ECONtribute
MARKETS & PUBLIC POLICY

ECONtribute
Discussion Paper No. 275

Ends versus Means: Kantians, Utilitarians, and Moral Decisions

Roland Bénabou Armin Falk Luca Henkel

January 2024 www.econtribute.de

UNIVERSITÄT BONN Universität zu Köln

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2126/1-390838866 is gratefully acknowledged.

Ends versus Means: Kantians, Utilitarians, and Moral Decisions

Roland Bénabou Armin Falk Luca Henkel

First version: November 2022
This version: January 15, 2024

Abstract

Choosing what is morally right can be based on the consequences (ends) resulting from the decision – the Consequentialist view – or on the conformity of the means involved with some overarching notion of duty – the Deontological view. Using a series of experiments, we investigate the overall prevalence and the consistency of consequentialist and deontological decision-making, when these two moral principles come into conflict. Our design includes a real-stakes version of the classical trolley dilemma, four novel games that induce ends-versus-means tradeoffs, and a rule-following task. These six main games are supplemented with six classical self-versus-other choice tasks, allowing us to relate consequential/deontological behavior to standard measures of prosociality. Across the six main games, we find a sizeable prevalence (20 to 44%) of non-consequentialist choices by subjects, but no evidence of stable individual preference types across situations. In particular, trolley behavior predicts no other ends-versus-means choices. Instead, which moral principle prevails appears to be context-dependent. In contrast, we find a substantial level of consistency across self-versus-other decisions, but individuals' degree of prosociality is unrelated to how they choose in ends-versus-means tradeoffs.

Keywords: morality, deontological, consequentialist, Kantian, ends-versus-means, trolley dilemma, prosocial, altruism, social preferences.
JEL classification: C91, D01, D64

Contact: Roland Bénabou: Princeton University, NBER, CEPR, IZA, BREAD, and briq; rbenabou@princeton.edu. Armin Falk: University of Bonn; afalk@uni-bonn.de. Luca Henkel: University of Chicago and University of CEMA, CESifo, JILAEE; luca.henkel@uchicago.edu.
Acknowledgements: We are grateful for valuable comments to Jean-Francois Bonnefon, Benjamin Enke, Brit Grosskopf, Claudia Senik, Jean Tirole, and participants in seminars and conferences at briq, PSE, SciencesPo, TSE, UCEMA, and the University of Chicago. Egshiglen Batbayar, Erik Ortiz Covarrubias, Luca Michels and Adam Rangel provided excellent research assistance. Bénabou gratefully acknowleges the hospitality of Sciences Po and the Paris School of Economics during the year 2021-2022.
Funding: Funded by the German Research Foundation (DFG) under Germany's Excellence Strategy - EXC 2126/1-390838866. Support by the German Research Foundation (DFG) through CRC TR 224 (Project A01) is gratefully acknowledged.
Ethics approval: The study obtained ethics approval from Princeton University (No. 12513 and 15939).
Research transparency: The study was preregistered at the AEA RCT Registry (#AEARCTR-0007714).

"Act in such a way that you always treat humanity, whether in your own person or in the person of any other, never simply as a means, but always at the same time as an end" (Kant, 1785).

1 Introduction

When providing a public good or engaging in reciprocal behavior, individuals trade off the costs and benefits to themselves against those to others. The question of what constitutes moral behavior then has a simple answer, given by the degree of prosociality of the chosen action. The preferences and image motives underlying such tradeoffs are by now fairly well understood.

Much less studied is another important class of decisions, in which the question of what is the right thing to do elicits far less agreement. These are situations in which achieving some socially desirable *end* requires the use of *means* considered inherently objectionable. These include "sacrificial dilemmas" where helping or saving a greater number of people requires placing some in harm's way (e.g., hostage situations, medical triage, or vaccine priority), but also many less dramatic choices such as industrial and product safety, or bargaining situations where a bribe must be paid, a threat made, a lie told, or deeply held values compromised.

Unlike in "self-versus-other" choices, the two main approaches dominating Western moral philosophy and psychology now come into conflict. Under the Consequentialist view, the propriety of an action should be judged solely by its consequences (Bentham, 1789; Mill, 1863; Sinnott-Armstrong, 2019). Economics, in particular, is predominantly utilitarian, even when consequences include emotions or social image. The consequences of an act are evaluated and traded off within a decision-maker's preferences, and in particular the social-welfare functions assumed for normative analysis are almost universally consequentialist.

According to Deontological ethics, on the other hand, whether an action is permissible or not should be based only on its conformity with a series of rules, independent of their consequences for the situation at hand (Kant, 1785; Alexander and Moore, 2016). Many tradeoffs are then proscribed, as violating higher values such as life or human dignity. While rarely absolute in practice, such principles do place major restrictions on both individual choices (lying, sense of duty, refusals to "sell out") and the organization of society (taboos, proscription of "repugnant markets"). Relatedly, in the public debate, self-proclaimed deontologists often criticize economists' use of cost-benefit analysis (e.g., value of a statistical life) as fundamentally out of touch with most people's moral intuitions.

Research questions. This paper aims to extend the empirical study of moral behavior and preferences beyond standard prosociality (altruism, reciprocity, fairness), into the other dimension of consequentialism versus deontological ethics. We do this by exploring three main questions.

First, how prevalent in actual behaviors are these two moral principles? If a significant fraction of people knowingly choose non-consequentialist options, representing their preferences in the standard manner will lead to inaccurate positive predictions, and standard social welfare functions may be misleading guides to normative decisions.

Second, are these two guiding principles or intuitions stable personal "traits", preferences similar to risk-aversion, impatience, or especially altruism? Moving from between-subject variation to individual-level consistency, we ask how correlated a person's behavior is between different situations in which consequentialist and deontological reasoning prescribe opposite actions.

Third, how is this other dimension of morality related to standard social preferences? Specifically, are deontologists generally more (or less) benevolent than consequentialists? Accordingly, we examine to what extent decisions pitting the two moral principles against each other may be correlated with those pitting self-interest against concern for others' welfare.

To address these questions, we design and confront subjects with a series of choice situations consisting of two main blocks: a first, novel one designed to identify consequentialist vs. deontological decision-making (and corresponding "types" if they exist) through *ends-versus-means* (EVM) dilemmas, and a second one consisting of standard tests of altruism and reciprocity involving *self-versus-other* (SVO) tradeoffs.

Ends-versus-means decisions. Our first block starts with the Trolley problem (Foot, 1967; Thompson, 1976), which asks whether it is permissible to sacrifice one life to save a larger number. Although our general methodology does not at all depend on the Trolley, addressing it is unavoidable, given how extensively it has dominated normative debates and empirical investigations of the topic. There are now large international surveys exploring different variants of it, aiming to provide guidance for decisions to be made by autonomous vehicles, robots, and other AI algorithms (e.g., Awad et al., 2018). Despite its popularity, the Trolley problem has almost exclusively been posed as a *hypothetical* question, and cast in very abstract situations (another version being the organ-transplant dilemma). Our paper's first contribution is thus to experimentally study the trolley problem's value as a guide to the real decisions and policy choices it is meant to represent, by implementing a version of it with actual, life-saving consequences. Does this make a difference for behavior?

Our principal interest, however, is in the prevalence (among individuals) and the consistency (for each of them) of consequentialist versus deontological behavior across a broader set of less momentous, more common choice settings. The paper's second and main contribution is therefore to design a series of experimental games creating ends-versus-means tensions between these two moral principles, while leaving the decision maker's material and social payoffs unaffected. We start from a baseline situation in which the two principles agree: if the choice is between directing a donation of 15 Euros to a recognized charity treating children suffering from cancer, or one of 2 Euros to some random other subject, the right thing to do is unambiguously the former, and 93% of our subjects indeed make that choice. We then modify the decision problem so that bringing about this preferred social consequence requires either: (i) *lying* to another subject, at their expense; (ii) *bribing* another subject, who controls greater resources; (iii) making a morally *repugnant* but entirely cheap-talk, anonymous, statement; (iv) refraining from "expressively" choosing a moral option when the choices of others have already made it *ineffective,* and sticking to principle will only cause a further social loss; (v) violating an explicitly stated *rule,* absent any enforcement or meaningful effects of disobedience.

Self-versus-other decisions. Complementing this main series of six choice situations, our second block consists of a rich set of familiar games measuring prosocial behavior, administered to the same subjects a week after the first block. These include a dictator game, with either giving to or taking from a charity; a trust game, both as first- and as second-mover, using the strategy method; a public-goods game; and a moral-luck game (rewarding either intentions or outcomes). Finally, at the end of this block, we administer two questionnaires widely used in the moral-psychology literature, namely the Oxford Utilitarian Scale (Kahane et al., 2018) and the Moral Foundations Questionnaire (Graham

et al., 2011), as well as the more recent Moral Universalism short module of Enke, Rodríguez-Padilla, and Zimmermann (2022); we also gather standard sociodemographic variables.

A key aspect of our design is that, throughout the twelve decision tasks, all aspects of behavior are measured *within subject,* allowing us to look for stable patterns (preference types) within each of the consequentialist/deontological and prosocial/selfish dimensions, as well as for potential correlations between the two. Our experiments were administered online, with 593 subjects (mostly students) completing the experiment. The procedures and main statistical hypotheses to be tested were preregistered.[1]

Results. We report four main results. First, across the various ends-versus-means (EVM) games, a significant fraction of subjects choose non-consequential options that are in line with deontological principles. Numbers range from 20% in the bribing game to 44% in the repugnant-statement game. These revealed preferences clearly differ from those of "standard" economic agents. We employ extensive robustness checks and an additional robustness experiment with 122 subjects, to verify that this type of behavior is not driven by limited attention or subjects' misunderstanding of the decision situations.[2]

Second, we deliver mixed news on the informativeness of the Trolley dilemma as a guide to actual decisions. On the one hand, implementing it with real (statistical) lives at stake does not change behavior, relative to the standard procedure involving only hypothetical lives. In both cases, about 25% of choices are deontological. On the other hand, Trolley choices predict close to no other behavior: across the eleven (EVM and SVO) other incentivized games, there is only a small correlation of 0.16 with behavior in the lying game, and no significant correlation with any of the other ten, for which all correlations are below 0.06.

Third, and more generally, we find very little evidence of individual-level consistency in deontological versus consequential decision-making: behaviors across the six games pitting one principle against the other are largely uncorrelated. Almost no subjects choose the deontological option in five or more of the six main games, only 10% in four or more, and only 10% never, behaving as standard economic agents; about 27% make one such choice, and 50% either two or three. Thus, there appears to be no single preference type or parameter that would robustly predict choices across ends-versus-means dilemmas. Instead, which moral principle prevails for an individual appears to be context-dependent. This is in stark contrast to standard social preferences, which exhibit substantial within-subject consistency across the six self-versus-other games (SVO) games in the second block. For instance, dictator-game giving reliably predicts behavior in all five other SVO games with correlations ranging between 0.2 and 0.65.

Fourth, ends-versus-means decision-making is unrelated to general prosocial preferences. Put differently, acting deontologically versus consequentially in the Trolley, or any of the other main games, is not predictive of whether a person is more or less prosocial. This result contrasts with people's perceptions, which tend to rate individuals who make deontological actions as more moral and trustworthy (e.g., Everett, Pizarro, and Crockett, 2016; Everett et al., 2018).

Finally, beyond its specific findings, our paper develops a flexible paradigm to study Consequen-

[1] https://www.socialscienceregistry.org/trials/7714

[2] In the robustness experiment, subjects face close variants of each original game, with the payoff structure simply reversed so that desirable ends and means are aligned. Almost all subjects ($88\% - 96\%$) now choose the option for which both moral principles agree.

tialist vs. Deontological decision making, based on a new class of ends-versus-means choice tasks. These are fundamentally distinct from classical self-versus-other tradeoffs, and indeed reveal an essentially independent dimension of moral preferences.

1.1 Related literature

We first contribute to the nascent literature in economics that investigates behavior motivated by consequentialist and deontological considerations. Starting with Laffont (1975), several papers have conceptualized Kant's categorical imperative for decision-making in strategic situations (Roemer, 2010; Alger and Weibull, 2013, 2016; Alger, Weibull, and Lehmann, 2020). In these models, Kantian agents act (in part) as if their decision would cause all other players to make a similar choice, in line with the categorical imperative.[3] Van Leeuwen and Alger (2021) find support for the existence of such a motive, mixed with standard Nash-consequentialist behavior, in the Prisoner's, Trust, and Ultimatum games. Closely related is the literature on ethical voters, who derive a fixed "duty" or "expressive" utility from casting a ballot for an alternative that is considered morally superior, even if it goes against their self-interest or if the probability of being pivotal would not justify the cost of voting. Feddersen and Sandroni (2006), Feddersen, Gailmard, and Sandroni (2009), and Feess, Kerzenmacher, and Timofeyev (2022) develop such models; the last two also conduct experiments that show, by varying the probability that a vote will be pivotal, that both consequentialist and deontological or expressive motives are at work among subjects.[4]

Compared to these papers, we study consequentialistic and deontological behavior in individual choice situations with neither strategic considerations nor any self-interest motive. Chen and Schonger (2022) also focus on situations without strategic interactions. They model deontological decision-making as governed by lexicographic preferences, first over acts *per se*, and secondarily over consequences. In two experiments, they show that when the consequences of a donation decision become increasingly hypothetical, subjects are more likely to donate money, suggesting again a mix of deontological and consequential motives. Bénabou and Tirole (2011) and Bénabou et al. (2023) show how behaviors displaying Kantian-like "sacred values" and "taboos tradeoffs" arise, even for consequentialist agents, in a model of moral identity maintained through self-signaling.

We add to this literature by providing evidence on the prevalence and, importantly, the consistency (or lack thereof) of consequential and deontological behavior across a range of decision situations. Furthermore, in contrast to the previous literature, we purposefully abstract from preferences based on self-other considerations. Apart from rule breaking, our ends-versus-means decision situations were designed so that subjects' self-interest plays *no role*, as their choices do not influence their own payoffs. This feature also allows us to investigate, in a second stage, the relation between deontological/consequential motives and prosociality. By analyzing patterns both within and across the EVM and the SVO blocks, our paper also connects to a nascent literature that studies components underlying individuals' behavior across different choice situations. Falk et al. (2018), Dean and Ortoleva (2019), and Chapman et al. (2023) study the relationships between social, risk and time

[3] *"Act only according to that maxim through which you can at the same time will that it become a universal law".* (Kant, 1785).

[4] Also related is the literature on the role of intention vs. consequences, showing that when judging the prosociality of an opponent's action, most individuals act non-consequentially by considering intentions: They reward or sanction actions yielding identical consequences differently if underlying intentions differ (Falk, Fehr, and Fischbacher, 2003, 2008).

preferences, and Stango and Zinman (2023) the correlation patterns between different behavioral biases.

We also contribute to the large literature that uses the trolley problem to investigate underlying moral principles (Greene et al., 2001; Hauser et al., 2007; Bartels, 2008; Lanteri, Chelini, and Rizzello, 2008; Rai and Holyoak, 2010; Costa et al., 2014; Gawronski and Beer, 2017; Awad et al., 2020). These studies typically focus on moral judgments instead of behavior and use hypothetical situations and questionnaires, with two exceptions. Gold, Pulford, and Colman (2015) create a trolley-like incentivized situation in which subjects could divert meals among children living in an orphanage. They compare behavior with moral judgments, finding no difference. Bostyn, Sevenhant, and Roets (2018) design a situation in which subjects believe they are distributing electrical shocks among mice. The experiment involves deception, as no shocks are actually delivered. Comparing subjects' behavior in this (believed to be) "real" situation with what others say they would do when the same question is presented hypothetically, they find that the former case leads to significantly fewer deontological choices. In contrast to these two papers, our SAL-trolley paradigm involves actual human lives (albeit statistical ones), and is therefore closest to the classical version of the dilemma. Importantly, we further add to the literature a comprehensive assessment of the predictive power of the trolley problem for behavior in related EVM dilemmas, as well as in SVO tradeoffs.

2 Experimental design

Our design consists of three main blocks. The first contains six experiments confronting subjects with "ends-versus-means" decisions. The second elicits standard prosocial preferences using well-established "self-versus-other" decision paradigms. The third block contains several questionnaires.

2.1 Ends-versus-means block

We designed six decision situations sharing the key feature that the consequences (ends) of the choice options create a moral conflict with the actions (means) required to bring them about, as shown in Table 1. A second aspect fundamentally distinguishing them from standard games studied in the literature on social preferences is that they do not contain any self-versus-other tradeoff.

Table 1: Ends-versus-means decision situations

Decision situation	Ends	Means
Trolley problem	Life of three people vs. one person	Sacrifice vs. spare one person
Lying game	Donation vs. money to other subject	Lie vs. tell truth to another subject
Bribe game	Donation vs. money to other subject	Bribe vs. not bribe another subject
Group donation game	Donation vs. money to other subject	Consequential vs. inconsequential voting
Statement choice	Donation vs. money to other subject	Make vs. not make repugnant statement
Rule-following task	No money vs. money to self	Follow vs. break rule

The first decision task is the trolley problem, the most studied situation where such an EVM conflict arises. At stake in our version of the problem are (statistical) human lives, and in order to save three, one has to actively sacrifice one. The next four tasks implement EVM dilemmas that all build on a simple baseline choice between two options, Option A and Option B. Choosing Option A generates

a donation of 15 Euros to a charity that supports children suffering from cancer. Choosing Option B instead increases the payoff of another (random) subject by 2 Euros. In this basic decision, Option A dominates Option B based on consequences, and no controversial means are needed to achieve the preferred end. Accordingly, deontological and consequentialist reasoning agree that Option A is the morally right choice, and indeed 93% of subjects in our experiment choose it when facing this decision.[5]

Having established a setting where the underlying moral prescriptions are largely undisputed, we extend the situation to four variants designed to induce ends-versus-means tradeoffs, for which deontological and consequentialist ethics make opposing prescriptions.

In the lying game, subjects must deceive another participant in order to trigger the 15 Euros donation. In the bribe game, they have to bribe another subject who controls a larger donation. In the statement-choice situation, they must make a morally repugnant cheap-talk, anonymous statement. In the group-donation game, the choices of others have already destroyed the donation, so choosing that option is a pure "expressive" (and again anonymous) act that only reduces social surplus. Lastly, we complement this set with a rule-following task. Here, achieving positive consequences is only possible by disobeying a type of rule that is generally important in society, but meaningless in the concrete situation at hand.

In addition, for each game in this block, we designed what we call an *aligned* version, administered to a separate set of subjects. In these variants, we switch the mapping between actions and consequences so that there is no longer an ends-versus-means tension. In the lying game, for instance, it is now the truthful message that triggers the donation; in the statement game, it is endorsing the virtuous rather than repugnant assertion. Deontological and consequentialist principles now prescribe the same option, while other features of each game, and in particular its complexity, remain constant. Comparing subjects' behavior between the original and aligned versions allows us to check whether they pay attention to, understand properly, and care about, the decision problems they face.

We now describe each of the EVM games, the respective ends-versus-means conflict, and the corresponding aligned version in detail. Appendix Section E contains the instructions given to subjects in the experiment.

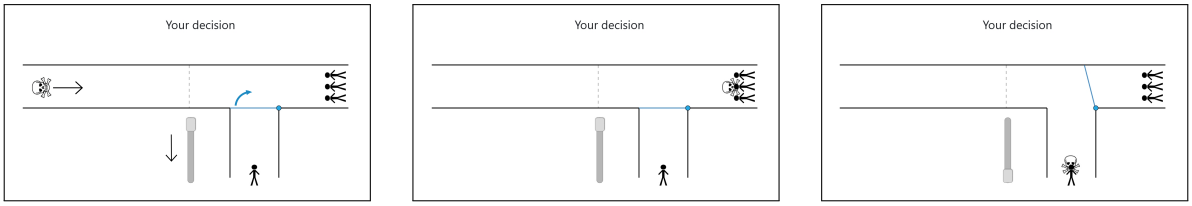
2.1.1 Trolley problem

The Trolley is one of the most popular mental experiments in moral philosophy and applied ethics, used extensively to distinguish between modes of moral decision-making. It is, however, frequently criticized for representing an artificial and implausible situation, detached from any decisions people face in reality (e.g., Bauman et al., 2014). In our main block of decision tasks, we implement the trolley problem with real, meaningful consequences using the *Saving a Life* Paradigm (Falk and Graeber, 2020; Bénabou et al., 2023). Key features of our version are that choices involve the saving of actual human lives, and that they occur in a realistic context that nonetheless contains all relevant features of the Trolley problem.

[5]Furthermore, when asked, 91% of subjects agree that from a moral point of view, there is a right and a wrong decision for them in this situation; and of those, 98% name Option A as the right decision. Similarly, 95% think that other people consider Option A to be the morally right choice.

Figure 1: Decision screen of the trolley problem

(a) Initial position (b) End position not pulling the slider (c) End position pulling the slider



Notes: The figure displays screenshots of the animation used for the trolley problem decision. Panel a) shows the starting position from which the skull symbol slowly moves to the right. The arrows are included for illustration and were not part of the experimental decision screen. Panel b) shows the end position if a subject decides not to pull the slider, indicating no redirection of the donation. Panel c) shows the end position if a subject decides to pull the slider, indicating a redirection of the donation.

Saving a Life (SAL) paradigm. We partnered with the non-profit organization *Operation ASHA*, which treats people in India suffering from tuberculosis using innovative methods and procedures. Tuberculosis is a highly lethal infectious disease if untreated, but curable with a high success rate if treated. An amount of 380 Euros allows the charity to treat five people suffering from tuberculosis, one of which on average would have otherwise died. Thus, each donation implemented saves one life in expectation. This calculation takes into account treatment success rates, other fatality rates, and alternative treatment possibilities, using peer-reviewed epidemiological studies and information about the cost structure of the charity. Subjects receive detailed information on the context of the SAL choice paradigm and all aspects of the computation.

Based on this setup, we implement a very similar dilemma as in the classical Trolley. Prior to the experiment, the charity identified people suffering from tuberculosis who could be treated in two distinct Indian states, A and B.[6] The situation described to subjects was then the following.

