

# Moral Judgments, Gender, and Antisocial Preferences: An Experimental Study\*

Juergen Bracht<sup>†</sup> and Adam Zylbersztejn<sup>‡</sup>

May 2018

## Abstract

We study questionnaire responses to situations in which sacrificing one life may save many other lives. We demonstrate gender differences in moral judgments: males are more supportive of the sacrifice than females. We investigate a source of the endorsement of the sacrifice: antisocial preferences. First, we measure individual proneness to spiteful behavior, using an experimental game with monetary stakes. We demonstrate that spitefulness can be sizable – a fifth of our participants behave spitefully – but it is not associated with gender. Second, we find that gender is consistently associated with responses even when we account for individual differences in the propensity to spitefulness.

Keywords: Moral Dilemma; Moral Judgments; Experiment; Gender; Antisocial Preferences; Individual Differences.

JEL Classification: C91; D03; D63

## 1 Introduction

A growing literature in psychology, economics and philosophy investigates people's moral judgments. For that purpose, participants are usually asked to engage in thought experiments in which they respond to hypothetical dilemmas – two possibilities neither of which is unambiguously acceptable. Typically, these stylized situations (such as the trolley dilemma outlined in the next section) involve the question of whether one should sacrifice one life to save many others, and thus highlight the conflict between utilitarianism (according to which an act is right if and only if it leads to the greatest total amount of well-being), and deontology (which prescribes that acts are inherently good or bad, regardless of their consequences). When making such judgments, people are driven by their feelings of agreement or disagreement with a certain course of action (Haidt, 2001, 2003). Hence,

---

\*We thank an associate editor and two anonymous referees for helping us improve the paper. This research has been supported by the French National Research Agency (ANR), through the program Investissements d'Avenir (ANR-10-LABX-93-01). The research has been performed within the framework of LABEX CORTEX (ANR-11-LABX-0042) of Universite de Lyon, within the program Investissements d'Avenir (ANR-11-IDEX-007) run by the French National Research Agency (ANR). Maxim Frolov, Ivan Ouss and Remi Yin assisted with the experimental sessions. We thank Daniel Bartels, Astrid Hopfensitz, Michal Krawczyk, David Masclet, Ernesto Reuben, seminar participants at the Department of Economics of the University of Crete, at the GATE CNRS Laboratory, and at the 2015 ASFEE meeting in Paris for valuable comments. Adam Zylbersztejn is grateful to the University of Lyon 2 for support through ASPRE 2015 and AIP 2016.

<sup>†</sup>University of Aberdeen Business School, Department of Economics, Edward Wright Building, Dunbar Street, Aberdeen, AB24 3QY, Scotland. Email: [juergen.bracht@abdn.ac.uk](mailto:juergen.bracht@abdn.ac.uk).

<sup>‡</sup>Corresponding author. Univ Lyon, Université Lumière Lyon 2, GATE L-SE UMR 5824, F-69130, Ecully, France. Email: [zylbersztejn@gate.cnrs.fr](mailto:zylbersztejn@gate.cnrs.fr).

moral dilemma questionnaires have become well-established methodology for eliciting people’s moral judgments.<sup>1</sup> A recent line of research on moral dilemmas investigates whether moral judgment is associated with gender, and how moral judgment is related to other-regarding preferences. A first set of papers concludes that there are gender differences in moral judgment; men are more prone to endorse the utilitarian solution in hypothetical moral dilemmas. The findings are reported in Petrinovich, O’Neill and Jorgensen (1993), Zamzov and Nichols (2009), Banerjee, Huebner, and Hauser (2010) and Buckwalter and Stich (2014). Furthermore, a large-sample investigation by Hauser et al. (2007) with over 2000 participants confirms a large and significant gender difference in a variety of moral dilemmas. Another large-sample study by Bourget and Chalmers (2014) finds a number of gender differences in judgments, including females being less likely to agree with the utilitarian solution in a moral dilemma. In addition, Hsu, Anen, and Quartz (2008) report that males tend to favor the utilitarian solution in a real-world moral dilemma more than females: a choice of actual allocation of meals among poor African kids through a charity organization.<sup>2</sup>

The second set of studies points to a – somewhat surprising – conclusion that people with anti-social attitudes are prone to make utilitarian judgments. Koenigs et al. (2007) conducted a study of brain-damaged patients with acquired sociopathy. These kind of emotional deficits are similar to those observed in psychopaths (Saver and Damasio, 1991). These kind of patients are found to display unusually high levels of endorsement of the utilitarian solution to moral dilemmas. In the same vein, Glenn et al. (2010) report that psychopathic traits – measured using Levenson’s Self-Report Psychopathy Scale – predict a stronger endorsement of utilitarian solutions across several moral dilemmas. Other related studies also point to a negative association between the endorsement of utilitarianism and an aversion to harming others (Crockett et al., 2010; Cushman et al., 2012) and empathic concerns (Choe and Min, 2011; Gleichgerrecht and Young, 2013). In a similar vein, Kahane et al. (2015) report that high rates of utilitarian judgments are associated with the endorsement of unfair business practices, rational egoism, (hypothetically) donating less to a charity, and a weaker identification with the whole humanity.

Finally, Bartels and Pizarro (2011) – in a study most closely related to ours – report gender difference in moral judgments as well (in a large set of moral dilemmas that we also employ herein). However, their report also highlights the importance of individual traits: there is a positive relation between endorsement of sacrifice-one-live-to-save-many-lives and measures of psychopathic personality (low empathy, callous affect, and thrill-seeking), Machiavellianism (cynical, emotionally detached from others, and manipulative), and perceived life meaninglessness (melancholic existential concerns). Furthermore, they find that men tend to score higher in all of those psychological questionnaires, and that the gender effect on moral judgments fades away when psychological traits are accounted for. Their finding, in turn, raises the question of whether the relationship between moral judgment and gender is spurious because the research failed to account for antisocial traits as a confounding factor.

