



Munich Personal RePEc Archive

## **Utilitarian Moral Judgments Are Cognitively Too Demanding**

Da Silva, Sergio and Matsushita, Raul and De Sousa, Maicon

2016

Online at <https://mpra.ub.uni-muenchen.de/69387/>

MPRA Paper No. 69387, posted 10 Feb 2016 06:02 UTC

# Utilitarian moral judgments are cognitively too demanding

Sergio Da Silva<sup>a\*</sup>, Raul Matsushita<sup>b</sup>, Maicon De Sousa<sup>a</sup>

<sup>a</sup> *Department of Economics, Federal University of Santa Catarina, Florianopolis, S.C., 88049-970, Brazil*

<sup>b</sup> *Department of Statistics, University of Brasilia, Brasilia, D.F., 70910-900, Brazil*

\* Corresponding author. Tel.: +55 48 3721 9901; fax: +55 48 3721 9901. Email address: professorsergiodasilva@gmail.com (S. Da Silva).

## Abstract

We evaluate utilitarian judgments under the dual-system approach of the mind. In the study, participants respond to a cognitive reflection test and five (sacrificial and greater good) dilemmas that pit utilitarian and non-utilitarian options against each other. There is judgment reversal across the dilemmas, a result that casts doubt in considering utilitarianism as a stable, ethical standard to evaluate the quality of moral judgments. In all the dilemmas, participants find the utilitarian judgment too demanding in terms of cognitive currency because it requires non-automatic, deliberative thinking. In turn, their moral intuitions related to the automatic mind are frame dependent, and thus can be either utilitarian or non-utilitarian. This suggests automatic moral judgments are about descriptions, not about substance.

**Subject areas:** cognitive psychology, moral philosophy

**Keywords:** cognitive reflection, utilitarianism, moral judgment

**How to cite this paper:** Da Silva, S., Matsushita, R. and De Sousa, M. (2016) Utilitarian Moral Judgments Are Cognitively Too Demanding. *Open Access Library Journal*, 3: e2380. <http://dx.doi.org/10.4236/oalib.1102380>

## **1. Introduction**

The ethical theory of utilitarianism [1, 2] is summarized in Bentham's aphorism that "the greatest happiness of the greatest number is the basis of morals and legislation." However simple it may appear to maximize the welfare of the group, this proposition allows for more than one interpretation. These will be discussed later. The major alternatives to utilitarian ethics are the Kantian deontological and the Aristotelian virtue-based ones. However, many favor the utilitarian perspective as the ethical yardstick against which the quality of people's moral judgment should be evaluated [3-5].

Current moral psychology is inclined to the type of interpretation called act utilitarianism [6-8]. However, there is also a rule utilitarianism [9]. Killing one person to save five may be justified under the act utilitarianism, but not under the rule utilitarianism. Act utilitarianism thus assesses utilitarian judgments through "sacrificial moral dilemmas," such as that of a runaway trolley that hurtles toward five unaware workers. The only way to save them is to push a heavy man, standing nearby on a footbridge, onto the track, where he will die stopping the trolley [10]. Facing such a dilemma, most people make the non-utilitarian choice. Worse, those who adopt the utilitarian judgment paradoxically exhibit antisocial characteristics that are at odds with the very social content of the utilitarian ethics [11]. For this reason, that way of representing utilitarianism has been criticized and "greater good" dilemmas have been proposed instead as a "genuine" utilitarian approach [12].

Automatic and deliberative cognitive processes should mediate moral judgments [13]. Most cognitive psychologists currently favor a dual-system approach [14] in which two systems compete for control of our inferences, actions and judgments. Intuitive decisions require little reflection and use "System 1." By scanning associative memory, we usually base our actions on experiences that worked well in the past. However, whenever we make decisions by constructing mental models or simulations of future possibilities, we engage "System 2." Individuals differ in their cognitive abilities, but a simple test can assess their differences. The cognitive reflection test (CRT) [15] gauges the relative powers of Systems 1 and 2: the individual ability to think quickly with little conscious deliberation (using System 1) and to think in a slower and more reflective way (using System 2). High scores on the CRT track how strong System 2 is for one individual.

Here, we evaluate utilitarian judgments under the dual-system approach in a study where participants respond to a CRT and to sacrificial and greater good dilemmas that pit utilitarian and non-utilitarian options against each other.

The next section describes the experiment details. Section 3 presents results and discusses them, and Section 4 concludes this report.

