*Research Article*

# The Role of Conscious Reasoning and Intuition in Moral Judgment

## Testing Three Principles of Harm

**Fiery Cushman,[1] Liane Young,[1] and Marc Hauser[1,2,3]**

[1]Department of Psychology, [2]Department of Organismic and Evolutionary Biology, and [3]Department of Biological Anthropology, Harvard University

ABSTRACT—*Is moral judgment accomplished by intuition or conscious reasoning? An answer demands a detailed account of the moral principles in question. We investigated three principles that guide moral judgments: (a) Harm caused by action is worse than harm caused by omission, (b) harm intended as the means to a goal is worse than harm foreseen as the side effect of a goal, and (c) harm involving physical contact with the victim is worse than harm involving no physical contact. Asking whether these principles are invoked to explain moral judgments, we found that subjects generally appealed to the first and third principles in their justifications, but not to the second. This finding has significance for methods and theories of moral psychology: The moral principles used in judgment must be directly compared with those articulated in justification, and doing so shows that some moral principles are available to conscious reasoning whereas others are not.*

A topic of recent concern in moral psychology is the extent to which conscious reasoning, as opposed to intuition, plays a role in determining moral judgment (Greene & Haidt, 2002; Haidt, 2001; Pizarro & Bloom, 2003). In terms common to social psychology, the question is whether moral judgment is a controlled or an automatic process (Bargh, 1999).

The conscious-reasoning perspective has been the central focus for students of moral development in the tradition of Kohlberg (1969). Kohlberg characterized children's moral development by focusing on the content of their justifications

rather than the source of their moral judgments. An implicit assumption of this perspective is that people generate moral judgments by consciously reasoning over the principles they articulate in moral justification.

One challenge to the conscious-reasoning perspective comes from work by Haidt in which subjects failed to articulate sufficient justifications for their moral judgments (Haidt & Hersh, 2001). Haidt (2001) proposed that moral judgments arise as intuitions generated by automatic cognitive processes, and that the primary role of conscious reasoning is not to generate moral judgments, but to provide a post hoc basis for moral justification. Although there is increasing support for the role of intuition in moral judgment, some researchers argue that both conscious reasoning and intuition play a role in judgment, as well as justification (Greene, in press; Pizarro & Bloom, 2003; Pizarro, Uhlmann, & Bloom, 2003).

A critical ingredient missing from the current debate is an experimental method that clearly links data on moral judgment with data on moral justification. Without establishing that an individual uses a specific moral principle, it makes little sense to ask whether the content of that principle is directly available to conscious reasoning. Therefore, in the present study, we first identified three moral principles used by subjects in the judgment of moral dilemmas, and then explored the extent to which subjects generated justifications based on these principles. Our approach, adopted in part from moral philosophy, was to compare judgments across tightly controlled pairs of scenarios. We parametrically varied each pair of scenarios to target only one factor at a time, holding all others constant. We use the term *principle* to denote a single factor that when varied in the context of a moral dilemma consistently produces divergent moral judgments. By using the term "principle" to refer to such factors, however, we make no assumptions about the nature of the psychological mechanisms that underlie sensitivity to them.

Address correspondence to Fiery Cushman, 984 William James Hall, 33 Kirkland St., Cambridge, MA 02138, e-mail: cushman@wjh.harvard.edu.

We investigated three principles:

- *The action principle*: Harm caused by action is morally worse than equivalent harm caused by omission.
- *The intention principle*: Harm intended as the means to a goal is morally worse than equivalent harm foreseen as the side effect of a goal.
- *The contact principle*: Using physical contact to cause harm to a victim is morally worse than causing equivalent harm to a victim without using physical contact.