In this paper, we provide a novel method of investigating the interaction between moral judgments, gender, and antisocial preferences. We elicit the propensity for other-damaging behavior (spitefulness) using an experimental game, and use this measure as a control when explaining gender difference in responses to a large moral dilemma questionnaire. The measure of spitefulness if

---

<sup>1</sup>Some experimenters have also investigated incentivized moral dilemmas in which people’s decisions bear real material consequences. See Hsu, Anen, and Quartz (2008), Gold, Colman, and Pulford (2014) and Gold, Colman, and Pulford (2015). For instance, Gold, Colman, and Pulford (2014) look at a version of a moral dilemma in which the harm is a small economic loss. The authors find some differences between actions and judgments, yet conclude that nothing definitive can be said about their scope and causes.

<sup>2</sup>On the other hand, a replication study by Seyedsayamdost (2015), and studies of Gold et al. (2014, 2015) report no systematic gender difference in moral judgments.

based on behavioral data collected in an economic experiment. We observe decisions in a laboratory game in which the decision maker is granted a flat payoff of 10 Euros, and – in addition – sets the payoff for another participant to any amount between 0 and 10 Euros. An amount lower than 10 Euros harms the other person and indicates a preference for spite.<sup>3</sup>

Following the previous studies on moral judgments and antisocial preferences, the main interest of our experimental game is to provide a direct way to classify participants into two categories: those who engage in antisocial behavior and those who do not. Antisocial behavior is widespread in our setting.<sup>4</sup> In line with the previous evidence from several economic experiments (reviewed in the next section), in our experiment roughly one participant in five purposefully reduces the other participant’s earnings. We use this behavioral indicator variable as a control in regression analysis. We find that gender is associated with moral judgment even when we control for participants’ spitefulness: males are found to be more prone than females to endorse the act of sacrificing one life to save many other lives. Thus, in line with the previous findings from the psychology literature, our results suggest that men indicate a greater endorsement of utilitarian solutions. However, notwithstanding the previous findings, our study suggests that this does not happen because men are more likely to display preferences that most would consider antisocial, and thus immoral.

## 2 Hypotheses

**Hypothesis 1.** 20% of players in the experimental game will reduce the other person’s payoff.

This hypothesis directly stems from the classic model of other-regarding preferences by Levine (1998) which captures pro- and antisocial motives (altruism and spitefulness) in a symmetric manner. Players are assumed to have a utility function that weights other player’s payoff with a parameter  $a$  ranging between -1 and 1: positive values of  $a$  indicate altruism and negative values indicate spitefulness. In addition, players may be more altruistic to an opponent who is more altruistic toward them (which is not, however, relevant for our one-shot distributional task). Levine estimates the model using the data from several economic experiments: the ultimatum game (Roth et al., 1991), competitive auction (Roth et al., 1991), the centipede game (McKelvey and Palfrey, 1992), and the public goods game (Isaac and Walker, 1988), and concludes that 20% of the population can be clearly characterized as spiteful (average value of  $a$  equal to -0.9).<sup>5</sup> Similar estimates are provided by Charness et al. (2014) in the context of sabotage behavior in a tournament experiment.<sup>6</sup>

---

<sup>3</sup>Charness and Grosskopf (2001) have conducted an experiment with games that are similar to our game. Their design focuses on the association between self-reported happiness and one’s concern for relative payoffs. They find little association; most participants disregarded relative payoffs and instead typically made choices resulting in higher social payoff. Another related game is the envy game (Bartling et al., 2009). However, the envy game allows for enforcement of an egalitarian outcome, whereas our spite game allows for enforcement of an unequal allocation. We suspect that individuals who are spiteful are envious as well.

<sup>4</sup>For richer taxonomies of other-regarding preferences in the laboratory see Fisman, Kariv and Markovits (2007), Murphy, Ackermann, and Handgraaf (2011) and Kerschbamer (2015).

<sup>5</sup>Interestingly, Levine also notes (p. 616): *Finally, we turn to other implications of the theory that could be tested in future experiments. For example, there is a set of implications of the theory for one-player games that has not been examined experimentally: The theory predicts that spiteful play should take place, even in a one-player setting. In other words, if a single player is given the option to deprive an opponent of money at a sufficiently modest cost to himself, then we should observe 20% or more of players availing themselves of this option.* Our experiment can be viewed as a direct out-of-the-sample test of this conjecture. In the light of our data, the 20% rate of spiteful behavior conjectured by Levine looks highly accurate.

<sup>6</sup>Zizzo and Oswald (2001) find an even larger fraction of spiteful behavior. In their experiment, participants may pay to ‘burn’ the money of the bigger earners. Together with a similar experiment Zizzo (2004) in which players have

The next three hypotheses stem from the literature reviewed in the opening section of the paper: a large body of papers showing a gender difference in moral judgments (Hypothesis 3), as well as the study by Bartels and Pizarro (2011) highlighting the interplay between gender, antisocial preferences, and moral judgments (Hypotheses 2 and 4).<sup>7</sup>

**Hypothesis 2.** Male participants are more spiteful than female participants.

**Hypothesis 3.** Male participants display a stronger support for utilitarianism in moral judgments than female participants.

**Hypothesis 4.** Individuals who act spitefully in the experimental game may display a stronger support for utilitarianism in moral dilemmas.

