Cognitive Science 34 (2010) 311-321

Copyright © 2009 Cognitive Science Society, Inc. All rights reserved.

ISSN: 0364-0213 print / 1551-6709 online DOI: 10.1111/j.1551-6709.2009.01088.x

Moral Principles or Consumer Preferences? Alternative Framings of the Trolley Problem

Tage S. Rai, Keith J. Holyoak

Department of Psychology, University of California, Los Angeles

Received 15 June 2009; received in revised form 29 August 2009; accepted 2 September 2009

Abstract

We created paired moral dilemmas with minimal contrasts in wording, a research strategy that has been advocated as a way to empirically establish principles operative in a domain-specific moral psychology. However, the candidate "principles" we tested were not derived from work in moral philosophy, but rather from work in the areas of consumer choice and risk perception. Participants were paradoxically less likely to choose an action that sacrifices one life to save others when they were asked to provide more reasons for doing so (Experiment 1), and their willingness to sacrifice lives depended not only on how many lives would be saved, but on the number of lives at risk (Experiment 2). The latter effect was also found in a within-subjects design (Experiment 3). These findings suggest caution in the use of artificial dilemmas as a key testbed for revealing principled bases for moral judgment.

Keywords: Moral judgment; Biases and heuristics; Decision making; Modularity

1. Introduction

...I believe that some philosophical analysis, based on exotic moral dilemmas, is inadvertently and even comically replicating the early work of Kahneman and Tversky: uncovering situations in which intuitions, normally quite sensible, turn out to misfire.

—Sunstein, 2005, p. 541

In attempting to understand the mental processes underlying our sense of morality, philosophers and psychologists have often implicitly or explicitly assumed that lay people's moral judgments are guided by reasoning or intuitions based on universal, normative

Correspondence should be sent to Tage S. Rai, Department of Psychology, UCLA, 405 Hilgard Ave., Los Angeles, CA 90095-1563. E-mail: raitriumphant@gmail.com

principles that are realized in a domain-specific set of psychological mechanisms (see Stich, 2006 for a review). The most ambitious of these approaches has been a proposed universal moral grammar (UMG) (Dwyer, 2006; Hauser, 2006; Mikhail, 2007). According to UMG, normative moral principles (e.g., prohibition of intentional battery) are realized psychologically within a "universal grammar" for a modular "moral faculty," analogous to the theory of universal grammar for human languages proposed by Chomsky (1964, 1995) (but see Dupoux & Jacob, 2007; Evans & Levinson, 2009). According to this approach, human moral psychology is based on innate principles (e.g., "do not inflict intentional battery on others"), modulated by cultural learning of more specific parameters (e.g., learning that the relevant "others" are members of a particular in-group).

This view has garnered support from experiments in which moral judgments are elicited for paired dilemmas designed to isolate the effects of potential moral principles. For example, Hauser, Cushman, Young, Jin, and Mikhail (2007) presented participants with one of two versions of the "trolley problem" (Foot, 1967; Thomson, 1985) in order to isolate the "doctrine of double effect," which states that harm is acceptable if it occurs as an unintended but foreseen consequence of action in service of a greater good, but not as a means toward an end (Foot, 1967). In both versions, a trolley headed toward five people can be redirected onto a side track where it will be stopped from looping back to the main track by a heavy object. Participants were less likely to judge the action as permissible if it was indicated that the object was a man than if it was indicated that a man was standing in front of the object, suggesting that the man's death is more acceptable when it is an unintended but foreseen side effect of action rather than a means to stop the trolley (but see Waldmann & Dieterich, 2007, for a failure to replicate using slightly different cover stories.) Given that the actions have the same utility (one person dies, five people live), and as there are no prior beliefs or preexisting societal rules governing how to respond to trolley problems, Hauser et al. (2007) argued that this pattern of responses results from innate computational principles akin to a doctrine of double effect, and that a domain-specific moral module guides moral judgment (for reviews see Hauser, 2006; Hauser, Young, & Cushman, 2008a; Mikhail, 2007).

