conducted by LY and FC according to the same criteria as scenarios 1 and 2. LY and FC achieved an inter-observer reliability of 76% on which cases to include in category 3. In order to limit the analysis to the clearest cases, any justification coded into category 3 by either observer was omitted from the analysis; of 68 subjects, 45 were coded into category 3 by at least one of the two observers. Of the remaining 23 justifications, LY's and FC's codes were perfectly correlated, with no exceptions. Twenty of the subjects provided an insufficient response (category 2), and three provided a sufficient response (category 1). These data suggest that, at a 95% level of confidence, between 2% and 34% of individuals who perceived a difference between these scenarios would be able to provide a sufficient justification for their judgments. Of the three individuals who provided sufficient justifications, only one had received any formal education in moral philosophy, one was a high school student, and the other two had attained bachelor degrees.

Discussion

The aim of the present study was to adapt the standard philosophical technique of contrasting pairs of similar moral dilemmas in order to answer two questions: whether there are widely shared principles that guide moral judgments in certain contexts and whether these principles are invoked when subjects justify their moral decisions. Our analyses generate two central conclusions: (1) in the context of the trolley problems we studied, all of the demographically defined groups tested within our sample showed the same pattern of judgments and (2) subjects generally failed to provide justifications that could account for the pattern of their judgments.

We contrasted two pairs of trolley problems. For the first pair, scenario 1 (Denise, turning the train) and scenario 2 (Frank, shoving the man), the observed pattern of judgments was consistent with at least three possible moral distinctions: (1) Foreseen versus intended harm (*Principle of the double effect):* it is less permissible to cause harm as an intended means to an end than as a foreseen consequence of an end; (2) *Redirection versus introduction of threat:* it is less permissible to cause harm by introducing a new threat (e.g. pushing a man) than by redirecting an existing threat (e.g. turning an out-of-control train onto a man); and (3) *Personal versus impersonal:* it is less permissible to cause harm by direct physical contact than by an indirect means. The first two distinctions have been discussed in the philosophical literature as the content of plausible moral principles, while the third has emerged from considerations of both behavioral and neurophysiological evidence (Greene *et al.*, 2001; Greene *et al.*, 2004). Scenario 3 (Ned) and scenario 4 (Oscar) were designed to probe just the first of these moral principles, the principle of the double effect.

Results, presented in section 2, show that across a variety of nationalities, ethnicities, religions, ages, educational backgrounds (including exposure to moral philosophy), and both genders, shared principles exist. That is, across every subpopulation tested, scenario 1 (turning the train) elicited a significantly higher proportion of permissibility judgments than scenario 2 (shoving the man), suggesting that one of the three principles described above, or their combination, guided the moral judgments made by each group. Consistency was also observed across several demographic groups in the contrast between scenarios 3 and 4, although fewer subpopulations were tested.

Results presented in section 3 suggest that, not only is the principle of the double effect used in each of the subpopulations, but also these potential sources of variation in fact do not result in significant differences in the extent to which the principle is employed. Even in those cases in which significant differences were identified between subpopulations, the extent of the difference between groups was small.

Taken together, the results presented in sections 2 and 3 indicate a surprisingly small role for gender, age, education, exposure to moral philosophy, ethnicity and nationality in shaping subjects' use of the principle of the double effect in scenarios 3 and 4, and of the principle of the double effect among other possible moral principles in scenarios 1 and 2.

These results present some problems for the conscious reasoning perspective. If moral judgments are the product of conscious reasoning from a set of moral principles, one might expect that those educated in moral philosophy would be more likely to invoke the principle of the double effect (a central focus of much modern moral philosophy) or a related principle than those lacking such an education. Likewise, from the rationalist perspective, one might expect differences in beliefs and attitudes that co-vary with other demographic characteristics to produce concordant differences in the use of moral principles. At least for the principles tested, and at least within the range of variation of our subject population, the conscious reasoning perspective cannot account for the pattern of results presented here.

