



Brief article

The mismeasure of morals: Antisocial personality traits predict utilitarian responses to moral dilemmas

Daniel M. Bartels^{a,*}, David A. Pizarro^b

^a Columbia University, Uris Hall 502, 3022 Broadway, New York, NY 10027, United States

^b Department of Psychology, Cornell University, 224 Uris Hall, Ithaca, NY 14853, United States

ARTICLE INFO

Article history:

Received 12 January 2011

Revised 27 May 2011

Accepted 29 May 2011

Available online 16 July 2011

Keywords:

Morality

Judgment

Decision making

Psychopathy

Values

Ethics

Intuition

Utilitarianism

Machiavellianism

Emotions

Reasoning

Moral rules

No Meaning

Moral dilemmas

ABSTRACT

Researchers have recently argued that utilitarianism is the appropriate framework by which to evaluate moral judgment, and that individuals who endorse non-utilitarian solutions to moral dilemmas (involving active vs. passive harm) are committing an error. We report a study in which participants responded to a battery of personality assessments and a set of dilemmas that pit utilitarian and non-utilitarian options against each other. Participants who indicated greater endorsement of utilitarian solutions had higher scores on measures of Psychopathy, machiavellianism, and life meaninglessness. These results question the widely-used methods by which lay moral judgments are evaluated, as these approaches lead to the counterintuitive conclusion that those individuals who are least prone to moral errors also possess a set of psychological characteristics that many would consider prototypically immoral.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Moral judgments are unique. Like many of our attitudes, (e.g., toward a favorite sports team) they are often central to our identity and are accompanied by strong emotions. Yet unlike these other attitudes, attitudes in the moral domain come with a strong sense that others should agree—a sense of *normativity* (Skitka, Bauman, & Sargis, 2005). In recent years, researchers have made a great deal of progress toward understanding these unique judgments by proposing frameworks that describe and explain various features of moral judgment (e.g., Baron & Spranca, 1997; Greene, Sommerville, Nystrom, Darley, &

Cohen, 2001; Haidt & Joseph, 2004; Iliev et al., 2009; Mikhail, 2007; Nichols & Mallon, 2006; Tetlock, 2003).

Recently, some theorists have adopted a strategy of comparing people's moral judgments to a normative ethical standard—that of utilitarianism—to evaluate the *quality* of moral judgment (e.g., Baron & Ritov, 2009; Greene et al., 2009; Sunstein, 2005). In this paper, we question the close identification of utilitarian responses with optimal moral judgment by demonstrating that the endorsement of utilitarian solutions to a set of commonly-used moral dilemmas correlates with a set of psychological traits that can be characterized as emotionally callous and manipulative—traits that most would perceive as not only psychologically unhealthy, but also morally undesirable. These results, we believe, give rise to an important methodological concern: namely, that the methods widely used as a yardstick for

* Corresponding author. Tel.: +1 212 854 1557.

E-mail address: dmb2199@columbia.edu (D.M. Bartels).

determining optimal morality (i.e., assessing responses to moral dilemmas that pit the death of one vs. the death of many) may be tracking what many would regard as its opposite—a muted aversion to causing a person's death.

1.1. Utilitarianism, deontology, and the error-and-bias approach in moral psychology

The question of how to determine which moral claims and decisions are correct has traditionally been the domain of normative ethics in philosophy. One of the biggest debates in the field has centered on the question of which principle(s) should guide our moral evaluations, with many philosophers defending one of two approaches to determine the morally right course of action. One the one hand, deontological approaches describe a set of rules or principles that serve as constraints on what kinds of actions are morally permissible (e.g., the constraint that it is morally forbidden to take an innocent life). On the other hand, utilitarianism argues that what is morally required is best determined by one simple rule—whether or not an action brings about the greatest total well-being.

For psychologists studying morality, this philosophical debate has provided a conceptual backdrop for the descriptive study of moral judgment. Using the moral dilemmas first introduced by philosophers engaged in this debate, psychologists have explored when lay moral intuitions appear to adhere to the prescriptions of deontological or utilitarian approaches. Increasingly, many psychologists have adopted these normative frameworks as a standard by which to evaluate the quality of the moral intuitions themselves, arguing that the study of bias in the moral domain can help improve moral decision making. Some view deontological judgments as cognitive errors, akin to the errors that result from using heuristics in other judgmental domains. Baron and Ritov (2009) make this assumption salient, stating that “decisions made on the basis of deontological principles usually lead to results that are not as good as the best that could be achieved.” (p. 136). Others have arrived at similar conclusions—that the use of non-utilitarian “heuristics” can lead to pervasive and dangerous errors in moral judgment, and even to judgments that border on absurdity (e.g., Sunstein, 2005).