In the present paper, we consider the extent to which the above method of using paired dilemmas actually provides a clear window onto the fundamental computations underlying moral judgments, by considering the role of domain-general biases of the sort investigated in the field of judgment and decision making. Framing effects and other biases have been studied extensively in a number of domains, including the evaluation of gambles, consumer preferences, and risk perception, where seemingly small changes in wording or order often lead to large differences in judgment (Gilovich & Griffin, 2002). Similar effects may guide many moral judgments (Machery, in press; Unger, 1996; Sunstein, 2005). Spranca, Minsk, and Baron (1991) found that sins of omission, in which a moral wrong is allowed to happen, are judged less severely than sins of commission, in which a moral wrong is committed. Also, when assessing multiple moral dilemmas, participants consider actions to be less morally permissible as they evaluate dilemmas later in the order (Haidt & Baron, 1996; Petrinovich & O'Neill, 1996). In the context of the trolley problem, Petrinovich and O'Neill (1996) found that judgments are affected by whether options are described in terms of the number of people who would be saved or in terms of those who would die under the

different options, just as in Kahneman and Tversky's (1984) "Asian disease" problem in which participants had to choose among risky options for preventing a disease outbreak (for a review, see Sinott-Armstrong, 2008).

Although the examples above suggest that domain-general judgment and decision-making biases may be prevalent in moral judgment, such biases have not typically been incorporated into theoretical conceptualizations of moral psychology (see Machery, in press; Sunstein, 2005, for similar arguments). Proponents of domain-specific theories of moral judgment argue that domain-general biases simply represent errors in moral judgment, whereas a theory of moral psychology should be primarily concerned with identifying evidence of moral competence (Hauser, 2005; Mikhail, 2005). The aim of the present paper is to determine if the method of presenting paired moral dilemmas to identify domain-specific moral principles, such as a doctrine of double effect, can in fact empirically distinguish a domain-specific principle from a domain-general bias; and consequently, whether this method can provide compelling evidence supporting claims of domain specificity in moral judgment.

As an example of the conceptual difficulties, consider the versions of the trolley problem tested by Hauser et al. (2007), in which the only difference between the paired cases was whether the man was the object or was standing in front of the object. Rather than attributing differences in judgments to a complex principle of moral psychology (Mikhail, 2007), the effect might reflect variation in attentional focus, perceptions of omission versus commission (allowing the trolley to hit the man when it is sent toward the object, versus sending it toward the man), salience of the objectification of the man (Waldmann & Dieterich, 2007), or perhaps some implicit hope that the man in front of the object might somehow escape.

Hauser, Young, and Cushman (2008b, p. 173) provide a particularly clear statement of the empirical method that has been advocated by proponents of a UMG: "to assess whether particular principles serve as the basis for our moral judgments" we must "develop a battery of paired dilemmas that isolate psychologically meaningful and morally relevant, principled distinctions" and "determine whether these targeted principles guide subjects' moral judgments." In the experiments reported here we adopt this approach; but rather than relying on prescriptive philosophy as the sole source of hypotheses about "psychologically meaningful and principled distinctions," we turned to research in the fields of consumer preferences and risk perception on domain-general biases. In particular, we investigated whether support for sacrificing potential victims in the trolley problem will paradoxically decrease if participants are given the option to express more reasons to take the action, and whether judgments in regard to the number of lives that must be saved to support sacrifice is dependent on the number of lives at risk.