The action principle has been well researched in psychology, where it is often called omission bias (Baron & Ritov, 2004; Spranca, Minsk, & Baron, 1991). The relevance of the action principle is also recognized in the philosophical literature (Quinn, 1989; Rachels, 1975). The intention principle, often identified as the doctrine of the double effect, has received intense scrutiny by philosophers (Foot, 1967; Nagel, 1986), but markedly less by psychologists (but see Mikhail, 2002; Royzman & Baron, 2002). The contact principle has been comparatively understudied in both psychology and philosophy; although it bears a superficial similarity to Greene's distinction between personal and impersonal moral dilemmas (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001), physical contact is neither a necessary nor a sufficient condition for a personal moral dilemma.

Having established that subjects make use of a principle, one can then ask whether this principle is available to conscious reflection during justification. On the one hand, we hypothesized that a hallmark of conscious reasoning is that the principles used in judgments are articulated in justifications. On the other hand, we hypothesized that intuitive responses are accompanied by insufficient justifications, uncertainty about how to justify, denial of a relevant principle, or confabulation of alternative explanations for judgments. Although it is possible that moral principles consistently cited during justification were nonetheless engaged without conscious reasoning during judgment, one may conclude that these principles are at least available for conscious processes of moral reasoning. By contrast, those principles that consistently cannot be cited during justification appear to be genuinely inaccessible to conscious processes of reasoning.

## METHOD

Subjects voluntarily logged on to the Moral Sense Test Web site, moral.wjh.harvard.edu. Previous work with a different set of dilemmas revealed no substantive differences in responses obtained from subjects who answered questions on the Web and those who completed more traditional pen-and-paper tests (Hauser, Cushman, & Young, in press). Subjects were 37 years old on average, and the sample had a small male bias (58%). We instructed subjects to participate only if they were fluent in English; 88% listed English as their primary language. Most subjects indicated they were from the United States, Canada, or the United Kingdom; 25% had been exposed to formal education in moral philosophy.

After completing a demographic questionnaire, subjects received 32 moral scenarios separated into two blocks of 16. Each block included 15 test scenarios and 1 control scenario. Order of presentation was counterbalanced between subjects, varying both within and between blocks. For each scenario, subjects rated the protagonist's harmful action or omission on a scale from 1 to 7, with 1 labeled "Forbidden," 4 labeled "Permissible," and 7 labeled "Obligatory." In a third block, subjects were asked to justify their pattern of responses for up to five pairs of scenarios. We asked subjects to justify only responses conforming to the three principles being tested (e.g., when an action was judged worse than an omission). Subjects were presented with the text of the two scenarios side by side, reminded which they judged more permissible, and asked to justify their pattern of responses.

All subjects had the opportunity to exit the testing session after any number of blocks. We analyzed data only from subjects who successfully completed all three blocks. Subjects were omitted from all analyses if they failed either of the two control scenarios (by judging permissible the killing or allowed death of five people despite a costless alternative), or if they completed any of the 32 scenarios in fewer than 4 s, deemed the minimum comprehension and response time on the basis of pilot research. Additionally, subjects were removed from the analyses of justifications if they misunderstood the task, provided a nonsensical response, or provided a judgment that made it clear they had misunderstood a scenario. These subjects were not removed from our judgment analyses because not every subject justified each judgment, thereby precluding the uniform application of this procedure. Of 591 justifications, 65 were removed from the analyses of justifications.

The test scenarios comprised 18 controlled pairs. What follows are brief descriptions of 4 scenarios; the actual text of all 32 (test and control) scenarios is available on the Web at moral.wjh.harvard.edu/methods.html.

"Evan" (action, intended harm, no contact): Is it permissible for Evan to pull a lever that drops a man off a footbridge and in front of a moving boxcar in order to cause the man to fall and be hit by the boxcar, thereby slowing it and saving five people ahead on the tracks?

"Jeff" (omission, intended harm, no contact): Is it permissible for Jeff not to pull a lever that would prevent a man from dropping off a footbridge and in front of a moving boxcar in order to allow the man to fall and be hit by the boxcar, thereby slowing it and saving five people ahead on the tracks?