## **2. Materials and methods**

The experimenter (MDS) sent Google Docs questionnaires to students from four large Brazilian universities: University of Brasilia, Federal University of Santa Catarina, Federal University of Rio Grande do Sul and Federal Institute of Minas Gerais. The students were enrolled in several disciplines, ranging from physics, chemistry, civil engineering, computer science, statistics, business administration, accounting, economics, physiotherapy, nutrition, design and pedagogy. The questionnaires were administered to the students by their own teachers using university intranet pages. The questionnaires were also sent to participants not related to the

universities. A total of 468 individuals participated (235 males, 233 females; age: mean = 24.21, standard deviation = 4.81). The participants were 393 university students and 75 young adults unaffiliated with the universities. All participants were instructed to respond to the CRT in Table 1 and to the utilitarian moral dilemmas displayed in Table 2. Two types of questionnaire were presented. In the type 1 questionnaires (204 respondents), the CRT came first, whereas in the type 2 questionnaires (264 respondents), the moral dilemmas came first.

The three simple questions of the CRT were originally designed to elicit automatic responses that are wrong [15]. The questionnaires instructed the participants to respond to the three questions in Table 1 in less than 30 seconds; this was to ensure an automatic choice. The correct answers in Table 1 are 5, 5 and 47. However, the most common, intuitive answers are 10, 100 and 24. The intuitive answer that springs quickly to mind in question 1 is “10 cents.” This is wrong because if  $x$  is the ball cost, then  $(1.00 + x) + x = 1.10 \rightarrow x = 0.05$ . The automatic answer to question 2 is 100. To see why this is wrong, take the first sentence and multiply 5 machines by 20; thus, it would take the same 5 minutes to make 100 widgets. The intuitive answer in question 3 is 24, but if exponential  $d^n$  represents the function where  $n$  is the number of days it takes the patch to cover the entire lake, thus, half is always  $n - 1$ . Then, the patch will cover half the lake in  $48 - 1 = 47$  days. The three questions on the CRT are easy because their solutions are easily understood when explained [15]. However, reaching one correct answer requires the suppression of an erroneous answer that impulsively comes to mind.

Table 1. Cognitive reflection test

<i>Question</i>	<i>Right answer</i>	<i>Automatic, wrong answer</i>
1 A bat and a ball cost \$1.10 in total. The bat costs \$1.00 more than the ball. How much does the ball cost?	5 cents	10 cents
2 If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?	5 minutes	100 minutes
3 In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take the patch to cover half the lake?	47 days	24 days

Source: Ref. [15]

At the end of the questionnaires, participants were asked whether they had really spent less than 30 seconds to answer to the CRT. There was also a question inquiring as to whether they already knew any of the three questions. We removed from the sample those who declared to have taken more than 30 seconds or to have known at least one of the questions. As a result, the sample was reduced to 398 participants (type 1 questionnaire:  $N = 173$ ; type 2 questionnaire:  $N = 225$ ).

The utilitarian moral dilemmas in Table 2 were selected to purposely contrast circumstantial or parochial dilemmas with humanity dilemmas. Personal dilemmas are associated with egocentric attitudes and no identification with the whole of humanity [12]. “Greater good” dilemmas consider the latter.

The benchmark sacrificial dilemma is that of the trolley (Table 2). We consider both the personal [10] and impersonal [16] versions because this distinction matters [17]. A further justification for administering two types of questionnaire is precisely to consider this difference. Type 1 questionnaires involve the personal sacrificial dilemma [10], and type 2 questionnaires spot the impersonal sacrificial dilemma [16]. In the personal version, the trolley hurtles toward five unaware workers and the only way to save them is to push a heavy man, standing nearby on a

footbridge, onto the track. In the impersonal version, throwing a switch diverts the trolley onto another track, killing one worker. In terms of simple utilitarian arithmetic, this sacrificial dilemma is still utilitarian (but see [12]). After all, one should consider the “greater good” of killing just one man rather than killing five. Although the logic of choice in both personal and impersonal versions is the same, its architecture changes. And this fact usually causes preference reversals on this moral judgment. Indeed, most people choose the utilitarian action in the impersonal version and the non-utilitarian one in the personal version [18]. The next section will show this finding is replicated in our own questionnaires. Moreover, we will provide an explanation for such moral preference reversals based on the model of two minds.