- A donation of 380 Euros has been *preset* to cover 5 people suffering from tuberculosis in State A. If it is actually implemented (after the experiment), none of the five will die from the disease. If not implemented, one will die, in expectation. Donation to State A is thus the default.

- In State B, 15 people are suffering from tuberculosis, for whom no donation is initiated. If the donation preset for State A is redirected to State B, the amount will be multiplied by three. With those 1,140 Euros, 15 people will be treated, thus saving (in expectation) three lives that would otherwise be lost to tuberculosis, but foregoing the saving of one in State A.

Decision. Subjects have the choice of whether to redirect the donation from State A to State B or not. They can do so within an animation, a screenshot of which is depicted in Figure 1 (arrows added for exposition). Absent any action, the skull symbol proceeds from left to right along the track and will eventually hit the three figures representing lives in State B. The subject can, however, redirect the skull toward the one figure representing a life in State A, by pulling a lever that will cause a gate (drawn in blue in Panel A) to pivot, provided this is done before the skull has passed the gate. Subjects thus have two options:

[6]We used the states of Maharashtra and Orissa, in which the charity operates. We randomized between subjects which state was State A and which State B.

- Not pulling the slider: no redirection of donation → three lives in State B are lost to tuberculosis.
- Pulling the slider: redirecting donation → one life in State A is lost to tuberculosis.

Tradeoff. The moral dilemma involved is exactly as in the classical Trolley problem. According to deontological ethics, it is not permissible to redirect the donation (or, equivalently, the skull), thereby causing the death of (more precisely, not saving) one person in State A in order to save three in State B. This represents active harm to a human being, and arguably using them as a means to an end. According to consequentialist cost-benefit analysis, on the other hand, saving three rather than one is the right thing to do, and no different from tradeoffs implicitly made every day in medicine, product or road safety, and public-budget allocations.

Treatments. Our SAL version of the Trolley problem differs from moral philosophers' classical one in two ways. First, the scenario is not about imagined train tracks and repairmen happening to be on them, but about actual patients and treatments in the context of a real and common disease. To investigate the effect of this added realism, subjects also faced the classical train-track trolley version at some other stage in the experiment (a week apart, and in randomized order). Second, and most importantly, subjects' choices are not hypothetical but can have real, important consequences. To properly isolate the effect of this latter feature, we ran two different between-subject treatments. In Treatment *SAL-Hypothetical*, the choice environment is exactly as described above, but all choices are, and are presented as, hypothetical: there is never any actual donation. In contrast, in Treatment *SAL-Real*, for each subject, there is a 10% probability that their decision will be implemented, resulting in either one life saved in State A or three in State B.[7]

***Aligned* version.** A donation sufficient to treat 15 people and thus save three lives in State A is preset, and none in State B. If the donation is redirected, it is reduced so that only five people will be treated in State B, thus saving one life there but sacrificing three in State A. Consequentialist cost-benefit analysis and the deontological directive against using a human being as a means to an end now both agree to not redirect the donation, thus prescribing to not pulling the slider.

2.1.2 Lying game

This game was adapted from the classical sender-receiver game in Gneezy (2005), which we modify from a self-versus-other to an ends-versus-means dilemma. The receiver must choose between options A and B without knowing anything about the consequences of either option. The only piece of information they receive is a message from the sender, who knows that if the receiver chooses option A, 15 Euros will be donated to the children's cancer charity, whereas choosing B will earn the receiver 2 Euros. As with all the EVM games, donations and transfers are paid by the experimenter, eliminating any self-versus-other tradeoff for the decision-maker. The sender can send either of the following messages:

- Message 1: "Option A will give you the higher personal payment" (lie).

[7] We ended up implementing the decisions of 23 subjects. As a result of the ensuing donations in both states, 265 patients were treated, and thus 53 lives were saved, in expectation.

- Message 2: “Option B will give you the higher personal payment” (truth).

The sender is informed that in more than 90% of cases, receivers choose the option mentioned in the message,[8] and that they will never know whether the message was true or false, nor what situation the sender faced. The outcome of interest concerns the decisions of senders, whereas the behavior of subjects playing the role of receivers is not part of the analysis. Hence, subjects of the main experiment take the sender’s role, while receivers are part of a separate sample.

Tradeoff. The game puts subjects in a situation in which they need to lie (to someone else’s minor detriment) in order to trigger the more socially valuable donation. Such “white lies” are justifiable by consequentialist principles but not under deontological ethics (Erat and Gneezy, 2012).[9] Accordingly, the former prescribes Message 1, the latter Message 2.

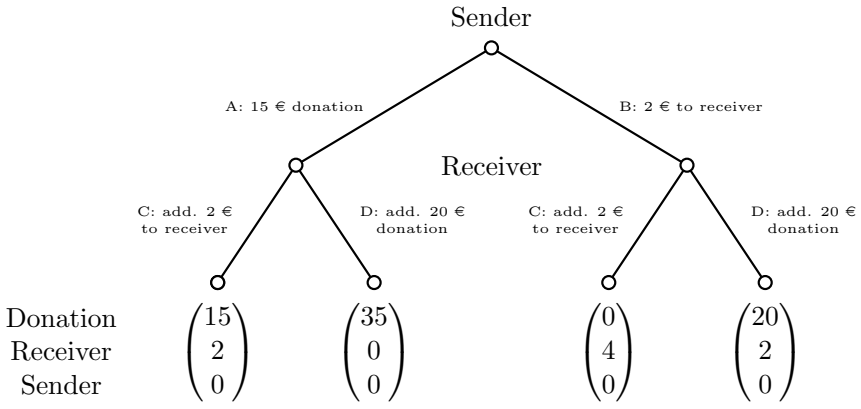
***Aligned* version.** The sender can either send a message that Option A triggers the higher donation (truth) or a message that Option B will trigger it (lie). Because almost all subjects follow this message, just as in the main game, sending the truthful message achieves the more desirable consequence. Accordingly, both moral principles prescribe sending this message.

2.1.3 Bribe game

Two subjects are paired together, one playing the role of a sender and the other playing a receiver. Payoffs for both players and the charity are determined by the sender’s choices, knowing the receiver’s predetermined conditional responses.

The decision situation unfolds in two stages, summarized by Figure 2. In the first stage, the sender chooses between Option A, which implements a 15 Euros donation to the charity, and Option B, which pays 2 Euros to the receiver. In the second stage, knowing what the sender chose, the receiver decides between Option C, which yields them 2 Euros, and Option D, which triggers a separate 20 Euros donation to the charity.

Figure 2: Bribe game



[8] We obtained this number through a pilot.

[9] *“To be truthful (honest) in all declarations is therefore a sacred command of reason prescribing unconditionally, one not to be restricted by any conveniences.”* (Kant, 1785)

Receivers' contingent decisions are elicited using the strategy method, and all senders are (truthfully) informed that their paired receiver was one of those who decided to make their choice between C and D hinge on their sender's decision, as follows:

- If the sender chooses A, the receiver will choose C and take 2 Euros for themselves.
- If the sender chooses B, the receiver will choose D, thus triggering a donation of 20 Euros to charity while keeping the 2 Euros sent by the other player.

Thus, the receiver, in effect, demands a bribe in order to implement the most socially desirable option (the word bribe is never used in the instructions). Again, we are only interested in the choices of the senders confronted with such a demand, who constitute the subjects of the main experiment. Receivers are part of a separate sample, and their choices are not of primary interest.[10]

Tradeoff. Note that the sender is not facing a strategic situation but, once again, a simple dilemma pitting a deontological approach, which entails sticking with Option A and making a donation of 15 Euros without regard to ultimate consequences, against a consequentialist one that will increase the total donation to 20 Euros but requires paying the bribe demanded by the other player in order to achieve this end.

***Aligned* version.** Instead of being matched with receivers who demand a bribe (choice of B over A) for donating (choosing D over C), senders are matched with receivers who demand that they donate (Option A) in order for themselves to choose their donation option (Option D). To implement the socially most desirable option, senders must thus choose to donate rather than bribe; consequentialist and deontological prescriptions coincide.

2.1.4 Statement choice

Subjects first receive information on the harmful effects of CO_2 on the environment and its contribution to climate change. They are subsequently informed that, as part of the study, the purchase of carbon offsets has been prearranged, each such certificate corresponding to offsetting 1 ton of CO_2 from the atmosphere. It is also explained to them that once they have completed their task, the computer will randomly destroy (i.e., not follow through with the purchase) one certificate with a probability of 50%. The destruction of the certificate thus means that the planned removal of 1 ton of CO_2 will not take place. This process is completely independent of any subject's actions, and the certificate feature is included only to make more salient both the climate-change problem and the fact that there exist ways to alleviate it. In this context, subjects can choose to submit one of the following declarations (pressing the corresponding button):

1. "I support the preservation and protection of the environment."
2. "I support the destruction of the environment."

[10] Among them, 60% demanded a bribe to choose the donation, while 40% chose either the donation (23%) or the money (17%) unconditionally. All senders were paired with one among the first group.

As explained to the subjects, the first statement leads to another subject receiving 2 Euros, whereas the latter triggers a donation of 15 Euros to the children's cancer charity. In addition to being fully anonymous (like all choices in our experiments), subjects are informed that the results of the experiment will not be used for any other purpose, such as an opinion poll, thus depriving the statements of any instrumental value.

Tradeoff. The dilemma is thus to stick to one's values and submit the first statement,[11] as at least strongly suggested by the deontological approach, or to submit the second one in order to achieve an unambiguously better outcome, in line with a consequentialist view. Examining people's willingness to (anonymously) make a statement that is antithetical to their moral identity is a procedure similar to that in Bursztyn et al. (2020), but in our case, the consequence of doing so is not a material reward for oneself (one fifth or a day's wage in that paper), but once again creating a positive social externality.

***Aligned* version.** Here, choosing the statement in support of the environment triggers the donation, while choosing the one supporting the destruction of the environment leads to another subject receiving 2 Euros. Hence, both moral principles prescribe endorsing the first statement.

2.1.5 Group decision game

For this task, adapted from Falk, Neuber, and Szech (2020), subjects are sorted into groups of size 6. Each member makes their decision autonomously from the others, but the consequences of the actions taken by any can affect everyone. At the start of the game, the group is entrusted with 15 Euros, which is preset to be donated to the children's cancer charity by the end of the study. The first five members (first-movers), who take action earliest, are part of a separate sample, while the sixth one (second-mover) is part of the main study. First-movers simultaneously choose between:

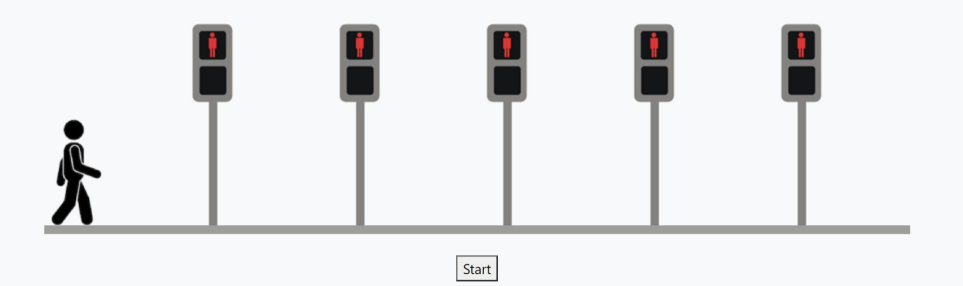
- Option A: This choice has no further consequence, but simply preserves the donation intact if it is still relevant.
- Option B: This choice grants the member who chooses it 2 Euros as an additional payment for themselves. If even just one group member chooses this option, however, the group's donation will be canceled.

The second-mover also has two choices: Option A entails no additional payment to anyone and preserves the donation in case it is still intact; Option B grants 2 Euros to some subject outside the group, but also destroys the donation in case it is still intact.

Before making their decision, second-movers learn whether the donation has already been destroyed by the choices of the first movers. Importantly, in our experiment, at least one first mover in every group opted for option B. As in Falk, Neuber, and Szech (2020), all second movers thus choose with full knowledge that the donation had, effectively, already been canceled.

[11] Our sample consists of very environmentally-conscious subjects: 96% agree that fighting climate change is important, and 99% agree that measures to protect the environment are important. For the link between Kantian-like behaviors and moral-identity maintenance, see Bénabou and Tirole (2011) and Bénabou et al. (2023).

Figure 3: Decision screen of the rule-following task



Tradeoff. Option A is thus entirely inconsequential, but choosing it allows the subject to (anonymously) express a moral preference for donating, an act aligned with deontological principles (such as the categorical imperative: one should will that everyone always chooses Option A). Choosing Option B, meanwhile, is clearly consequentialist, as it generates a positive outcome for another participant who took no part in the task, without affecting the already foregone charitable donation.

***Aligned* version.** A group of 6 subjects is now entrusted with a 2 Euro payment destined for another subject who is not a group member. As in the main experiment, Option A has no further consequence but simply preserves that payment, if still intact. If anyone chooses Option B, on the other hand, the 2 Euro payment is canceled and a 15 Euro donation to the charity is triggered (each time Option B is chosen). As in the main game, our subjects of interest are second-movers, who decide last and are informed about whether the 2 Euro payment has been destroyed by previous choices. Again, in all groups, at least one first mover chose Option B, thus rendering Option A for the second mover inconsequential. As a consequence, both moral principles are aligned in their prescription of Option B, as this option is both consequential and morally desirable under deontological principles.

2.1.6 Rule-following task

The design is taken from Kimbrough and Vostroknutov (2016). The subject controls a stick figure walking across the computer screen along a series of traffic lights and decides how long to wait at each, see Figure 3 for a visualization. Initially, the figure is at the left of the screen, and all lights are red. Once the animation starts, the figure "walks" towards the end of the path (right of the screen), automatically stopping and waiting at every red light. Each time, however, the subject can decide to press a button that causes the figure to proceed through the red light without waiting.

Subjects receive an endowment of 8 Euros and incur a deduction of 0.08 Euros for each second it takes the figure to walk across the screen. Without stopping at any red light, it takes four seconds to complete the track, costing the subject about 2 Euros in total. Waiting at each red light roughly doubles these losses to about 4 Euros. In the instructions received by subjects, they are told that the "rule of the game" is to stop at each red light until it turns green; there is no enforcement or incentive to follow the rule, however.

Tradeoff. The design creates a tradeoff between a deontological approach to the problem, which entails following the stated rule of stopping at each traffic light (or the meta-rule that "one should play by the rules") and incurring losses, versus a consequentialist calculus, which favors maximizing the total payoff by breaking the non-instrumental rule (walking through red lights).

***Aligned* version.** Subjects now start with 0 Euro and receive 0.08 for each second that the figure is walking across the screen, up to a maximum of 8 Euros. Accordingly, following the rule by waiting at the traffic lights maximizes the payoff that subjects can receive from the game, and is thus the aligned choice.

2.2 Self-versus-other block

2.2.1 Altruism, reciprocity, and prosociality

Our second set of decision situations, which subjects take either one week after or one week before the main block (with the order randomized), consists of a series of standard choice tasks from the literature on social preferences, listed in Table 2. Their defining feature is that subjects now face tradeoffs between their own monetary outcomes and those of other subjects, which is why we label these situations as self-versus-other (SVO). Most are very familiar, and we implement them using standard procedures. Therefore, we leave the description of procedural details to Section A of the Appendix. More novel is the moral-luck game, which we describe next.

Table 2: Prosociality (self-versus-other) decision situations

Variable	Elicitation method	Definition
Altruism	Dictator game with charity as the recipient, giving frame, 20 Euros endowment	Amount allocated to charity
Altruism Taking	Dictator game with charity as the recipient, taking frame, 20 Euros endowment	Amount allocated to charity
Trust	First mover in trust game, 5 Euros endowment	Amount send to second mover
Pos. Recipr. Low	Second mover in trust game when having received 1 Euros out of 5 Euros (strategy method)	Amount send back to first mover
Pos. Recipr. High	Second mover in trust game when having received 5 Euros out of 5 Euros (strategy method)	Amount send back to first mover
Public goods game contribution	Public goods game, 5 Euros endowment, group size of 3, 1.5 marginal per capita return from contributing	Amount contributed
Rewarding intentions	Moral luck game	Amount allocated to first mover in $S4 - S3$
Rewarding consequences	Moral luck game	Amount allocated to first mover in $S4 - S2$
Rewarding consequences - intentions	Moral luck game	Amount allocated to first mover in $(S4 + S3) - (S2 + S1) - ((S4 - S3) + (S2 - S1))$

Notes: For details on the first six variables, see Appendix A. For details on the last three, see Section 2.2.2.

2.2.2 Moral-luck game

There are two players: a first-mover who chooses between two lotteries with different payoff distributions for themselves and for the charity, and a dictator who allocates additional money between themselves and the first-mover. These features make the game one of conditional altruism, which is why it is included in this block. At the same time, it is designed to measure and compare the extent to which dictators reward socially desirable intentions versus socially desirable outcomes, meaning the game could also be informative about deontological versus consequentialist preferences, or norms.

First-movers, whose behavior is not *per se* the object of interest, choose between two lotteries: (i) Lottery M, which yields 10 Euros to self with 70% probability and a 15 Euros donation to the charity with 30% probability; (ii) Lottery D, which yields 10 Euros to self with 30% probability and a 15 Euros donation to the charity with 70% probability.

Subjects in the role of the dictator then make, using the strategy method, allocation choices to the first-mover they are paired with, for each of the four possible choice-outcome:

S1 First-mover chooses M, 10 Euros payment realizes

S2 First-mover chooses D, 10 Euros payment realizes

S3 First-mover chooses M, 15 Euros donation realizes

S4 First-mover chooses D, 15 Euros donation realizes

The dictator is endowed with 10 Euros and the amount x they allocate to the first mover is tripled (they keep the remaining $10 - x$), in order to induce positive giving in each of the four possible situations. The first-mover thus potentially receives money from the allocation task and the lotteries. The dictator's choices then allow us to ask: (i) fixing the lottery outcome, to what extent do dictators take the first-mover's lottery choice ("intention") into account when allocating money to them? (ii) fixing the lottery choice, to what extent do dictators take the outcome into account in their allocation? (iii) is their decision more responsive to the first-mover's *ex-ante* lottery choice or to its *ex-post* realization? Table 2 specifies exactly how each of these propensities is measured.

2.3 Questionnaires

To complement our incentivized behavioral tasks, we administer several widely used questionnaires, which constitute the third block of our design. First, we employ the Oxford Utilitarianism Scale (OUS, Kahane et al., 2018), which elicits subject's "permissive attitudes toward instrumental harm", very much as in the Trolley dilemma (OUS-IH), and their "impartial concern for the greater good", a utilitarian-like tendency to judge the well-being of every individual as equally important (OUS-IB). Second, we administer the Moral Foundations Questionnaire (MFQ, Graham et al., 2011), which aims to measure five distinct dimensions of people's moral concerns: care/harm of others, fairness/reciprocity, in-group/loyalty, authority/respect, and purity/sanctity. Third, we include the Moral Universalism short module of Enke, Rodríguez-Padilla, and Zimmermann (2022). Using hypothetical allocation games, the module measures the extent to which subjects exhibit the same level of altruism towards strangers as towards in-group members (MU scale). Finally, we confront subjects with the classical, hypothetical train-track Trolley dilemma, and also include a module on political

attitudes and religiosity. For sociodemographic variables, we collected age, gender, subject of studies, final high-school grade, and gross monthly income.

2.4 Procedural details

The study was run online as a virtual-lab experiment using oTree (Chen, Schonger, and Wickens, 2016), with the subject pool of the BonnEconLab. Subjects were invited using hroot (Bock, Baetge, and Nicklisch, 2014) and had to log in at a specified date and time, where an experimenter was available throughout to answer questions and address any issues. Subjects had to complete two sessions, separated by one week, each lasting about 45 minutes. One session contained the main ends-versus-means decision situations, including the SAL Trolley, and the other one the self-versus-other situations and the questionnaires. The order of sessions, as well as the order of decision tasks within each session, was randomized. Subjects earned a 12 Euros show-up fee for each session, and one decision in each session was selected for real implementation –except for the *Real* treatment of the SAL trolley, in which every subject's choice was implemented with 10% probability.

Overall, 626 subjects took part in the main experiment, of which 593 completed both sessions.[12] Based on the preregistration, we excluded the top 1% fastest subjects, as well as subjects who preferred giving 2 Euros to another participant over a 15 Euros donation to the children-with-cancer charity in the baseline task (41 subjects). The reason for excluding them is that no opposing predictions in the ends-versus-means games exist for them. The final sample thus consisted of 548 subjects (339 women, mean age = 27, SD = 8). Results are robust when considering the full sample.

In order to ensure that all participants fully understood the decision rules and consequences, we included extensive comprehension checks throughout the experiment. For instance, each of the major decision situations (except rule-following and statement choice) featured a quiz after their introduction, and subjects were only allowed to proceed once they answered all questions correctly.[13] For the group-donation game and the statement choice, we also implemented attention and memory checks. At the end of the experiment (approximately 15-25 minutes after their decisions), subjects were asked, unannounced: (i) in the group-donation game, whether the donation was destroyed by the choices of the other group members; (ii) in the statement choice, what were the consequences of making the repugnant statement. Proper recall at that later time constitutes a lower bound on comprehension and attentiveness during the choice phase, which we will use to test the robustness of our results.

To further ascertain subjects' comprehension and attention to the tasks, we ran the *aligned* robustness experiments. A separate sample of 135 subjects (82 women, mean age 24, SD = 5) faced the aligned versions of the EVM games, in a single session. As in the main experiment, we randomized the order of decision tasks, and one decision in each session was selected for real implementation, excluding the aligned version of the trolley problem, which was stated hypothetically. We again excluded the top 1% fastest subjects and subjects who preferred giving 2 Euros to another participant

[12] We pre-registered a sample size of 600 completes. This sample size was chosen based on power calculations for the comparison between the *Hypothetical* and *Real* treatments of the SAL trolley. Assuming that 20% of subjects choose the deontological option in *Hypothetical* and a significance level of 5%, we have 80% power to detect a 10 pp. treatment effect (two-sample test of proportions) –that is, a reduction in the fraction of subjects choosing the deontological option to 10% or less, or an increase to 30% or more.