The characterization of non-utilitarian moral decisions as errors of judgment is especially pronounced in research on the role of emotion in moral judgment. Such investigations have increasingly relied on the method of recording participants' responses to “sacrificial” dilemmas, where the question of whether to kill a person to prevent others from dying is posed. For example, consider Thomson's (1985) footbridge case:

In the path of a runaway train car are five railway workmen who will surely be killed unless you, a bystander, do something. You are standing on a pedestrian walkway that arches over the tracks next to a large stranger. Your body would be too light to stop the train, but if you push the stranger onto the tracks, killing him, his large body will stop the train. In this situation, would you push him?

Adopting a dual-process approach to moral judgment, Greene and colleagues have collected evidence that when evaluating moral dilemmas that are especially emotional

(like the footbridge case), individuals are likely to favor the utilitarian option when the “deliberative” mental system is recruited (Greene et al., 2001). Consistent with the moral heuristics approach described above, Greene et al. (2009) equate the tendency to make non-utilitarian moral judgments while under the influence of the “intuitive” system to the tendency to stereotype racial minorities under similar conditions (p. 1145), arguing that non-utilitarian judgments are not only less-than-ideal, but potentially damaging.

One implication of adopting a utilitarian framework as a normative standard in the psychological study of morality is the inevitable conclusion that the vast majority of people are often morally wrong. For instance, when presented with Thomson's footbridge dilemma, as many as 90% of people reject the utilitarian response (Mikhail, 2007). Many philosophers have also rejected utilitarianism, arguing that it is inadequate in important, morally meaningful ways, and that it presents an especially impoverished view of humans as “locations of utilities [and nothing more]. . .” and that “persons do not count as individuals. . . any more than individual petrol tanks do in the analysis of the national consumption of petroleum” (Sen & Williams, 1982, p. 4). For those who endorse utilitarianism, the ubiquitous discomfort toward its conclusions points to the pessimistic possibility that human moral judgment is even more prone to error than many other forms of judgment, and that attempting to improve the quality of moral judgment will be a steep uphill battle.

Before drawing those conclusions, it might prove useful to investigate individuals who are more likely to endorse utilitarian solutions and perhaps use them as a psychological prototype of the “optimal” moral judge. What do those 10% of people who are comfortable with the utilitarian solution to the footbridge dilemma look like? Might these utilitarians have other psychological characteristics in common? Recently, consistent with the view that rational individuals are more likely to endorse utilitarianism (e.g., Greene et al., 2001), a variety of researchers have shown that individuals with higher working memory capacity and those who are more deliberative thinkers are, indeed, more likely to approve of utilitarian solutions (Bartels, 2008; Feltz & Cokely, 2008; Moore, Clark, & Kane, 2008). In fact, one well-defined group of utilitarians likely shares these characteristics as well—the subset of philosophers and behavioral scientists who have concluded that utilitarianism is the proper normative ethical theory.

Yet in addition to the link between deliberative thinkers and utilitarian judgments, there is another possible psychological route to utilitarian preferences—the ability to inhibit emotional reactions to harm (or the inability to experience such emotions in the first place). For instance, patients with damage to the ventromedial pre-frontal cortex, who have emotional deficits similar to those observed in psychopaths (leading some researchers to refer to this type of brain damage as “acquired sociopathy”; Saver & Damasio, 1991), are more likely to endorse utilitarian solutions to sacrificial dilemmas (Koenigs et al., 2007). Yet it is always questionable to generalize from clinical populations, as their deficits might lead to utilitarian judgments through qualitatively different psychological mechanisms than those at work in non-clinical populations.

In this paper, we provide evidence that utilitarian preferences are associated with a variety of psychological traits associated with those of the clinical populations mentioned above, and in doing so provide a critique to the method of closely identifying utilitarian responses with optimal moral judgment. Specifically, we predicted that variability on personality traits that reflect the devaluation of life, emotional callousness, and manipulateness would predict utilitarian preference. If such associations are found in non-clinical populations, it would suggest that there are at least two distinct routes that may give rise to an endorsement of utilitarian options in these moral dilemmas—one characterized by a tendency to favor rational deliberation when making moral decisions and one characterized by a muted aversion to causing a person's death, as observed in certain brain-damaged patients.