Table 2. Utilitarian moral dilemmas

<i>Dilemma</i>	<i>Type</i>	<i>Description</i>
D1 (type 1 questionnaire)	Personal sacrificial dilemma (other- benefit dilemma) [10]	<i>Trolley problem: Footbridge.</i> In the path of a runaway train car are five railway workers 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. <i>In this situation, would you push the man?</i>
D1 (type 2 questionnaire)	Impersonal sacrificial dilemma (other-benefit dilemma) [16]	<i>Trolley problem: Bystander.</i> In the path of a runaway train car are five railway workers 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 worker. <i>In this situation, would you flip the switch?</i>
D2	Personal self-benefit dilemma [17]	<i>Preventing Ebola.</i> You are a Peace Corps health-worker who is volunteering in a rural African village. A prominent man from a nearby village has contracted the Ebola virus that is extremely contagious, incurable, and almost always fatal within a week. Miraculously, this man has survived for a month, and so he must be a rare carrier who is immune to the virus’ deadly effects. However, this man wrongly believes that your health center can cure his disease. You see him approaching and you know that if he enters the village he will spread the virus to hundreds of innocent people who, unlike him, will die. There is a loaded gun in the health center. You realize that the only way you can prevent him from entering the village and spreading the virus to you and the rest of the village is to shoot and kill him as he approaches. <i>Should you kill the man in order to save yourself and the rest of the village?</i>
D3	Greater good dilemma [12]	<i>Firefighter.</i> Albert is a firefighter who is trying to rescue people from a burning building. The building is about to collapse and, in the time left, Albert will only be able to rescue one more person. In the room he has entered, Albert finds two trapped people that he immediately recognizes. One is a famous peace negotiator well known for his work resolving armed conflicts around the world and who is likely to continue doing this important work if he survives. The other is a poor, uneducated housekeeper. The housekeeper is Albert’s mother. Albert has to decide which of these to save, and the one he does not save will die. <i>Should Alberto save the peace negotiator rather than his mother?</i>
D4	Greater good dilemma [12]	<i>One vs. many donation.</i> Mark is a U.S. businessman who would like to give \$1,000 to charity, and is deciding between two charities. The first focuses on preventing disease in the U.S., and Mark’s donation would save the life of one child. The second focuses on preventing disease in a distant foreign country Mark has never been to, and his donation would save the lives of several children. Mark has already made up his mind to donate to one of the two charities, and just needs to decide to which one. <i>Should Mark donate to the charity in the foreign country?</i>
D5	Greater good dilemma [12]	<i>Benjamin’s car or donation.</i> Benjamin is a college student who has always wanted to have his own car. He works after school jobs, saves all of his money, and rarely goes out, until finally he has \$7,000. On his way to the used car dealership, he reads about a tsunami in Southeast Asia that has left thousands dead, wounded, and homeless. A number is given where donations can be processed to help these victims. Benjamin knows that if he were to donate a

---

	significant amount to this charity, his donation would make a real difference to the lives of several of the victims of the tsunami. <i>Should Benjamin donate any of the money rather than buy the car?</i>
--	---

---

The previous sacrificial dilemmas in which one is asked whether to sacrifice a single individual to save a group of strangers is referred to as “other-benefit dilemmas.” In contrast, whenever the sacrifice also benefits the participant, he or she faces a “self-benefit dilemma” [17]. Dilemma D2 is of the self-benefit variety (Table 2). It is equally criticized on the basis that it is not genuinely utilitarian, because it lacks the identification with the whole of humanity [12]. However, this standpoint has not reached a consensus in literature, partially because it is too new. We then decided to include the self-benefit dilemma D2 in the questionnaires for the same reason we included the other-benefit dilemmas D1. This will prove the right thing to do because our results will apply to utilitarian dilemmas of all types. In a broader sense, D2 is still utilitarian because it considers an arithmetic that pits the options of having less- or more-harmed individuals against each other. In particular, D2 is more akin to the D1 of type 1 questionnaires, in that D2 is personal.

In the three greater good dilemmas in Table 2, participants are left a choice to identify themselves with humanity as a whole. This contrasts with either a parochial attachment to one’s own community or country [12], or a focus on one’s particular circumstance. It is argued that such all-encompassing, impartial concern is a core feature of classical utilitarianism [19].

For each dilemma D1-D5, participants were also asked to rate their perceived wrongness of not adopting the utilitarian action (“a little wrong” = 1 to “very wrong” = 5) [12]. This question thus asks for a conscious, deliberate response that engages System 2.

### 3. Results and discussion

Figure 1 shows the results regarding CRT performance. The marked difference between the two types of questionnaire occurs for the performance of those who does not score at all on the test. In type 2 questionnaires, where the CRT is posted after the moral dilemmas, participants who scored none of the questions were reduced by nearly 8 percent. This result is statistically significant. A one-sided, two-sample bootstrap  $t$ -test with 10,000 replicates shows a  $p$ -value less than 0.03, and thus we conclude the two distributions are not identical. We explain this improved performance by the fact that in type 2 questionnaires, participants became more cognitively alert as a consequence of previously answering the questions related to the moral dilemmas.