[13] Upon giving a wrong answer in the comprehension questions, subjects were informed that one or more of their choices was incorrect and that they needed to check the instructions again.

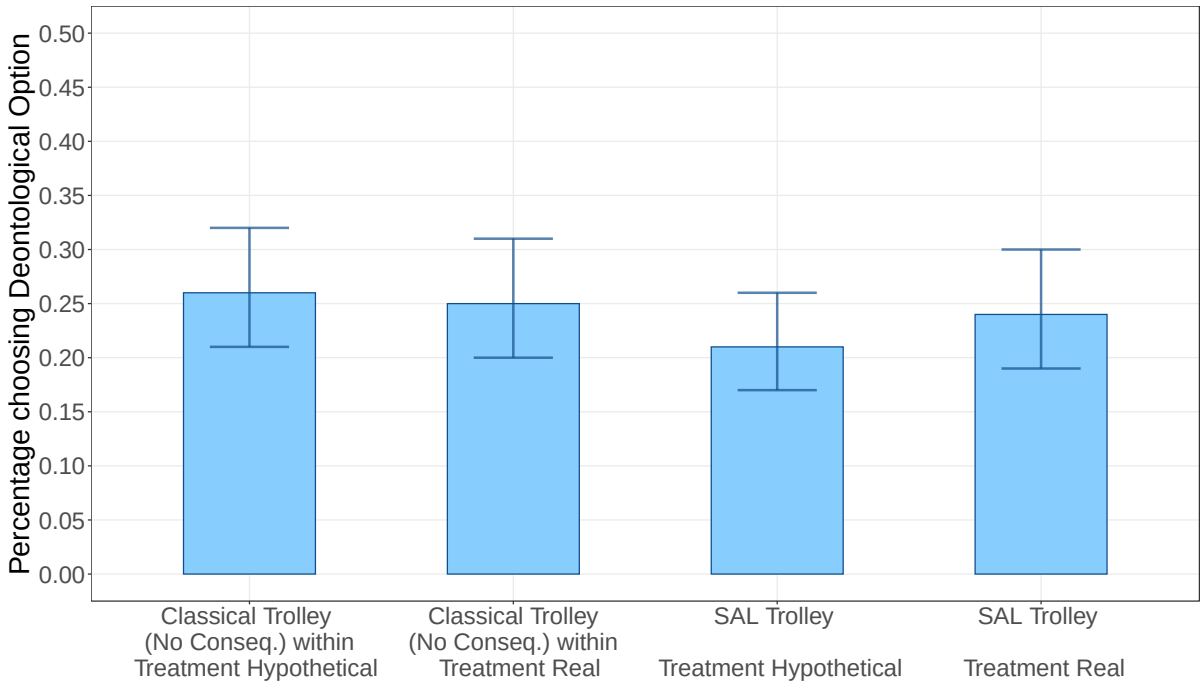
over a 15 Euros donation to the children-with-cancer charity in the baseline task (11 subjects), leaving us with 122 subjects for the robustness experiment.

3 Results

3.1 How prevalent are deontological and consequential decision-making?

We begin our empirical analysis by investigating the prevalence of the two types of decision-making. We first test whether deontologically motivated behavior is an artifact of hypothetical decisions and abstract situations, or whether subjects display similar choices in realistic situations with significant consequences, using the SAL trolley problem. We then investigate decision-making in our full set of ends-versus-means decision situations.

Figure 4: The effect of changing realism and consequences on behavior in the trolley dilemma



Notes: Each bar displays the fraction of subjects choosing the deontological option - not pulling the lever - in the trolley problem. The first two bars display behavior in the classical train-track version, the last two behavior in the Saving A Life paradigm. Treatment *Real* is that in which the SAL trolley has real consequences, whereas in *Hypothetical* decisions are without consequences. Error bars indicate 95% confidence intervals using one-sample tests of proportions.

Trolley. Figure 4 displays the trolley-problem results. In answering the classical train-track dilemma, 26% of subjects assigned to the *Hypothetical* treatment of the SAL-Trolley choose the deontological option – a fraction quite similar to what other studies investigating this version of the trolley have found.[14] An indistinguishable proportion of 25% among those assigned to the *Real* treatment made the same choice ($p = 0.82$, two-sample test of proportions). This result indicates that randomization across treatments was successful, because the treatments concern the SAL version of the trolley, not the train track version.

[14]For instance, across 42 countries and 70,000 participants in Awad et al. (2020), 19% choose the deontological option.

More interesting is the comparison between behavior in the classical trolley and our hypothetical SAL version. In total, 21% of subjects choose the deontological option in the SAL trolley, a small and insignificant difference compared to the classical trolley ($p = 0.17$, two-sample test of proportion). Thus, introducing a more realistic but still hypothetical setting does not appear to significantly influence decision-making.

Turning to the influence of introducing real consequences, we find that 24% of subjects in the *SAL-Real* treatment choose the deontological option. This fraction is not significantly different from that in the *Hypothetical* treatment ($p = 0.48$, two-sample test of proportions). Accordingly, real consequences neither substantially reduce nor increase the extent of deontological or consequential decision-making by our subjects. Overall, our treatments demonstrate that deontological behavior in sacrificial dilemmas is a robust phenomenon, being chosen by about one-fourth of respondents.

Appendix Section B reports a series of checks showing that our results are not driven by the preregistered exclusion restrictions, nor by potential confusion among subjects. The observed differences between classical and SAL-trolley, as well as between treatments, become even smaller using the full sample of subjects (Figure B.1), or only subjects who scored high on the SAL trolley comprehension test (Figure B.2). Another potential concern could be that our treatment with real consequences was too weak to induce behavioral effects, since only 1 in 10 subjects had their choices implemented. In Appendix Section B.2, we report the results of a robustness test addressing this concern. Using the same SAL paradigm but now with a self-versus-other tradeoff (taking money vs. triggering the donation), we find that implementing real consequences with a similar probability drastically changes the extent of prosocial decision-making, compared to a hypothetical situation. These results suggest that the absence of a real-consequences treatment effect in the trolley is specific to ends-versus-means situations, where two moral principles conflict.

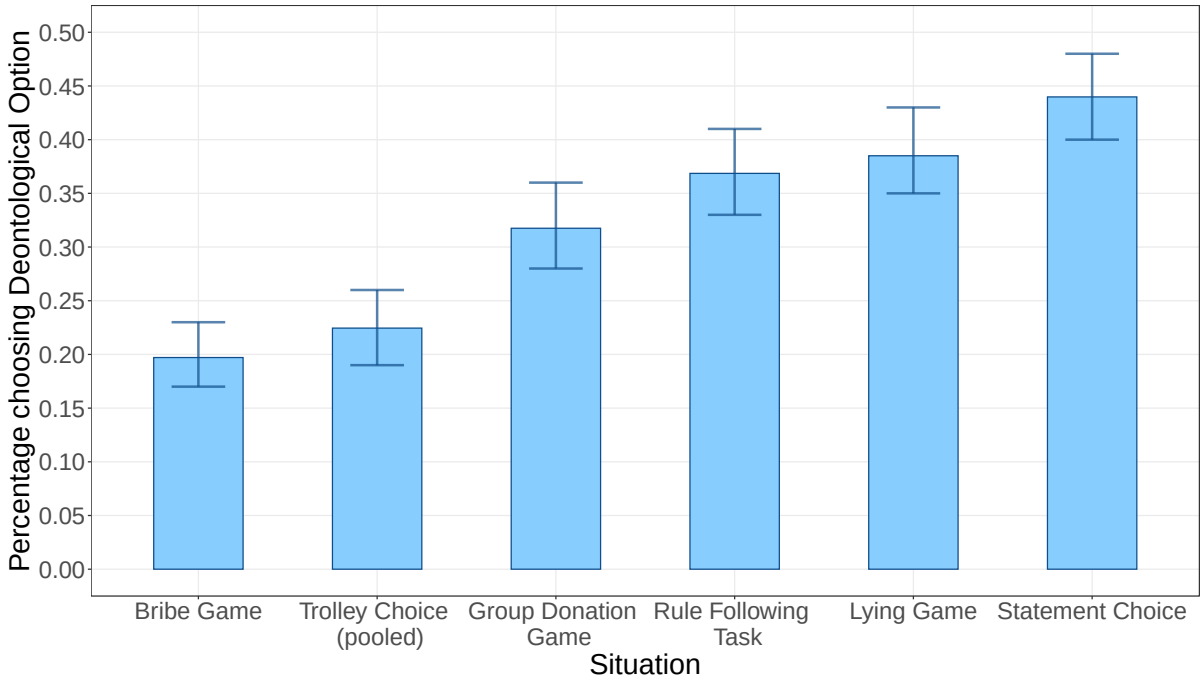
We also investigate whether our aggregate null finding might mask heterogeneity in the types of people choosing either option in the different versions. For instance, it could be that male subjects (say) are more likely to choose the deontological option when consequences are hypothetical, but less likely once real consequences are introduced. However, we find no evidence of heterogeneous treatment effects across a wide range of characteristics; see Appendix Figure B.3 for details.

Because we neither find differences in the aggregate nor in individual characteristics between treatments, we will pool decisions between treatments for the subsequent analysis. The results in the following sections remain unchanged if we use data from only one of the treatments.

Other ends-versus-means decisions. We now turn to decision-making in our other five EVM games. Figure 5 displays, for each game and for the (pooled) SAL trolley, the fraction of subjects choosing the deontological option. In every game, we find a substantial number of subjects making such a choice. Thus, 20% of subjects refuse to "bribe" the receiver to achieve the higher donation, 32% choose the inconsequential option in the group donation game, and 37% follow the rule in the rule-following task. In the lying game, 39% refuse to lie to the receiver, and 44% are unwilling to make the repugnant statement. These fractions are statistically different from zero in every instance ($p < 0.001$, one-sample test of proportion).

Could these results be driven by subjects' confusion about the consequences of their decisions? In our instructions, we extensively covered the implications of each situation, administered comprehension quizzes, and implemented a memory check for both the group-donation game and the statement

Figure 5: Distribution of Deontological and Consequentialist decision-making in ends-versus-means tradeoff situations



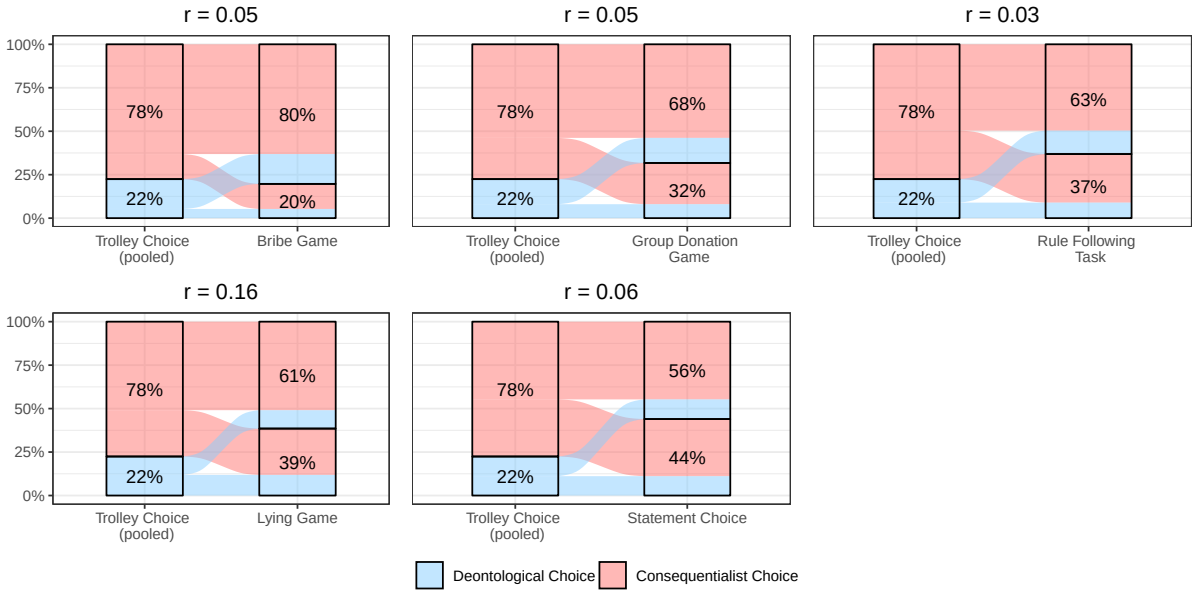
Notes: Each bar displays the fraction of subjects choosing the deontological option. See Section 2.1 for details on the situations and the definition of the deontological option. Error bars indicate 95% confidence intervals using one-sample tests of proportions.

choice. In total, 80% of subjects for the statement choice and 81% for the group-donation game correctly recalled the consequences at the end of the experiment, so they must have understood them when making the decision. Similar fractions were able to correctly answer all comprehension questions for the other games on their first try, and less than 5% needed more than one try. In Appendix Figure C.1, for each decision situation, we restrict the sample to subjects who answered all comprehension questions correctly on their first try and correctly answered the recall questions, respectively. We again see a significant fraction of subjects choosing the deontological option in every one of the six EVM decision situations, with the fraction of subjects choosing the deontological option ranging from 17% to 38%. Likewise, in Appendix Figure C.2, we replicate this finding among the full set of subjects without the preregistered exclusion restrictions, with fractions ranging from 20% to 45%.

To provide further evidence that subjects pay close attention to the actions and consequences of the tradeoffs we study, we turn to the *aligned* robustness experiment. If a significant fraction of subjects choosing non-consequential options in the original ends-versus-means situations was a result of random choice or errors, we should observe a similar pattern in the aligned counterparts, which share the same type of instructions and the same degree of complexity.

Instead, in line with our preregistered hypothesis, we observe in each case that almost all subjects choose the aligned option, from 88% in the lying game to 96% in the statement choice situation (see Appendix Figure C.3). Thus, as soon as the structure of the decision situations is changed so that both principles agree on what the moral choice is, almost everyone chooses the aligned option. Moreover, there is no indication that the games for which we find a high degree of non-consequential decision-making are those where subjects choose the non-aligned option to a higher degree. For instance,

Figure 6: Relationship of trolley behavior with behavior in the other ends-versus-means situations



Notes: Each diagram displays decision-making in the respective ends-versus-means game, conditional on behavior in the SAL-trolley dilemma. Percentages shown are the fractions of subjects choosing the deontological and consequential option in each case. Above each plot is the Pearson correlation coefficient between the SAL trolley and the respective decision situation displayed.

the statement choice has the highest degree of non-consequential choices (44%), yet coincidentally, the highest share of aligned choices (96%). Together with the fact that we observe similar behavior when focusing on high-comprehension subjects, these results support the view that the variations in the share of non-consequential choices between games are not a result of differential understanding of their instructions.

We conclude that the behavior we observe results from deliberate decision-making, leading us to state our first result.

Result 1. *Both consequential and deontological decision-making are pervasive in our experiment, with significant fractions (20% to 44%) of subjects behaving deontologically in ends-versus-means tradeoffs. Moreover, this behavior appears to be independent of whether the tradeoffs are embedded in a realistic or abstract setting, and whether they entail real consequences or not.*

3.2 How consistent are deontological and consequential decision-making?

Having established the *prevalence* of both moral principles in the previous section, we next investigate their *consistency* across decision situations. We first ask whether the trolley dilemma, as the principal one traditionally used to elicit moral intuitions, is *predictive* of behavior in other ends-versus-means situations. Figure 6 displays, in pairwise flow charts, how choices in the SAL trolley are related to those in the other EVM games. We find no evidence that trolley choices are predictive in other settings. In each case, large fractions of subjects switch from choosing the deontological option in the trolley to the consequential version in the other situation, and vice versa. This is also reflected in the correlations (Pearson correlation tests), which are close to zero and insignificant ($p > 0.1$)

Table 3: Correlation matrix for ends-versus-means decision situations

	Group Don. Game	Bribe Game	Statement Choice	Lying Game	Rule foll. Task
SAL Trolley	0.05 [-0.04, 0.13]	0.05 [-0.03, 0.14]	0.06 [-0.02, 0.14]	0.16 [0.08, 0.24]	0.03 [-0.05, 0.12]
Group Don. Game		0.14 [0.06, 0.23]	0.15 [0.07, 0.23]	-0.09 [-0.17, 0.00]	-0.06 [-0.14, 0.03]
Bribe Game			0.19 [0.11, 0.27]	0.04 [-0.04, 0.12]	-0.08 [-0.17, 0.00]
Statement Choice				-0.06 [-0.14, 0.02]	0.04 [-0.04, 0.12]
Lying Game					0.00 [-0.08, 0.09]

Notes: The table displays Pearson correlation coefficients between the six ends-versus-means decision situations. Brackets display 95% confidence intervals.

for all decision situations except the *lying game,* where a small but significant correlation does exist ($p < 0.01$).[15]

In Table 3, we extend the analysis to all pairwise correlations between the six EVM games. In each case, the indicator variable is equal to one if the subject chooses the deontological option (as defined in Section 2.1), and zero if they choose the consequential one. We find that no decision situation consistently predicts behavior, because no situation is significantly correlated with more than two others. Moreover, since we observe no correlation larger than 0.2 and sometimes even negative ones, there appears to be no detectable consistency in behavior in general. Instead, whether an individual makes the consequentialist or the deontological choice appears to be situation-specific.

A complementary way to analyze the consistency of choices is to examine the number of times a subject chooses either the consequential or deontological option; Figure 7 displays the results. If subjects behave consistently according to one of the two moral principles, we should observe a (skewed) U form. Instead, the empirical pattern displays an opposite, inverse-U shape. Essentially no subject chooses six times the deontological option, and only a small minority of just above 10% behave as standard economic agents, choosing six times the consequential option.

A potential concern might be that subjects being confused or paying limited attention could attenuate correlations between variables, artificially decreasing consistency. In Section 3.4 below, we address this issue in a series of robustness checks and show that our results remain unchanged when we exclude confused and inattentive subjects.

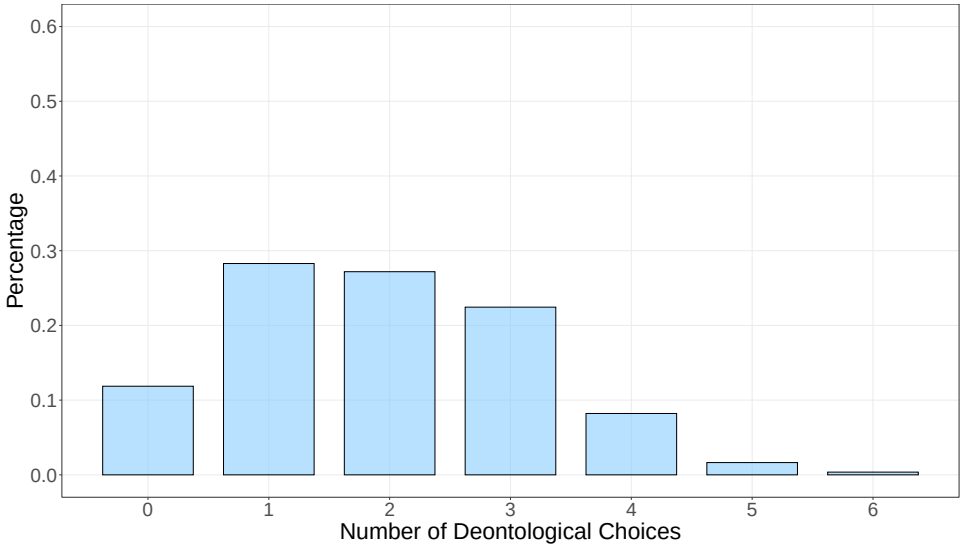
Result 2. *Subjects exhibit essentially no consistency in deontological/consequentialist choices across ends-versus-means tradeoff situations. Neither the trolley dilemma, nor any other ends-versus-means decision, consistently predicts behavior in the other EVM games.*

3.3 How are deontological and consequential decision-making related to prosocial behavior and morality-questionnaire measures?

We now turn to the comparison of behavior in ends-versus-means dilemmas with that in self-versus-other tradeoffs. We examine whether behavior in these two decision situation blocks is distinct, or

[15]Note that that these p-values are not corrected for multiple hypothesis testing. Applying such a correction would move the p-values even further away from significance.

Figure 7: Consistency of deontological and consequential decision-making



Notes: Histogram of the number of times a subject chose the deontological option in the ends-versus-means decision situations.

systematically related to EVM decision-making. Additionally, we analyze the relation of EVM decision-making with measures obtained through hypothetical questionnaires. Table 4 displays as main result the correlation matrix between these different measures.

Behavior within the self-versus-other block. We start by analyzing behavior within prosocial preferences. In stark contrast to behavior in EVM situations, we find that behavior in SVO decisions is rather consistent across games. We code variables so that higher values indicate higher degrees of giving to others, i.e., higher donations in the dictator game, more money sent in the trust game, etc. Almost all measures within this block are significantly correlated with each other, with most correlations falling between 0.30 and 0.40. For instance, altruism, as defined by the extent of giving to charity in a simple allocation game, consistently predicts the other eight measures, with correlations ranging from 0.20 to 0.65. In fact, the lowest correlation of altruism with any other measure is higher than the highest correlation within the ends-versus-means block. Thus, subjects show a relatively high degree of consistency in the way they resolve self-versus-other tradeoffs.[16]

Relation between ends-versus-means block and self-versus-other block. Next, we turn to comparing behavior between EVM and SVOs games. As displayed in the second part of rows in Table 4, we find that the two blocks are largely unrelated to each other. All pairwise correlations of measures between the two blocks are below 0.20, with the large majority being close to zero. Accordingly, behavior in one block does not predict behavior in the other. For example, subjects choosing the deontological option in the trolley donate on average 6.5 of their 20 Euros endowment to the charity, while those choosing the consequentialist option donate a mere 0.13 Euros more (0.7 pp.). We

[16]For altruism, trust, and both reciprocity measures, we can compare our individual-level correlations with those in Chapman et al. (2023). They report for altruism (measurement-error corrected) correlations of 0.34 with both reciprocity measures and 0.60 with trust. Their correlation coefficient of trust with both reciprocity measures is 0.49, and the correlation between low and high reciprocity is 0.86. Hence, with the exception of the relation of altruism with trust, our correlations are very similar to theirs, even though they use a representative sample of the US population while our sample contains mostly German students.

observe similarly minute differences for the other games. Hence, acting deontologically versus consequentially in any of our EVM blocks is not at all predictive of whether a subject is more or less prosocial, and vice versa.