2. Current study

To test our predictions about one set of factors underlying the preference to sacrifice an innocent person for the sake of a greater good, we gave people a battery of 14 footbridge-like moral dilemmas, as well as a set of three individual difference measures that reflected the tendencies described above—psychopathic personality, Machiavellianism, and perceived life meaningfulness. Psychopathy refers to a personality style characterized by low empathy, callous affect, and thrill-seeking. In addition to this measure of psychopathy, we included a measure of Machiavellianism, which refers to the degree to which people are cynical, emotionally detached from others, and manipulative. Both psychopathy and Machiavellianism share the aspects of emotional coldness, aggression, and willingness to engage in or rationalize deceit, but while correlated they have been found to be distinct in previous studies (Paulhus & Williams, 2002). Finally, we included the No Meaning Scale, which measures people's melancholic existential concerns—how meaningful they perceive life to be. If life is perceived to be meaningless, we reasoned, then the question of whether to kill one to save five may well reduce to a simple “math problem”—a dispassionate comparison of utilities. A high score on this scale indicates that a perception that life is relatively meaningless, which has been found to correlate with measures of depression (Kunzendorf & Maguire, 1995).

3. Methods

Two-hundred and eight undergraduates (101 female) participated in exchange for \$3 each. They were tested in a small group setting (typically one to four participants per session). Participants responded to 14 sacrificial dilemmas and a battery of individual difference measures. The presentation of the dilemmas and individual difference measures was counterbalanced across participants and did not affect the results.

The dilemmas (drawn from Bartels (2008)) were presented in random order. Each pitted utilitarian and deontological options against each other, as in the footbridge case presented earlier (see Appendix A). Participants indicated their preferences by clicking on a box, as below:

In this situation, would you push the man?

NO YES

Higher scores indicate greater preference for the utilitarian solution.

The individual differences battery included an adapted version of a 30-item psychopathy scale (SRP-III, Paulhus, Neumann, & Hare, *in press*), the 18-item No Meaning scale (Kunzendorf, Moran, & Gray, 1995), and the 20-item Machiavellianism scale (Mach-IV, Christie & Geis, 1970). We also included two additional items in order to control for potential confounds: A 10-item social desirability scale (MC-1, Strahan & Gerbasi, 1972), a standard measure of a participant's tendency to respond in a manner that would be perceived favorably by others. This was included in order to control for the possibility that responses to the moral dilemmas were a reflection of this tendency. We also collected information about the gender of our respondents, as recent research has demonstrated systematic gender differences in not only the traits of psychopathy and Machiavellianism (Paulhus & Williams, 2002), but also in a variety of philosophical intuitions, including responses to ethical dilemmas (Buckwalter & Stich, 2010).

Participants responded to a randomized ordering of all 78 items, including “I like to see fist-fights” (psychopathy), “When you really think about it, life is not worth the effort of getting up in the morning” (No Meaning), and “The best way to handle people is to tell them what they want to hear” (Machiavellianism). Participants rated their level of agreement by clicking on a continuum bounded by “strongly disagree” and “strongly agree.”

4. Results

Participants who scored higher on the psychopathy ($\alpha = .86$), no meaning ($\alpha = .90$), and Machiavellianism ($\alpha = .78$) scales indicated a greater preference for utilitarian options in the ethical dilemmas. This was true for the overall analysis, where we collapse responses across all 14 dilemmas ($r_s = .38, .21, .35, p_s < .05$), as well as for the vast majority of the individual ethical dilemmas (see Table 1). As can be seen in Fig. 1, only the minority of participants who scored high on each of our focal personality measures indicated a general overall utilitarian preference.

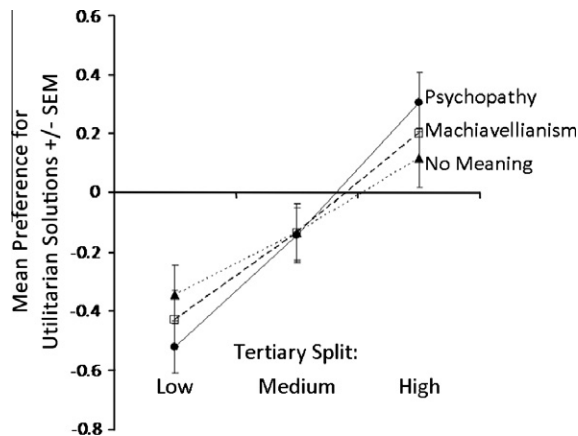
Table 2 reports the correlations between our predictor variables and average preference for the utilitarian solutions. Social desirability was significantly correlated with each of the predictor variables, and male participants scored significantly higher on psychopathy, no meaning, and Machiavellianism, scored lower on social desirability, and indicated more utilitarian preferences than female participants (Table 2 reports these biserial correlations).