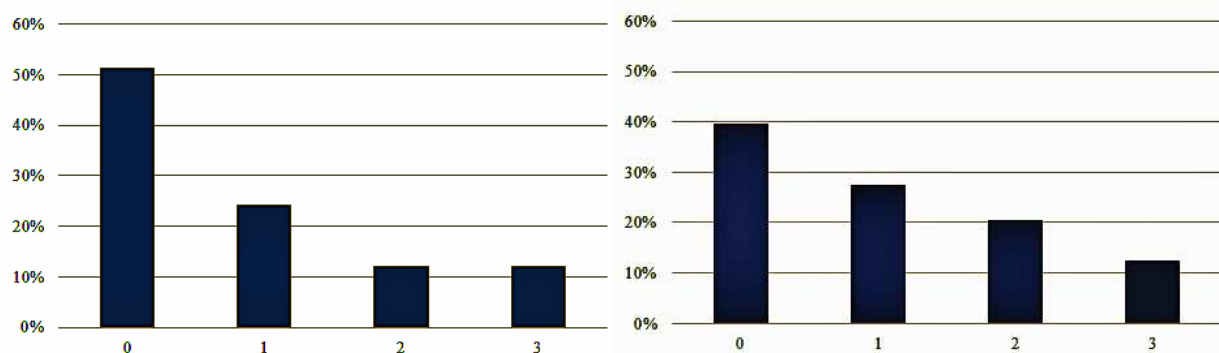


Figure 1. CRT performance. Left: type 1 questionnaires; right: type 2 questionnaires. Those answering type 2 questionnaires became more cognitively alert after deciding on the moral dilemmas, and thus improved the performance on the CRT.

Indeed, Table 3 shows the average CRT performance for type 2 questionnaires was higher (1.06) than that for type 1 questionnaires (0.85) ( $p$ -value < 0.03). University students scored higher than non-students did on both types of questionnaire (statistically significant for type 2 questionnaires). Moreover, in both types of questionnaire, those answering within 30 seconds scored higher on the CRT than those answering in more than 30 seconds, but the difference in performance between the two groups was not statistically significant. Those who provided non-utilitarian answers considering the five dilemmas scored higher on the CRT, but this result is to be taken with caution because the dilemmas are not cognitively similar. For example, in D1 and D5 non-utilitarians beat utilitarians on CRT performance in both types of questionnaire, but the reverse is true considering dilemmas D2, D3 and D4 (Figure 2).

Table 3. Average CRT and type of questionnaire

Average CRT	Type 1 questionnaires	Type 2 questionnaires	$p$ -value*
general	0.85	1.06	0.026
for university students	0.88	1.15	0.012
for non-students	0.70	0.66	<b>0.575</b>
difference between the two groups: $p$ -value	<b>0.234</b>	0.004	-
for those answering within 30 seconds	0.95	1.12	<b>0.148</b>
for those answering in more than 30 seconds	0.78	1.00	0.055
difference between the two groups: $p$ -value	<b>0.151</b>	<b>0.199</b>	-
for utilitarian answers considering dilemma D1	0.40	0.98	0.009
D2	0.91	1.16	0.058
D3	0.88	1.08	<b>0.286</b>
D4	0.85	1.06	0.036
D5	0.68	0.97	<b>0.117</b>
for non-utilitarian answers considering dilemma D1	0.89	1.27	0.009
D2	0.81	0.90	<b>0.261</b>
D3	0.85	1.06	0.032
D4	0.81	1.05	<b>0.258</b>
D5	0.91	1.07	0.092

\*  $p$ -values from one-sided, two-sample bootstrap  $t$ -test with 10,000 replicates