Analyses of justifications presented in section 4 have even more direct implications for the competing predictions of the two target perspectives. In our sample, a large majority of subjects failed to sufficiently justify their moral judgments, including a majority of those subjects who had been exposed to readings in moral philosophy. The dissociation between judgment and justification was most striking in the second pair of dilemmas, scenarios 3 and 4, which isolated the distinction between intended and foreseen effects. Thus, for subjects who perceived a difference between these cases, only the principle of double effect can account for their pattern of judgments. On the view that conscious reasoning accounts for the nature of our moral judgments, subjects who judged these cases to be different should have appealed to the principle of double effect (or the gist of the principle) in justifying their judgments. Although not all subjects judged these cases differently, those that did generally failed to appeal to the principle of

double effect or its central distinction between intended means and foreseen side effect.

We acknowledge that while subjects failed to provide sufficient justifications for their judgments under the present testing conditions, other methods may reveal that subjects are indeed able to retrieve such principles. For example, if subjects had more time to answer or were given a set of alternative principles, they might have derived the correct answer. We believe that there are strong arguments against both of these possibilities. First, we did not pressure subjects into giving justifications as fast as possible; subjects had as much time as they needed to reply. Although it is possible that subjects would express the correct principle if we supplied several alternatives, this result would not necessarily count as evidence in favor of the rationalist perspective. Rather, as discussed by Haidt (2001) and others, pointing to a principled reason may well count as evidence of post-hoc rationalization or an attempt to find a principle that is consistent with one's judgment. We therefore conclude that under the conditions employed, intuition drives subjects' judgments, and with little or no conscious access to the principles that distinguish between particular moral dilemmas.

Our results challenge the strong thesis that when we deliver moral verdicts we do so by appealing to consciously accessible principles of moral right and wrong. As in other work in moral psychology, our study does not deny the role which conscious reasoning plays in moral judgments. What we suggest is that when people make certain kinds of moral judgments, they may do so without consciously applying explicitly understood principles. Although the most dominant alternative to the conscious reasoning view is the emotion account, whereby intuitions are said to be emotionally-mediated (Damasio, 1994; de Waal, 1996; Greene and Haidt, 2002; Haidt, 2001; Haidt, 2003; Hume, 1739/1978; Nichols, 2004; Prinz, 2004), results from the present study, together with other empirical and theoretical work (Chomsky, 1986; Cushman *et al.*, in press; Dwyer, 1999; Dwyer, 2004; Hauser, 2006; Mikhail *et al.*, 1998; Mikhail, 2000; Rawls, 1971), provide a necessary amendment to this account. If, as Greene and colleagues (2001; 2004) suggest, our moral judgments are guided by a personal-impersonal distinction, mediated by emotional processes, then all subjects should have judged scenarios 3 and 4 to be the same. However, even though both scenarios involve an impersonal act—indeed the same act of throwing a switch—a subset of subjects judged these scenarios differently, perceiving an act in which one intends harm as impermissible and an act in which one merely foresees the harm that one causes as permissible. The conclusion we draw from these results is that the personal-impersonal dimension, while of potential importance in explaining some moral judgments, does not tell the entire story. The missing piece of the theoretical puzzle is the part of our psychology that evaluates the causes and consequences of action, especially its intentional structure.

In conclusion, our results challenge the view that moral judgments are solely the product of conscious reasoning on the basis of explicitly understood moral principles. Though we sometimes deliver moral judgments based on consciously

accessed principles, often we fail to account for our judgments. When we fail, it appears that operative, but not expressed principles, drive our moral judgments. Future work aims to extend the range of moral dilemmas and cultures sampled in order to further refine our understanding of the nature of moral judgments.

Marc Hauser
Departments of Psychology, Organismic and Evolutionary Biology, Biological Anthropology,
Harvard University,

Fiery Cushman, Liane Young and R. Kang-Xing Jin
Department of Psychology
Harvard University

John Mikhail
Georgetown University Law Center, Washington DC

Appendix

Text of scenarios:

Scenario 1: Denise is a passenger on a train whose driver has just shouted that the train's brakes have failed, and who then fainted of the shock. On the track ahead are five people; the banks are so steep that they will not be able to get off the track in time. The track has a side track leading off to the right, and Denise can turn the train onto it. Unfortunately there is one person on the right hand track. Denise can turn the train, killing the one; or she can refrain from turning the train, letting the five die.

Is it morally permissible for Denise to switch the train to the side track?