"Frank" (action, intended harm, contact): Is it permissible for Frank to push a man off a footbridge and in front of a moving boxcar in order to cause the man to fall and be hit by the boxcar, thereby slowing it and saving five people ahead on the tracks?

"Dennis" (action, foreseen harm as side effect, no contact): Is it permissible for Dennis to pull a lever that redirects a moving
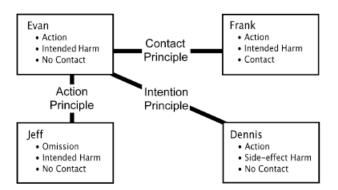
**Fig. 1.** Principle-based contrasts for four scenarios arranged into three pairs.

boxcar onto a side track in order to save five people ahead on the main track if, as a side effect, pulling the lever drops a man off a footbridge and in front of the boxcar on the side track, where he will be hit?

Some scenarios belonged to more than one pair; for instance, "Evan" was contrasted with "Jeff" to yield an action-principle comparison, with "Frank" to yield a contact-principle comparison, and with "Dennis" to yield an intention-principle comparison (Fig. 1). Six pairs varied across the action principle, six varied across the intention principle, and six varied across the contact principle.

The methods used were in accordance with the regulations of the institutional review board at Harvard University.

## RESULTS

### Judgments

Paired-sample $t$ tests were performed on each of the 18 controlled pairs of scenarios to determine whether subjects rated one scenario in the pair significantly more permissible than the other in the direction predicted by the relevant principle. Statistical significance was achieved in 17 out of 18 pairs at .05, two-tailed ($N = 332$); in the remaining pair, mean permissibility ratings trended in the appropriate direction but fell short of significance, $p = .144$ (Table 1). Across scenarios with different content, subjects judged action as worse than omission, intended harm as worse than foreseen harm, and harm via contact as worse than harm without contact.

### Justifications

A total of 526 justifications were coded for five nonexclusive attributes. The attributes were as follows:

- *Sufficiency*: The subject mentioned a factual difference between the two cases and either claimed or implied that it was the basis of his or her judgments. It was not necessary for the subject to identify the target principle, so long as the principle generated by the subject could adequately account for his or her pattern of responses on the scenario pair in question.
- *Failure*: The subject suggested an alternative principle, but this alternative could not account for his or her actual pattern of judgments.

**TABLE 1**

*Differences in Permissibility for Pairs of Moral Scenarios*

| Scenario pair | Mean difference | SD | $t(331)$ | Effect size ($d$) | $p$ (two-tailed) | $p_{\text{rep}}$ |
|---|---|---|---|---|---|---|
| Action-principle pairs | | | | | | |
| Boxcar | 0.70 | 2.03 | 6.32 | 0.34 | <.001 | >.999 |
| Pond | 1.69 | 2.00 | 15.34 | 0.84 | <.001 | >.999 |
| Ship | 0.83 | 2.01 | 7.56 | 0.41 | <.001 | >.999 |
| Car | 0.90 | 1.77 | 9.26 | 0.50 | <.001 | >.999 |
| Boat | 0.98 | 1.98 | 8.97 | 0.49 | <.001 | >.999 |
| Switch | 0.26 | 1.87 | 2.56 | 0.13 | <.01 | >.94 |
| Intention-principle pairs | | | | | | |
| Speedboat | 0.29 | 1.15 | 4.65 | 0.25 | <.001 | >.999 |
| Burning | 1.12 | 1.58 | 12.90 | 0.70 | <.001 | >.999 |
| Boxcar | 0.50 | 1.68 | 5.38 | 0.29 | <.001 | >.999 |
| Switch | 0.28 | 1.77 | 2.92 | 0.15 | <.005 | >.96 |
| Chemical | 0.24 | 1.51 | 2.91 | 0.15 | <.005 | >.96 |
| Shark | 0.30 | 1.77 | 3.14 | 0.16 | <.003 | >.978 |
| Contact-principle pairs | | | | | | |
| Speedboat | 0.89 | 1.44 | 11.36 | 0.62 | <.001 | >.999 |
| Intended burning | 0.24 | 1.40 | 3.18 | 0.17 | <.003 | >.978 |
| Boxcar | 1.07 | 1.72 | 11.28 | 0.62 | <.001 | >.999 |
| Foreseen burning | 0.37 | 1.22 | 5.50 | 0.30 | <.001 | >.999 |
| Aquarium | 0.17 | 1.35 | 2.31 | 0.12 | <.022 | >.947 |
| Rubble | 0.10 | 1.27 | 1.47 | 0.07 | <.145 | >.77 |