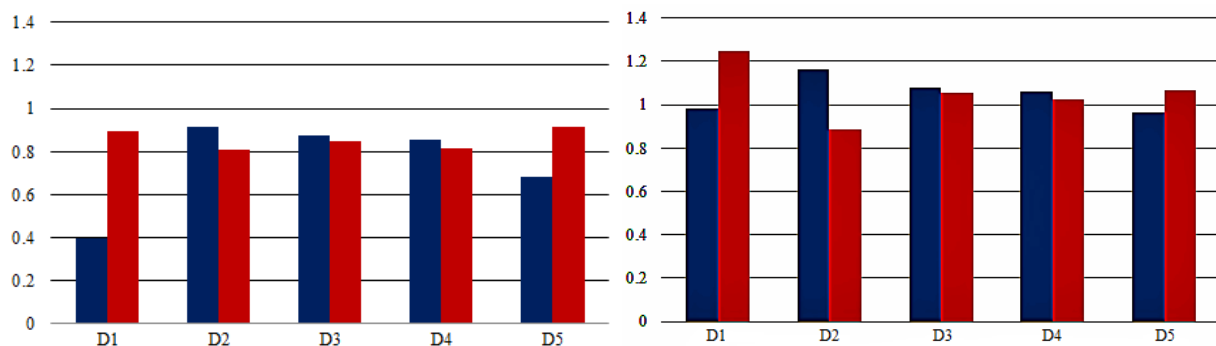


Figure 2. Average CRT for utilitarian (blue) and non-utilitarian (red) answers given to the five moral dilemmas. Left: type 1 questionnaires; right: type 2 questionnaires. The different dilemmas are not cognitively similar in that in D1 and D5 non-utilitarians beat utilitarians on CRT performance, but the reverse is true considering dilemmas D2, D3 and D4. This occurs for both types of questionnaire (Table 4).

Figure 3 displays the utilitarian and non-utilitarian answers considering the five dilemmas. As for the sacrificial dilemma D1, in its personal version (type 1 questionnaires) most respondents made the non-utilitarian choice. There was reversal in the impersonal version of D1 (type 2 questionnaires): most respondents now favored the utilitarian choice. In other words, most respondents were non-utilitarians when the dilemma was framed as pushing a person, but they reverted their judgment and became utilitarians when the dilemma was presented as moving an object. The  $p$ -values in Table 3 show this result is statistically significant for both types of questionnaires. This result is in line with the previous literature [18].

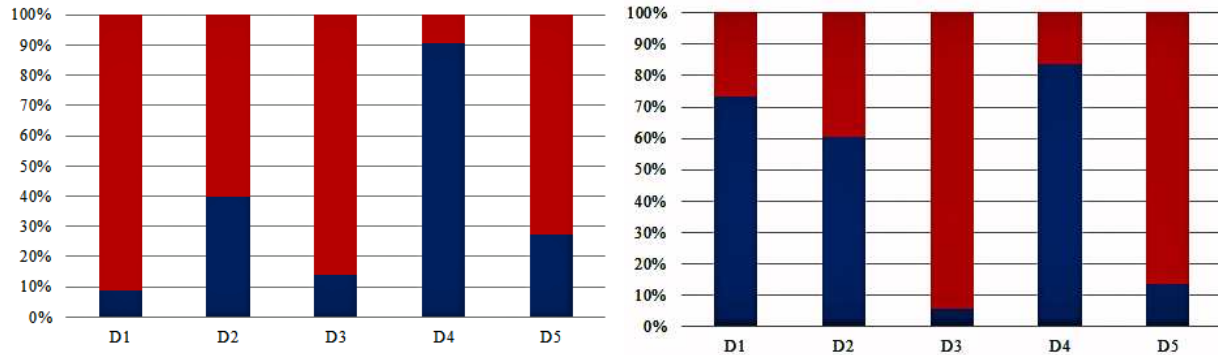


Figure 3. Utilitarian (blue) and non-utilitarian (red) answers given to the five moral dilemmas changed across the dilemmas.

In both types of questionnaire, respondents found it more wrong not to make the utilitarian choice in dilemma D1 (Figure 4). When they deliberately think (using System 2), they are capable of making the utilitarian arithmetic. This does not mean, however, participants are automatic non-utilitarians in dilemma D1 because the default, automatic decision using System 1 is frame dependent. And, this is compatible with the view that people do not carry any predetermined moral preferences. Indeed, a result from the literature of framing effects is that “there is no underlying preference that is masked or distorted by a frame. Our preferences are about framed problems, and our moral intuitions are about descriptions, not about substance” [20]. Moreover, we found no correlation between perceived wrongness and CRT (available upon request), which means perceived wrongness carries independent information not already conveyed in the CRT.