Relation with hypothetical questionnaires. How is behavior in our incentivized decision situations correlated with measures obtained from hypothetical questionnaires? The third group of rows with the header "Questionnaires" in Table 4 displays our results. Focusing first on the relation between behavior in the SAL trolley and questionnaire measures, we find a high correlation of 0.50 with behavior in the classical trolley dilemma. This result confirms our earlier finding that sacrificial-dilemma choices are robust to variations in the displayed situation (real-life donations versus train track, real or hypothetical stakes). It is also further evidence that our findings are not just reflecting inattention by subjects or a predominance of noise in their decisions. We also find a substantial correlation of -0.43 with the Oxford Utilitarian scale instrumental-harm measure (OUS-IH), in the expected direction: subjects choosing the deontological option in the SAL trolley are less likely to endorse instrumental harm to obtain a greater good. Since the OUS-IH items describe situations purposefully designed to be similar to the trolley, a high correlation is reassuring. Next, replicating Kahane et al. (2018), we find a lower correlation of -0.21 of trolley behavior with the measure capturing subjects' instrumental benefit concerns (OUS-IB). The classical trolley and both OUS measures are also not predictive of behavior in any of the five other ends-versus-means decision situations with correlations not higher than 0.15, providing further evidence for the lack of consistency in behavior across this class of tradeoffs.

Behavior in the trolley, or any other EVM game, is also not correlated with any measure from the Moral Foundations Questionnaire (MFQ) or Moral Universalism scale (MU) scales. These scales appear to measure largely orthogonal dimensions of moral preferences or intuitions. Turning to self-versus-other situations, moral universalism (the slope of prosociality as social distance increases) is *negatively* correlated with the degree of prosociality identified in all SVO situations, albeit with low values.[17] The results for the different MFQ subscales are more mixed. The Harm and Fairness subscales, which capture the weight subjects put on values concerning protecting the weak and upholding equality, are correlated with the altruism and trust measure, but not with any other behavior. The other three subscales capturing subjects' weight on loyalty (MFQ Ingroup), authority, and purity appear to be behaviorally distinct from behavior in the SVO situations.

Principal component analysis. An important question is how well each of our experimental parts – the two decision blocks and the questionnaires - can be explained by a subset of variables. To this end, we perform a principal component analysis (PCA) separately on the ends-versus-means and the self-versus-other block. As a reference, we also perform a PCA on the Oxford Utilitarian scale. This type of analysis helps to assess the existence of components underlying subjects' behavior, complementing our previous analysis investigating pairwise correlations.

Figure 8 shows the results by displaying, for each block, the eigenvalues of the respective components. The higher the eigenvalue, the more variance is explained by the respective component. By construction, the first component captures the most variance, the second the most among the remaining variance and so on. We find that the eigenvalues of the EVM components are almost uniformly

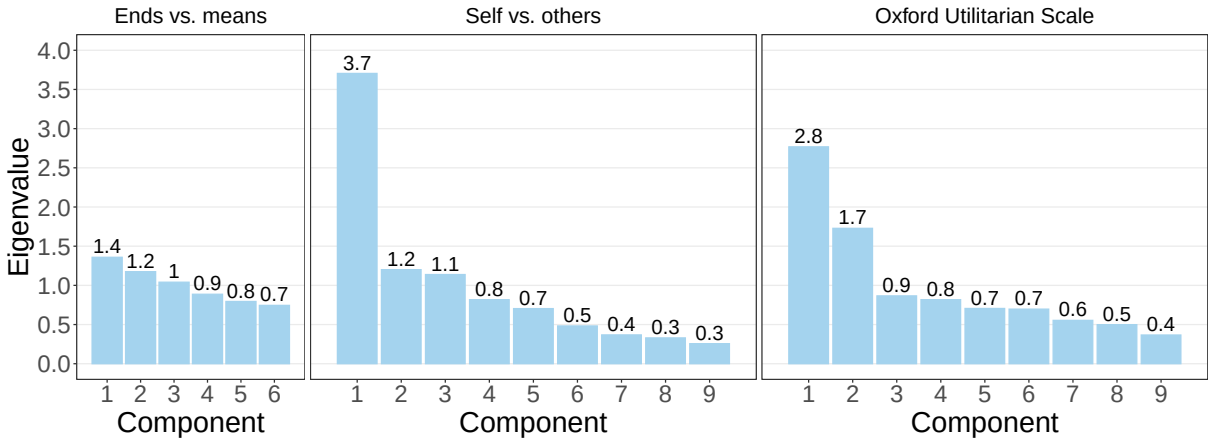
[17] Enke, Rodríguez-Padilla, and Zimmermann (2022) similarly find that moral universalism (weakly) negatively predicts the total amount given to various charities.

Table 4: Correlations of ends-versus-means, self-versus-other, and questionnaire measures

	Ends-versus-means decision situations						Self-versus-others decision situations								Questionnaires								
	SAL Trolley	Group Don.	Bribe	Statement Choice	Lying	Rule following	Altruism	Altruism Taking	Trust	Pos. Recipr. Low	Pos. Recipr. High	PGG Cont.	Reward Cons.- Int.	Reward Int.	Reward Cons.	Classical Trolley	OUS IB	OUS IH	MU Scale	MFQ Harm	MFQ Fairness	MFQ Ingroup	MFQ Authority	MFQ Purity
SAL Trolley		0.05	0.05	0.06	0.16	0.03	-0.01	-0.03	0.02	-0.05	-0.03	0.00	-0.02	0.05	-0.02	0.50	-0.12	-0.29	-0.04	-0.06	0.00	-0.06	0.00	0.05
Group Don.			0.14	0.15	-0.09	-0.06	0.05	0.03	-0.04	0.14	0.05	-0.02	0.09	-0.03	-0.02	0.02	0.09	0.02	-0.02	0.10	0.02	0.08	0.03	0.12
Bribe				0.19	0.04	-0.08	-0.04	-0.02	-0.11	-0.06	-0.15	-0.06	-0.06	-0.18	-0.18	-0.01	-0.03	-0.04	0.01	0.02	-0.04	0.08	0.05	0.14
Statement Choice					-0.06	0.04	0.00	0.01	-0.12	0.04	-0.07	0.00	0.05	-0.15	-0.04	-0.01	0.03	-0.07	-0.02	0.03	-0.03	0.15	0.13	0.17
Lying						0.00	-0.03	-0.04	-0.03	-0.03	0.00	-0.05	0.00	-0.02	-0.01	0.14	-0.15	-0.11	-0.05	-0.03	-0.02	-0.08	-0.04	-0.10
Rule following							0.11	0.06	0.03	0.11	0.10	0.03	0.11	-0.01	0.13	-0.02	0.08	0.00	-0.07	0.04	0.00	-0.08	-0.02	-0.02
Altruism								0.65	0.20	0.31	0.38	0.40	0.36	0.23	0.23	0.00	0.22	0.00	-0.13	0.19	0.17	0.01	-0.02	0.00
Altruism Taking									0.16	0.26	0.35	0.36	0.28	0.21	0.21	0.01	0.20	0.00	-0.13	0.20	0.15	0.00	-0.02	-0.04
Trust										0.35	0.45	0.47	0.38	0.30	0.23	0.00	0.05	0.02	-0.13	0.00	0.06	-0.06	-0.13	-0.17
Pos. Recipr. Low											0.71	0.36	0.35	0.13	0.14	-0.02	0.17	0.03	-0.16	0.08	0.09	-0.02	-0.08	-0.04
Pos. Recipr. High												0.45	0.41	0.30	0.21	0.01	0.18	0.04	-0.18	0.09	0.11	-0.12	-0.16	-0.15
PGG Cont.													0.40	0.31	0.22	0.03	0.11	-0.13	-0.26	0.06	0.12	-0.02	-0.07	-0.09
Reward Cons.- Int.														0.24	0.55	-0.02	0.08	-0.01	-0.15	0.06	0.07	-0.07	-0.12	-0.11
Reward Int.															0.40	0.07	0.02	0.00	-0.12	-0.01	0.04	-0.10	-0.08	-0.14
Reward Cons.																0.04	0.01	0.01	-0.04	0.05	0.06	-0.09	-0.08	-0.09
Classical Trolley																	-0.21	-0.43	-0.04	-0.04	0.01	-0.08	-0.06	-0.01
OUS IB																		0.23	-0.06	0.35	0.24	0.08	-0.04	0.05
OUS IH																			0.15	-0.03	-0.03	0.16	0.15	0.07
MU Scale																				0.02	0.02	0.22	0.16	0.20
MFQ Harm																					0.58	0.23	0.17	0.26
MFQ Fairness																						0.16	0.08	0.14
MFQ Ingroup																							0.57	0.51
MFQ Authority																								0.55
MFQ Purity																								

Notes: Colors in heatmap change with each 0.1 correlation magnitude.

Figure 8: Principal component analyses



Notes: The figure displays the results of principal component analyses conducted on the six ends-versus-means and nine self-versus-other decision situations, as well as the nine items of the Oxford Utilitarian scale. The y-axis shows the eigenvalues of each component. Higher eigenvalues indicate the respective component explains more variance in the data.

concentrated around one. In contrast, for the SVO block, the first component has a substantially higher eigenvalue than the other components. Accordingly, one component can explain the SVO choice well, whereas no such component exists for the ends-versus-means choices. This result is not specific to the fact that we have six EVM situations and nine SVO measures. We also conducted a PCA for every combination of six situations out of the nine SVO measures: in every one of the resulting 84 combinations, the variance explained by the first component of these six situations is twice as high as the first component can explain in the six EVM situations. The resulting eigenvalues from this exercise range from 2.36 to 3.02. Hence, no matter which SVO situations are taken, a dimensionality reduction always performs better among SVO tradeoffs than among EVM situations.

Result 3. *Subjects exhibit a relatively high degree of consistency across self-versus-other tradeoff situations. Behavior in these prosociality decisions is unrelated to behavior in ends-versus-means tradeoff situations.*

3.4 Robustness of correlations between blocks

Confusion or limited attention. As mentioned previously, a potential concern might be that noise due to subjects' confusion or limited attention could drive some of our results, especially the lack of consistency among the ends-versus-means situations. Note, however, that we do observe a relatively high degree of consistency among the variables in the self-versus-other block. Furthermore, behavior in our SAL version of the trolley is well correlated with psychological questionnaires designed to capture such reasoning. Therefore, a general lack of attention or misunderstanding among subjects is unlikely to explain our results. Consider next the possibility of differential confusion between the blocks – subjects having more problems understanding the EVM decision situations than the SVO ones. Instead, we find that subjects make more errors on average in the comprehension questions of the SVO block than in those of the EVM block, suggesting that their understanding is actually higher in the latter.

To further address these concerns, we conduct a series of robustness checks. First, we exclude subjects who fail more than one of the comprehension quizzes (Appendix Table D.1). Second, we

exclude those who fail one or both of the attention-memory checks we employed for the Group Donation and Statement Choice situation (Appendix Table D.2). Third, we exclude the Top 15% fastest subjects (Appendix Table D.3). In every case, we replicate our main findings: almost zero correlations within the ends-versus-means block, high correlations within the SVO block, and very low correlations between the two blocks.

Measurement. The variables in the EVM block are dichotomous, while those in the SVO block are more fine-grained. Accordingly, relying on standard Pearson correlation coefficients could potentially confound the differences we find between the two blocks, as well as those within the first one relative to the second.

To alleviate this concern, we show that the documented consistency within the SVO block prevails even when all its variables are dichotomized, in two different ways. First, we split subjects based on whether they fully behave according to the classical model of maximizing self-interest. That is, if a subject donates nothing to the charity, does not cooperate at all in the public goods game, sends nothing in the trust game, etc., we code the respective variable as zero, and one otherwise. Second, we categorize subjects based on a median split for each respective variable. For example, the altruism variable is then zero if a subject donates less than the median, and one if they donate more. Under both approaches, we again find a high degree of consistency within the SVO block, and no correlation with the ends-versus-means block; see Appendix Table D.4 for the first approach, and D.5 for the second.

To further verify that the low correlations between the two blocks are not an artifact of our measurements, we also compute point-biserial correlations, which are designed for pairwise comparisons between a binary and a continuous variable. Again, we find low correlations between the two blocks; with no apparent pattern, see Appendix Table D.6 for details. Hence, our main result also appears not to be driven by measurement.

4 Conclusion

Our results bring substantial nuance to the classical dichotomy opposing Consequentialism and Deontologism. On one hand, deontological decision-making is a robust phenomenon, not an artifact of hypothetical choices or unrealistic dilemmas: we find it in a real-stakes version of the trolley and in five other games pitting socially desirable ends against arguably (and, for our subjects, *de facto*) objectionable means. Treating all economic agents as solely consequentialist thus appears to be an inaccurate representation of their preferences and decisions.

On the other hand, in our subject population, there appear to be no consistently deontological "types" and only a small fraction of consistently consequentialist ones. Most individuals' choices across the six ends-versus-means dilemmas are largely uncorrelated, which is at odds with any stable mixture of two preference types, invariant across choice situations. Subjects clearly feel the pull of both moral principles, but appear to resolve it in a very situation-specific manner.

The high degree of context-specific behavior in ends-versus-means tradeoffs means that for such tradeoffs, findings in one setting may not generalize to other settings. Moreover, behavior in one setting cannot readily be used to predict behavior in other settings. As we show, this holds particularly for the workhorse dilemma used in the literature to understand moral behavior – the trolley

dilemma. As a consequence, in order to understand behavior in a given setting, behavior in the specific setting needs to be measured. For instance, in order to understand how people resolve ends-versus-means tradeoffs in the context of autonomous driving, it is not sufficient to measure behavior in other settings and extrapolate. In contrast, this is less of an issue in self-versus-other tradeoffs. Here, the workhorse game used in the literature – the dictator game – predicts behavior in other self-versus-other settings well. As such, extrapolation to other settings or decision situations appears to be possible. For instance, experimentally measured altruism predicting prosocial behavior in the field (e.g., Rustagi, Engel, and Kosfeld, 2010; Falk et al., 2018; Kosse and Tincani, 2020; Chen et al., 2021; Lades, Laffan, and Weber, 2021).

As for potential explanations for the lack of consistency between ends-versus-means tradeoffs, subjects act as if they had very flat priors on what their preference "should" be, and even minor contextual signals or cues could sway them. These results point to interesting avenues for further research, both experimental and theoretical – to better understand what seems to be a new type of internal preference conflict. Clearly, behavior in such tradeoffs is not well captured by existing economic preferences or well-behaved preference orderings, for instance those orderings used for self-versus-other tradeoffs. Modelling such behavior may require a more fundamental departure from standard economic preference models.

On the methodological side, the paper introduces a new experimental paradigm – ends-versus-means games, as fundamentally distinct from the self-versus-other games used to measure prosociality. It is quite flexible, and in particular independent of the trolley dilemma: while another contribution of the paper is to make trolley choices "real," this particular decision task could be dropped without changing the general architecture of the paradigm, or the overall results found in our study. Similarly, a researcher could replace one or more of our proposed tests of Consequentialism vs. Deontologism (lying game, bribing game, non-pivotal group choice, aversive statement, rule following) with some variant that they prefer. The methodology of examining the consistency of decisions within that ends-versus-means block, and potential correlations with behavior in standard prosociality games, would remain the same.

References

Alexander, Larry, and Michael Moore (2016). *Deontological Ethics*. Edward N. Zalta, ed. The Stanford Encyclopedia of Philosophy. [1]

Alger, Ingela, and Jörgen W. Weibull (2013). "Homo Moralis–Preference Evolution Under Incomplete Information and Assortative Matching". *Econometrica* 81 (6): 2269–302. [4]

Alger, Ingela, and Jörgen W. Weibull (2016). "Evolution and Kantian Morality". *Games and Economic Behavior* 98 (7): 56–67. [4]

Alger, Ingela, Jörgen W. Weibull, and Laurent Lehmann (2020). "Evolution of preferences in structured populations: Genes, guns, and culture". *Journal of Economic Theory* 185: 1–52. [4]

Awad, Edmond, Sohan Dsouza, Richard Kim, Jonathan Schulz, Joseph Henrich, Azim Shariff, Jean François Bonnefon, and Iyad Rahwan (2018). "The Moral Machine experiment". *Nature* 563 (7729): 59–64. [2]

Awad, Edmond, Sohan Dsouza, Azim Shariff, Iyad Rahwan, and Jean François Bonnefon (2020). "Universals and variations in moral decisions made in 42 countries by 70,000 participants". *Proceedings of the National Academy of Sciences of the United States of America* 117 (5): 2332–37. [5, 16]

Bartels, Daniel M. (2008). "Principled moral sentiment and the flexibility of moral judgment and decision making". *Cognition* 108 (2): 381–417. [5]

Bauman, Christopher W., A. Peter Mcgraw, Daniel M. Bartels, and Caleb Warren (2014). "Revisiting external validity: Concerns about trolley problems and other sacrificial dilemmas in moral psychology". *Social and Personality Psychology Compass* 8 (9): 536–54. [6]

Bénabou, Roland, Armin Falk, Luca Henkel, and Jean Tirole (2023). "Eliciting Moral Preferences under Image Concerns: Theory and Experiment". *Working Paper*, [4, 6, 11, 34]

Bénabou, Roland, and Jean Tirole (2011). "Identity, morals, and taboos: Beliefs as assets". *Quarterly Journal of Economics* 126 (2): 805–55. [4, 11]

Bentham, Jeremy (1789). *An Introduction to the Principles of Morals*. London: Athlone. [1]

Bock, Olaf, Ingmar Baetge, and Andreas Nicklisch (2014). "Hroot: Hamburg Registration and Organization Online Tool". *European Economic Review* 71: 117–20. [15]

Bostyn, Dries H., Sybren Sevenhant, and Arne Roets (2018). "Of Mice, Men, and Trolleys: Hypothetical Judgment Versus Real-Life Behavior in Trolley-Style Moral Dilemmas". *Psychological Science* 29 (7): 1084–93. [5]

Bursztyn, Leonardo, Bruno Ferman, Ali Hasanain, Michael Callen, Saad Gulzar, and Noam Yuchtman (2020). "Political identity: Experimental evidence on Anti-Americanism in Pakistan". *Journal of the European Economic Association* 18 (5): 2532–60. [11]

Chapman, Jonathan, Mark Dean, Pietro Ortoleva, Erik Snowberg, and Colin Camerer (2023). "Econographics". *Journal of Political Economy Microeconomics* 1 (1): 115–61. [4, 21]

Chen, Daniel L., and Martin Schonger (2022). "Social preferences or sacred values? Theory and evidence of deontological motivations". *Science Advances* 8 (19): [4]

Chen, Daniel L., Martin Schonger, and Chris Wickens (2016). "oTree-An open-source platform for laboratory, online, and field experiments". *Journal of Behavioral and Experimental Finance* 9: 88–97. [15]

Chen, Zihua, Ximeng Fang, Timo Freyer, Lorenz Goette, and Chui Yee Ho (2021). "Prosociality predicts individual behavior and collective outcomes in the COVID-19 pandemic". *Working Paper*, [26]

Costa, Albert, Alice Foucart, Sayuri Hayakawa, Melina Aparici, Jose Apesteguia, Joy Heafner, and Boaz Keysar (2014). "Your morals depend on language". *PLoS ONE* 9 (4): [5]

Dean, Mark, and Pietro Ortoleva (2019). "The empirical relationship between nonstandard economic behaviors". *Proceedings of the National Academy of Sciences* 116 (33): 16262–67. [4]

Enke, Benjamin, Ricardo Rodríguez-Padilla, and Florian Zimmermann (2022). "Moral Universalism: Measurement and Economic Relevance". *Management Science* 68 (5): 3590–603. [3, 14, 22]

Erat, Sanjiv, and Uri Gneezy (2012). "White lies". *Management Science* 58 (4): 723–33. [9]

Everett, Jim A.C., Nadira S. Faber, Julian Savulescu, and Molly J. Crockett (2018). "The costs of being consequentialist: Social inference from instrumental harm and impartial beneficence". *Journal of Experimental Social Psychology* 79 (August): 200–16. [3]

Everett, Jim A.C., David A. Pizarro, and Molly J. Crockett (2016). "Inference of Trustworthiness From Intuitive Moral Judgments". *Journal of Experimental Psychology: General* 145 (6): 772–87. [3]

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde (2018). "Global Evidence on Economic Preferences". *Quarterly Journal of Economics* 133 (4): 1645–92. [4, 26]

Falk, Armin, Ernst Fehr, and Urs Fischbacher (2003). "On the nature of fair behavior". *Economic Inquiry* 41 (1): 20–26. [4]

Falk, Armin, Ernst Fehr, and Urs Fischbacher (2008). "Testing theories of fairness-Intentions matter". *Games and Economic Behavior* 62 (1): 287–303. [4]

Falk, Armin, and Thomas Graeber (2020). "Delayed Negative Effects of Prosocial Spending on Happiness". *Proceedings of the National Academy of Sciences* 117 (12): 6463–68. [6]

Falk, Armin, Thomas Neuber, and Nora Szech (2020). "Diffusion of being pivotal and immoral outcomes". *Review of Economic Studies* 87 (5): 2205–29. [11]

Feddersen, Timothy, Sean Gailmard, and Alvaro Sandroni (2009). "Moral bias in large elections: Theory and experimental evidence". *American Political Science Review* 103 (2): 175–92. [4]

Feddersen, Timothy, and Alvaro Sandroni (2006). "A Theory of Participation in Elections". *American Economic Review* 96 (4): 1271–82. [4]

Feess, Eberhard, Florian Kerzenmacher, and Yuriy Timofeyev (2022). "Utilitarian or Deontological Models of Moral Behavior - What Predicts Morally Questionable Decisions?" *European Economic Review* 149 (forthcoming): 104264. [4]

Foot, Philippa (1967). "The Problem of Abortion and the Doctrine of the Double Effect". *Oxford Review* 5: 5–15. [2]