### 3 Experimental Design

We conducted the sessions at the Experimental Economics Laboratory Paris (LEEP). There were 12 sessions, each with between 11 and 20 participants. From LEEP’s database, we recruited participants who had completed LEEP’s registration process. We stratified the sampling to ensure balance, with 99 female and 99 male participants. The participants’ average age was 24. Most participants had previously taken part in an experiment at LEEP. Roughly four-fifths of our participants were still enrolled in university studies. For recruitment, we used the ORSEE software (Greiner, 2004). For the sessions, we used the Regate software (Zeilinger, 2000). We asked participants to complete two experimental tasks. Specifically, participants were asked to both play a game whose outcome determined the participant’s monetary gains, and to answer a questionnaire consisting of 14 questions with different moral dilemmas. We accounted for potential order effects in two ways. In six sessions, participants first played the game and then completed the questionnaire. In the other six sessions, the order was reversed. Furthermore, we arranged the order of the presentation of the 14 dilemmas to ensure balance across sessions.<sup>8</sup>

At the beginning of each session, instructions were distributed and read aloud. Furthermore, participants were informed that additional instructions would be displayed on their computer screens later on. The instructions and questions, translated from French into English, are available as supplementary material in the appendix.

---

a chance to steal money from others, these results suggests that one could manipulate the amount of spitefulness if the other player is given a larger endowment than the decision-maker. See also Abbink et al. (2018) for related evidence. In addition, Abbink and Sadrieh (2009) and Abbink and Herrmann (2011) study the joy-of-destruction game and find that people do not use the opportunity to destroy others’ wealth when they are observed by their potential victims (which might expose them to a direct retaliation), but people do so if such behavior can be concealed behind a veil of random destruction. Indeed, Abbink and Sadrieh (2009) find that the frequency of a spiteful act increases from less than 10% in the “open” treatment to almost 40% in the “hidden” treatment and the destroyed amounts differ considerably between the two conditions as well. In another study, Nishimura et al. (2011) build on an auction model in which spiteful low-value types may overbid so as to reduce the high-value types’ gains, and confirm the presence of such an aggressive behavior in the laboratory (similar results are also provided by Kimbrough and Reiss, 2012).

<sup>7</sup>Hypothesis 2 is also corroborated by the findings in Croson and Gneezy (2009) and Engel (2006). Croson and Gneezy (2009) review the literature on gender differences in other-regarding preferences, and conclude that females are more sensitive to social cues in determining appropriate behavior. In a meta-analysis of dictator game experiments, Engel (2006) reports that females act more prosocially, i.e. give significantly more than males. Note, however, that the dictator game is well-suited for capturing prosocial behavior (sharing with others), but is silent on why people do not share with others: both selfishness and spitefulness preferences predict to the same behavior. Our spite game is a way to tease apart selfishness from antisocial preferences.

<sup>8</sup>Petrinovich and O’Neill (1996), Alexander and Weinberg (2007), and Zamzow and Nichols (2009) show that the order of presentation affects moral responses.

### 3.1 Moral Judgment Questionnaire

We elicit participants' moral judgments based on their responses to a 14-item questionnaire with hypothetical moral dilemmas based on Bartels (2008). This questionnaire provides a contextually rich and comprehensive set of dilemma questions previously used in the psychology literature to study moral judgments.<sup>9</sup> Our judgment elicitation method comes from Bartels and Pizarro (2011). Below, we illustrate this procedure for The Trolley (Footbridge) Dilemma.

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 next to a larger stranger on a pedestrian walkway that arches over the tracks. 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?

Please, indicate your answer by ticking a box on the scale displayed below (the leftmost box corresponds to the strongest disagreement, the rightmost box corresponds to the highest agreement):

NO     -2     -1     1     2    YES

The dilemmas were presented one by one and in random order. The whole set of moral dilemmas used in this study, along with the experimental instructions, is provided in the appendix. Note that the integer  $-2$  signifies the strongest disagreement of the sacrifice. The integer  $2$  signifies the strongest agreement. We refer to the integers as agreement points.

### 3.2 Experimental Game

Two players, A and B, play a simple game. Player A's payoff is 10 Euros. Player A chooses player B's payoff by picking an integer between 0 and 10 Euros. Player B makes no decision. In our sessions, pairs of participants were formed. Each participant made a decision as a Player A. A random draw at the end of each session determined participants' actual roles and actual payoffs. In addition, each participant received a show-up fee of 5 Euros.<sup>10</sup>

## 4 Results

This section outlines our main findings. Our experimental evidence supports Hypotheses 1 and 2 (see Findings 1-2), but not Hypotheses 3 and 4 (see Findings 3-5). Let us first turn to participants'

---

<sup>9</sup>The references for the dilemmas are: Submarine (Greene et al. 2001), Trespassers (Greene et al. 2001), Hostages (Greene et al. 2001), Bystander (Foot 1967), Life Raft (Regan 1983), Plane Crash (Marshall 1993, Greene et al. 2001), Prisoners of War (Baron 1992), Fumes (Thompson 1986), Spelunkers ([http://news.bbc.co.uk/2/hi/uk\\_news/magazine/4954856.stm](http://news.bbc.co.uk/2/hi/uk_news/magazine/4954856.stm)), Surgery (Foot 1967), Derailment (Unger 1996), Footbridge (Thompson 1985), and Baby (Alda et al. 1983, Greene et al. 2001).

<sup>10</sup>We used this procedure to collect data on each participant's game decision. Note that our procedure makes it plain that each participant's choice could affect another participant's payoff. Our procedure asks participants to make contingent choices before learning about their actual role in the game (Selten, 1967). An alternative procedure could have first let each participant know her role in the game ex ante, and then ask the participant to make choices at her actual information set. The literature shows that both methods yield similar results with simple distributional games that involve very few contingent choices (for the dictator game see Cason and Mui (1998) and for a solidarity game see Büchner et al. (2007). For a survey of further experimental evidence on the two methods, see Charness and Brandts (2011).