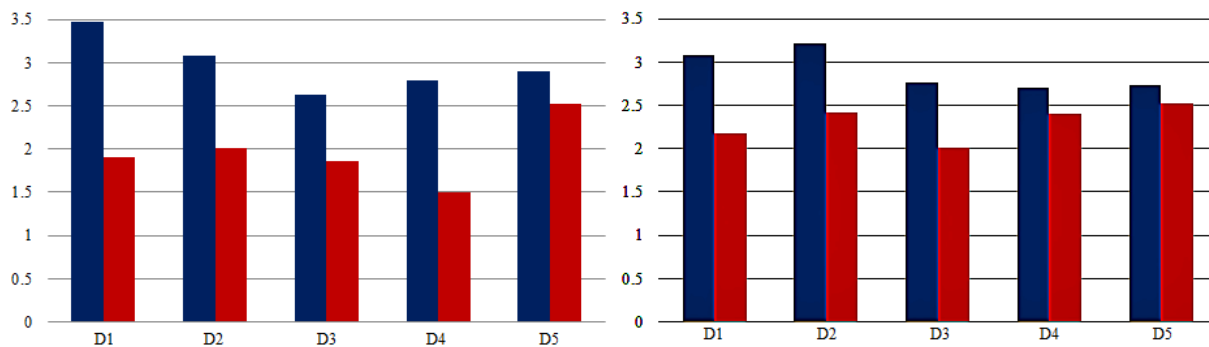


Figure 4. Average perceived wrongness of not adopting the utilitarian action for utilitarian (blue) and non-utilitarian (red) answers given to the five moral dilemmas. Left: type 1 questionnaires; right: type 2 questionnaires.



There was also judgment reversal in the self-benefit dilemma D2 across the two types of questionnaire. In type 1 questionnaires, most respondents adopted the non-utilitarian choice, and in type 2 questionnaires, most respondents were utilitarians (which meant killing a man to save oneself and the others). However, the  $p$ -values in Table 3 show this result is statistically significant for type 1 questionnaires only. This judgment reversal may be related to the architecture of choice participants faced. In type 1 questionnaires, the CRT was answered prior to the moral dilemmas. We speculate that that may have generated “ego depletion” [21], in which case System 1 takes control. Killing a man is bad, and for System 1 “What You See Is All There Is” [20]. Performing the utilitarian arithmetic (killing one to save oneself and the others) is cognitively demanding. In type 2 questionnaires, there was no possibility of ego depletion because the CRT was posted at the end. This may have freed up System 2 to perform the utilitarian arithmetic and explain why most respondents adopted the utilitarian judgment. Interestingly, in both types of questionnaire, respondents found it more wrong not to make the utilitarian choice of killing a man. Here, the utilitarian choice is difficult and requires the use of System 2. However, as in the previous dilemmas D1, the automatic choice was also frame dependent in dilemma D2.

As for the great good dilemmas D3, D4 and D5, the type of questionnaire did not matter for the judgments made by the participants. However, most respondents were non-utilitarians in D3 and D5, and adopted a utilitarian judgment in D4. There was judgment reversal even across the three greater good dilemmas. This may be related to the roles System 1 and System 2 played in each.

In dilemma D3, only 14 percent (in type 1 questionnaires) and 6 percent (in type 2 questionnaires) made the utilitarian choice (Table 4). Most respondents were not prepared to save the peace negotiator rather than Alberto’s mother. They did not identify themselves with humanity as a whole, and preferred to make a judgment based on genetic ties. This strongly suggests the involvement of System 1 in mediating the non-utilitarian answers. Most respondents preferred saving the mother despite perceiving the action of not saving the pacifier as more wrong (Figure 4). This shows a conflict between the two minds in the judgment making of dilemma D3. Being utilitarian in this dilemma D3 was highly costly in terms of cognitive currency. Incidentally, those giving the utilitarian answer scored slightly higher on the CRT than non-utilitarians did (Table 4), but the differences between the groups was not statistically significant.

Table 4. Answers given to the five utilitarian moral dilemmas

Utilitarian moral dilemma	Type 1 questionnaire					Type 2 questionnaire				
	D1	D2	D3	D4	D5	D1	D2	D3	D4	D5
Utilitarian answers	9%	40%	14%	91%	27%	73%	60%	6%	84%	13%
Non-utilitarian answers	91%	60%	86%	9%	73%	27%	40%	95%	17%	87%
Average CRT for utilitarians	0.40	0.91	0.88	0.85	0.68	0.98	1.16	1.08	1.06	0.97
Average CRT for non-utilitarians	0.89	0.81	0.85	0.81	0.91	1.27	0.89	1.05	1.05	1.07
Average CRT difference between the two groups: $p$ -value*	0.03	<b>0.26</b>	<b>0.43</b>	<b>0.47</b>	0.10	0.04	0.03	<b>0.46</b>	<b>0.49</b>	<b>0.31</b>
Average perceived wrongness of not adopting the utilitarian action (utilitarians)	3.47	2.45	2.33	2.11	2.28	3.07	2.86	3.31	2.82	2.83
Average perceived wrongness of not adopting the utilitarian action (non-utilitarians)	1.91	1.77	1.99	1.31	1.95	2.22	2.82	2.82	2.97	2.84
Average perceived wrongness difference between the two groups: $p$ -value*	0.00	0.00	<b>0.12</b>	0.00	0.09	0.00	<b>0.41</b>	0.03	<b>0.25</b>	<b>0.49</b>

\*  $p$ -value from one-sided, two-sample bootstrap  $t$ -test with 10,000 replicates