Gawronski, Bertram, and Jennifer S. Beer (2017). "What makes moral dilemma judgments "utilitarian" or "deontological"?" *Social Neuroscience* 12 (6): 626–32. [5]

Gneezy, Uri (2005). "Deception: The role of consequences". *American Economic Review* 95 (1): 384–94. [8]

Gold, Natalie, Briony D. Pulford, and Andrew M. Colman (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. [5]

Graham, Jesse, Brian A. Nosek, Jonathan Haidt, Ravi Iyer, Spassena Koleva, and Peter H. Ditto (2011). "Mapping the Moral Domain". *Journal of Personality and Social Psychology* 101 (2): 366–85. [2, 14]

Greene, Joshua D., R. Brian Sommerville, Leigh E. Nystrom, John M. Darley, and Jonathan D. Cohen (2001). "An fMRI investigation of emotional engagement in moral judgment". *Science* 293 (5537): 2105–8. [5]

Hauser, Marc, Fiery Cushman, Liane Young, R. Kang-Xing Jin, and John Mikhail (2007). "A dissociation between moral judgments and justifications". *Mind and Language* 22 (1): 1–21. [5]

Kahane, Guy, Jim A.C. Everett, Brian D. Earp, Lucius Caviola, Nadira S. Faber, Molly J. Crockett, and Julian Savulescu (2018). "Beyond sacrificial harm: A two-dimensional model of utilitarian psychology". *Psychological Review* 125 (2): 131–64. [2, 14, 22]

Kant, Immanuel (1785). *Grundlegung zur Metaphysik der Sitten*. J. F. Hartknoch. [1, 4, 9]

Kimbrough, Erik O., and Alexander Vostroknutov (2016). "Norms Make Preferences Social". *Journal of the European Economic Association* 14 (3): 608–38. [12]

Kosse, Fabian, and Michela M. Tincani (2020). "Prosociality Predicts Labor Market Success around the World". *Nature Communications* 11 (1): 5298. [26]

Lades, Leonhard K., Kate Laffan, and Till O. Weber (2021). "Do Economic Preferences Predict Pro-Environmental Behaviour?" *Ecological Economics* 183 (5): 106977. [26]

Laffont, Jean-Jacques (1975). "Macroeconomic Constraints, Economic Efficiency and Ethics: An Introduction to Kantian Economics". *Economica* 42 (168): 430. [4]

Lanteri, Alessandro, Chiara Chelini, and Salvatore Rizzello (2008). "An experimental investigation of emotions and reasoning in the trolley problem". *Journal of Business Ethics* 83 (4): 789–804. [5]

Mill, John Stuart (1863). *Utilitarianism*. Son & Bourn Parker, ed. 1 ed. London: West Strand. [1]

Rai, Tage S., and Keith J. Holyoak (2010). "Moral principles or consumer preferences? Alternative framings of the trolley problem". *Cognitive Science* 34 (2): 311–21. [5]

Roemer, John E. (2010). "Kantian equilibrium". *Scandinavian Journal of Economics* 112 (1): 1–24. [4]

Rustagi, Devesh, Stefanie Engel, and Michael Kosfeld (2010). "Conditional Cooperation and Costly Monitoring Explain Success in Forest Commons Management". *Science* 330 (6006): 961–65. [26]

Sinnott-Armstrong, Walter (2019). *Consequentialism*. Edward N. Zalta, ed. The Stanford Encyclopedia of Philosophy. [1]

Stango, Victor, and Jonathan Zinman (2023). "We Are All Behavioral, More or Less: A Taxonomy of Consumer Decision Making". *Review of Economic Studies* 90 (3): 1470–98. [5]

Thompson, Judith Jarvis (1976). "Killing, Letting Die, and the Trolley Problem". *Monist* 59 (2): 204–17. [2]

Van Leeuwen, Boris, and Ingela Alger (2021). "Estimating Social Preferences and Kantian Morality in Strategic Interactions". *Working Paper*, [4]

Appendix

A Details on self versus others decision situations

A.1 Dictator game

Subjects play two variations of a dictator game, a giving and a taking game. In each, they can allocate 20 Euros between themselves and a charity helping children suffering from cancer. In the taking game, 20 Euros are to be donated to the charity, and subjects can decide to take money for themselves. In the giving game, they are endowed with 20 Euros and can donate an amount to the charity.

A.2 Trust game

Subjects play both roles of the standard trust game with an endowment of 5 Euros and a multiplication factor of 3. Senders can send integers from 0 to 5 Euros. Recipient choices are elicited using the strategy method, i.e., subjects decide how much to send back conditionally on each of the six possible amounts send.

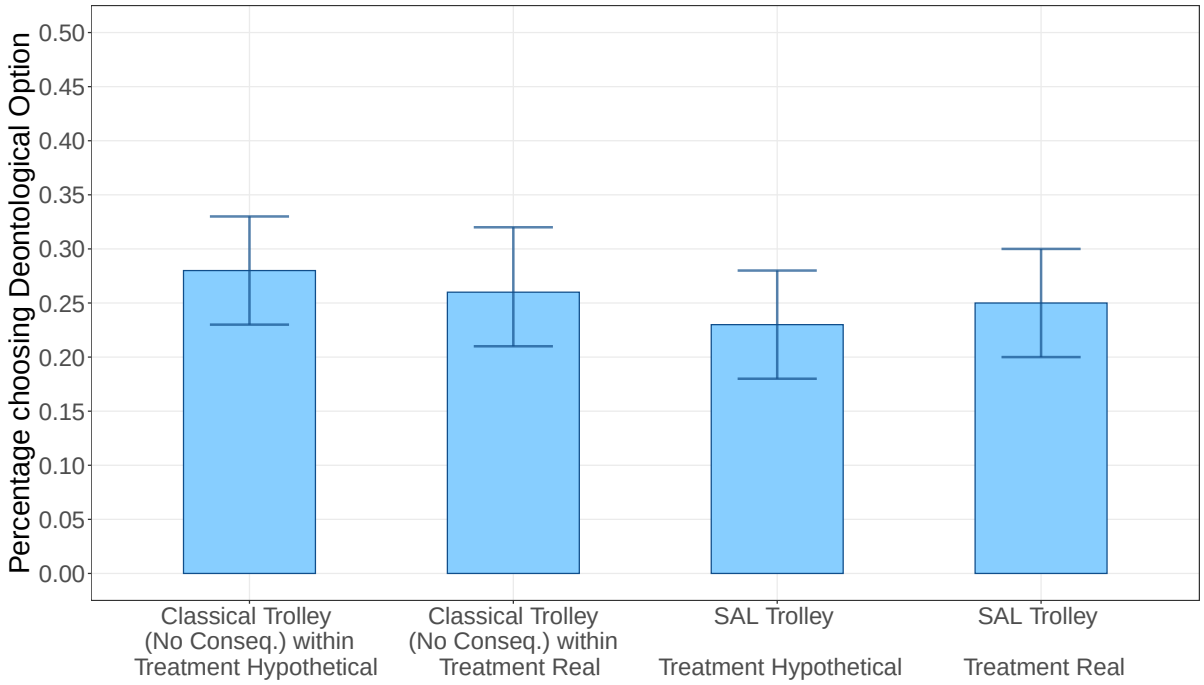
A.3 Public goods game

Subjects play the standard public goods game with an endowment of 5 Euros, a group size of 3, and a multiplication factor of 1.5.

B Robustness analysis trolley dilemma

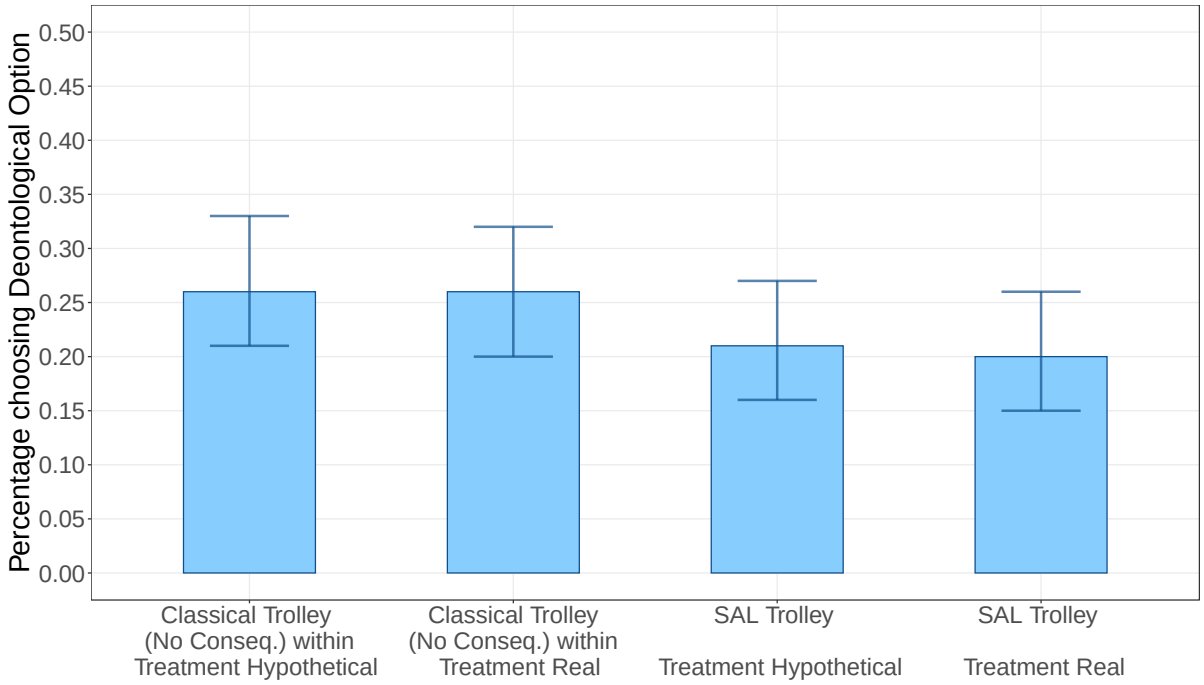
B.1 Figures

Figure B.1: The effect of changing realism and consequences on behavior in the trolley dilemma using the full sample



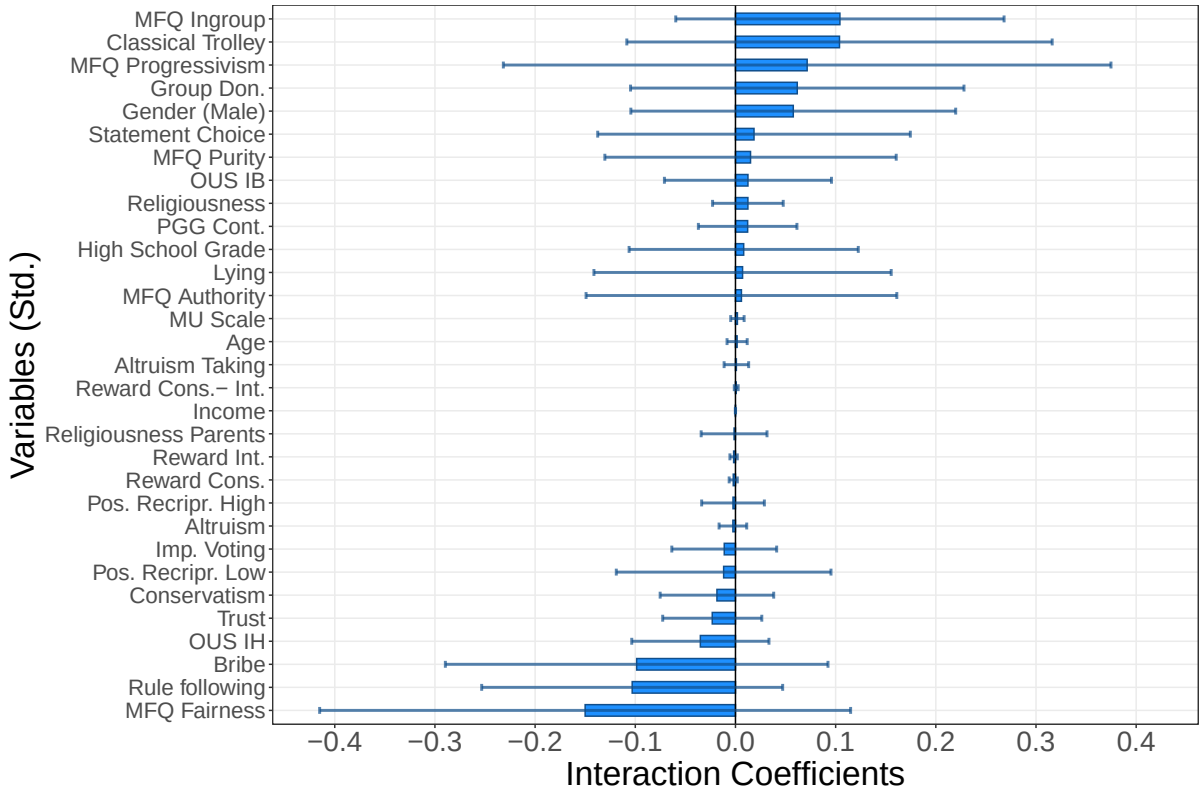
Notes: Each bar displays the fraction of subjects choosing the deontological option - not switching the lever - in the trolley problem using the full sample of 593 subjects. The first two bars display behavior in the classical train track trolley, the last two behavior using the Saving A Life paradigm. Treatment *Real* indicates the treatment in which the SAL trolley has real consequences, in treatment *Hypothetical* are decisions without consequences. Error bars indicate 95% confidence intervals using one-sample tests of proportions.

Figure B.2: The effect of changing realism and consequences on behavior in the trolley dilemma among high comprehension subjects



Notes: Each bar displays the fraction of subjects choosing the deontological option - not switching the lever - in the trolley problem using only subjects that make no more than one mistake in the trolley comprehension quiz. The first two bars display behavior in the classical train track trolley, the last two behavior using the Saving A Life paradigm. Treatment *Real* indicates the treatment in which the SAL trolley has real consequences, in treatment *Hypothetical* are decisions without consequences. Error bars indicate 95% confidence intervals using one-sample tests of proportions.

Figure B.3: Differences in characteristics of subjects between the *Hypothetical* and *Real* treatments



Notes: The graph displays interactions term coefficients obtained from an OLS regression with an indicator that equals 1 for observations from the Real treatment. Bars indicate 95% confidence intervals constructed from robust standard errors.

B.2 Robustness experiment

This section describes a robustness experiment addressing the concern that the probabilistic implementation employed in the treatment *Real* introduced an insufficient real consequence incentive. Specifically, in treatment *Real*, each choice in the Saving a Life trolley has a 10% chance of being implemented with real consequences. In comparison, in treatment *Hypothetical*, all choices had no real consequences. We find that subjects' behavior does not differ between the two treatments. One potential explanation is that subjects perceive a 10% implementation probability as essentially hypothetical, rendering the treatment comparison problematic. In the robustness experiment, we show that subjects are highly responsive to a chance from 0% implementation probability to a positive probability in self-versus-other tradeoffs. We use the design and results of one treatment of Bénabou et al. (2023), to which we add a hypothetical treatment.

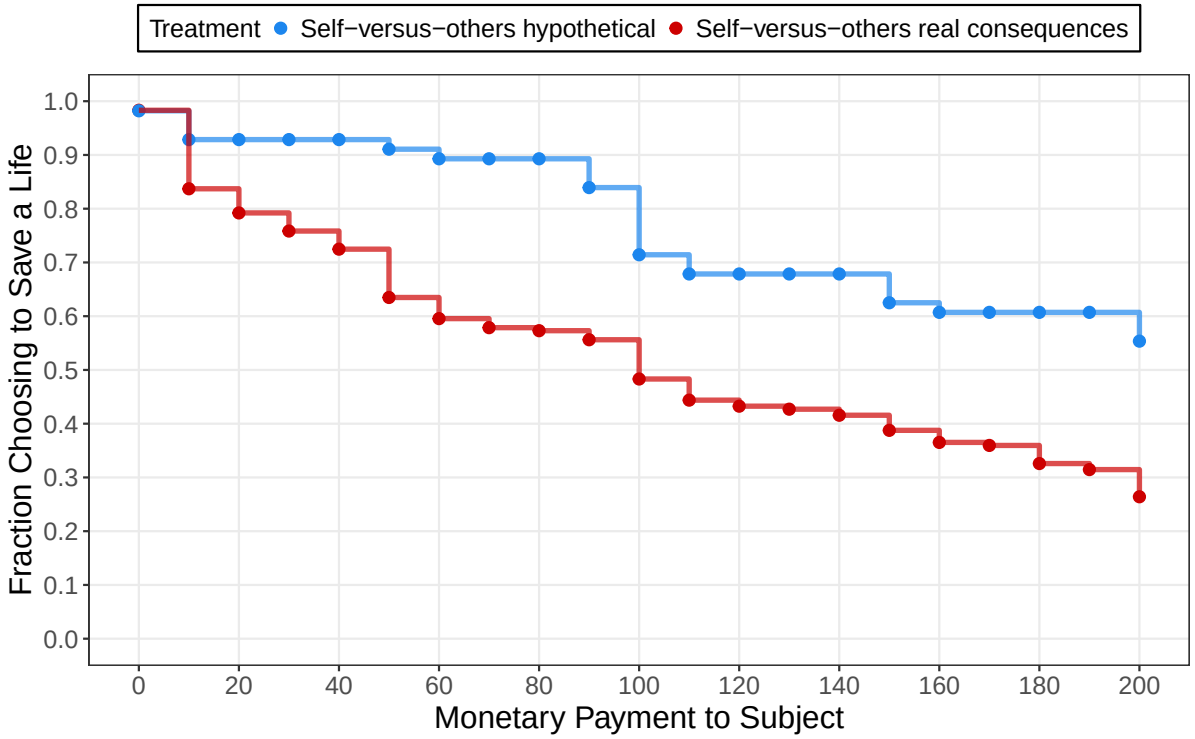
Design. We used the same Saving a Life paradigm as employed in the main experiment. However, instead of the trolley's ends-versus-means tradeoff, we employed a self-versus-other tradeoff: subjects could choose whether to trigger a donation sufficient to save one human life or take money for themselves as additional payment. Subjects faced this tradeoff multiple times in the form of a price-list design. In 21 contingent choices starting with 0 Euro and proceeding in 10 Euros increments up to 200 Euros, subjects could indicate whether they want to save a life or take the respective money for themselves.

Treatment. There were two between-subject treatments. In the *self-versus-other real* treatment[18], subjects' choices had the chance to be implemented with real consequences. Out of each session (consisting of 20-24 subjects), two subjects were drawn. For these selected subjects, one price from the price list was randomly drawn, and their pre-stated choices for the drawn price were implemented. Therefore, in this treatment, with an 8% to 10% chance, subjects either triggered the donation or received up to 200 Euros through their decisions. We subsequently ran another treatment, called *self-versus-other hypothetical*, in which subjects' choices were purely hypothetical. In total, 178 subjects took part in the *self-versus-other real* treatment (Bénabou et al., 2023), and we recruited 56 subjects for the treatment *self-versus-other hypothetical*.

Results. Introducing real consequences to this self-versus-other tradeoff resulted in significantly different choice distributions ($p < 0.001$, Kolmogorov-Smirnov test). As displayed in Figure B.4, for all prices except 0 and 10 Euros, real consequences lead to a significant decrease in the likelihood that subjects choose to save a life instead of taking the offered amount. For instance, at 100 Euros, 71% of subjects choose to save a life when consequences are hypothetical, while only 48% do so when choices might have real consequences ($p = 0.003$, two-sided Fisher's exact test). Consequently, introducing real consequences through a probabilistic payment system leads to markedly different results in self-versus-other tradeoffs.

[18] In Bénabou et al. (2023), this treatment is called *MPL Low Image*.

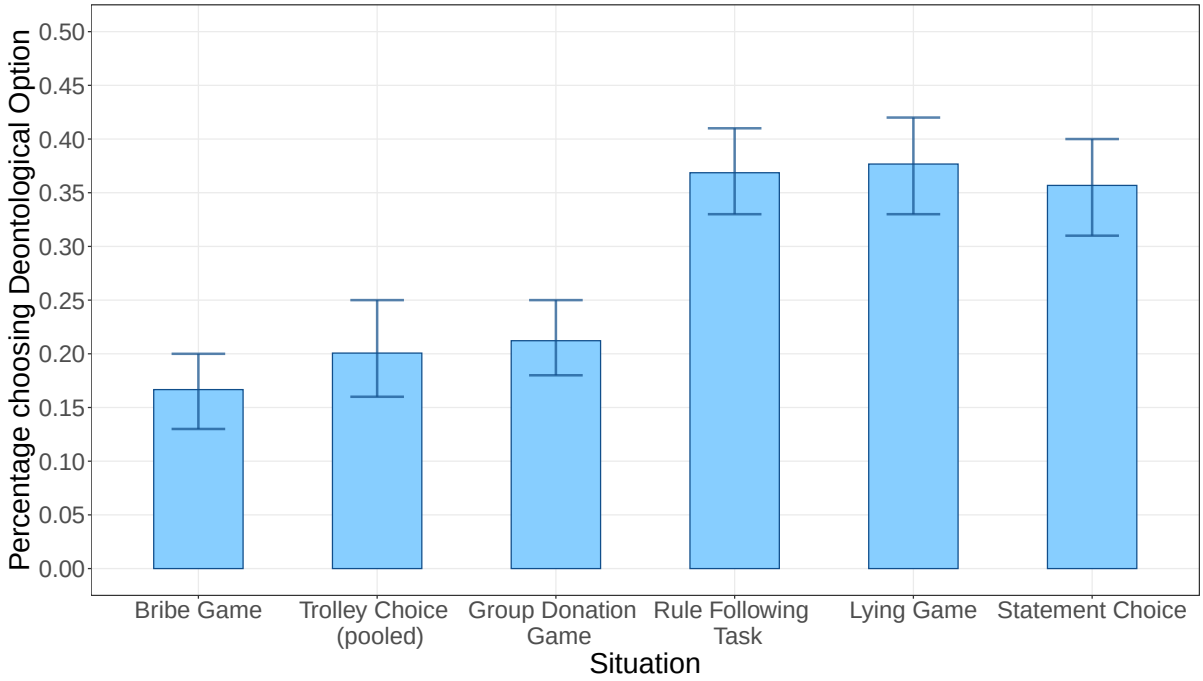
Figure B.4: The effect of changing consequences on behavior in self versus other tradeoff situations



Notes: The figure displays the fraction of subjects choosing to save a life for each offered price separately for the *self-versus-other hypothetical* and *self-versus-other real consequences* treatments.