To control for the observed effects of gender and social desirability on utilitarian preferences, we conducted separate multiple regressions for each of our focal individual difference variables—using each as a predictor of utilitarian

Table 1

Correlations between Individual differences and preferences.

	Psychopathy	No meaning	Machiavellianism
Average preference	0.38*	0.21*	0.35*
Submarine	0.33*	0.25*	0.37*
Trespassers	0.37*	0.17*	0.31*
Hostages	0.33*	0.14*	0.27*
Bystander	0.16*	0.08	0.20*
Liferaft	0.24*	0.12†	0.28*
Plane crash	0.26*	0.20*	0.20*
Prisoners of war	0.26*	0.12†	0.20*
Fumes	0.09	−0.01	0.12†
Spelunkers	0.21*	0.07	0.27*
Soldiers	0.21*	0.11	0.22*
Surgery	0.22*	0.21*	0.04
Derailment	0.06	0.04	0.04
Footbridge	0.20*	0.18*	0.06
Baby	0.26*	0.17*	0.30*

* $p < .05$.† $p < .10$.**Fig. 1.** Degree of utilitarian preference for participants scored as low, medium, and high on psychopathy, Machiavellianism, and No Meaning (numbers greater than zero on the y-axis indicate utilitarian preference across dilemmas).

¹ For a more discerning test of the relationship between our traits of interest and utilitarian preference, we ran a second set of analyses on a subset of our items. The subset was selected on the basis of a methodological critique offered by Kahane and Shackel (2008), who asked philosophers to code the dilemmas used in Greene et al. (2001) and concluded that some dilemmas did not reliably pit utilitarianism against deontology. Of the 14 dilemmas we used, eleven were close variants of those used in Greene et al. (2001), and seven of these were judged by the philosophers polled by Kahane and Shackel as pitting utilitarianism against deontology. We re-ran each of the analyses reported in the paper, restricting the focus of these analyses to these seven dilemmas (Bystander, Liferaft, Fumes, Soldiers, Surgery, Footbridge, and Baby), and found no difference between these analyses and the overall analysis presented in the paper. Each of the three personality traits of interest significantly correlated with utilitarian preference (psychopathy = 0.30, no meaning = 0.21, Machiavellianism = 0.19, $ps < .05$), and each of these relationships remained significant when controlling for gender and social desirability (Std/bs: psychopathy = 0.30, no meaning = 0.15, Machiavellianism = 0.29, $ps < .05$). We thank an anonymous reviewer for suggesting this analysis.

preferences while controlling for gender and social desirability. As Table 3 shows, the relationships between utilitarian preferences and psychopathy, no meaning, and Machiavellianism are robust.¹ We also ran a multiple regression using our three focal factors, social desirability, and gender as predictors, and found that psychopathy and Machiavellianism each uniquely predict average utilitarian preferences (Std/bs = .23), while no meaning, social desirability, and gender dropped to non-significance (Std/bs = .01, .09, and −.11).²

5. Discussion

Our study illustrates that the widely adopted use of sacrificial dilemmas in the study of moral judgment fails to distinguish between people who are motivated to endorse utilitarian moral choices because of underlying emotional deficits (such as those captured by our measures of psychopathy and Machiavellianism) and those who endorse it out of genuine concern for the welfare of others and a considered belief that utilitarianism is the optimal way of achieving the goals of morality. Consistent with what is known about the emotional deficits in a clinical population that endorses utilitarian responses to moral dilemmas (i.e., patients with damage to their VMPFC), we found that non-clinical individuals who indicated utilitarian preferences scored higher on measures of psychopathy and Machiavellianism. In addition, these participants also appear to perceive less meaning in life. Although these results are consistent with a recent study reporting a relationship between psychopathic traits and moral judgment utilizing a different measure of psychopathy (Glenn, Koleva, Iyer, Graham, & Ditto, 2010), a number of previous researchers have failed to find a reliable association between the two (Cima, Tonnaer, & Hauser, 2010; Dolan & Fullam, 2010; Glenn, Raine, Schug, Young, & Hauser, 2009). One possibility for the discrepancy between these studies and ours is that differences in the population from which the subjects were drawn—the above-mentioned studies have sampled directly from a psychiatric population or a population of criminal offenders. One possibility is that individuals diagnosed with psychopathy may be highly motivated to