The great majority of participants chose the utilitarian answer for dilemma D4, regardless of questionnaire type. Greater good dilemma D4 is conceived to allow for less personal

involvement and emotional load, and for more deliberative control in moral judgments [12]. Thus, System 2 judgments are more likely to rule and it is not so surprising that respondents favored the utilitarian choice in dilemma D4. Most participants judged it right for an American executive to donate to a charity in a foreign country rather than one in his own country. The emotional distance was even greater for our sample of Brazilian respondents evaluating the moral attitude of an American individual. Utilitarians in dilemma D4 scored higher on the CRT in our sample (Table 4 and Figure 2), though this difference was not statistically significant (Table 4). However, the role System 2 played in this decision can be evaluated by considering the perceived wrongness of not making the utilitarian choice. Not giving the utilitarian answer was more heavily perceived as wrong for this dilemma D4 (statistically significant for type 1 questionnaires).

Most respondents rejected the utilitarian choice in greater good dilemma D5 (either donate money or buy oneself a car; Figure 3), despite perceiving the wrongness of doing that (statistically significant for type 1 questionnaires). Again, being utilitarian is cognitively too demanding due to the self-interest component of D5 [12]. Judgments aimed at maximizing the greater good are thus weakened by considerations of self-interest. However, self-interest is supposed to be out of place if one were genuinely aiming to promote the greater good from an impartial utilitarian standpoint [12]. The results on the CRT even show the choice of maximizing self-interest over maximizing the greater good is thoughtful, in that non-utilitarians for this dilemma D5 scored higher on the test (statistically significant for type 1 questionnaires).

Data from an fMRI (functional magnetic resonance imaging) study provide evidence for the conflicting roles System 1 and System 2 play in utilitarian moral judgments [22]. The dorsolateral prefrontal cortex and anterior cingulate cortex are brain regions associated with cognitive control, and therefore with the role System 2 plays. Such areas are recruited to resolve moral dilemmas in which a utilitarian judgement has to overcome a personal, automatic judgment. When System 1 rules in personal judgments, there is increase in amygdala activity. This neuroscientific evidence matches our experimental evidence that utilitarian judgments are mediated by an underlying tension between System 1 and System 2.

## 5. Conclusion

This study experimentally evaluated utilitarian judgments under the dual-system approach. Participants responded to a cognitive reflection test as well as to sacrificial and greater good dilemmas that pit utilitarian and non-utilitarian options against each other. We considered utilitarian moral dilemmas that purposely contrast circumstantial or parochial dilemmas with humanity dilemmas. While personal dilemmas are associated with egocentric attitudes, greater good dilemmas consider the whole of humanity.

Performance on the cognitive reflection test improved when participants responded to questionnaires that posted the moral dilemmas at the beginning. This possibly has occurred because participants became more cognitively alert after answering the questions related to the dilemmas.

The different dilemmas D1-D5 are not cognitively similar. Non-utilitarians beat utilitarians on CRT performance in dilemmas D1 and D5, but the reverse is true considering dilemmas D2, D3 and D4.

Most respondents were non-utilitarians when the sacrificial moral dilemma D1 was framed as pushing a person, but they reverted their judgment and became utilitarians when the dilemma

was framed as moving an object. This moral judgment reversal replicated previous findings in literature [18].

Respondents found it more wrong not to make the utilitarian choice in D1. When they deliberately thought (using System 2), they were capable of making the utilitarian arithmetic, though this does not mean participants were automatic non-utilitarians. The default, automatic decision using System 1 was frame dependent. This is compatible with the view that people do not carry any predetermined moral preferences. Moral intuitions are about descriptions, not about substance.

Respondents also found it more wrong not to make the utilitarian choice of killing a man in the self-benefit dilemma D2. However, their moral judgment was also frame dependent. Being utilitarian was costly in cognitive currency. Performing the utilitarian arithmetic required the use of System 2, and participants made the utilitarian judgment in D2 because System 2 was at work as the arithmetic involved protecting oneself. However, under ego depletion, System 1 ruled and participants reverted their moral judgments and became non-utilitarians. After all, for System 1 killing is bad and WYSIATI (What You See Is All There Is).

Most participants were not utilitarians when facing the greater good dilemma D3. Most respondents were not prepared to save a peace negotiator rather than one's mother. They did not identify themselves with humanity as a whole, and preferred to make a judgment based on genetic ties. Being utilitarian by using System 2 in this dilemma D3 was considered by the participants of the experiment as highly costly in terms of cognitive currency.