*Note.* All $t$ tests were within subjects. Probability of replication ($p_{\text{rep}}$) was calculated according to Killeen (2005).

**TABLE 2**

*Proportion of Justifications Exhibiting Each Attribute and Differences in Proportions Across Principles, With Trials as the Unit of Analysis*

| | Proportion | | | Chi-square analysis | | | Interobserver reliability | |
|---|---|---|---|---|---|---|---|---|
| Attribute | Action-principle pairs | Intention-principle pairs | Contact-principle pairs | $\chi^2(2, N = 526)$ | $p$ | Cramer's $\phi$ | Overall percentage agreement | Kappa |
| Sufficiency | .81 | .32 | .60 | 91.149 | <.001 | 0.42 | 89 | .77 |
| Failure | .06 | .16 | .10 | 10.869 | .004 | 0.14 | 89 | .32 |
| Uncertainty | .05 | .22 | .04 | 39.058 | <.001 | 0.27 | 97 | .78 |
| Denial | .02 | .17 | .13 | 28.781 | <.001 | 0.23 | 95 | .67 |
| Alternative explanation | .10 | .29 | .32 | 34.344 | <.001 | 0.26 | 90 | .69 |

Note. For this analysis, the proportion of justifications exhibiting each attribute was calculated as the proportion of the total number of justifications provided for each principle. For each attribute, the chi-square analysis tested whether the proportions differed significantly by principle.

- *Uncertainty*: The subject explicitly referenced his or her own uncertainty about how to justify the scenario, or directly stated that he or she could not justify the responses.
- *Denial*: The subject stated that, at the point of justification, he or she did not consider there to be any moral difference between the two scenarios.
- *Alternative explanation*: The subject appealed to an alternative explanation of his or her responses, either (a) invoking facts that were not present in the scenarios or (b) claiming that he or she made a mistake in selecting the appropriate response, for instance, by clicking on the wrong button.

Our complete coding criteria, including examples from the data, are available on the Web at moral.wjh.harvard.edu/methods.html. Justifications were coded by one of the authors (F.C.) and by a colleague familiar with the research, and ties were broken by another author (L.Y.). Table 2 provides the overall percentage agreement between coders for individual attributes, along with Cohen's kappa, a statistic of interobserver reliability for which values between .60 and .70 are considered fair, from .70 to .75 are considered good, and above .75 are considered excellent (Fleiss, 1981). For one of the attributes, failure, the low Cohen's kappa of .32 warrants caution in the interpretation of results, although the overall agreement, 89%, was quite high.

Differences between principles in the proportion of justifications meeting criteria for each attribute were tested by chi-square analysis (Table 2). For all five attributes, the proportion of justifications meeting criteria differed significantly between principles. Critically, subjects readily provided sufficient justifications for the action principle, rarely did so for the intention principle, and showed an intermediate level for the contact principle. Although justifications were coded as sufficient even when subjects provided logically adequate justifications for their behavior other than the target principle in question, we observed that 95% of sufficient justifications depended on the target principle. Subjects were more likely to provide failed

justifications or to express uncertainty for intention-principle cases than for action-principle and contact-principle cases, and were more likely to deny the moral relevance of the distinction or to appeal to alternative explanations for intention-principle and contact-principle cases than for action-principle cases.