² To further examine the discriminant and predictive validity of Psychopathy, No Meaning, and Machiavellianism, we conducted two additional analyses. First, we compared the fit of a confirmatory factor analysis that imposed a one-factor solution to capture the scale variance in the 68 scale items of our three focal variables to a second model that assigned the scale items each to their respective construct. The three-factor model fit significantly better, suggesting that these constructs have reasonable discriminant validity. Next, we ran a structural equation model that was similar in form to the multiple regression, except that instead of using simple averages for the five variables for which we have multiple responses, it estimated latent constructs. The inferences the structural equation model yields concerning the three focal personality variables mirror those from the regression: psychopathy and Machiavellianism each uniquely predicted utilitarian preference. Whereas Paulhus and Williams (2002) found these two traits to be distinctively related to external constructs (e.g., psychopathy negatively correlates with neuroticism and openness, whereas Machiavellianism is positively correlated with both), we found unique predictive validity for each construct in predicting utilitarian preference. However, as did Paulhus and Williams, we found the two traits to be significantly correlated, which may suggest that these constructs reflect a common underlying source of individual differences.

Table 2

Intercorrelations among variables.

	No meaning	Machiavellianism	Social desirability	Gender (1 = M, 2 = F)	Avg. preference
Psychopathy	0.50*	0.61*	–0.43*	–0.42*	0.38*
No meaning	–	0.40*	–0.21*	–0.16*	0.21*
Machiavellianism	–	–	–0.54*	–0.24*	0.35*
Social desirability	–	–	–	0.14*	–0.15*
Gender (M = 1, 2 = F)	–	–	–	–	–0.25*

* $p < .05$.**Table 3**

Relationships between individual differences and preferences, controlling for social desirability and gender—standard betas.

Factor	Std beta	Social desirability	Gender (1 = M, 2 = F)
Psychopathy	0.34*	0.02	–0.11
No meaning	0.16*	–0.08	–0.21*
Machiavellianism	0.34*	0.06	–0.18*

* $p < .05$.

report in a manner that they believe will make them seem like an “average” individual because, among other reasons, they may be concerned that their responses may have consequences for their treatment or incarceration. By contrast, our respondents, whom we have reason to believe share similar emotional tendencies as psychopaths in a clinical population, may not only have no concern over such consequences, they may not even be aware that their responses deviate from that of “normal” respondents. Nonetheless, the current findings may be sufficient to question the conclusion some researchers have reached given the overall lack of evidence between psychopathy and moral judgment in previous studies—namely, that the role of emotion in shaping moral judgment has been overestimated (e.g., Glenn et al., 2009).

Importantly, these results also give rise to a methodological concern in the study of moral judgment—namely, that we should be wary of favoring a method that equates the quality of moral judgment with responses that are endorsed primarily by individuals who are likely perceived as *less* moral (because they possess traits like callousness and manipulativeness). In other words, adopting such a method can lead to the counterintuitive inference that “correct” moral judgments are most likely to be made by the individuals least likely to possess the character traits generally perceived as moral.

We should note that our results do not speak to whether utilitarianism (or deontology) is the correct normative ethical theory, as the characteristics of a theory's proponents cannot determine its normative status. In addition, favoring a utilitarian or deontological solution to a sacrificial moral dilemma does not necessarily indicate that a participant endorses (or understands) utilitarianism or deontology as a full-blown ethical theory—just because an individual responds *like* a utilitarian would be not sufficient evidence that she is a utilitarian. Consider, for instance, a man who finds the thought of pushing a fat man off of a footbridge to his death to be intrinsically appealing—whether or not the action saved more lives. Concluding on the basis of his response to a sacrificial di-

lemma that he must be convinced that utilitarianism is the best ethical theory makes little sense.

Nor do our results show that endorsing utilitarianism is pathological, as it is unlikely that the personality styles measured here would characterize all (or most) proponents of utilitarianism as an ethical theory (nor is the measure of psychopathic personality traits we used sufficient to conclude that any respondents reach clinical levels of psychopathy). It is also possible that possessing these sub-clinical psychopathic traits may be of moral value inasmuch as individuals who are capable of such emotional detachment, while appearing to possess a questionable moral character in some situations, may be better able to act for the greater good in ways that would prove difficult for many (such as the very situations described in our target dilemmas). Nonetheless the relative infrequency of such events would seem, at the very least, to undermine the validity of using these measures as a metric for optimal moral judgment in everyday life.

Finally, our empirical demonstration points to the problematic nature of studying moral judgment by identifying “errors” in how subjects respond to moral dilemmas. As Pizarro and Uhlmann (2005) argued, it may be sufficient to simply document how, when, and why individuals make the moral judgments that they do without relying on the adoption of a normative standard.³ We believe psychologists can make progress by developing accurate descriptive theories that explain why individuals favor deontological judgments in some situations and utilitarian judgments in others (or whether moral judgment is even adequately captured by these philosophical frameworks; cf. Bauman & Skitka, 2009) without having to rely on the claim that individuals are making an “error” in some cases.