Allocated Amount (in Euros)	0	1	2	3	4	5	6	7	8	9	10	$\Sigma$
Number of Choices by Females	0	1	0	1	1	7	2	1	8	4	74	99
Number of Choices by Males	0	1	0	1	1	11	0	1	1	2	81	99
Total Number of Choices	0	2	0	2	2	18	2	2	9	6	155	198

Table 1: Distribution of Transfers in the Experimental Game

behavior in our incentivized game. Player A is asked to choose player B’s payoff. We call this choice the allocated amount. Table 1 shows the distribution of transfers; by female participants, by male participants, and by the pooled sample.

**Finding 1 (Frequency of Spiteful Choices)**

**Twenty-two percent of the participants chose a transfer less than the maximal transfer.**

**Support.** Table 1, row 4 shows the overall frequency of transfers. Most participants (155) selected the maximal transfer, 10. A sizable fraction of the participants allocated lower amounts; the most frequent among those were 5 (18 participants), 8 (9 participants), and 9 (6 participants).

**Finding 2 (No Gender Difference in allocated Amounts)**

**We do not reject the null hypothesis that the distribution of transfers is equal for male participants and female participants.**

**Support.** Table 1, row 2 shows the frequency of transfers by female participants. Table 1, row 3 shows the frequency for male participants. A Wilcoxon rank-sum test of no difference of the distributions shows  $p = 0.347$ .

Next, we turn to moral judgment. Table 2 shows the score for each of the 14 items of the moral-dilemma questionnaire. The panel to the left shows the statistics sorted by gender, the panel in the middle shows the statistics sorted by the allocated amount, and the panel to the right shows the statistics for pooled sample.

Note that the moral situations are different in various dimensions. All those situations have something in common, and our questionnaire is designed to measure that common factor. The internal consistency of the questionnaire is very high: Cronbach’s  $\alpha$  is around 0.8 for the 14 situations (see the final row of Table 2)<sup>11</sup>. The high correlations between the response to different situations show that the questionnaire is well suited for measuring a commonality – the degree of endorsement of the utilitarian solution. Note that the commonality does allow for a reasonable discussion of moral principles – utilitarianism versus ontology. Furthermore, after partitioning the sample by gender and allocated amount, the high internal consistency of the questionnaire is upheld. In table 2 (final row), we now report Cronbach’s  $\alpha$  for male ( $N = 99$ ) and female participants ( $N = 99$ ), and for other-damaging participants ( $N = 43$ ) and participants who are not other-damaging ( $N = 155$ ). These results suggest that moral judgment is robust across fourteen situations in the four subgroups.

**Finding 3 (Gender Difference in Moral Judgment)**

**We reject the null hypothesis that the distribution of moral judgments is equal for male and female participants.**

<sup>11</sup>Tavakol and Dennick (2011) describe internal consistency as the extent to which all the items in a test measure the same concept. Internal consistency is connected to the inter-relatedness of the items within the test (p. 53). Commonly, values of at least 0.7 are considered as acceptable (p. 54).

Type of Dilemma Observations	Gender			Allocated Amount			Pooled
	Female 99	Male 99	$p$	Amount=10 155	Amount<10 43	$p$	
<i>Submarine</i>	-0.071	0.535	0.004	0.271	0.093	0.491	0.232
<i>Trespassers</i>	-1.222	-0.707	0.009	-0.981	-0.907	0.760	-0.965
<i>Hostages</i>	-0.606	-0.030	0.012	-0.342	-0.233	0.697	-0.318
<i>Bystander</i>	0.394	0.828	0.036	0.535	0.884	0.168	0.611
<i>Life Raft</i>	0.222	1.010	0.001	0.568	0.791	0.385	0.616
<i>Plane Crash</i>	-1.586	-1.404	0.198	-1.548	-1.302	0.150	-1.495
<i>Prisoners of War</i>	-0.010	0.202	0.350	0.077	0.163	0.757	0.096
<i>Fumes</i>	0.121	0.636	0.010	0.316	0.605	0.275	0.379
<i>Spelunkers</i>	-0.596	0.253	0.001	-0.213	-0.023	0.471	-0.172
<i>Soldiers</i>	0.222	0.485	0.243	0.316	0.488	0.528	0.354
<i>Surgery</i>	-1.818	-1.455	0.003	-1.697	-1.419	0.062	-1.636
<i>Derailment</i>	-0.222	0.061	0.188	-0.058	-0.163	0.688	-0.081
<i>Footbridge</i>	-1.606	-1.576	0.818	-1.594	-1.581	0.939	-1.591
<i>Baby</i>	-0.869	-0.172	0.001	-0.548	-0.419	0.619	-0.520
Overall	-0.546	-0.095	0.001	-0.350	-0.216	0.315	-0.321
Internal Consistency							
Cronbach's $\alpha$	0.765	0.829	–	0.824	0.791	–	0.818

Table 2: Points of agreement for sacrifice in the moral-judgement questionnaire. Range of agreement points:  $-2$  (strongest disagreement),  $-1$ ,  $1$ ,  $2$  (strongest agreement).  $p$  values from two-sided  $t$ -tests.

**Support.** Table 2, column 4 shows the  $p$  values of the test of no difference in the distribution of agreement points of female participants and male participants for each of the 14 situations. Male participants' declared support for the sacrifice differs from that of the female participants. In 9 out of 14 situations, the difference is significant at the 5% level with a  $t$ -test. Table 2, third to final row shows the  $p$  values of the test of no difference in the distribution of overall scores of female and male participants. Male participants' average score is not equal to female participants' average score. The two-sided  $t$ -test shows a highly significant difference ( $p < 0.001$ ).