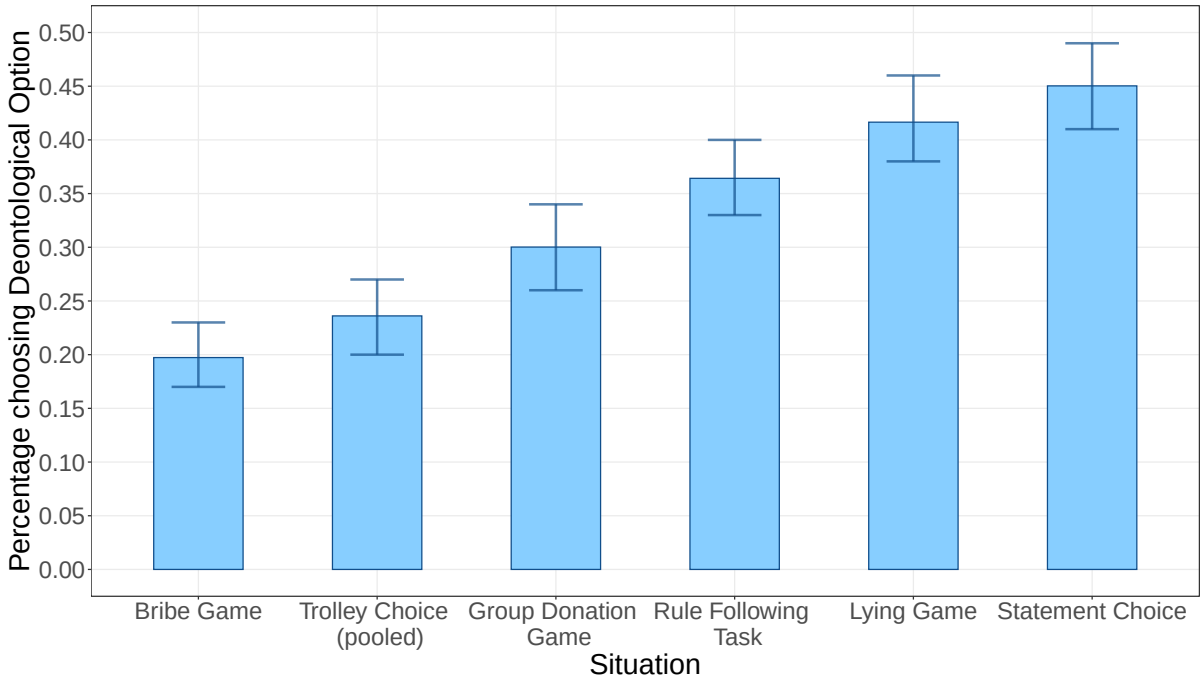
C Robustness analysis ends-versus-means decisions

Figure C.1: Distribution of Deontological and Consequentialist decision-making in ends-versus-means tradeoff situations among high comprehension subjects



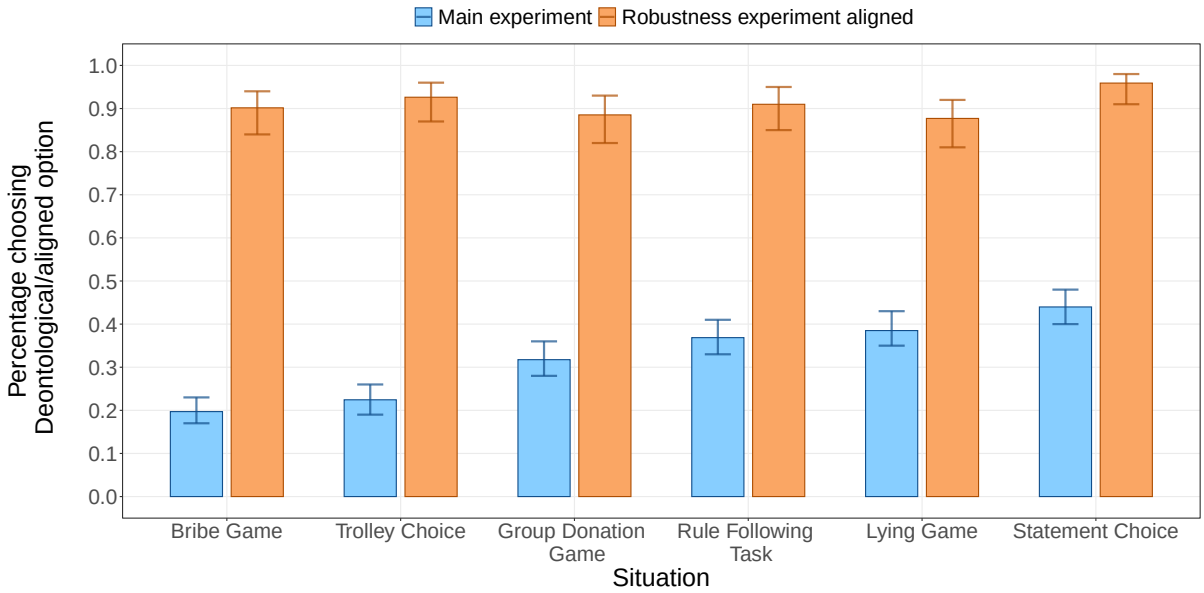
Notes: Each bar displays the fraction of subjects choosing the deontological option in ends-versus-means tradeoff situations using only subjects that answer all comprehension and recall questions correctly. See Section 2.1 for details on the situations and the definition of the deontological option. Error bars indicate 95% confidence intervals using one-sample tests of proportions.

Figure C.2: Distribution of Deontological and Consequentialist decision-making in ends-versus-means tradeoff situations using the full sample of subjects



Notes: Each bar displays the fraction of subjects choosing the deontological option in ends-versus-means tradeoff situations using the full sample of 593 subjects. See Section 2.1 for details on the situations and the definition of the deontological option. Error bars indicate 95% confidence intervals using one-sample tests of proportions.

Figure C.3: Distribution of choices in the ends-versus-means tradeoff situations and aligned situations of the robustness experiment



Notes: Each bar displays the fraction of subjects choosing the deontological option in the main experiment (left bar in blue) and the fraction of subjects choosing the aligned choice in the robustness experiment (right bar in orange). See Section 2.1 for details on the situations and the definition of the deontological and aligned option. Error bars indicate 95% confidence intervals using one-sample tests of proportions.

D Robustness between blocks

Table D.1: Correlations of ends-versus-means, self-versus-other and questionnaire measures comprehension sample

	Ends-versus-means decision situations						Self-versus-others decision situations								Questionnaires								
	SAL Trolley	Group Don.	Bribe	Statement Choice	Lying	Rule following	Altruism	Altruism Taking	Trust	Pos. Recripr. Low	Pos. Recripr. High	PGG Cont.	Reward Cons.- Int.	Reward Int.	Reward Cons.	Classical Trolley	OUS IB	OUS IH	MU Scale	MFQ Harm	MFQ Fairness	MFQ Ingroup	MFQ Authority	MFQ Purity
SAL Trolley		0.00	-0.02	0.05	0.17	0.10	-0.01	0.00	0.01	-0.07	-0.05	0.00	0.00	0.03	-0.03	0.56	-0.13	-0.30	0.00	-0.07	-0.05	-0.04	0.02	0.06
Group Don.			0.08	0.00	-0.14	-0.05	0.03	0.05	0.06	0.13	0.08	0.01	0.14	0.04	0.07	0.02	0.11	0.03	-0.01	0.10	0.00	0.02	-0.04	0.06
Bribe				0.09	0.04	-0.08	0.00	0.01	-0.07	-0.01	-0.08	-0.07	-0.09	-0.08	-0.11	0.01	-0.07	0.07	0.00	-0.06	-0.07	0.03	0.00	0.06
Statement Choice					-0.08	0.07	0.03	0.05	-0.13	0.02	-0.03	0.00	0.03	-0.12	-0.06	0.01	0.04	-0.15	-0.08	0.03	0.05	0.09	0.06	0.10
Lying						0.04	-0.05	-0.09	-0.04	-0.04	0.01	-0.03	0.03	-0.03	0.03	0.15	-0.22	-0.10	-0.05	-0.06	-0.05	-0.05	-0.02	-0.04
Rule following							0.08	0.04	-0.01	0.09	0.10	0.04	0.10	0.01	0.11	-0.02	0.05	-0.05	-0.12	0.00	-0.03	-0.14	-0.06	-0.07
Altruism								0.66	0.16	0.32	0.32	0.43	0.30	0.26	0.20	-0.01	0.24	0.01	-0.16	0.21	0.16	-0.05	-0.02	0.00
Altruism Taking									0.11	0.25	0.34	0.37	0.25	0.27	0.22	0.00	0.26	0.01	-0.15	0.23	0.12	-0.06	-0.01	-0.04
Trust										0.32	0.42	0.46	0.31	0.31	0.18	-0.01	0.06	0.06	-0.10	-0.04	0.05	-0.01	-0.10	-0.16
Pos. Recripr. Low											0.72	0.40	0.33	0.21	0.18	-0.04	0.16	0.00	-0.18	-0.04	0.05	-0.06	-0.12	-0.07
Pos. Recripr. High												0.46	0.37	0.33	0.19	0.00	0.18	0.06	-0.20	0.03	0.02	-0.13	-0.16	-0.18
PGG Cont.													0.39	0.38	0.27	0.03	0.11	-0.10	-0.23	0.03	0.08	-0.06	-0.06	-0.11
Reward Cons.- Int.														0.35	0.61	0.03	0.12	-0.03	-0.18	0.02	0.03	-0.13	-0.12	-0.10
Reward Int.															0.42	0.06	0.07	0.01	-0.17	0.02	0.01	-0.10	-0.07	-0.12
Reward Cons.																0.09	0.04	-0.05	-0.04	0.05	0.07	-0.09	-0.13	-0.09
Classical Trolley																	-0.22	-0.46	0.00	-0.06	-0.04	-0.06	-0.05	0.01
OUS IB																		0.24	-0.05	0.38	0.22	0.07	-0.02	0.01
OUS IH																			0.19	-0.01	-0.03	0.17	0.14	0.04
MU Scale																				0.09	0.05	0.24	0.14	0.20
MFQ Harm																					0.51	0.22	0.17	0.26
MFQ Fairness																						0.18	0.08	0.11
MFQ Ingroup																							0.55	0.50
MFQ Authority																								0.53
MFQ Purity																								

Notes: Colors in heatmap change with each 0.1 correlation magnitude.

Table D.2: Correlations of ends-versus-means, self-versus-other and questionnaire measures memory check sample

	Ends-versus-means decision situations						Self-versus-others decision situations								Questionnaires								
	SAL Trolley	Group Don.	Bribe	Statement Choice	Lying	Rule following	Altruism	Altruism Taking	Trust	Pos. Recripr. Low	Pos. Recripr. High	PGG Cont.	Reward Cons.- Int.	Reward Int.	Reward Cons.	Classical Trolley	OUS IB	OUS IH	MU Scale	MFQ Harm	MFQ Fairness	MFQ Ingroup	MFQ Authority	MFQ Purity
SAL Trolley		0.05	0.01	0.05	0.16	0.04	0.05	0.00	0.05	0.00	0.02	0.05	0.02	0.10	0.00	0.54	-0.09	-0.31	-0.07	-0.08	0.02	-0.02	-0.01	0.06
Group Don.			0.07	0.07	-0.07	0.00	0.12	0.12	0.04	0.19	0.11	-0.01	0.14	0.04	0.03	0.03	0.06	0.03	-0.03	0.08	0.04	-0.02	-0.05	0.04
Bribe				0.13	0.08	-0.09	-0.01	0.05	-0.05	-0.02	-0.14	-0.06	-0.05	-0.15	-0.12	-0.02	-0.02	-0.10	0.02	0.12	0.05	0.10	0.02	0.05
Statement Choice					-0.05	0.04	0.04	0.08	-0.04	0.06	-0.02	0.09	0.09	-0.07	0.04	0.01	0.04	-0.07	-0.11	0.09	0.09	0.13	0.15	0.13
Lying						0.07	-0.06	-0.10	-0.04	0.03	0.00	-0.05	0.01	-0.03	-0.01	0.11	-0.14	-0.11	-0.10	-0.07	-0.02	-0.08	-0.05	-0.10
Rule following							0.13	0.11	0.01	0.10	0.15	0.04	0.16	0.01	0.11	0.00	0.11	0.00	-0.08	0.11	0.08	-0.12	-0.02	0.01
Altruism								0.65	0.18	0.32	0.38	0.40	0.35	0.30	0.22	0.02	0.28	-0.05	-0.10	0.24	0.25	-0.06	-0.02	-0.01
Altruism Taking									0.12	0.28	0.38	0.33	0.29	0.24	0.24	0.01	0.27	-0.04	-0.08	0.23	0.18	-0.02	0.00	-0.01
Trust										0.31	0.41	0.45	0.39	0.24	0.21	0.01	0.05	0.01	-0.08	0.03	0.08	0.01	-0.13	-0.09
Pos. Recripr. Low											0.72	0.34	0.38	0.12	0.16	-0.01	0.16	0.02	-0.13	0.13	0.12	-0.02	-0.05	-0.02
Pos. Recripr. High												0.46	0.48	0.29	0.25	0.03	0.21	0.03	-0.14	0.15	0.16	-0.13	-0.14	-0.10
PGG Cont.													0.44	0.34	0.25	0.02	0.14	-0.15	-0.21	0.06	0.11	0.00	-0.06	-0.07
Reward Cons.- Int.														0.33	0.56	-0.02	0.16	-0.01	-0.12	0.12	0.13	-0.11	-0.12	-0.08
Reward Int.															0.46	0.13	0.09	0.00	-0.07	0.03	0.05	-0.05	-0.08	-0.08
Reward Cons.																0.07	0.08	0.03	-0.01	0.08	0.09	-0.14	-0.11	-0.07
Classical Trolley																	-0.25	-0.47	-0.05	-0.08	0.02	-0.09	-0.05	0.02
OUS IB																		0.23	-0.08	0.40	0.24	0.06	-0.06	0.02
OUS IH																			0.17	-0.03	-0.04	0.14	0.13	0.05
MU Scale																				-0.02	-0.01	0.19	0.14	0.20
MFQ Harm																					0.54	0.16	0.12	0.26
MFQ Fairness																						0.08	0.00	0.12
MFQ Ingroup																							0.54	0.48
MFQ Authority																								0.52
MFQ Purity																								

Notes: Colors in heatmap change with each 0.1 correlation magnitude.

Table D.3: Correlations of ends-versus-means, self-versus-other and questionnaire measures excluding speeders

	Ends-versus-means decision situations						Self-versus-others decision situations								Questionnaires								
	SAL Trolley	Group Don.	Bribe	Statement Choice	Lying	Rule following	Altruism	Altruism Taking	Trust	Pos. Recipr. Low	Pos. Recipr. High	PGG Cont.	Reward Cons.- Int.	Reward Int.	Reward Cons.	Classical Trolley	OUS IB	OUS IH	MU Scale	MFQ Harm	MFQ Fairness	MFQ Ingroup	MFQ Authority	MFQ Purity
SAL Trolley		0.03	0.06	0.06	0.17	0.06	-0.01	-0.03	0.02	-0.05	-0.03	-0.02	-0.02	0.07	-0.01	0.52	-0.15	-0.29	-0.02	-0.07	-0.03	-0.06	-0.01	0.06
Group Don.			0.13	0.14	-0.06	-0.08	0.07	0.04	-0.04	0.13	0.05	-0.01	0.07	-0.02	-0.01	0.01	0.09	0.03	0.00	0.10	0.03	0.09	0.02	0.09
Bribe				0.16	0.06	-0.09	0.01	0.03	-0.09	-0.05	-0.12	-0.07	-0.02	-0.16	-0.14	-0.01	-0.03	-0.04	0.03	0.06	0.00	0.09	0.06	0.16
Statement Choice					-0.04	0.05	0.04	0.05	-0.12	0.04	-0.08	0.00	0.07	-0.16	0.00	0.02	0.03	-0.08	-0.01	0.08	0.01	0.17	0.15	0.18
Lying						0.02	-0.03	-0.05	-0.05	-0.02	-0.01	-0.05	0.01	-0.02	-0.02	0.15	-0.17	-0.10	-0.06	-0.06	-0.03	-0.08	-0.03	-0.09
Rule following							0.10	0.06	0.00	0.10	0.09	0.00	0.08	-0.03	0.11	-0.01	0.07	0.00	-0.07	0.02	-0.03	-0.08	0.01	-0.03
Altruism								0.64	0.16	0.30	0.35	0.40	0.33	0.22	0.18	-0.03	0.21	0.01	-0.13	0.19	0.17	0.04	0.02	0.02
Altruism Taking									0.12	0.25	0.33	0.37	0.26	0.18	0.17	-0.03	0.19	0.01	-0.12	0.20	0.14	0.02	0.01	-0.04
Trust										0.32	0.42	0.46	0.36	0.27	0.21	-0.01	0.01	0.05	-0.12	-0.03	0.06	-0.04	-0.11	-0.17
Pos. Recipr. Low											0.69	0.33	0.32	0.11	0.11	-0.03	0.16	0.04	-0.14	0.06	0.07	-0.03	-0.08	-0.05
Pos. Recipr. High												0.43	0.39	0.28	0.16	0.00	0.17	0.07	-0.16	0.06	0.10	-0.14	-0.15	-0.16
PGG Cont.													0.40	0.31	0.23	0.01	0.12	-0.10	-0.26	0.09	0.13	-0.01	-0.05	-0.06
Reward Cons.- Int.														0.22	0.54	-0.02	0.06	-0.01	-0.13	0.03	0.05	-0.09	-0.11	-0.11
Reward Int.															0.39	0.08	0.00	-0.01	-0.10	-0.04	0.04	-0.11	-0.07	-0.16
Reward Cons.																0.05	-0.01	0.01	-0.01	0.02	0.04	-0.10	-0.06	-0.10
Classical Trolley																	-0.23	-0.44	-0.02	-0.03	0.02	-0.10	-0.06	-0.02
OUS IB																		0.25	-0.06	0.35	0.25	0.08	-0.04	0.02
OUS IH																			0.16	-0.03	-0.04	0.16	0.15	0.06
MU Scale																				0.02	0.03	0.23	0.15	0.20
MFQ Harm																					0.60	0.24	0.16	0.24
MFQ Fairness																						0.18	0.08	0.15
MFQ Ingroup																							0.56	0.51
MFQ Authority																								0.54
MFQ Purity																								

Notes: Colors in heatmap change with each 0.1 correlation magnitude.

Table D.4: Correlations between ends-versus-means and self-versus-other blocks using behavior categorization for the later variables

	Ends-versus-means decision situations						Self-versus-others decision situations								
	SAL Trolley	Group Don.	Bribe	Statement Choice	Lying	Rule following	Altruism	Altruism Taking	Trust	Pos. Recipr. Low	Pos. Recipr. High	PGG Cont.	Reward Cons. comp Int.	Reward Int.	Reward Cons.
SAL Trolley		0.05	0.05	0.06	0.16	0.03	-0.04	-0.09	-0.03	-0.09	-0.04	-0.04	-0.07	-0.06	-0.06
Group Don.			0.14	0.15	-0.09	-0.06	0.05	0.06	-0.03	0.06	0.07	0.01	0.10	0.10	0.10
Bribe				0.19	0.04	-0.08	-0.02	-0.03	-0.08	-0.09	-0.08	-0.04	-0.12	-0.12	-0.12
Statement Choice					-0.06	0.04	0.05	0.07	-0.04	0.03	0.02	0.03	0.01	0.02	0.00
Lying						0.00	-0.06	-0.08	-0.04	-0.01	0.00	-0.05	-0.05	-0.05	-0.05
Rule following							0.07	0.06	0.06	0.07	0.09	0.02	0.04	0.05	0.05
Altruism								0.66	0.25	0.35	0.40	0.40	0.39	0.40	0.40
Altruism Taking									0.27	0.34	0.35	0.42	0.39	0.40	0.40
Trust										0.45	0.44	0.43	0.42	0.41	0.42
Pos. Recipr. Low											0.78	0.44	0.52	0.51	0.51
Pos. Recipr. High												0.45	0.55	0.54	0.54
PGG Cont.													0.49	0.48	0.48
Reward Cons.- Int.														0.98	0.99
Reward Int.															0.97
Reward Cons.															

Table D.5: Correlations between ends-versus-means and self-versus-other blocks using median splits for the later variables

	Ends-versus-means decision situations						Self-versus-others decision situations								
	SAL Trolley	Group Don.	Bribe	Statement Choice	Lying	Rule following	Altruism	Altruism Taking	Trust	Pos. Recipr. Low	Pos. Recipr. High	PGG Cont.	Reward Cons. comp Int.	Reward Int.	Reward Cons.
SAL Trolley		0.05	0.05	0.06	0.16	0.03	-0.02	-0.01	-0.02	-0.08	-0.02	0.01	-0.03	-0.03	-0.07
Group Don.			0.14	0.15	-0.09	-0.06	0.02	0.05	-0.03	0.13	0.00	-0.02	0.09	-0.04	-0.03
Bribe				0.19	0.04	-0.08	0.00	0.01	-0.10	-0.06	-0.14	-0.03	-0.07	-0.16	-0.20
Statement Choice					-0.06	0.04	0.01	0.01	-0.10	0.05	-0.08	0.01	0.05	-0.13	-0.05
Lying						0.00	0.00	-0.05	-0.02	0.01	0.04	-0.04	0.01	-0.03	-0.02
Rule following							0.13	0.01	0.03	0.14	0.06	0.06	0.05	-0.02	0.08
Altruism								0.52	0.17	0.28	0.26	0.40	0.32	0.23	0.26
Altruism Taking									0.12	0.22	0.21	0.27	0.28	0.22	0.24
Trust										0.22	0.34	0.42	0.32	0.09	0.17
Pos. Recipr. Low											0.55	0.24	0.31	0.18	0.23
Pos. Recipr. High												0.30	0.32	0.27	0.25
PGG Cont.													0.38	0.25	0.21
Reward Cons.- Int.														0.15	0.45
Reward Int.															0.38
Reward Cons.															