³ For discussion and debate concerning how best to study and understand morality in the context of normative standards, see Bazerman & Greene, 2010; Bennis, Medin, & Bartels, 2010a, 2010b; Schwartz, 2010; Tetlock & Mitchell, 2010.

Acknowledgements

For helpful comments, we acknowledge without implicating Jonathan Baron, Chris Bauman, Eugene Caruso, Ben Converse, Kamel Jedidi, Craig Joseph, Ye Li, Pete McGraw, Paul Bloom, Steven Sloman, David Tannenbaum, and Oleg Urminsky.

Appendix A. Sacrificial dilemmas

Submarine (Greene et al., 2001)

You are the captain of a small military submarine traveling underneath a large iceberg. An onboard explosion has dramatically decreased the oxygen supply and has injured a member of the crew. He may die from his injuries. The remaining oxygen is not sufficient for the entire six-man crew to make it to the surface. If you shoot the injured crewmember, there will be just enough oxygen for everyone else to survive.

In this situation, would you shoot the crewmember?

Trespassers (Greene et al., 2001)

You and a group of friends are backpacking, and while crossing a rocky plain in the Guatemalan highlands, you inadvertently camp on a local clan's burial ground. The leader of the clan says that his clan's laws command him to kill all six of you as a sacrifice to the gods, but that if you kill one of your own, the gods will be appeased, and the rest will be let go.

In this situation, would you kill one of your friends?

Hostages (Greene et al., 2001)

You are traveling with five other peace advocates in a war-torn country when your group is taken hostage by a group of terrorists. After a week, the terrorists' demands have not been met, and they offer you a choice: either they will kill all of you, or if you execute one of your fellow hostages, you and the others are free to go.

In this situation, would you execute one of your fellow hostages?

Bystander (Foot, 1967)

In the path of a runaway train car are five railway workmen who will surely be killed unless you, a bystander, do something. If you flip a switch, the train will be diverted onto another track, where it will kill a single railway workman.

In this situation, would you flip the switch?

Liferaft (Regan, 1983)

You are on a small ship, a fire breaks out, and the ship has to be abandoned. Because your tiny liferaft is carrying more than its capacity, it is sitting dangerously low in the water. The seas get rough and the raft begins to fill with water. Unless you do something, all six of you will drown. There is an injured person onboard who may die either way. If you throw him overboard, everyone else will be saved.

In this situation, would you throw him overboard?

Plane Crash (Marshall, 1993; Greene et al., 2001)

Your plane has crashed in the Himalayas. The only survivors are you, some other men, and a young boy.

The six of you travel for days, battling extreme cold and wind. Your only chance of survival is to make it to a village a few days away. The boy cannot move very quickly. Without food, you and the other men will surely die. One of the men suggests killing the boy and eating his remains over the next few days.

In this situation, would you sacrifice the boy?

Prisoners of War (Baron, 1992)

You and some other soldiers were captured. After a year in a prison camp, your group tried to escape but was caught. The warden has decided to hang your group in front of the other prisoners of war. At the gallows, he releases the noose from your neck and announces that if you pull the chair from underneath one man in your group, the remaining five will be set free, otherwise you all die. He means what he says.

In this situation, would you remove the chair?

Fumes (Thomson, 1986)

You are the late-night watchman in a hospital where an accident has occurred in one of the on-site testing labs, and now there are deadly fumes rising up through the hospital's ventilation system. The fumes are headed to a certain area where there are five patients who will surely die. If you flip a switch, the ventilation system will cause the fumes to bypass this room and enter a room containing a single patient, killing him.

In this situation, would you flip the switch?

Spelunkers (<http://news.bbc.co.uk/1/hi/magazine/4954856.stm>)

You and five others are exploring a seashore cave. A rock falls and blocks your exit. The tide is rising. You spot a hole elsewhere and let a large man in your group out first. He gets stuck, and unless you get out soon, everyone but this man (whose head is sticking out of the cave) will drown. You have a stick of dynamite that will not move the rock, but will blast the man out of the hole. He does not want to die; neither do you or the others.

In this situation, would you blast him out?

Soldiers (Greene et al., 2001)

You are leading a group of soldiers returning from a completed mission in enemy territory when one of your men steps in a trap. He is injured, and the trap is connected to a device that alerts the enemy to your presence. If the enemy finds your group, all six of you will die. If you leave him behind, he will be killed, but the rest of the group will escape safely.

In this situation, would you leave him behind?