#### **Finding 4 (No Difference in Moral Judgment Between Spiteful and Non-Spiteful Players)**

**We do not reject the null hypothesis that the distribution of moral judgments is equal for spiteful and non-spiteful players.**

**Support.** Table 2, columns 5 and 6 show the scores of participants who act prosocially as Player A in the experimental game (i.e., the allocated amount equals 10), and those who act antisocially (by transferring less than 10). Among 14 items, we find no instance for which the difference in moral judgments between these two groups is significant at the 5% level, and one instance in which it is weakly significant at the 10% level. The  $p$  values from two-sided  $t$ -tests are shown in column 7. We conclude that the overall average scores are not statistically different at the conventional levels.

#### **Finding 5 (Gender Differences in Moral Judgments)**

**We find a gender difference in moral judgments even when controlling for the allocated amount.**

**Support.** Table 3 reorganizes the data presented in Table 2. It summarizes the odds ratios from an ordered logistic regression of the score on moral dilemma question on a set of explanatory

variables used in Table 2. Those predictors include two main indicator variables –  $1[Male]$  (set to 1 for male participants, and to 0 for female participants) and  $1[Amount < 10]$  (set to 1 if the allocated amount is less than 10, and to 0 otherwise), as well an indicator variable for each moral dilemma question (13 variables, *Submarine* is omitted as reference) to control for the contextual fixed effects.<sup>12</sup> Since each participant responds to a series of 14 moral dilemma questions, standard errors are clustered at the individual level.<sup>13</sup> This analysis points to the robustness of gender difference in moral judgments – the odds ratio associated to  $1[Male]$  is greater than 1 and statistically significant ( $p < 0.001$ ). On the other hand, antisocial behavior observed in the experimental game is not a statistically significant predictor of the endorsement of utilitarianism in moral judgments ( $p = 0.136$ ).<sup>14</sup>

## 5 Conclusions

We conduct a questionnaire that allows us to measure consistently moral judgment in fourteen well-studied one-life-for-five-lives situations in which inflicting harm on one person allows many others to escape suffering. Echoing previous empirical findings, we report systematic gender differences in moral judgments: male participants are more supportive of the sacrifice than female participants. We also investigate a simple allocation game in which an advantaged player simply decides about the payoff of the disadvantaged player, while her own payment is held constant. In line with the theoretical conjecture by Levine (1998), we find that a fifth of the participants transfer less than the maximum, and that male and female players behave spitefully in equal proportion. Finally, combining the data on moral judgment with the data from the experimental game points to the robustness of the gender difference in moral judgments to the individual heterogeneity in other-regarding preferences.

The first important point to note is that the results reported herein differ at the face value from the previous psychological studies examining the association between utilitarian judgments and antisocial motives. At the same time, it should be also stressed that our methodology differs from theirs. Our approach relies on measuring antisocial preferences through an incentivized, context-free economic game aimed at eliciting the willingness to engage into spiteful behavior towards others. The previous studies on the topic, in turn, employ the standard methodology of psychological research – surveys with multiple hypothetical, context-loaded items aimed at identifying a range of antisocial traits (such as psychopathy, Machiavellism, life meaninglessness, low empathy, low aversion for harming others, endorsement for unethical business practices, low declared donation to a hypothetical charity). In the light of recent developments in psychometrics resulting in new scales for measuring spitefulness (Marcus et al., 2014) and its association with moral values such as (lack of care for) avoiding harm and being fair to others (Zeigler-Hill et al., 2015), we believe that future research should combine hypothetical and incentivized measures to shed light on the link between antisocial psychological traits and proneness to antisocial behavior, and further our understanding of the association between the endorsement for utilitarianism and other-regarding motives.

We also reckon that numerous policies – such as the distribution of wealth or punishment – can be viewed as moral dilemmas with the key feature that the decision maker has to weigh the present

---

<sup>12</sup>Those contextual fixed effects turn out to be important controls in the model: their joint insignificance is strongly rejected with  $p < 0.001$ .

<sup>13</sup>As noted by Vossler (2012), this technique allows for valid inference in a regression model for any form of unobserved within-unit serial correlation.

<sup>14</sup>We emphasize that the interpretation of this result should be purely correlational. Like some studies before, we are interested in whether and how antisocial preferences can predict moral judgments. We than an anonymous referee for pointing this out.



Predictor	Odds ratio / (SE)	<i>p</i>
1[ <i>Male</i> ]	1.931 (0.284)	< 0.000
1[ <i>Amount</i> < 10]	1.279 (0.211)	0.136
Dilemma indicator variables:		
<i>Trespassers</i>	0.220 (0.036)	< 0.000
<i>Hostages</i>	0.523 (0.075)	< 0.000
<i>Bystanders</i>	1.602 (0.236)	0.001
<i>Life Raft</i>	1.660 (0.218)	< 0.000
<i>Plane Crash</i>	0.097 (0.017)	< 0.000
<i>Prisoners of War</i>	0.879 (0.118)	0.338
<i>Fumes</i>	1.198 (0.181)	0.234
<i>Spelunkers</i>	0.636 (0.081)	< 0.000
<i>Soldiers</i>	1.172 (0.174)	0.284
<i>Surgery</i>	0.063 (0.012)	< 0.000
<i>Derailment</i>	0.693 (0.108)	0.018
<i>Footbridge</i>	0.069 (0.013)	< 0.000
<i>Baby</i>	0.410 (0.063)	< 0.000
Cutoff 1	-1.039 (0.146)	
Cutoff 2	0.072 (0.144)	
Cutoff 3	1.376 (0.152)	

Table 3: Odds ratios from an ordered logistic regression model. Dependant variable: individual score in a moral dilemma question.  $N = 2772$ . To account for multiple (14) responses per individual, standard errors are clustered at the individual level (198 clusters).

harm to one person with the future benefits to other persons. Our analysis suggests that males could be better enforcers of societal goals than females: men are more willing to harm one person. Furthermore, men show willingness to harm another person not because males are more spiteful than females and like to harm another person, but because males are willing to assess and weigh the benefits of different outcomes. So far, this issue lacks systematic investigation in experimental economic (Croson and Gneezy, 2009).