Table D.6: Biserial correlations between ends-versus-means and self-versus-other blocks

	Altruism	Altruism Taking	Trust	Pos. Recipr. Low	Pos. Recipr. High	PGG Cont.	Reward Cons. comp Int.	Reward Int.	Reward Cons.
SAL Trolley	-0.01	-0.04	0.03	-0.07	-0.04	0.01	-0.03	0.07	-0.03
Group Don.	0.06	0.04	-0.05	0.18	0.06	-0.03	0.11	-0.04	-0.02
Bribe	-0.05	-0.03	-0.16	-0.08	-0.21	-0.09	-0.08	-0.26	-0.25
Statement Choice	0.00	0.01	-0.15	0.05	-0.09	-0.01	0.07	-0.19	-0.05
Lying	-0.03	-0.05	-0.03	-0.03	0.00	-0.06	0.00	-0.03	-0.01
Rule following	0.15	0.08	0.04	0.14	0.13	0.04	0.15	-0.01	0.16

E Instructions

E.1 General information

Welcome to the study

Welcome, and thank you for your interest in today's online study! Please note that you may only participate in this study once. Also, you may only participate if you have registered for this study in our participation database.

As already announced during registration, this study consists of two parts. You will complete the second part exactly in one week at the same time. For completing both parts in full, you will receive a lump sum of 12 EUR upon completion of the study.

You will make decisions on the computer in this study. Your decisions will allow you to earn extra money. All payments, i.e., both the compensation for your participation and any additional payments based on your decisions, will be sent to you by bank transfer following the second part.

Please note: You will receive the payment for your participation of 12 EUR as well as any additional money based on your decisions only if you successfully complete both parts. Therefore, please make sure to show up for the second part next week as well and complete it entirely.

Structure of the study and your payoff

In the following, you will make different decisions in each of several successive sections. The decisions in each section are completely independent of each other. This means your decisions in one section do not affect the consequences or possible payoffs of another section. Furthermore, similar-sounding decision situations in different sections do not mean that your decisions should necessarily be similar as well. You will see the number of the current section in the upper right corner of each section.

From all the decisions you will make today, a computer will randomly select one decision. Each of the decisions has the same chance of being selected. This selection is independent of your decisions.

The consequence of the selected decision will then be implemented exactly as described in the decision. Since one of your decisions will actually be implemented, you should consider each decision carefully and treat each decision as if it were actually implemented.

In what follows, this is always implicitly assumed. For example, suppose a decision is described as giving you an amount of money as a consequence. In that case, you will receive that amount exactly as described if that decision is indeed randomly selected.

Donation to the *Förderkreis*

Some decisions in today's part are about the possibility of making a donation to the *Förderkreis für krebskranke Kinder und Jugendliche Bonn e.V.* (in short: the *Förderkreis*). The *Förderkreis* is a regional aid organization from Bonn.

Information about the Förderkreis

The *Förderkreis* comprehensively supports young people with cancer and their families in dealing with the disease. The organization offers psychological support, organizes leisure activities as well as aftercare, and offers school support. The projects and participations of the *Förderkreis* at a glance are:

- Team Bärenstark: Practical help during the treatment period for patients and relatives.
- Psychological and psychosocial counseling for patients and relatives
- Pedagogical support in the hospital
- Financing of hospital clowns and music therapy
- Follow-up care by the KoCkPiT team (Children's Palliative Care Team Bonn)
- Support for parents of sick children and adolescents
- Support for orphaned parents

So if in the following, a donation to the *Förderkreis* is mentioned, it is always a donation to the organization described above, with which the listed projects are supported.

If in the following a donation is initiated by your decisions, it will always be paid by the scientists involved from study funds. So you do not have to donate (or pay) any amount of money personally. If you decide against a donation, the amount will not be donated accordingly.

Consequences for other study participants

In some sections, your decisions have consequences for other study participants. For example, in some choices, you can send money to a randomly selected other person who is also participating in this study.

Important: The other people will not learn about your decisions or personal information about you at any time. Notably, the other individuals will also receive payouts through other situations, which are random in amount. This ensures that the other people cannot draw any conclusions about your decisions from the amount of the payoffs at the end of the trial.

On the next page, the individual sections begin.

E.2 Ends-versus-means block

E.2.1 Trolley problem

The consequences of the next decision

[Treatment *Hypothetical*]

For the next section of this study, the decisions are hypothetical in nature. This means that the consequence of your decision in the next section will not be implemented in real terms. Thus, the following section is purely a thought experiment.

[Treatment *Real*]

For your decision in the next section, there is a probability of 1 in 10, i.e., 10%, that it will be implemented as described. At the end of the study, a computer-generated random number will determine whether your decision will actually be implemented. If so, the consequences of your decision will be realized exactly as described. Therefore, since your decision may actually be implemented, you should think carefully about the decision.

Information

Below you will read important information that is relevant to your decision later.
After you have made your decision, we will ask you four simple questions about the information and instructions presented bellow. For each question you then answer correctly, you will receive an additional 0.5 EUR.
Please read all the information carefully. Not only is the information important for your decision, but you will also potentially receive a higher payout.

Information about tuberculosis

What is tuberculosis?
Tuberculosis – also called Phthisis or White Death – is an infectious disease, which is caused by bacteria. Roughly one-third of all humans are infected with the pathogen of tuberculosis. Active tuberculosis breaks out among 5 to 10% of all those infected. Tuberculosis is primarily airborne. This is also why quick treatment is necessary.

What are the symptoms and consequences of tuberculosis?
Tuberculosis patients often suffer from very unspecific symptoms like fatigue, the feeling of weakness, lack of appetite, and weight loss. At an advanced stage of lung tuberculosis, the patient coughs up blood, leading to the so-called rush of blood. Without treatment, a person with tuberculosis dies with a probability of 43%.

How prevalent is tuberculosis?

In the year 2018, about 10 million people have been recorded as falling ill with active tuberculosis. Almost 1.5 million people die of tuberculosis each year. This means that tuberculosis causes more annual deaths than HIV or malaria.

Is tuberculosis curable?
According to the World Health Organization (WHO), the United Nations agency for international public health, “tuberculosis is preventable and curable”. Treatment takes place by taking antibiotics several times a week over a period of 6 months. It is important to take the medication consistently. Since 2000, an estimated 53 million lives have been saved through effective diagnosis and treatment of tuberculosis.

The success rate of treatment for a new infection is usually over 85%.
The preceding figures and information were provided by WHO and are freely available. For more details click here (link opens in new tab).

Operation ASHA

Operation ASHA is a charity organization specialized since 2005 on treating tuberculosis in disadvantaged communities. The work of *Operation ASHA* is based on the insight that the biggest obstacle for the treatment of tuberculosis is the interruption of the necessary 6-month-long regular intake of medication.

For a successful treatment, the patient has to come to a medical facility twice a week – more than 60 times in total – to take the medication. Interruption or termination of the treatment is fatal because this strongly enhances the development of a drug-resistant form of tuberculosis. This form of tuberculosis is much more difficult to treat and almost always leads to death.

The concept of *Operation ASHA*
To overcome this problem, *Operation ASHA* developed a concept that guarantees regular treatment through immediate spatial proximity to the patient. A possible non-adherence is additionally prevented by visiting the patient at home.

By now, *Operation ASHA* runs more than 360 treatment centers, almost all of which are located in the poorer regions of India. More than 60,000 sick persons have been identified and treated that way.

Operation ASHA is an internationally recognized organization, and its success has been covered by the New York Times, BBC, and Deutsche Welle, for example. The MIT and the University College London have already conducted research projects about the fight against tuberculosis in cooperation with *Operation ASHA*. The treatment method employed by *Operation ASHA* is described by the World Health Organization (WHO) as “highly efficient and cost-effective”.

The impact of a donation to *Operation ASHA*

By donating money to *Operation ASHA*, it is possible to save people from death by tuberculosis.

To save a person's life means here to successfully cure a person with tuberculosis, who otherwise would die because of the tuberculosis. A donation of 380 EUR ensures that at least one human life can be expected to be saved. The information used to calculate the donation amount is obtained from public statements from the World Health Organization (WHO), peer-reviewed research studies, Indian Government statistics, and published figures from *Operation ASHA*.

In the calculation, information was conservatively interpreted, or a pessimistic number was used so that the donation amount of 380 EUR is in the case of doubt higher than the actual costs to save a human life. In addition, in the calculation of the treatment success rate of *Operation ASHA*, the mortality rate for alternative treatment by the state tuberculosis program in India and the different detection rates for new cases of tuberculosis are included.

In the context of this study, an agreement made with *Operation ASHA* will ensure that 100% of the donation will be used exclusively for the diagnosis and treatment of tuberculosis patients. This means that every Euro of the donation amount goes directly to saving human lives, and no other costs will be covered. Based on a very high number of cases, the contribution of a donation of 380 EUR can be summarized as follows:

With a donation of 380 EUR 5 additional patients infected with tuberculosis can be treated through *Operation ASHA*.

If these 5 persons are not treated through *Operation ASHA*, it is expected that one patient will die.

If, through the donation of 380 EUR all 5 patients are treated, it is expected that no patient will die.

Based on these calculations it can be concluded that the donation of 380 EUR will save one human life. Going further, this also means that two lives are saved by a donation of 760 EUR, and three lives are saved by a donation of 1140 EUR.

The relationship of a donation of 380 EUR to Operation ASHA and a human life is illustrated in the following graph: [Figure here]

So, in summary, the following can be said about the effect of a donation for a given number of 5 patients suffering from tuberculosis:

- Donation results in saving one person from death.
- Failure to donate results in the death of one person.

Whenever the life or death of a person is mentioned in connection with a donation, the above-mentioned facts are implicitly meant. The same applies when the life or death of several people

is involved.

The concrete situation

As mentioned before, *Operation ASHA* operates numerous treatment centers in India. In particular, *Operation ASHA* operates in the two Indian states of [State A] and [State B].

[State A]
In this experiment, 5 people were identified in [State A] who are infected by tuberculosis. For these people, the donation amount required to save a life was initiated to *Operation ASHA*. Through this, only the people who are sick in [State A] will be treated. This will result in a person from [State A] being saved from death by tuberculosis.

If the initiated donation amount were not realized, a person in [State A] would die from tuberculosis.

[State B]
After the donation was initiated, 15 people suffering from tuberculosis were identified in [State B]. Without donation and subsequent treatment by *Operation ASHA*, three of the people from [State B] who became ill would die from tuberculosis.

In the event that the people in [State B] are helped through the donation to *Operation ASHA* necessary to save three lives, these three people would be saved from death by tuberculosis.

Redirecting the donation

The initiated donation to patients in [State A] can be redirected to help save patients in [State B] from death by tuberculosis. The amount would be multiplied to save three lives in [State B].

Summary

Tuberculosis is a worldwide common bacterial infectious disease. The success rate of medical treatment of a new disease is very high. Nevertheless, close to 1.5 million people die every year from tuberculosis. The biggest obstacle to the curing of tuberculosis is the potential stopping of continuous treatment with antibiotics. The concept of *Operation ASHA* is therefore based on the immediate proximity to the patient as well as the control and recording of the regular intake of medication. A donation of 380 EUR to Operation ASHA can save a human life. A donation of 760 EUR can save two lives, and a donation of 1140 EUR can save three lives.

How is the donation connected to the saving of a life?
The donation of 380 EUR already accounts for the fact that someone inflicted with the illness could have survived without treatment by *Operation ASHA*; i.e., instead of through *Operation ASHA*, they could have received treatment through other actors (such as the public health system). The amount is, therefore, sufficient for the diagnosis and complete treatment of multiple sufferers.

What does it mean to "save a life"?
To save a life means here the successful curing of a person suffering from tuberculosis, who otherwise would die because of the Tuberculosis. This means: The amount donated is sufficient to identify and treat enough people with tuberculosis that there is at least one person among them who would otherwise have died from tuberculosis in expectation. To summarize:

- A donation has the effect of saving a person from death.
- Failure to donate results in the death of a person.

The initial situation
In [State A], 5 people suffering from tuberculosis were identified. For these, the necessary donation amount was initiated to save one person from death due to tuberculosis. Thus, without further intervention, none of these diseased people will die of tuberculosis.

After initiating this donation, an additional 15 people suffering from tuberculosis were identified in [State B]. Without further intervention, three of these diseased people would die of tuberculosis. If the initiated donation amount is redirected, it will be multiplied to save three people in [State B] from dying of tuberculosis.

Your decision

So, in terms of the context just described, the situation is as follows:

Without intervention, three people in [State B] will die of tuberculosis. By intervening, these three people can be saved. However, as a direct consequence of this intervention, one person in [State A] will die of tuberculosis.

You will have the opportunity to choose between intervening and not intervening in this situation:

Do not intervene
Do not intervene: do not divert the donation.
Consequence: You decide not to intervene. This means that you decide against an intervention on your end that would cause a patient in [State A] to die from tuberculosis. In return, three people from [State B] will not be saved from dying from tuberculosis.

Intervene
Intervene: redirect the donation.
Consequence: You decide to intervene. This means that you decide in favor of an intervention that causes a patient in [State A] to die from tuberculosis. In return, your redirected donation will save three people from [State B] from dying of tuberculosis.

In a moment, you will have the opportunity to make this decision using an animation. The next page will explain how this animation works.

The decision animation

In a moment, you will have the opportunity to make your decision with the help of an animation. The initial situation of the animation is shown in the following figure:

[animation figure]
The animation depicts the decision situation just described. As soon as the animation starts, the skull shown moves slowly to the right towards the three people. In case you don't see three people on the right or they are outside the white border, please click [here].

You can change the path of the skull using the slider located to the left of the junction (blue line). You do this by dragging the slider with your computer mouse all the way down and then releasing it. This is necessary because otherwise, the slider will move back to its original position. If the slider is dragged all the way down and released, the downward junction will open so that the skull on the track will move down to the single person.

You can make your decision before the skull reaches the dashed gray line in front of the junction. After arriving at the line, moving the slider will have no effect.

Hence, you have two choices, with the consequences described earlier:

Do not intervene
Do not intervene: do not move the slider, the junction will not open.
Consequence: You decide not to intervene. This means that you decide against an intervention on your end that would cause a patient in [State A] to die from tuberculosis. In return, three people from [State B] will not be saved from dying from tuberculosis.

Intervene
Intervene: move the slider, the junction will open.
Consequence: You decide to intervene. This means that you decide in favor of an intervention that causes a patient in [State A] to die from tuberculosis. In return, your redirected donation will save three people from [State B] from dying of tuberculosis.

On the following screen, you will have the opportunity to familiarize yourself with the controller. You will be able to make your actual decision afterward.

Testing the slider

You will now have the opportunity to try out the slider of the animation. Note that you need to drag the slider all the way down and then release it to open the branch.

[slider test]

Click “Next” when you are sufficiently familiar with the slider.

On the next screen you can now make your decision. Note that the animation will start directly.

As a reminder:

[Treatment *Hypothetical*]
As previously described, this is a hypothetical decision that is not actually implemented.

[Treatment *Real*]
As previously described, your decision may actually be implemented, resulting in actual consequences.

[Trolley decision]

E.2.2 Trolley problem – aligned version

The consequences of the next decision

[Treatment *Hypothetical*]

For the next section of this study, the decisions are hypothetical in nature. This means that the consequence of your decision in the next section will not be implemented in real terms. Thus, the following section is purely a thought experiment.

[Treatment *Real*]

For your decision in the next section, there is a probability of 1 in 10, i.e., 10%, that it will be implemented as described. At the end of the study, a computer-generated random number will determine whether your decision will actually be implemented. If so, the consequences of your decision will be realized exactly as described. Therefore, since your decision may actually be implemented, you should think carefully about the decision.

Information

Below you will read important information that is relevant to your decision later.
After you have made your decision, we will ask you four simple questions about the information and instructions presented bellow. For each question you then answer correctly, you will receive an additional 0.5 EUR.
Please read all the information carefully. Not only is the information important for your decision, but you will also potentially receive a higher payout.

Information about tuberculosis

What is tuberculosis?

Tuberculosis – also called Phthisis or White Death – is an infectious disease, which is caused by bacteria. Roughly one-third of all humans are infected with the pathogen of tuberculosis. Active tuberculosis breaks out among 5 to 10% of all those infected. Tuberculosis is primarily airborne. This is also why quick treatment is necessary.

What are the symptoms and consequences of tuberculosis?
Tuberculosis patients often suffer from very unspecific symptoms like fatigue, the feeling of weakness, lack of appetite, and weight loss. At an advanced stage of lung tuberculosis, the patient coughs up blood, leading to the so-called rush of blood. Without treatment, a person with tuberculosis dies with a probability of 43%.

How prevalent is tuberculosis?
In the year 2018, about 10 million people have been recorded as falling ill with active tuberculosis. Almost 1.5 million people die of tuberculosis each year. This means that tuberculosis causes more annual deaths than HIV or malaria.

Is tuberculosis curable?
According to the World Health Organization (WHO), the United Nations agency for international public health, “tuberculosis is preventable and curable”. Treatment takes place by taking antibiotics several times a week over a period of 6 months. It is important to take the medication consistently. Since 2000, an estimated 53 million lives have been saved through effective diagnosis and treatment of tuberculosis.

The success rate of treatment for a new infection is usually over 85%.
The preceding figures and information were provided by WHO and are freely available. For more details click here (link opens in new tab).

Operation ASHA

Operation ASHA is a charity organization specialized since 2005 on treating tuberculosis in disadvantaged communities. The work of *Operation ASHA* is based on the insight that the biggest obstacle for the treatment of tuberculosis is the interruption of the necessary 6-month-long regular intake of medication.

For a successful treatment, the patient has to come to a medical facility twice a week – more than 60 times in total – to take the medication. Interruption or termination of the treatment is fatal because this strongly enhances the development of a drug-resistant form of tuberculosis. This form of tuberculosis is much more difficult to treat and almost always leads to death.

The concept of *Operation ASHA*
To overcome this problem, *Operation ASHA* developed a concept that guarantees regular treatment through immediate spatial proximity to the patient. A possible non-adherence is additionally pre-

vented by visiting the patient at home.

By now, *Operation ASHA* runs more than 360 treatment centers, almost all of which are located in the poorer regions of India. More than 60,000 sick persons have been identified and treated that way.

Operation ASHA is an internationally recognized organization, and its success has been covered by the New York Times, BBC, and Deutsche Welle, for example. The MIT and the University College London have already conducted research projects about the fight against tuberculosis in cooperation with *Operation ASHA*. The treatment method employed by *Operation ASHA* is described by the World Health Organization (WHO) as “highly efficient and cost-effective”.

The impact of a donation to *Operation ASHA*

By donating money to *Operation ASHA*, it is possible to save people from death by tuberculosis.

To save a person’s life means here to successfully cure a person with tuberculosis, who otherwise would die because of the tuberculosis. A donation of 380 EUR ensures that at least one human life can be expected to be saved. The information used to calculate the donation amount is obtained from public statements from the World Health Organization (WHO), peer-reviewed research studies, Indian Government statistics, and published figures from *Operation ASHA*.

In the calculation, information was conservatively interpreted, or a pessimistic number was used so that the donation amount of 380 EUR is in the case of doubt higher than the actual costs to save a human life. In addition, in the calculation of the treatment success rate of *Operation ASHA*, the mortality rate for alternative treatment by the state tuberculosis program in India and the different detection rates for new cases of tuberculosis are included.

In the context of this study, an agreement made with *Operation ASHA* will ensure that 100% of the donation will be used exclusively for the diagnosis and treatment of tuberculosis patients. This means that every Euro of the donation amount goes directly to saving human lives, and no other costs will be covered. Based on a very high number of cases, the contribution of a donation of 380 EUR can be summarized as follows:

With a donation of 380 EUR 5 additional patients infected with tuberculosis can be treated through *Operation ASHA*.

If these 5 persons are not treated through *Operation ASHA*, it is expected that one patient will die.

If, through the donation of 380 EUR all 5 patients are treated, it is expected that no patient will die.

Based on these calculations it can be concluded that the donation of 380 EUR will save one human life. Going further, this also means that two lives are saved by a donation of 760 EUR, and three lives are saved by a donation of 1140 EUR.

The relationship of a donation of 380 EUR to Operation ASHA and a human life is illustrated in the following graph: [Figure here]

So, in summary, the following can be said about the effect of a donation for a given number of 5 patients suffering from tuberculosis:

- Donation results in saving one person from death.
- Failure to donate results in the death of one person.

Whenever the life or death of a person is mentioned in connection with a donation, the above-mentioned facts are implicitly meant. The same applies when the life or death of several people is involved.

The concrete situation

As mentioned before, *Operation ASHA* operates numerous treatment centers in India. In particular, *Operation ASHA* operates in the two Indian states of [State A] and [State B].

[State A]
In this experiment, 15 people were identified in [State A] who are infected by tuberculosis. For these people, the donation amount required to save a life was initiated to *Operation ASHA*. Through this, only the people who are sick in [State A] will be treated. This will result in three people from [State A] being saved from death by tuberculosis.

If the initiated donation amount were not realized, three people in [State A] would die from tuberculosis.

[State B]
After the donation was initiated, 5 people suffering from tuberculosis were identified in [State B]. Without donation and subsequent treatment by *Operation ASHA*, one person from [State B] who became ill would die from tuberculosis.

In the event that the people in [State B] are helped through the donation to *Operation ASHA* necessary to save one live, these one person would be saved from death by tuberculosis.

Redirecting the donation

The initiated donation to patients in [State A] can be redirected to help save patients in [State B] from death by tuberculosis. The amount would be reduced to save one live in [State B].