Scenario 2: Frank is on a footbridge over the train tracks. He knows trains and can see that the one approaching the bridge is out of control. On the track under the bridge there are five people; the banks are so steep that they will not be able to get off the track in time. Frank knows that the only way to stop an out-of-control train is to drop a very heavy weight into its path. But the only available, sufficiently heavy weight is a large man wearing a backpack, also watching the train from the footbridge. Frank can shove the man with the backpack onto the track in the path of the train, killing him; or he can refrain from doing this, letting the five die.

Is it morally permissible for Frank to shove the man?

Scenario 3: Ned is taking his daily walks near the train tracks when he notices that the train that is approaching is out of control. Ned sees what has happened: the

driver of the train saw five men walking across the tracks and slammed on the brakes, but the brakes failed and they will not be able to get off the tracks in time. Fortunately, Ned is standing next to a switch, which he can throw, that will temporarily turn the train onto a side track. There is a heavy object on the side track. If the train hits the object, the object will slow the train down, thereby giving the men time to escape. Unfortunately, the heavy object is a man, standing on the side track with his back turned. Ned can throw the switch, preventing the train from killing the men, but killing the man. Or he can refrain from doing this, letting the five die.

Is it morally permissible for Ned to throw the switch?

Scenario 4: Oscar is taking his daily walk near the train tracks when he notices that the train that is approaching is out of control. Oscar sees what has happened: the driver of the train saw five men walking across the tracks and slammed on the brakes, but the brakes failed and the driver fainted. The train is now rushing toward the five men. It is moving so fast that they will not be able to get off the track in time. Fortunately, Oscar is standing next to a switch, which he can throw, that will temporarily turn the train onto a side track. There is a heavy object on the side track. If the train hits the object, the object will slow the train down, thereby giving the men time to escape. Unfortunately, there is a man standing on the side track in front of the heavy object, with his back turned. Oscar can throw the switch, preventing the train from killing the men, but killing the man. Or he can refrain from doing this, letting the five die.

Is it morally permissible for Oscar to throw the switch?

Control 1: Dr. Irwin is in charge of a patient who is dying. All this patient needs in order for his good health to be restored is a small dose of drug X. Fortunately, Dr. Irwin happens to have an unlimited amount of this drug X. Dr. Irwin can save his patient if he administers the necessary dosage at once.

Is it morally permissible for Dr. Irwin to give his patient the drug?

Control 2: David is driving a train when the brakes fail. Ahead of him, five people are working on the track with their backs turned. They cannot see or hear the train approaching. Fortunately, David can switch the train to a side track, which is completely clear, if he acts immediately. If David switches his train to the side track, he will save the five people working on the track. If he does not switch his train, the train will run over the five people.

Is it morally permissible for David to switch his train to the side track?

References

Baron, J. and Siepmann, M. 2000: Using web questionnaires for judgment and decision making research. In: M.H. Birnbaum (ed.), *Psychological Experiments on the Internet*. New York: Academic Press, pp. 235–265.

Berthoz, S., Armony, J. L., Blair, R. J. R. and Dolan, R. J. 2002: An fMRI study of intentional and unintentional (embarassing) violations of social norms. *Brain*, 125, 1696–1708.

Chomsky, N. 1986: *Knowledge of Language: Its Nature, Origin, and Use*. New York, NY: Praeger.

Cushman, F., Young, L. and Hauser, M. D. in press: The role of conscious reasoning and intuition in moral judgments: testing three principles of harm. *Psychological Science*.

Damasio, A. 1994: *Descartes' Error*. Boston, MA: Norton.

de Waal, F. B. M. 1996: *Good Natured*. Cambridge: Harvard University Press.

Dworkin, R. M. 1998: Darwin's new bulldog. *Harvard Law Review*, 111, 1718–1738.

Dwyer, S. 1999: Moral competence. In: K. Murasugi and R. Stainton (eds), *Philosophy and Linguistics*. Boulder: Westview Press, pp. 169–190.

Dwyer, S. 2004: How good is the linguistic analogy. Available at http://www.umbc/du/philosophy/dwyer

Fischer, J. M. and Ravizza, M. 1992: *Ethics: Problems and Principles*. New York: Holt, Rinehart and Winston.

Foot, P. 1967: The problem of abortion and the doctrine of double effect. *Oxford Review*, 5, 5–15.

Greene, J. D. and Haidt, J. 2002: How (and where) does moral judgment work? *Trends in Cognitive Science*, 6, 517–523.