In our experiment, we elicit participants' moral judgment of an action, not the actual moral action. We emphasize that both aspects – judgment of an action and the action itself – are important. Consider, for instance, a question of judgment: ‘Which life-or-death decisions does our society want to make?’ On the contrary, consider now the question of action: ‘Whom do we want to make those decisions?’ For instance, society may have to decide on *who* serves in the security forces that tackle terrorist incidences, and *in which way* do we want our forces to tackle those threats to our lives. Altogether, we believe that more evidence is needed to understand the relationship between moral judgment of an action and actual moral behavior. Our paper is a step in this direction.

## References

- [1] Abbink, K., and Herrmann, B., 2011, The moral costs of nastiness, *Economic Inquiry*, 49(2), 631–633.
- [2] Abbink, K., Masclet, D., and Mirza, D., 2018, Inequality and inter-group conflicts: experimental evidence, *Social Choice and Welfare*, 50(3), 387–423.
- [3] Abbink, K., and Sadrieh, A., 2009, The pleasure of being nasty, *Economics Letters*, 105(3), 306–308.
- [4] Adleberg, T., Thompson, M., and Nahmias, E., 2015, Do Men and Women have Different Philosophical Intuitions? Further Data, *Philosophical Psychology*, 28, 615–641.
- [5] Alda, A. et al. 1983, Goodbye, Farewell, and Amen (Television Series Episode). In B. Metcalfe (producer), M\*A\*S\*H, Los Angeles, CA: 20th Century Fox Television.
- [6] Alexander, J. and Weinberg, J., 2007, Analytic Epistemology and Experimental Philosophy, *Philosophy Compass*, 2, 56–80.
- [7] Banerjee, K., Huebner, B., and Hauser, M.D., 2010, Intuitive Moral Judgments are Robust Across Demographic Variation in Gender, Education, Politics, and Religion: A Large-scale Web-based Study, *Journal of Cognition and Culture*, 10, 253–281.
- [8] Baron, J., 1992, The Effect of Normative Beliefs on Anticipated Emotions, *Journal of Personality And Social Psychology*, 63, 320–330.
- [9] Bartels, D. M., 2008, Principled Moral Sentiment and the Flexibility of Moral Judgment and Decision Making, *Cognition*, 108, 381–417.
- [10] Bartels, D. M. and Pizarro, D. A., 2011, The Mismeasure of Morals: Antisocial Personality Traits Predict Utilitarian Responses to Moral Dilemmas, *Cognition*, 121, 154–161.
- [11] Bartling, B., Fehr, E., Marechal, M.A., and Schunk, D., 2009, Egalitarianism and Competitiveness, *American Economic Review*, 99, 93–98.

- [12] Brandts, J. and Charness, G., 2011, The Strategy Versus the Direct-Response Method: A First Survey of Experimental Comparisons, *Experimental Economics*, 14, 375–398.
- [13] Bourget, D. and Chalmers, D.J., 2014, What Do Philosophers Believe?, *Philosophical Studies*, 170, 465–500.
- [14] Büchner, S., Coricelli, G., and Greiner, B., 2007, Self-centered and Other-regarding Behavior in the Solidarity Game, *Journal of Economic Behavior and Organization*, 62, 293–303.
- [15] Buckwalter, W. and Stich, S., 2014, Gender and Philosophical Intuition, In J. Knobe & S. Nichols (eds.), *Experimental Philosophy*, Vol.2., Oxford: Oxford University Press, 307–346.
- [16] Cason, T. and Mui, V.-L., 1998, Social Influence in the Sequential Dictator Game, *Journal of Mathematical Psychology*, 42, 248–265.
- [17] Charness, G. and Grosskopf, B., 2001, Relative Payoffs and Happiness: An Experimental Study, *Journal of Economic Behavior & Organization*, 45, 301–328.
- [18] Charness, G., Masclet, D., and Villeval, M.C., 2011, The Dark Side of Competition for Status, *Management Science*, 60, 38–55.
- [19] Choe, S. Y., and Min, K. H., 2011, Who makes utilitarian judgments? The influences of emotions on utilitarian judgments, *Judgment and Decision Making*, 6(7), 580–592.
- [20] Crockett, M. J., Clark, L., Hauser, M. D., and Robbins, T. W., 2010, Serotonin selectively influences moral judgment and behavior through effects on harm aversion, *Proceedings of the National Academy of Sciences*, 107(40), 17433–17438.
- [21] Croson, R. and Gneezy, U., 2009, Gender Differences in Preferences, *Journal of Economics Literature*, 47, 448–474.
- [22] Cushman, F., Gray, K., Gaffey, A., and Mendes, W. B., 2012, Simulating murder: The aversion to harmful action, *Emotion*, 12(1), 2–7.
- [23] Engel, C., 2011, Dictator Games: A Meta Study, *Experimental Economics*, 14, 583–610.
- [24] Fisman, R., Kariv, S., and Markovits, D., 2007, Individual Preferences for Giving, *American Economic Review*, 97, 1858–1876.
- [25] Foot, D., 1967, The Problem of Abortion and the Doctrine of the Double Effect, *Oxford Review*, 5, 5–15.
- [26] Gleichgerrcht, E., and Young, L., 2013, Low levels of empathic concern predict utilitarian moral judgment, *PLoS One*, 8(4), e60418.
- [27] Glenn, A.L., Spassena, K., Iyer, R., Graham, J., and Ditto, P.H., 2010, Moral Identity in Psychopathy, *Judgment and Decision Making*, 5, 497–505.
- [28] Gold, N., Colman, A., and Pulford, B., 2014, Cultural Differences in Responses to Real-life and Hypothetical Trolley Problems, *Judgment and Decision Making*, 9, 65–76.
- [29] Gold, N., Colman, A., and Pulford, B., 2015, Do as I Say, Don't Do as I Do: Differences in Moral Judgments Do not Translate Into Differences in Decisions in Real-life Trolley Problems, *Journal of Economic Psychology*, 47, 50–61.