The present studies are based on well-established findings in the field of judgment and decision making, and there is no a priori reason why similar effects would not be found with moral judgments about trolley problems. However, observing such phenomena would undermine the rationale for using paired moral dilemmas to establish psychological moral principles. The effects we sought to show cannot be attributed to prior beliefs, emotions, or preexisting societal rules any more than putative moral principles, such as that of double effect. Thus, if domain-general biases can be demonstrated with moral judgments using the

same method employed to identify moral principles within a UMG, then this method is insufficient to distinguish performance errors due to domain-general biases from competence based on domain-specific moral principles. Proponents of employing dilemmas to identify potential moral principles that would support the existence of a UMG or similar domain-specific moral psychology would either have to argue that the phenomena we investigate in the present paper actually do represent moral principles, or else exercise substantial caution in claiming that moral dilemmas are especially well-suited to reveal such principles.

2. Experiment 1

To make an initial determination as to whether domain-general biases studied in other domains, such as consumer preferences, might also apply to moral judgments, Experiment 1 tested a basic finding in the judgment and decision-making literature regarding processing fluency. People expect that difficulty in generating reasons for a choice is a reliable (inverse) index of the quality of the choice; thus, paradoxically, asking participants to generate more reasons for a choice can make them less likely to endorse it, in spite of having generated more reasons (Schwarz et al., 1991; Schwarz, 1998). Experiment 1 examined whether choice of preferred moral action in the trolley problem is similarly influenced by the number of reasons participants are asked to provide in favor of redirecting the train. This manipulation also provides a test of a dual process theory of moral judgment that has been applied to trolley problems (Greene & Haidt, 2002, Greene, Sommerville, Nystrom, Darley, & Cohen, 2001). According to the dual process view, refusal to redirect the train is based on emotional engagement; hence, any task that requires a greater degree of "cold" reasoning, such as stating reasons to redirect the train, should lead to greater endorsement of the utilitarian option of redirecting the train.

2.1. Method

Participants in all experiments were recruited via the Internet, and completion of the questionnaires was voluntary. Questionnaires were posted on the Craigslist site under their volunteers section in Los Angeles, New York, Chicago, and Washington, D.C. After giving consent to participate, each participant was randomly assigned to a condition. The IP addresses of participants' computers were recorded to ensure that participants did not participate in the study multiple times.

Experiment 1 employed a between-subjects design in which participants were presented with a trolley problem adapted from Waldmann and Dieterich (2007). Participants were asked to provide reasons that would justify redirecting the train onto a sidetrack:

Please try to list two (seven) different reasons for the employee to take the proposed action. Try your best to come up with reasons that are genuinely different. If you find that coming up with two (seven) reasons for the employee to take the action is difficult, then only write down as many reasons as you can.

Participants were either told to provide two reasons (n = 60) or seven reasons (n = 64). Note that the instructions specified that participants could provide whatever number of responses they wished. After writing down reasons to take the action, participants were asked to rate "to what extent do you think the proposed action should be taken?" using a four-point scale that ranged from "definitely disagree" (0) to "definitely agree" (3).

2.2. Results and discussion

Participants given the option to provide up to seven reasons for redirecting the train provided more reasons (M = 3.20) than did those who were given the option to provide up to two responses (M = 1.67), t(122) = 5.46, p < .001. Yet paradoxically, level of agreement with taking action to redirect the train was significantly lower for participants in the "seven reasons" condition (M = 1.83) than for those in the "two reasons" condition (M = 2.17), t(122) = 2.11, p < .05. Thus, a bias previously identified in the field of consumer preferences guides moral intuitions in the trolley problem. The size of this effect, though modest, is comparable to those found in similar studies outside of the moral domain (Schwarz et al., 1991). The fact that greater engagement in reasoning (i.e., giving more reasons for redirecting the train) leads to fewer rather than more utilitarian judgments contradicts the prediction of a dual process theory of moral judgment (Greene & Haidt, 2002; Greene et al., 2001).