Surgery (Foot, 1967)

You are a surgeon with a number of patients. Five of them need organ transplants. Each of them needs a different organ or they will surely die. You have another patient who is healthy and would be an ideal organ donor for the others. If you transplant his organs (against his will) into the bodies of the other patients, they will live but he will die.

In this situation, would you perform this transplant?

(continued on next page)

Derailement (Unger, 1996)

In the path of a runaway train car are five railway workmen who will surely be killed unless you, a bystander, do something. If you flip a switch, the train will be diverted onto a set of tracks in need of repair. The train will be derailed and go down a hill, across a road, and into a man's yard. The owner, sleeping in his hammock, will be killed

In this situation, would you flip the switch?

Footbridge (Thomson, 1985)

In the path of a runaway train car are five railway workmen who will surely be killed unless you, a bystander, do something. You are standing on a pedestrian walkway that arches over the tracks next to a large stranger. Your body would be too light to stop the train, but if you push the stranger onto the tracks, killing him, his large body will stop the train

In this situation, would you push the man?

Baby (Alda & et al., 1983; Greene et al., 2001)

Enemy soldiers have taken over your village and will kill all remaining civilians. You and five others are hiding in the cellar of a large house. Soldiers have come to search the house for valuables. A baby in your group begins to cry. So, you cover her mouth, but she cannot breathe. If you remove your hand, the baby can breathe, but her crying will summon the soldiers who will kill everyone in the cellar

In this situation, would you smother the baby?