Not choosing the utilitarian answer was perceived as much more wrong for the greater good dilemma D4. This was not so surprising because the design of the dilemma makes room for System 2 to take control. D4 allows for less personal involvement and emotional load and for more deliberative control in moral judgments. For this reason, most participants judged it right for an American executive to donate to a charity in a foreign country rather than to one in his own country.

Most respondents facing the dilemma to either donate money or buy oneself a car, made the non-utilitarian choice despite perceiving the wrongness of doing that. As before, being utilitarian here was cognitively too demanding and self-interest prevailed over the greater good.

Taken as a whole, all the utilitarian dilemmas showed the utilitarian judgment was too demanding in terms of cognitive currency. And the moral intuitions related to System 1 were sometimes non-utilitarian or utilitarian, depending on how a dilemma was framed. Automatic moral judgments are about descriptions, not about substance.

There is fMRI neuroscientific evidence for our experimental results [22]. Judgments in moral utilitarian dilemmas are mediated by a tension between System 1 and System 2. Being utilitarian by using System 2 is too demanding because the brain areas associated with cognitive control (the dorsolateral prefrontal cortex and anterior cingulate cortex) increase activity. In contrast, when System 1 moral judgments occur, the amygdala shows greater activity.

## References

- [1] Bentham J (1789) *An Introduction to the Principles of Morals and Legislation*. Oxford: Clarendon Press.
- [2] Mill JS (1861) *Utilitarianism and Other Essays*. London: Penguin Books.

- [3] Baron J, Ritov I (2009) Protected values and omission bias as deontological judgments. In Bartels DM, Bauman CW, Skitka LJ, Medin DL (Eds.). *Moral Judgment and Decision Making: The Psychology of Learning and Motivation* (pp. 133-167). San Diego: Elsevier.
- [4] Greene JD, Cushman FA, Stewart LE, Lowenberg K, Nystrom LE, Cohen JD (2009) Pushing moral buttons: The interaction between personal force and intention in moral judgment. *Cognition* 111 (3), 364-371.
- [5] Sunstein CR (2005) Moral heuristics. *Behavioral and Brain Sciences* 28, 531-542.
- [6] Singer P (1979) *Practical Ethics*. Cambridge: Cambridge University Press.
- [7] Cushman F, Young L, Greene JD (2010) Our multi-system moral psychology: Towards a consensus view. In Doris JM (Ed.), *Oxford Handbook of Moral Psychology*. Oxford: Oxford University Press
- [8] Greene JD (2008) The secret joke of Kant's soul. In Sinnott-Armstrong W (Ed.), *Moral Psychology: The Neuroscience of Morality* (pp. 35-79). Cambridge: MIT Press.
- [9] Kahane G, Shackel N (2010) Methodological issues in the neuroscience of moral judgment. *Mind and Language* 25, 561-582.
- [10] Thomson JJ (1985) The trolley problem. *Yale Law Journal* 94 (6), 1395-1415.
- [11] Bartels DM, Pizarro DA (2011) The mismeasure of morals: Antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition* 121 (1), 154-161.
- [12] Kahane G, Everett JAC, Earp BD, Farias M, Savulescu J (2015) 'Utilitarian' judgments in sacrificial moral dilemmas do not reflect impartial concern for the greater good. *Cognition* 134 (1), 193-209.
- [13] Greene JD, Morelli SA, Lowenberg K, Nystrom LE, Cohen JD (2008) Cognitive load selectively interferes with utilitarian moral judgment. *Cognition* 107 (3), 1144-1154.
- [14] Evans JSBT (2008) Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology* 59 (1), 255-278.
- [15] Frederick S (2005) Cognitive reflection and decision making, *Journal of Economic Perspectives* 19 (4), 25-42.
- [16] Foot P (1967) The problem of abortion and the doctrine of double effect. *Oxford Review* 5, 5-15.
- [17] Moore AB, Clark BA, Kane MJ (2008) Who shalt not kill? Individual differences in working memory capacity, executive control, and moral judgment. *Psychological Science* 19 (6), 549-557.

- [18] Hauser M, Cushman F, Young L, Jin R, Mikhail J (2007) A dissociation between moral judgments and justifications. *Mind & Language* 22 (1), 1-21.
- [19] Hare RM (1981) *Moral thinking*. Oxford: Oxford University Press.
- [20] Kahneman D (2011) *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.
- [21] Hagger MS, Wood C, Stiff C, Chatzisarantis NL (2010) Ego depletion and the strength model of self-control: A meta-analysis. *Psychological Bulletin* 136 (4), 495-525.
- [22] Greene JD, Nystrom LE, Engell AD, Darley JM, Cohen JD (2004) The neural bases of cognitive conflict and control in moral judgment. *Neuron* 44 (2), 389-400.