It is possible that the attribute "alternative explanation" represents a baseline error rate for the task of judging pairs of moral dilemmas; that is, subjects' alternative explanations might actually be true accounts of their behavior. If this interpretation is accurate, then subjects should have been equally likely to make an error in judgment on any trial, regardless of the principle that the scenario was used to target during justification. Therefore, the proportion of the total number of pairs of judgments made that resulted in an alternative explanation during justification should have been equal across the three principles. (In the analyses described in the previous paragraph, the proportion of justifications that relied on alternative explanations was calculated over the total number of justifications of paired scenarios, whereas in the analyses reported here, it was calculated over the total number of judgments of paired scenarios.) Analyses revealed statistically significant differences among these proportions, $\chi^2(2, N = 1,502) = 10.163$, $p = .006$. Subjects were almost twice as likely to arrive at an alternative explanation after judging a pair of intention-principle (9%) or contact-principle (9%) scenarios than after judging a pair of action-principle scenarios (5%). Alternative explanations therefore seem to be driven by the principle demanded in justification, and do not exclusively represent a baseline error rate for the task of judging moral dilemmas.

For a more stringent test of the differences in the attributes elicited by the three principles, we treated controlled pairs of scenarios, rather than individual justifications, as the unit of analysis. After calculating for each scenario pair the proportion of justifications meeting criteria for a given attribute, we used an analysis of variance (ANOVA) to determine whether the mean proportion for each attribute differed by principle. Though this

**TABLE 3**

*Proportion of Justifications Exhibiting Each Attribute and Differences in Proportions Across Principles, With Scenario Pairs as the Unit of Analysis*

| Attribute | Mean proportion | | | $F(2, 15)$ | $p$ | $\eta_p^2$ |
|---|---|---|---|---|---|---|
| | Action-principle pairs | Intention-principle pairs | Contact-principle pairs | | | |
| Sufficiency | .79 | .30 | .59 | 22.44 | <.001 | .749 |
| Failure | .07 | .18 | .09 | 5.75 | .014 | .434 |
| Uncertainty | .05 | .23 | .04 | 12.28 | .001 | .621 |
| Denial | .02 | .20 | .14 | 8.14 | .004 | .520 |
| Alternative explanation | .11 | .29 | .34 | 8.60 | .003 | .534 |

*Note.* For this analysis, the proportion of justifications exhibiting each attribute was calculated separately for each scenario pair. The mean proportions of the six scenario pairs that tested each principle are listed under ''Mean proportion.'' The analysis of variance tested whether the mean proportions differed by principle.

analysis reduced the sample size from 526 justifications to 18 scenario pairs, each ANOVA confirmed the previous set of analyses (Table 3). To rule out the hypothesis that differences between principles were driven by the size of the judgment difference between items in a scenario pair, we repeated the ANOVAs with mean judgment difference as a covariate. All five ANOVAs still yielded significant differences by principle.

### DISCUSSION

This study presents evidence for three principles that guide judgments of moral dilemmas, each focused on trade-offs of life and death, but varying in content. Subjects rated harmful actions as morally worse than harmful omissions (the action principle), harm intended as the means to an end as morally worse than harm foreseen as the side effect of an end (the intention principle), and harm involving physical contact as morally worse than harm without contact (the contact principle).

The content of subjects' justifications differed greatly by principle. In the case of the action principle, a large majority of subjects were able to provide sufficient justifications for their judgments, whereas relatively few provided failed justifications, denied any moral difference between the scenarios, or expressly doubted their ability to justify their responses. These data are consistent with the conscious-reasoning model of moral judgment. Although it is possible that subjects constructed the action principle post hoc upon having to justify their responses, a large majority of subjects at least possessed the requisite explicit knowledge required by the conscious-reasoning account.