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.

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.

Greenwald, A. G., Nosek, B. and Banaji, M. R. 2003: Understanding and using the Implicit Association Test: 1. An improved scoring algorithm. *Journal of Personality and Social Psychology*, 85, 197–216.

Haidt, J. 2001: The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108, 814–834.

Haidt, J. 2003. The moral emotions. In: R. J. Davidson, K. R. Scherer and H. H. Goldsmith (eds), *Handbook of Affective Sciences*. Oxford: Oxford University Press, pp. 852–870.

Hauser, M. D. 2006: *Moral Minds: How Nature Designed Our Universal Sense of Right and Wrong*. New York: Ecco/Harper Collins.

Hauser, M. D., Young, L. and Cushman, F. in press: Reviving Rawls' linguistic analogy. In: W. Sinnott-Armstrong (ed.), *The Biology and Psychology of Morality*. New York: Oxford University Press.

Hume, D. 1739/1978: *A Treatise of Human Nature*. Oxford: Oxford University Press.

Kamm, F. M. 1992: *Creation and Abortion: A Study in Moral and Legal Philosophy*. New York: Oxford University Press.

Kamm, F. M. 1998a: Moral intuitions, cognitive psychology, and the harming-versus-not-aiding distinction. *Ethics*, 108, 463–488.

Kamm, F. M. 1998b: *Morality, Mortality: Death and Whom to Save From It*. New York: Oxford University Press.

Kamm, F. M. 2000: Nonconsequentialism. In: H. LaFollette (ed.), *Ethical Theory*. Malden, CT: Blackwell Publishing, pp. 205–226.

Kaplow, L. and Shavell, S. 2002: *Fairness versus Welfare*. Cambridge: Harvard University Press.

Kohlberg, L. 1981: *Essays on Moral Development, Volume 1: The Philosophy of Moral Development*. New York: Harper Row.

Korsgaard, C. M. 1996: *Sources of Normativity*. Cambridge: Cambridge University Press.

Kraut, R., Olson, J., Banaji, M. R., Bruckman, A., Cohen, J. and Cooper, M. 2004: Psychological research online. *American Psychologist*, February/March, 105–117.

Mikhail, J. 2000: Rawls' linguistic analogy: A study of the 'generative grammar' model of moral theory described by John Rawls in 'A theory of justice'. PhD thesis, Cornell University, p. 375.

Mikhail, J. in press: *Rawls' Linguistic Analogy*. New York: Cambridge University Press.

Mikhail, J., Sorrentino, C. and Spelke, E. S. 1998. Toward a universal moral grammar. In: *Proceedings of the Cognitive Science Society*, pp. 1250.

Moll, J., Oliveira-Souza, R., Bramai, I. E. and Grafman, J. 2002: Functional networks in emotional moral and nonmoral social judgments. *Neuroimage*, 16, 696–7073.

Nichols, S. 2004: *Sentimental Rules*. New York: Oxford University Press.

Nosek, B., Banaji, M. R. and Greenwald, A. G. 2002: Harvesting intergroup attitudes and stereotypes from a demonstration website. Group Dynamics, 6, 101–115.

O'Neill, P. and Petrinovich, L. 1998: A preliminary cross cultural study of moral intuitions. *Evolution and Human Behavior*, 19, 349–367.

Petrinovich, L., O'Neill, P. and Jorgensen, M. J. 1993: An empirical study of moral intuitions: towards an evolutionary ethics. *Ethology and Sociobiology*, 64, 467–478.

Piaget, J. 1932/1965: *The Moral Judgment of the Child*. New York: Free Press.

Prinz, J. J. 2004: *Gut Reactions*. New York: Oxford University Press.

Rawls, J. 1971: *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Royzman, E. B. and Baron, J. 2002: The preference for indirect harm. *Social Justice Research*, 15, 165–184.

Schmidt, W. C. 1997: World-Wide-Web survey research: Benefits, potential problems, and solutions. *Behavior Research Methods and Computers*, 29, 274–279.

Thomson, J. J. 1970: Individuating actions. *Journal of Philosophy*, 68, 774–781.

Wheatley, T. and Haidt, J. 2005: Hypnotic disgust makes moral judgments more severe. *Psychological Science*, 16, 780–785.