3. Experiment 2

Studies of mental accounting (Thaler, 1999) have found that willingness to incur a cost in order to save on a product depends on the relative savings one will experience rather than the absolute savings. For example, people are more likely to report being willing to drive 20 min in order to save \$5 on a \$15 calculator than to save \$5 on a \$125 jacket (Kahneman & Tversky, 1984). Fetherstonhaugh, Slovic, Johnson, and Friedrich (1997) found similar effects in a study of risk perception, in which participants preferred to send aid that could save a fixed number of lives to a refugee camp when it had a smaller number of refugees. Experiment 2 examined whether people's choice of preferred moral action in the trolley problem is similarly influenced by the "relative savings" of moral action, defined as the number of lives saved relative to the number of lives at risk, rather than the absolute savings, defined simply as the number of lives saved.

Note that a strictly deontological principle in which killing is wrong and cannot be justified by the ends it achieves would imply that no number of lives saved can justify sacrificing others. In contrast, a simple utilitarian calculus would imply that as long as one more life is saved compared to those that are lost, the sacrifice is justified. Finally, a looser consequentialist framework, in which some price is assigned to the "wrong" committed by sacrificing a life unjustly, would imply that some number of lives (one or more) must be saved per life that is sacrificed. Critically, none of these established moral frameworks would predict that the number of lives that people believe must be saved in order to justify a sacrifice should vary as a function of the number of lives at risk.

3.1. Method

Experiment 2 employed a between-subjects design in which participants were presented with the following version of the trolley problem:

A high-speed train is about to hit a large railway car with 10 (40) people. An employee of the train company is sitting in an office at the station. He could press a button on a control panel which would move a small railway car with two people into the path of the train. This will slow the train down and give 8 of the 10 (40) people on the large car time to escape. However, the action will kill the two people on the small car. To what extent is sacrificing the two people on the small railway car the right thing to do?

Participants were either told that sacrificing the railway car with two people would allow enough time for 8 out of 10 people to escape from the larger car (n = 83), or 8 out of 40 (n = 81). Participants were then asked to report their level of agreement with the proposed action on a 20-point scale, where 0 was identified as "definitely disagree," 10 was identified as "agree somewhat," and 20 was identified as "definitely agree."

3.2. Results and discussion

Support for taking action to redirect the trolley was significantly higher for participants who were told that 8 out of 10 people would escape (M = 8.58) than for those who were told that 8 out of 40 people would escape (M = 6.07), t(162) = 2.63, p < .01. Thus, moral intuitions, just like consumer preferences and risk perception, are guided by relative rather than absolute savings.

4. Experiment 3

One step in assessing whether an influence on moral judgments should count as a "principle," rather than a mere bias, might be to consider the coherence of people's judgments. The basic idea of a coherence test, which underlies Rawls' (1971) notion of "reflective equilibrium" within a system of justice, is to "examine whether judgments made in isolation are still retained when explicitly compared" (Sunstein, Kahneman, Schkade & Ritov, 2002, p. 1159). Experiment 3 tested whether the influence of relative number of lives saved on moral judgments, identified in Experiment 2, persists in a more stringent within-subjects design.

Experiment 3 was modeled on a study by Fetherstonhaugh et al. (1997), who found that when assessing how many lives must be saved by a disease intervention in order to be chosen for a 10 million dollar grant, 65% of participants reported that a greater number of individuals needed to be saved for diseases in which a greater number of individuals were at risk. Experiment 3 examined whether or not judgments of the number of lives that would need to be saved by sacrificing two individuals in the trolley problem is similarly influenced

by the number of lives at risk. Importantly, the "price" that needs to be justified in the present experiment is not money, but the loss of two lives. Thus, participants were not asked to equate what might be considered incommensurable goods (money and lives), a possible source of performance error (Sunstein et al., 2002).

4.1. Method

Experiment 3 employed a 2×2 (wording x order) within-subjects design in which participants were sequentially presented with the two versions of the trolley problem used in Experiment 2, which differed solely in whether the number of lives at risk was stated to be 10 or 40. The order of the two versions was counterbalanced, so that in the first problem participants were either told that there were 10 people on the other train (n = 52) or that there were 40 people on it (n = 55). Participants were then asked how many people would need to be saved in order to make sacrificing the two people on the small train "the right thing to do." Participants were then told that another train was headed toward another railway car, and assessed the alternative version.