- [30] Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., and Cohen, J. D., 2001, An fMRI Investigation of Emotional Engagement in Moral Judgement, *Science*, 293, 2105–2108.
- [31] Greiner, B., 2004, An Online Recruitment System for Economic Experiments, University of Cologne, Working Paper Series in Economics No. 10, 79–93.
- [32] Haidt, J., 2001, The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment, *Psychological Review*, 108.4, 814–834.
- [33] Haidt, J., 2003, The Moral Emotions. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of Affective Sciences*. Oxford: Oxford University Press, 852–870.
- [34] Harsanyi, J.C., 1976, *Essays on Ethics, Social behavior, and Scientific Explanation*, Volume 12, Springer Science & Business Media.
- [35] Hauser, M., Cushman, F., Young, L., Kang-Xing Jin, R., and Mikhail, J., 2007. A Dissociation between Moral Judgments and Justifications, *Mind & Language*, 22, 1–21.
- [36] Hsu, M., Anen, C., and Quartz, S., 2008, The Right and The Good: Distributive Justice and Neural Encoding of Equity and Efficiency, *Science*, Volume 320, 1092–1095.
- [37] Isaac, R.M. and Walker, J.M., 1988, Group Size Effects in Public Goods Provision: The Voluntary Contribution Mechanism, *The Quarterly Journal of Economics*, 103, 179–200.
- [38] Kahane, G., Everett, J. A., Earp, B. D., Farias, M., and Savulescu, J., 2015, ‘Utilitarian’ judgments in sacrificial moral dilemmas do not reflect impartial concern for the greater good, *Cognition*, 134, 193–209.
- [39] Kerschbamer, R., 2015, The Geometry of Distributional Preferences and a Non-Parametric Identification Approach: The Equality Equivalence Test, *European Economic Review*, 76, 85–103.
- [40] Kimbrough, E. O., and Reiss, J. P., 2012, Measuring the distribution of spitefulness, *PLoS One*, 7(8), e41812.
- [41] Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., and Damasio, A., 2007, Damage to The Prefrontal Cortex Increases Utilitarian Moral Judgments, *Nature*, 446, 908–911.
- [42] Levine, D. K., 1998, Modeling Altruism and Spitefulness in Experiments, *Review of Economic Dynamics*, 1, 593–622.
- [43] Lyubomirsky, S. and Ross, L., Hedonic Consequences of Social Comparison: A Contrast of Happy and Unhappy People, *Journal of Personality and Social Psychology*, 71, 1141–1157.
- [44] Marcus, D. K., Zeigler-hill, V., Mercer, S. H., and Norris, A. L., 2014, The psychology of spite and the measurement of spitefulness, *Psychological assessment*, 26(2), 563–574.
- [45] Marshal, F. (director), *Alive* (motion picture), Los Angeles, CA, United States, Paramount Pictures.
- [46] McKelvey, R. and Palfrey, T., 1992, An Experimental Study of the Centipede Game, *Econometrica*, 60, 803–836.

- [47] Murphy, R. O., K. A. Ackermann, and Handgraaf, M.J.J., 2011, Measuring Social Value Orientation, *Judgment and Decision Making*, 6, 771–781.
- [48] Nishimura, N., Cason, T. N., Saijo, T., and Ikeda, Y., 2011, Spite and reciprocity in auctions, *Games*, 2(3), 365–411.
- [49] Petrinovich, L. and O’Neill, P., 1996, Influence of Wording and Framing Effects on Moral Intuitions, *Ethology and Sociobiology*, 17, 145–171.
- [50] Petrinovich, L., O’Neill, P., and Jorgensen, M.J., 1993, An Empirical Study of Moral Intuitions: Toward an Evolutionary Ethics, *Journal of Personality and Social Psychology*, 64, 467–478.
- [51] Regan, T., 1983, *The Case of Animal Rights*, Berkeley, CA: University of California Press.
- [52] Roth, A.E., Prasnikar, V., Okuno-Fujiwara, and Zamir, S., 1991, Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An Experimental Study, *American Economic Review*, 81, 1068–1095.
- [53] Saver, J. and 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.
- [54] Selten, R., 1967, Die Strategiemethode zur Erforschung des Eingeschränkt Rationalen Verhaltens im Rahmen eines Oligopolexperiments, In H. Sauermann (Ed.), *Beiträge zur experimentellen Wirtschaftsforschung*, Tübingen: Mohr, 136–168.
- [55] Seyedsayamdost, H., 2015, On Gender and Philosophical Intuition: Failure of Replication and Other Negative Results, *Philosophical Psychology*, 28, 642–673.
- [56] Tavakol M. and Dennick R., 2011, Making sense of Cronbach’s alpha, *International Journal of Medical Education*, 53-55.
- [57] Thomson, J. J., 1985, The Trolley Problem, *Yale Law Journal*, 94, 1395–1415.
- [58] Thomson, J. J., 1986, *Rights, Restitution, and Risks*, Cambridge, MA, Harvard University Press.
- [59] Unger, P., 1996, *Living High and Letting Die: Our Illusion of Innocence*, New York: Oxford University Press.
- [60] Vossler, C. A., 2013, Analyzing repeated-game economics experiments: robust standard errors for panel data with serial correlation, In John A. List and Michael K. Price (Ed.) *Handbook on Experimental Economics and the Environment*, Northampton, MA, Edward Elgar Publishing Limited, 89–112.
- [61] Zamzow, J. L. and Nichols, S., 2009, Variations in Ethical Intuitions, In E. Sosa and E. Villanueva (eds.), *Metaethics*, Wiley Periodicals Inc., 368–388.
- [62] Zeiliger, R., 2000, A Presentation of Regate, Internet-based Software for Experimental Economics, <http://regate-ng.gate.cnrs.fr/sferriol/r>
- [63] Zeigler-Hill, V., Noser, A. E., Roof, C., Vonk, J., and Marcus, D. K., 2015, Spitefulness and moral values, *Personality and Individual Differences*, 77, 86–90.