Quite the opposite is true for the intention principle. Less than a third of subjects were able to provide sufficient justifications for their pattern of judgments regarding the intention-principle pairs. In 22% of the justifications, subjects specifically indicated uncertainty about how to justify their responses; in 17% of the justifications, subjects denied that there was any morally relevant difference between the cases; and 16% of the justifi-

cations in fact failed to account for the subject's pattern of judgments. These data are more consistent with an intuitionist model of moral judgment. Subjects reliably generated a pattern of moral judgments consistent with the intention principle, but were generally incapable of articulating it, to the point of expressing confusion when confronted with their own judgments or even denying their judgments altogether. The intention principle clearly plays a role in moral judgment, but most likely in the form of unconscious—or at least inexpressible—knowledge.

It is notable that subjects were almost three times as likely to invoke alternative explanations for their responses to intention-principle cases as to invoke alternative explanations for their responses to action-principle cases, either by claiming to have made an error (e.g., by pushing the wrong button) or by invoking unwarranted assumptions about the scenarios (e.g., assuming that a harm would actually not occur, even when the scenario explicitly stated that it would). There is no reason to believe that the particular scenarios used in intention-principle pairs were any more likely to invite either errors or assumptions than those used in action- or contact-principle pairs—the scenarios were identically constructed, and indeed, in several cases individual scenarios were used in contrasts of both types. Thus, what determines the use of alternative explanations appears not to be the scenario presented during judgment, but rather the principle targeted in justification. We conclude that in some cases, subjects' appeals to alternative explanations were confabulations generated at the moment of justification and prompted by the inability to justify the pattern of judgments. Parallel evidence comes from Wheatley and Haidt's (2005) recent study of priming under hypnosis: Subjects' confabulation accompanied their inability to provide a principled justification of moral judgment.

Subjects' justifications of their responses to contact-principle cases occupied an intermediate position between justifications for action-principle and intention-principle cases. Subjects were typically able to articulate the relevant principle used, but were relatively unwilling to endorse it as morally valid. Sixty

percent of justifications in contact-principle cases were sufficient, nearly twice the proportion for intention-principle cases. Additionally, only 10% of justifications were classified as failed, and 4% indicated uncertainty. These statistics are all comparable to those for action-principle cases. Yet in 13% of contact-principle justifications, subjects denied that physical contact made a moral difference, a proportion much closer to that for the intention-principle cases (17%) than for the action-principle cases (2%). Subjects were also about three times more likely to appeal to alternative explanations in contact-principle cases than in action-principle cases—a pattern of results nearly identical to that for the intention principle. Thus, although subjects were able to articulate the principle behind their reasoning in contact-principle cases, they often rejected it as morally invalid, or confabulated alternative explanations for their behavior. For instance, one subject wrote, "I guess I was fooled by the line-pulling seeming more passive than the man-pushing, but that view is hard to justify now."

Although a conscious-reasoning interpretation of subjects' justifications for contact-principle cases cannot be rejected definitively, the data favor the intuitionist view. The observation that many subjects used, but were unwilling to endorse, the contact principle is not readily explained by the conscious-reasoning model. Why would a subject reason consciously from an explicit principle about physical contact during judgment, but then disavow the same principle during justification? A more plausible explanation is that the contact principle guides moral judgments according to the intuitionist model during judgment, and that a process of post hoc reasoning at justification allows subjects to deduce the principle behind their judgments. Once deduced, the principle is regarded as morally irrelevant. Just such a process of post hoc reasoning has been proposed by Haidt (2001). This interpretation of contact-principle cases raises the question of why a similar process was not observed for the intention-principle cases. One possible explanation is that the factors over which the intention principle operates are less salient than those over which the contact principle operates; this speculation requires additional research.

In summary, our results show that although some moral principles are available for conscious reasoning in a large majority of subjects, others are not available and appear to operate in intuitive processes. The extent to which conscious reasoning or intuition plays a more dominant role depends on the particular moral principles triggered: The intention principle seems best characterized by the intuitionist model, the contact principle by intuitive judgment followed by rational reflection, and the action principle either by conscious reasoning or, at a minimum, by post hoc conscious reasoning.