4.2. Results and discussion

Fig. 1 depicts the mean estimate of the number of people who would need to be saved to justify sacrificing two people as a function of the total number at risk (10 or 40) and the order in which the two versions were encountered. When estimating the number of lives that would need to be saved in order to justify sacrificing two people, 42% (n = 45) of participants provided non-zero estimates that increased as the number of lives at risk increased, as

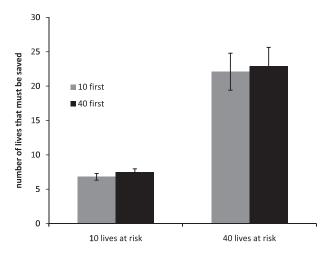


Fig. 1. Mean estimate of number of lives that need to be saved to justify taking action as a function of the number of lives at risk (10 or 40) and the order of the two versions of the problem (10 first, or 40 first) (Experiment 3). Error bars indicate 1 standard error of the mean.

predicted by sensitivity to relative savings. Just 2% (n=2) of participants gave estimates that decreased as the number of lives at risk increased, a highly reliable difference [Yates-corrected $\chi^2(1, n=47)=37.53, p<.001$]. Another 30% (n=32) of participants indicated that sacrificing two people to save others could not be justified, regardless of the number of lives at risk, as expected by a strict deontological view; and 21% (n=22) required that the same number be saved, regardless of the number of lives at risk, as expected by a consequentialist view. Responses from six participants could not be categorized and were dropped from analyses.

A 2×2 within-subjects analysis of variance revealed that the number of people who needed to be saved in order to justify sacrificing two people in the smaller car increased with the number of people at risk, F(1, 67) = 91.93, p < .001. Order had no influence either as a main effect or as an interaction, F < 1. Participants who indicated that sacrificing two people to save others could never be justified were excluded from this analysis because their responses could not be coded numerically. Between-subject analysis of first responses indicated that participants who were told that 40 people were at risk demanded that more people be saved (M = 22.54) than those who were told that 10 people were at risk (M = 7.11), t(67) = 6.31, p < .001, mirroring the results from Experiment 2.

Participants' judgments of how many people would need to be saved in order to justify a sacrifice were thus strongly influenced by the total number of lives at risk. Even in a within-subjects design in which participants made judgments in immediate succession, the threshold number of lives that needed to be saved to justify taking action roughly tripled when the number of potential victims was 40 as opposed to 10. These findings suggest that participants actually believed that more people needed to be saved in the condition in which more lives were at risk. Thus, a phenomenon of sensitivity to "relative saving," previously observed in studies of risk perception, appears to pass a coherence test.

5. General discussion

The present findings demonstrate that eliciting judgments of paired moral dilemmas cannot empirically distinguish whether influences on moral judgment are based on domain-specific moral principles, or on biases of the sort found in other decision contexts that require active weighing of costs and benefits. In our experiments, the research strategy that has been used to identify principles guiding judgments in a UMG or similar domain-specific moral psychology (Hauser, 2006; Mikhail, 2007) was extended to test novel candidates to be considered potential "moral principles." We found that people are paradoxically less likely to choose an action that sacrifices one life to save others if they are asked to provide more reasons for doing so (Experiment 1). In addition, people's willingness to sacrifice one life depended not just on how many lives would be saved, but on how many lives were at risk. This influence of "relative number of lives saved" was found both in a between-subjects design (Experiment 2) and in a within-subjects design (Experiment 3). Experiments 2 and 3 involved minimal manipulations within the dilemmas.