- [64] Zizzo, D. J., 2004, Inequality and procedural fairness in a money-burning and stealing experiment, In Frank Cowell (Ed.), *Inequality, Welfare and Income Distribution: Experimental Approaches*, Emerald Group Publishing Limited, 215–247.
- [65] Zizzo, D.J. and Oswald, A., 2001, Are People Willing to Pay to Reduce Others' Incomes?, *Annales d'Economie et de Statistique*, 39–65.

## A Appendix: Design and Instructions

In each session, participants were asked to answer a **standard questionnaire**. We used the questionnaire to elicit demographic information. In addition, participants were required both to play an **experimental game** and to answer a **moral judgment questionnaire**. Below we present the paper instructions (subsection 1), describe the instructions displayed on the participants' computer screens (subsection 2), and present the moral judgment questionnaire (subsection 3). Sentences in brackets were not part of the instructions but rather descriptions of what happened.

### A.1 Instructions On Paper

You are taking part in an experiment in which you can earn money. Your gains may depend on the decision made by another participant. Before we begin we would like you to answer a few standard questions concerning your age, education, profession, etc.. These questions will help us to get to learn something about your characteristics. Your identity and your monetary gains will remain confidential and anonymous.

[Participants filled out the standard questionnaire.]

Thank you for answering the questions.

#### **What happens in the Session**

The experiment consists of two separate parts. In **Part 1** your payment in the session will be determined. In **Part 2** you will be asked to answer questions that will allow us to learn more about you. Further instructions will be displayed on your screen before the beginning of each part.

#### **Payment of your earnings**

Your total payment will be the payoff you earn in Part 1 and a bonus of 5 Euros for completing the session. Payments are made individually and in cash.

**You are not allowed to talk during the experiment.** Participants who violate this rule will be excluded from the experiment and payments. It is important that you perfectly understand the rules of this experiment. Should you have any questions, please raise your hand.

**Thank you again for your participation.**

### A.2 Computer Screen

#### **[Experimental Game: Information on the First Screen]**

In this part of the experiment, your additional payment will be determined. The game has two players: player A and player B. **Only player A makes a decision, and this decision affects player B's payoff only.** You will be paired randomly with another participant. You will then be asked about the decision you would make as player A. Finally, **at the end of the experiment, a random draw will determine your actual role** in the pair. If you are player A, your decision will determine the other group member's payoff. If you are player B, your payoff will be determined by player A.

### [Experimental Game – Information on the Second Screen]

The game is played by two persons, player A and player B, and runs as follows. Player A's payoff is fixed and equals 10 Euros. Player A decides about player B's payoff. More precisely, player A selects an amount between 0 and 10 Euros which is then attributed to player B. Player A's choice does not affect her own payoff. Player B makes no decision and her payoff corresponds to the amount determined by player A.

Suppose that you are player A. What amount would you like to give to player B?

[Participants entered amount using the computer's keyboard.]

### A.3 Moral Judgment Questionnaire

In this part of the session, we ask you to answer a few questions. The answers will enable us to learn about some of your personal characteristics. We remind you that we are not able to link your answers with your identity.

[Participants used a computer mouse to tick the appropriate box on a scale of -2, -1, 1, 2. The leftmost box represents the strongest disagreement and the rightmost box corresponds to the highest agreement.]

#### 1. Submarine

You are the captain of a small military submarine traveling underneath a large iceberg. An on-board 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 crew member, there will be just enough oxygen for everyone else to survive.

In this situation, would you shoot the crew member?

#### 2. Trespassers

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?

#### 3. Hostages

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 will be free to go.

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



#### **4. Bystander**

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?

#### **5. Life Raft**

You are on a small ship, a fire breaks out, and the ship has to be abandoned. Because your tiny life raft 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 on board who may die either way. If you throw him overboard, everyone else will be saved.

In this situation, would you throw him overboard?

#### **6. Plane Crash**

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?

#### **7. Prisoners of War**

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?

#### **8. Fumes**

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?

#### **9. Spelunkers**

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?

### **10. Soldiers**

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?

### **11. Surgery**

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?

### **12. Derailment**

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?

### **13. Footbridge**

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 next to a larger stranger on a pedestrian walkway that arches over the tracks. 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?

### **14. Baby**

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?

[Participants were thanked for their participation and paid in private.]