The present findings bear in important ways on issues central to the study of moral cognition. We briefly raise three such issues. First, our results underscore the methodological importance of linking subjects' expressed moral principles to the operative principles underlying their patterns of judgment

(Hauser, in press; Mikhail, 2000). Controlled pairs of scenarios have long been used in philosophy and psychology to probe moral judgments (Kamm, 2001; Petrinovich, O'Neill, & Jorgensen, 1993; Spranca et al., 1991; Thompson, 1985), but have not been extended to moral justification. Without using tightly controlled scenario pairs to target narrowly defined principles, it is not possible to determine whether a subject's justification can account for his or her judgments. It has been demonstrated that under uniform testing conditions, this method can yield data that distinguish principles available to conscious reasoning from those that are not. Having demonstrated divergent results under a common paradigm, researchers are better equipped to detect genuine differences in the application of moral principles.

This methodological point has been underappreciated by advocates of both conscious-reasoning and intuitionist models. On the one hand, within the tradition established by Kohlberg (1969), the moral principles cited in justification were simply assumed to be identical to the moral principles used in judgment. Our data showing a high rate of failed justifications for intention-principle cases suggest that this assumption is not always valid. On the other hand, Haidt (2001) demonstrated moral dumbfounding (the inability to justify a moral judgment) by asking subjects why, for instance, it is wrong for a brother and sister to have intercourse, but did not demonstrate that subjects' patterns of judgments necessitate a principle more complex than "it is wrong for a brother and a sister to have intercourse." In the paradigm we used in the present study, dumbfounding might arise if subjects were asked *why* a harmful action is worse than a harmful omission. But this is an altogether different question from whether subjects have explicit knowledge of the principle to which their pattern of judgments conforms, which is simply that a harmful act *is* less permissible than a harmful omission. From either a conscious-reasoning or an intuitionist perspective, a proper analysis of subjects' justifications cannot be conducted in the absence of a precise accounting of their judgments. However, because of the possibility of post hoc reasoning in cases in which subjects provide sufficient justifications, the present method is better suited to rejecting the role of conscious reasoning than supporting it. A task for future studies is to design methodologies that provide strong evidence in favor of consciously reasoned moral judgments.

Second, our results provide evidence for moral principles that are formulated over detailed representations of the causal and intentional aspects of harm. To date, many of the moral principles proposed by psychologists have far broader content than those we have presented. Kohlberg's (1969) moral stages were formulated over concepts like authority, cooperation, and autonomy. Greene has proposed an emotion-based appraisal system with a general prohibition of actions fitting the semantic structure "me hurt you" (Greene, Nystrom, Engell, Darley, & Cohen, 2004). Haidt has formulated moral principles over the concepts of suffering, reciprocity, hierarchy, purity, and group identity (Haidt, in press). Ultimately, these useful generalizations will have to be married to

a more detailed analysis of basic concepts such as action, intention, and contact (Hauser, Cushman, & Young, in press; Hauser, Young, & Cushman, in press; Knobe, 2003; Mikhail, Sorrentino, & Spelke, 2002; Pizarro et al., 2003).

Finally, our results support the view that moral judgment can be accomplished by multiple systems: Some moral principles are available to conscious reflection—permitting but not guaranteeing a role for conscious reasoning—whereas others are better characterized by an intuitionist model. On the basis of neuroimaging data, Greene et al. (2004) have proposed a two-system model for moral judgment in which one system is characterized by the engagement of affective systems and the other is characterized by the engagement of cognitive systems. The current findings suggest that, regardless of where the division between affect and cognition is placed, a multisystem model of moral judgment is warranted. Such a multisystem model stands in contrast to Kohlberg's (1969) perspective, in which all moral reasoning is assumed to be the product of conscious reasoning. It also stands in contrast to the recent intuitionist proposal by Mikhail (Mikhail et al., 2002) to incorporate the sorts of moral principles discussed here into a single evaluative mechanism that accomplishes moral judgment in a manner loosely analogous to the way in which syntactic structure is analyzed in language.