To the extent that domain-general biases, triggered by small changes in wording, are ubiquitous in judgments about moral dilemmas (as our experiments suggest they may well be), then the task of isolating moral principles will prove to be more difficult. The different versions compared in the present experiments were no more likely to differ in prior beliefs or emotional salience than those used in similar prior studies (Hauser et al., 2007). The present evidence of domain-general biases in moral judgment of dilemmas does not rule out the possibility that there are also domain-specific principles, but it does highlight the theoretical problem facing advocates of using moral dilemmas to isolate principles hypothesized to exist in a domain-specific moral psychology. If experimenters demonstrate the impact of some carefully controlled variation of wording on moral judgments for a dilemma, by what criteria are we to decide whether this variation counts as a moral principle in a domain-specific moral module, rather than a domain-general heuristic or bias (Machery, in press; Mallon, 2008)?

In practice, it seems that theorists have used their own a priori assumptions to decide what kinds of variations could possibly be considered "moral" (Mallon, 2008; Sunstein, 2005). The consequence is that experimenters with differing entering theories of moral psychology can reach starkly different conclusions from the same observations. For example, Petrinovich, O'Neill, and Jorgensen (1993) hypothesized that evolutionary forces would have programmed into human psychology a moral preference for close kin over strangers. Using variants of the trolley problem, these investigators determined that people are less likely to redirect the trolley if they imagine the victim on the side track to be their brother. But other moral psychologists implicitly assume that such a predisposition could not possibly be part of the computational structure of our moral psychology (Hauser et al., 2008a, p. 127). Different theorists might thus interpret the phenomenon of kin preference as an innate moral principle (for good evolutionary reasons, the wrongness of a death is directly proportional to the victim's relatedness to the decision maker), or perhaps a parameter setting (given an innate principle of avoiding intentional battery, experience has taught us to whom the principle is to be selectively applied), or simply a performance error (people have an emotional bias towards saving close kin) that needs to be eliminated in a "clean" experimental design, by ensuring that all the lives at risk remain anonymous to the decision maker (Mallon, 2008).

The studies reported in the present paper demonstrate that judgments elicited in the trolley problem are vulnerable to the same biases as judgments in other decision contexts that require active weighing of costs and benefits, such as consumer choice and risk perception. It is certainly possible that much of moral judgment emerges out of a domain-specific moral psychology, and the present study does not discount this enticing theoretical possibility. However, the present experiments do demonstrate that eliciting judgments to moral dilemmas in which variation cannot be attributed to prior knowledge or social considerations is not sufficient for evaluating the claim that moral judgments emerge out of a domain-specific moral psychology. Variation in moral judgments could be due to either domain-specific moral principles or to nonmoral domain-general biases, and the method of eliciting judgments of moral dilemmas cannot distinguish between the two without a priori assumptions regarding what will be considered "moral." Progress in understanding the computational

bases for moral judgments will thus require more careful evaluation of the linkage between empirical phenomena and theoretical interpretations.

Acknowledgments

Preparation of this paper was supported by the UCLA NSF IGERT Interdisciplinary Relationship Science Program (TR) and AFOSR grant FA9550-08-1-0489 (KH). We thank Edouard Machery and Stephen Stich for helpful comments on earlier drafts.

References

Chomsky, N. (1964). Current issues in linguistic theory. New York: Pantheon.

Chomsky, N. (1995). The minimalist program. Cambridge, MA: MIT Press.

Dupoux, E., & Jacob, P. (2007). Universal moral grammar: A critical appraisal. Trends in Cognitive Sciences, 11, 373–378.

Dwyer, S. (2006). How good is the linguistic analogy? In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind (Vol. 2): Culture and cognition* (pp. 237–256). Oxford, England: Oxford University Press.

Evans, N., & Levinson, S. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Brain and Behavioral Sciences*, 32, 429–448.

Fetherstonhaugh, D., Slovic, P., Johnson, S., & Friedrich, J. (1997). Insensitivity to the value of human life: A study of psychophysical numbing. *Journal of Risk and Uncertainty*, 14, 283–300.

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

Gilovich, T., & Griffin, D. (2002). Introduction—heuristics and biases: Then and now. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 1–18). New York: Cambridge University Press.