References

- Alda, A., et al. (1983). Goodbye, farewell, and amen [Television series episode]. In B. Metcalfe (Producer), *M*A*S*H*. Los Angeles: 20th Century, Fox Television.
- Baron, J. (1992). The effect of normative beliefs on anticipated emotions. *Journal of Personality and Social Psychology*, 63, 320–330.
- Baron, J., & Ritov, I. (2009). Protected values and omission bias as deontological judgments. In D. M. Bartels, C. W. Bauman, L. J. Skitka, & D. L. Medin (Eds.), *Moral judgment and decision making: The psychology of learning and motivation* (Vol. 50, pp. 133–167). San Diego: Elsevier.
- Baron, J., & Spranca, M. (1997). Protected values. *Organizational Behavior and Human Decision Processes*, 70, 1–16.
- Bartels, D. M. (2008). Principled moral sentiment and the flexibility of moral judgment and decision making. *Cognition*, 108, 381–417.
- Bauman, C. W., & Skitka, L. J. (2009). In the mind of the perceiver: Psychological implications of moral conviction. In D. M. Bartels, C. W. Bauman, L. J. Skitka, & D. L. Medin (Eds.), *Moral judgment and decision making: The psychology of learning and motivation* (Vol. 50, pp. 339–359). San Diego: Elsevier.
- Bazerman, M. H., & Greene, J. D. (2010). In favor of clear thinking: Incorporating moral rules into a wise cost-benefit analysis—Commentary on Bennis, Medin, & Bartels (2010). *Perspectives on Psychological Science*, 5, 209–212.
- Bennis, W. M., Medin, D. L., & Bartels, D. M. (2010a). The costs and benefits of calculation and moral rules. *Perspectives on Psychological Science*, 5, 187–202.
- Bennis, W. M., Medin, D. L., & Bartels, D. M. (2010b). Perspectives on the ecology of decision modes: Reply to comments. *Perspectives on Psychological Science*, 5, 213–215.
- Buckwalter, W., Stich, S. (2010). Gender and philosophical intuition. Available at SSRN: <<http://ssrn.com/abstract=1683066>> (26.09.10).
- Christie, R., & Geis, F. L. (1970). *Studies in Machiavellianism*. New York: Academic Press.
- Cima, M., Tonnaer, F., & Hauser, M. D. (2010). Psychopaths know right from wrong but don't care. *Social Cognitive and Affective Neuroscience*, 5, 59–67.
- Dolan, M. C., & Fullam, R. S. (2010). Moral/conventional transgression distinction and psychopathy in conduct disordered adolescent offenders. *Personality and Individual Differences*, 49, 995–1000.
- Feltz, A., & Cokely, E. T. (2008). The fragmented folk: More evidence of stable individual differences in moral judgments and folk intuitions. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th annual conference of the Cognitive Science Society* (pp. 1771–1776). Austin, TX: Cognitive Science Society.
- Foot, P. (1967). The problem of abortion and the doctrine of the double effect. *Oxford Review*, 5, 5–15.
- Glenn, A. L., Koleva, S., Iyer, R., Graham, J., & Ditto, P. H. (2010). Moral identity in psychopathy. *Judgment and Decision Making*, 5, 497–505.
- Glenn, A. L., Raine, A., Schug, R. A., Young, L., & Hauser, M. (2009). Increased DLPFC activity during moral decision-making in psychopathy. *Molecular Psychiatry*, 14, 909–911.
- Greene, J. D., Cushman, F. A., Stewart, L. E., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2009). Pushing moral buttons: The interaction between personal force and intention in moral judgment. *Cognition*, 111(3), 364–371.
- Greene, J. D., 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., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, 133, 55–66.
- Iliev, R., Sachdeva, S., Bartels, D. M., Joseph, C. M., Suzuki, S., & Medin, D. L. (2009). Attending to moral values. In D. M. Bartels, C. W. Bauman, L. J. Skitka, & D. L. Medin (Eds.), *Moral judgment and decision making: The psychology of learning and motivation* (Vol. 50, pp. 169–190). San Diego: Elsevier.
- Kahane, G., & Shackel, N. (2008). Do abnormal responses show utilitarian bias? *Nature*, 452, 7185.
- Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., et al. (2007). Damage to the prefrontal cortex increases utilitarian moral judgments. *Nature*, 446, 908–911.
- Kunzendorf, R. G., Maguire, D. (1995). *Depression: The reality of 'no meaning' versus the delusion of negative meaning*. Unpublished manuscript, University of Massachusetts, Lowell.
- Kunzendorf, R. G., Moran, C., & Gray, R. (1995). Personality traits and reality-testing abilities, controlling for vividness of imagery. *Imagination, Cognition, and Personality*, 15, 113–131.
- Marshall, F. (Director). (1993). *Alive [Motion picture]*. United States: Paramount Pictures.
- Mikhail, J. (2007). Universal moral grammar: Theory, evidence, and the future. *Trends in Cognitive Sciences*, 11, 143–152.
- Moore, A. B., Clark, B. A., & Kane, M. J. (2008). Who shalt not kill? Individual differences in working memory capacity, executive control, and moral judgment. *Psychological Science*, 19, 549–557.
- Nichols, S., & Mallon, R. (2006). Moral dilemmas and moral rules. *Cognition*, 100, 530–542.
- Paulhus, D. L., Neumann, C. F., & Hare, R. D. (in press). *Manual for the self-report psychopathy scale*. Toronto: Multi-Health Systems.
- Paulhus, D. L., & Williams, K. M. (2002). The dark triad of personality: Narcissism, Machiavellianism, and psychopathy. *Journal of Research in Personality*, 36, 556–563.
- Pizarro, D. A., & Uhlmann, E. (2005). Do normative standards advance our understanding of moral judgment? *Behavioral and Brain Sciences*, 28, 558–559.
- Regan, T. (1983). *The case for animal rights*. Berkeley: University of California Press.
- Saver, J. L., & Damasio, A. R. (1991). Preserved access and processing of social knowledge in a patient with acquired sociopathy due to ventromedial frontal damage. *Neuropsychologia*, 29, 1241–1249.
- Schwartz, B. (2010). The limits of cost-benefit calculation: Commentary on Bennis, Medin, & Bartels (2010). *Perspectives on Psychological Science*, 5, 203–205.
- Sen, A., & Williams, B. (1982). *Utilitarianism and beyond*. Cambridge, UK: Cambridge University Press.
- Skitka, L. J., Bauman, C. W., & Sargis, E. G. (2005). Moral conviction: Another contributor to attitude strength or something more? *Journal of Personality and Social Psychology*, 88, 895–917.
- Strahan, R., & Gerbasi, K. C. (1972). Short, homogeneous versions of the Marlowe-Crowne social desirability scale. *Journal of Clinical Psychology*, 28, 191–193.
- Sunstein, C. R. (2005). Moral heuristics. *Behavioral and Brain Sciences*, 28, 531–542.
- Tetlock, P. E. (2003). Thinking the unthinkable: Sacred values and taboo cognitions. *Trends in Cognitive Sciences*, 7, 320–324.

- Tetlock, P. E., & Mitchell, G. (2010). Situated social identities constrain morally defensible choices: Commentary on Bennis, Medin, & Bartels (2010). *Perspectives on Psychological Science*, 5, 206–208.
- Thomson, J. J. (1985). The trolley problem. *Yale Law Journal*, 94, 1395–1415.
- Thomson, J. J. (1986). *Rights, restitution, and risk*. Cambridge, MA: Harvard University Press.
- Unger, P. (1996). *Living high and letting die: Our illusion of innocence*. New York: Oxford University Press.