In conclusion, this article has outlined a novel methodological approach to the study of moral psychology, highlighting the interaction of intuition and conscious reasoning and emphasizing the distinction between the principles that people use and the principles that people articulate.

## REFERENCES

Bargh, J.A. (1999). The unbearable automaticity of being. *American Psychologist*, *54*, 462–479.

Baron, J., & Ritov, I. (2004). Omission bias, individual differences, and normality. *Organizational Behavior and Human Decision Processes*, *94*, 74–85.

Fleiss, J.L. (1981). *Statistical methods for rates and proportions* (2nd ed.). New York: Wiley.

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

Greene, J.D. (in press). The secret joke of Kant's soul. In W. Sinnott-Armstrong (Ed.), *Moral psychology: Vol. 3. The neuroscience of morality*. Cambridge, MA: MIT Press.

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

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.

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

Haidt, J. (in press). Social intuitionists answer six questions about moral psychology. In W. Sinnott-Armstrong (Ed.), *Moral psychology: Vol. 2. The cognitive science of morality*. Cambridge, MA: MIT Press.

Haidt, J., & Hersh, M.A. (2001). Sexual morality: The cultures and emotions of conservatives and liberals. *Journal of Applied Social Psychology*, *31*, 191–221.

Hauser, M.D. (in press). *Moral minds: How nature designed a universal sense of right and wrong*. New York: Harper Collins.

Hauser, M.D., Cushman, F.A., & Young, L. (in press). A dissociation between moral judgment and justification. *Mind and Language*.

Hauser, M.D., Young, L., & Cushman, F.A. (in press). Reviving Rawls' linguistic analogy. In W. Sinnott-Armstrong (Ed.), *Moral psychology: Vol. 2. The cognitive science of morality*. Cambridge, MA: MIT Press.

Kamm, F.M. (2001). *Morality, mortality: Rights, duties, and status*. New York: Oxford University Press.

Killeen, P.R. (2005). An alternative to null-hypothesis significance tests. *Psychological Science*, *16*, 345–353.

Knobe, J. (2003). Intentional action in folk psychology: An experimental investigation. *Philosophical Psychology*, *16*, 309–324.

Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D.A. Goslin (Ed.), *Handbook of socialization theory and research* (pp. 151–235). New York: Academic Press.

Mikhail, J. (2002). *Aspects of the theory of moral cognition: Investigating intuitive knowledge of the prohibition of intentional battery and the principle of double effect* (Georgetown Public Law Research Paper 762385). Washington, DC: Georgetown University Law Center.

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, Ithaca, NY.

Mikhail, J.M., Sorrentino, C., & Spelke, E. (2002). *Aspects of the theory of moral cognition: Investigating intuitive knowledge of the prohibition of intentional battery, the rescue principle, the first principle of practical reason, and the principle of double effect*. Unpublished manuscript, Stanford University, Stanford, CA.

Nagel, T. (1986). *The view from nowhere*. Oxford, England: Oxford University Press.

Petrinovich, L., O'Neill, P., & Jorgensen, M.J. (1993). An empirical study of moral intuitions: Towards an evolutionary ethics. *Journal of Personality and Social Psychology*, *64*, 467–478.

Pizarro, D.A., & Bloom, P. (2003). The intelligence of the moral intuitions: Comment on Haidt (2001). *Psychological Review*, *110*, 193–198.

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.

Royzman, E., & Baron, J. (2002). The preference for indirect harm. *Social Justice Research*, *15*, 165–184.

Spranca, M., Minsk, E., & Baron, J. (1991). Omission and commission in judgment and choice. *Journal of Experimental Social Psychology*, *27*, 76–105.

Thompson, J.J. (1985). The trolley problem. *The Yale Law Journal*, *94*, 1395–1415.

Wheatley, T., & Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. *Psychological Science*, *16*, 780–784.