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

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

Haidt, J., & Baron, J. (1996). Social roles and the moral judgment of acts and omissions. European Journal of Social Psychology, 26, 201–218.

Hauser, M. (2006). Moral minds: How nature designed our universal sense of right and wrong. New York: Ecco Press

Hauser, M. (2005). Sunstein's heuristics provide insufficient descriptive and explanatory adequacy. *Brain and Behavioral Sciences*, 28, 553–554.

Hauser, M., Cushman, F., Young, L., Jin, R., & Mikhail, J. (2007). A dissociation between moral judgment and justification. *Mind and Language*, 22, 1–21.

Hauser, M., Young, L., & Cushman, F. (2008a). Reviving Rawls' linguistic analogy. In W. Sinnott-Armstrong (Ed.), *Moral psychology, Vol. 1: Evolution of morality* (pp. 107–144). Cambridge, MA: MIT Press.

Hauser, M., Young, L., & Cushman, F. (2008b). On misreading the linguistic analogy: Response to Jesse Prinz and Ron Mallon. In W. Sinnott-Armstrong (Ed.), *Moral psychology, Vol. 1: Evolution of morality* (pp. 171– 179). Cambridge, MA: MIT Press.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. American Psychologist, 39, 341-350.

Machery, E. (in press). The bleak implications of moral psychology. Neuroethics.

Mallon, R. (2008). Reviving Rawls' linguistic analogy inside and out. In W. Sinnott-Armstrong (Ed.), *Moral psychology, Vol. 1: Evolution of morality* (pp. 145–155). Cambridge, MA: MIT Press.

- Mikhail, J. (2005). Moral heuristics or moral competence: Reflections on Sunstein. *Brain and Behavioral Sciences*, 28, 557–558.
- Mikhail, J. (2007). Universal moral grammar: Theory, evidence, and the future. *Trends in Cognitive Sciences*, 11, 143–152.
- Petrinovich, L., O'Neill, P., & Jorgensen, M. J. (1993). An empirical study of moral intuitions: Towards an evolutionary ethics. *Ethology and Sociobiology*, 64, 467–478.
- Petrinovich, L., & O'Neill, P. (1996). Influence of wording and framing effects on moral intuitions. *Ethology and Sociobiology*, 17, 145–171.
- Rawls, J. (1971). A theory of justice. Cambridge, MA: Harvard University Press.
- Schwarz, N. (1998). Accessible content and accessibility experiences: The interplay of declarative and experiential information in judgment. *Personality and Social Psychology Review*, 2, 87–99.
- Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenaur-Schatka, H., & Simons, A. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology*, 61, 195–202.
- Sinott-Armstrong, W. (2008). Framing moral intuitions. In W. Sinott-Armstrong (Ed.), *Moral psychology, Vol.* 2: The cognitive science of morality (pp. 47–76). Cambridge, MA: MIT Press.
- Spranca, M., Minsk, E., & Baron, J. (1991). Omission and commission in judgment and choice. *Journal of Experimental Social Psychology*, 27, 76–105.
- Stich, S. (2006). Is morality an elegant machine or a kludge? Journal of Cognition and Culture, 6, 181-189.
- Sunstein, C. (2005). Moral heuristics. Behavioral and Brain Sciences, 28, 531–573.
- Sunstein, C., Kahneman, D., Schkade, D., & Ritov, I. (2002). Predictably incoherent judgments. Stanford Law Review, 54, 1153–1215.
- Thaler, R. H. (1999). Mental accounting matters. Journal of Behavioral Decision Making, 12, 183-206.
- Thomson, J. (1985). The trolley problem. The Yale Law Journal, 94, 1395–1415.
- Unger, P. (1996). Living high and letting die: Our illusion of innocence. New York: Oxford University Press.
- Waldmann, M. R., & Dieterich, J. (2007). Throwing a bomb on a person versus throwing a person on a bomb: Intervention myopia in moral intuitions. *Psychological Science*, *18*, 247–253.