Summary

Tuberculosis is a worldwide common bacterial infectious disease. The success rate of medical treatment of a new disease is very high. Nevertheless, close to 1.5 million people die every year from

tuberculosis. The biggest obstacle to the curing of tuberculosis is the potential stopping of continuous treatment with antibiotics. The concept of *Operation ASHA* is therefore based on the immediate proximity to the patient as well as the control and recording of the regular intake of medication. A donation of 380 EUR to Operation ASHA can save a human life. A donation of 760 EUR can save two lives, and a donation of 1140 EUR can save three lives.

How is the donation connected to the saving of a life?
The donation of 380 EUR already accounts for the fact that someone inflicted with the illness could have survived without treatment by *Operation ASHA*; i.e., instead of through *Operation ASHA*, they could have received treatment through other actors (such as the public health system). The amount is, therefore, sufficient for the diagnosis and complete treatment of multiple sufferers.

What does it mean to "save a life"?
To save a life means here the successful curing of a person suffering from tuberculosis, who otherwise would die because of the Tuberculosis. This means: The amount donated is sufficient to identify and treat enough people with tuberculosis that there is at least one person among them who would otherwise have died from tuberculosis in expectation. To summarize:

- A donation has the effect of saving a person from death.
- Failure to donate results in the death of a person.

The initial situation
In [State A], 15 people suffering from tuberculosis were identified. For these, the necessary donation amount was initiated to save three people from death due to tuberculosis. Thus, without further intervention, none of these diseased people will die of tuberculosis.

After initiating this donation, an additional 5 people suffering from tuberculosis were identified in [State B]. Without further intervention, one of these diseased people would die of tuberculosis. If the initiated donation amount is redirected, it will be reduced to save one person in [State B] from dying of tuberculosis.

Your decision

So, in terms of the context just described, the situation is as follows:

Without intervention, one person in [State B] will die of tuberculosis. By intervening, this person can be saved. However, as a direct consequence of this intervention, three people in [State A] will die of tuberculosis.

You will have the opportunity to choose between intervening and not intervening in this situation:

Do not intervene
Do not intervene: do not divert the donation.

Consequence: You decide not to intervene. This means that you decide against an intervention on your end that would cause three patients in [State A] to die from tuberculosis. In return, one person from [State B] will not be saved from dying from tuberculosis.

Intervene
Intervene: redirect the donation.
Consequence: You decide to intervene. This means that you decide in favor of an intervention that causes three patients in [State A] to die from tuberculosis. In return, your redirected donation will save one person from [State B] from dying of tuberculosis.

In a moment, you will have the opportunity to make this decision using an animation. The next page will explain how this animation works.

The decision animation

In a moment, you will have the opportunity to make your decision with the help of an animation. The initial situation of the animation is shown in the following figure:

[animation figure]
The animation depicts the decision situation just described. As soon as the animation starts, the skull shown moves slowly to the right towards the three people. In case you don't see three people on the right or they are outside the white border, please click [here].

You can change the path of the skull using the slider located to the left of the junction (blue line). You do this by dragging the slider with your computer mouse all the way down and then releasing it. This is necessary because otherwise, the slider will move back to its original position. If the slider is dragged all the way down and released, the downward junction will open so that the skull on the track will move down to the single person.

You can make your decision before the skull reaches the dashed gray line in front of the junction. After arriving at the line, moving the slider will have no effect.

Hence, you have two choices, with the consequences described earlier:

Do not intervene
Do not intervene: do not move the slider, the junction will not open.
Consequence: You decide not to intervene. This means that you decide against an intervention on your end that would cause three patients in [State A] to die from tuberculosis. In return, one person from [State B] will not be saved from dying from tuberculosis.

Intervene
Intervene: move the slider, the junction will open.
Consequence: You decide to intervene. This means that you decide in favor of an intervention that

causes three patients in [State A] to die from tuberculosis. In return, your redirected donation will save one person from [State B] from dying of tuberculosis.

On the following screen, you will have the opportunity to familiarize yourself with the controller. You will be able to make your actual decision afterward.

Testing the slider

You will now have the opportunity to try out the slider of the animation. Note that you need to drag the slider all the way down and then release it to open the branch.

[slider test]

Click "Next" when you are sufficiently familiar with the slider.

On the next screen you can now make your decision. Note that the animation will start directly.

As a reminder:

[Treatment *Hypothetical*]
As previously described, this is a hypothetical decision that is not actually implemented.

[Treatment *Real*]
As previously described, your decision may actually be implemented, resulting in actual consequences.

[Trolley decision]

E.2.3 Baseline choice

In this section, you have a choice between two options, Option A and Option B. The two options differ in whether a sum of money is donated to a charity or another study participant receives a sum of money. The two options are as follows:

Option A

If you choose option A, 15 EUR will be donated to the *Förderkreis für krebskranke Kinder und Jugendliche Bonn e.V.*

Option B

If you choose option B, another study participant will receive 2 EUR, and no money will be donated to the *Förderkreis*. The other study participant will not receive any additional information about your decision.

E.2.4 Lying game

In this section, you have been paired with another person. This person, like you, is a participant in this study.

Options

In this section, there are two possible payout options that you and the other person can achieve at the end of the study. The two options are:

Option A
15 EUR will be donated to the *Förderkreis* für krebskranke Kinder und Jugendliche Bonn e.V.

Option B
The other person receives 2 EUR and no money is donated to the *Förderkreis* für krebskranke Kinder und Jugendliche Bonn e.V.

The other person can choose between these two options, that is, between option A or option B. **However, the other person does not know what the exact consequences of the options are. In particular, the other person does not know which option leads to which payout.** Only you know what the consequences of the options are. The only information that the other person will have prior to their decision is a single message that you can send to the person.

Your decision in this section

You have a choice between the following two messages that you can send to the other person:

Message 1
"Option A gives you a higher personal cash payment than Option B".

Message 2
"Option B gives you a higher personal cash payment than Option A".

The sent message will be displayed to the other person exactly as shown above. In addition, the person will not learn which payoffs were associated with which option (also not afterward). Since the person will receive additional payments from other tasks in this study, as described at the beginning of the study, the person will also not learn whether your message was right or wrong.

The person was told that the message received could be true or false and that he or she would not receive any information about what the consequences of the options were. In such a situation, the other person follows the action prescribed in the message in almost all cases (>90%). Thus, you can assume with a very high probability the following:

- If you send message 1, the other person will choose option A (without knowing the exact consequences).

- If you send message 2, the other person will choose option B (without knowing the exact consequences).

Hence, the sequence of events is the following: You decide which of the two messages you want to send. Based on that, the other person will choose the respective option. The other person's decision will then determine the corresponding payments that will be made at the end of the study.

Summary

In summary, by choosing the message, you can have an impact on the payments and donations that are triggered. The other person is aware that (1) the message may be right or wrong, (2) they will not know which payoffs were associated with which option and whether the message was right or wrong.

[Comprehension questions]

E.2.5 Lying game -aligned version

In this section, you have been paired with another person. This person, like you, is a participant in this study.

Options

In this section, there are two possible payout options that you and the other person can achieve at the end of the study. The two options are:

Option A
15 EUR will be donated to the *Förderkreis* für krebskranke Kinder und Jugendliche Bonn e.V.

Option B
The other person receives 2 EUR and no money is donated to the *Förderkreis* für krebskranke Kinder und Jugendliche Bonn e.V.

The other person can choose between these two options, that is, between option A or option B. **However, the other person does not know what the exact consequences of the options are. In particular, the other person does not know which option leads to which payout.** Only you know what the consequences of the options are. The only information that the other person will have prior to their decision is a single message that you can send to the person.

Your decision in this section

You have a choice between the following two messages that you can send to the other person:

Message 1
"Option A triggers a higher donation to a charity than Option B".

Message 2
"Option B triggers a higher donation to a charity than Option A".

The sent message will be displayed to the other person exactly as shown above. In addition, the person will not learn which payoffs were associated with which option (also not afterward). Since the person will receive additional payments from other tasks in this study, as described at the beginning of the study, the person will also not learn whether your message was right or wrong.

The person was told that the message received could be true or false and that he or she would not receive any information about what the consequences of the options were. In such a situation, the other person follows the action prescribed in the message in almost all cases (>90%). Thus, you can assume with a very high probability the following:

- If you send message 1, the other person will choose option A (without knowing the exact consequences).
- If you send message 2, the other person will choose option B (without knowing the exact consequences).

Hence, the sequence of events is the following: You decide which of the two messages you want to send. Based on that, the other person will choose the respective option. The other person's decision will then determine the corresponding payments that will be made at the end of the study.

Summary

In summary, by choosing the message, you can have an impact on the payments and donations that are triggered. The other person is aware that (1) the message may be right or wrong, (2) they will not know which payoffs were associated with which option and whether the message was right or wrong.

[Comprehension questions]

E.2.6 Bribe game

In this section, you have been matched with another person. Like you, this person is a participant in the current study. In this section, you have a choice between two options, Option A and Option B. The two options are as follows:

Option A
If you choose Option A, you will help children with cancer. Your choice of Option A will trigger a donation of 15 EUR to the *Förderkreis für krebskranke Kinder und Jugendliche Bonn e.V.*
Option B
If you choose option B, you will send the other person 2 EUR. However, in this case, the donation otherwise triggered by option A will not be given to *Förderkreis für krebskranke Kinder und Jugendliche Bonn e.V.*

Additional consequence

The other person also has the choice between two options, namely between option C and option D. If the other person chooses option C, an additional 20 EUR will be donated to the *Förderkreis* If the person chooses option D, the person will receive 2 EUR for themselves as an additional payout.

The person has chosen to make a choice between the two options, C and D, dependent on **your decision** between A and B. This means the other person has **obligatorily** specified the following: The person will choose option D and thus the additional payout for himself. The choice of option D can only be prevented if you choose option B and send the other person 2 EUR. In this and only in this case, the other person will choose option C and thus trigger the donation. Thus, the consequences are as follows:

- You choose option A $\Rightarrow$ The other person chooses option D.
- You choose option B $\Rightarrow$ The other person chooses option C.

In other words, if you choose option A and thus trigger a donation of 15 EUR, the other person will stick with option D and receive 2 EUR. If you choose to send the 2 EUR to the other person by choosing option B, the other person will choose option C and choose the donation of 20 EUR.

[Comprehension questions]

E.2.7 Bribe game - aligned version

In this section, you have been matched with another person. Like you, this person is a participant in the current study. In this section, you have a choice between two options, Option A and Option B. The two options are as follows:

Option A
If you choose Option A, you will help children with cancer. Your choice of Option A will trigger a donation of 15 EUR to the *Förderkreis für krebskranke Kinder und Jugendliche Bonn e.V.*
Option B
If you choose option B, you will send the other person 2 EUR. However, in this case, the donation otherwise triggered by option A will not be given to *Förderkreis für krebskranke Kinder und Jugendliche Bonn e.V.*

Additional consequence

The other person also has the choice between two options, namely between option C and option D. If the other person chooses option C, an additional 20 EUR will be donated to the *Förderkreis* If the person chooses option D, the person will receive 2 EUR for themselves as an additional payout.

The person has chosen to make a choice between the two options, C and D, dependent on **your decision** between A and B. This means the other person has **obligatorily** specified the following:

The person will choose option D and thus the additional payout for himself. The choice of option D can only be prevented if you choose option A and thus the donation of 15 EUR to the *Förderkreis für krebskranke Kinder und Jugendliche Bonn e.V.* . In this and only in this case, the other person will choose option C and thus trigger the donation. Thus, the consequences are as follows:

- You choose option A ⇒ The other person chooses option C.
- You choose option B ⇒ The other person chooses option D.

In other words, if you choose option A and thus trigger a donation of 15 EUR, the other person will choose option C and thus the 20 EUR donation. If you choose to send the 2 EUR to the other person by choosing option B, the other person will stick to choosing option D and thereby receive 2 EUR.

[Comprehension questions]

E.2.8 Group decision game

Information

Your group

Together with 5 other people, you will form a group for your decisions. These people also participate in this study, so together, you will form a group of 6 individuals.
You can make all decisions autonomously and independently of the other members of the group. However, as described below, the consequences of your decision will depend on the decisions of the other group members.

The donation

Your group has been entrusted with a donation of 15 EUR, which is intended for the *Förderkreis* für krebskranke Kinder und Jugendliche Bonn e.V. and is meant to be donated to the charity after the study.

Your options

You can choose between two options: Option A and Option B. Depending on which option you and the other group members choose, different consequences will be realized. In particular, both your decision and the decisions of the other group members affect what happens to the donation of 15 EUR to the Föderkreis.

The donation to the Föderkreis will be destroyed (that is, not executed) if at least one member of your group chooses option B. That is, if either you or at least one other person with whom you form a group chooses option B, the donation will be destroyed. Only if no member of your group chooses option B, i.e., **all** group members choose option A, will the donation not be destroyed.
To summarize, the consequences of choosing Option A and Option B are as follows:

Option A
You choose to help children with cancer through the donation. However, the donation will only realize if all other members of your group choose Option A.

Option B
You decide against helping children with cancer through the donation. Since the donation will only be made if all group members choose option A, this will destroy the donation in any case.

Furthermore, by choosing option B, you generate an additional payoff of 2 EUR for another participant in this study who is not part of your group. Any other person in your group who chooses option B will receive a sum of money of 2 EUR for themselves.

Decisions of the other group members

Hence, the consequences of your choice depend not only on you but also on the decisions of the other 5 members of your group: only if none of the members of your group chooses option B, will the donation be made. On the other hand, the payouts of 2 EUR to other people will be made regardless of the decisions of other group members.

You and the other 5 Members of your group decide one after the other. You are the **last** person to decide, so the other group members have already decided at this point. When it's your turn, you will learn if anyone among the people who decided before you has already chosen option B. In this case, the donation will be destroyed.

Summary

If you choose option A, you will not generate any additional payout. If you choose option B, you generate an additional payout of 2 EUR for another participant. Whether the donation of 15 EUR to the *Förderkreis* is destroyed depends on whether at least one member of your group has chosen option B. So the consequences of your choice depend not only on you but also on the choices of the other group member.

[Comprehension questions]

E.2.9 Group decision game -aligned version

Information

Your group

Together with 5 other people, you will form a group for your decisions. These people also participate in this study, so together, you will form a group of 6 individuals.
You can make all decisions autonomously and independently of the other members of the group. However, as described below, the consequences of your decision will depend on the decisions of the other group members.

The donation

Your group has been entrusted with a payment of 2 EUR, which is intended for another person, who is not part of your group.

Your options

You can choose between two options: Option A and Option B. Depending on which option you and the other group members choose, different consequences will be realized. In particular, both your decision and the decisions of the other group members affect what happens to the payment of 2 EUR to the other person.

The payment will be destroyed (that is, not executed) if at least one member of your group chooses option A. That is, if either you or at least one other person with whom you form a group chooses option A, the donation will be destroyed. Only if no member of your group chooses option A, i.e., **all** group members choose option B, will the donation not be destroyed.
To summarize, the consequences of choosing Option A and Option B are as follows:

Option A
You choose to help children with cancer through the donation. Furthermore, by choosing option A, you generate an additional donation of 15 EUR to the *Förderkreis*. Since the payment to the other person will only be made if all group members choose option B, this will destroy the payment in any case.

Option B
You decide against helping children with cancer through the donation and for the additional payment to another person, who is not part of your group. However, this payment only be made if all other members of your group choose option B.

Decisions of the other group members

Hence, the consequences of your choice depend not only on you but also on the decisions of the other 5 members of your group: only if none of the members of your group chooses option A, will the payment be made. On the other hand, the donation of 15 EUR to the charity will be made regardless of the decisions of other group members.

You and the other 5 Members of your group decide one after the other. You are the **last** person to decide, so the other group members have already decided at this point. When it's your turn, you will learn if anyone among the people who decided before you has already chosen option A. In this case, the payment will be destroyed.

Summary

If you choose option A, you generate a donation of 15 EUR to the *Förderkreis*. If you choose option B, you will not generate any donation. Whether the payment of 2 EUR to another person is destroyed depends on whether at least one member of your group has chosen option A. So the consequences

of your choice depend not only on you but also on the choices of the other group member.

[Comprehension questions]

E.2.10 Statement choice

Context

Carbon dioxide (CO_2) is a gas that is notably produced as a by-product of human-controlled combustion processes of fossil fuels such as crude oil, natural gas, or coal. One example of this is the generation of electricity in coal-fired power plants or driving a car. As a result, CO_2 enters the atmosphere, which leads to global warming (greenhouse effect). The increasing CO_2 concentration in the atmosphere, therefore, contributes significantly to man-made climate change and, thus, to the advancing destruction of the environment.

In the context of this study, it was planned to buy CO_2 certificates, which ensure that **one ton of** CO_2 is removed from the atmosphere. More precisely, this entails an amount of 25 EUR, which will be used for reforestation. The reforestation afforded by this amount of money results in the removal of one ton of CO_2 from the atmosphere.

Consequences

At the end of the study, a computer will randomly decide what will happen to the planned CO_2 certificates. The following two possibilities exist:

- With a probability of 50%, the certificates will be purchased as planned, thus helping the environment by removing one ton of CO_2 from the atmosphere.
- With a probability of 50%, the certificates will be destroyed, i.e., not acquired. This contributes to the destruction of the environment by leaving one ton of CO_2 in the atmosphere.

This selection is made by the computer in a completely random fashion. Hence it is not influenced by any actions or decisions of yours at any time.

In this context, you can make a statement of intent about how you feel about the destruction of the environment. Namely, you can make one of the following two statements:

- I support the preservation and protection of the environment.
- I support the destruction of the environment.

If you make the statement “I support the preservation and protection of the environment” this has no further consequence.

However, if you declare “I support the destruction of the environment” in the case that the certificates are destroyed, a part of the amount will be donated to the *Förderkreis für krebskranke Kinder und Jugendliche Bonn e.V.* instead. Namely, in this case, 15 EUR will be donated to the *Förderkreis*.

Thus, selecting a statement does not influence whether the CO_2 allowances are destroyed or not (since this happens randomly). But the statement influences whether 15 EUR will be donated to the *Förderkreis*. Please note that selecting a statement is not an opinion poll. Specifically, regardless of which statement you choose, the results will not be used to influence opinions about the environment.

[Comprehension questions]

E.2.11 Statement choice -aligned version

Context

Carbon dioxide (CO_2) is a gas that is notably produced as a by-product of human-controlled combustion processes of fossil fuels such as crude oil, natural gas, or coal. One example of this is the generation of electricity in coal-fired power plants or driving a car. As a result, CO_2 enters the atmosphere, which leads to global warming (greenhouse effect). The increasing CO_2 concentration in the atmosphere, therefore, contributes significantly to man-made climate change and, thus, to the advancing destruction of the environment.

In the context of this study, it was planned to buy CO_2 certificates, which ensure that **one ton of** CO_2 is removed from the atmosphere. More precisely, this entails an amount of 25 EUR, which will be used for reforestation. The reforestation afforded by this amount of money results in the removal of one ton of CO_2 from the atmosphere.

Consequences

At the end of the study, a computer will randomly decide what will happen to the planned CO_2 certificates. The following two possibilities exist:

- With a probability of 50%, the certificates will be purchased as planned, thus helping the environment by removing one ton of CO_2 from the atmosphere.
- With a probability of 50%, the certificates will be destroyed, i.e., not acquired. This contributes to the destruction of the environment by leaving one ton of CO_2 in the atmosphere.

This selection is made by the computer in a completely random fashion. Hence it is not influenced by any actions or decisions of yours at any time.

In this context, you can make a statement of intent about how you feel about the destruction of the environment. Namely, you can make one of the following two statements:

- I support the preservation and protection of the environment.
- I support the destruction of the environment.

If you make the statement “I support the destruction of the environment” this has no further consequence.

However, if you declare “I support the preservation and protection of the environment” in the case that the certificates are destroyed, a part of the amount will be donated to the *Förderkreis für krebskranke Kinder und Jugendliche Bonn e.V.* instead. Namely, in this case, 15 EUR will be donated to the *Förderkreis*.
Thus, selecting a statement does not influence whether the CO_2 allowances are destroyed or not (since this happens randomly). But the statement influences whether 15 EUR will be donated to the *Förderkreis*. Please note that selecting a statement is not an opinion poll. Specifically, regardless of which statement you choose, the results will not be used to influence opinions about the environment.

[Comprehension questions]

E.2.12 Rule following game

On the next screen, you will control a character that runs across the screen in an animation.
As soon as the animation starts, you can start running by pressing the "Start" button in the middle of the screen. Your character will move towards a series of traffic lights and will wait at each one. To move your character again, press the "Walk" button in the middle of your screen.
The rule is that you should wait at each traffic light until it turns green.
Your payout depends on how much time you need to run across the screen with your character. You start with an initial payout of 8 EUR, which decreases by 0.08 EUR every second.
The next screen takes you to the animation. If the animation is too large for your browser window, please change the zoom factor so that the animation is displayed in its entirety.

E.2.13 Rule following game -aligned version

On the next screen, you will control a character that runs across the screen in an animation.
As soon as the animation starts, you can start running by pressing the "Start" button in the middle of the screen. Your character will move towards a series of traffic lights and will wait at each one. To move your character again, press the "Walk" button in the middle of your screen.
The rule is that you should wait at each traffic light until it turns green.
Your payout depends on how much time it takes you to walk your character across the screen. You start with an initial payout of 0 EUR. Your payout increases by 0.08 EUR every second, up to a maximum of 8 EUR. The longer you wait at the traffic lights, the higher your payout will be. You will reach the maximum payout if you wait at each traffic light until it turns green.
The next screen takes you to the animation. If the animation is too large for your browser window, please change the zoom factor so that the animation is displayed in its entirety.

E.3 Self-versus-others block

E.3.1 Dictator game giving

In this section, you will receive 20 EUR as an additional payment. If you take no further action, you will receive the full amount. You can choose to donate a certain amount of the 20 EUR to the *Förderkreis für krebskranke Kinder und Jugendliche Bonn e.V.* You can choose any amount between 0

EUR and 20 EUR. This amount will then not be paid out to you but donated to the *Förderkreis*.

Thus, you will receive an additional payment 20 EUR minus the amount you donate.

E.3.2 Dictator game taking

This section is about a donation of 20 EUR to the *Förderkreis für krebskranke Kinder und Jugendliche Bonn e.V.* If you take no further action, the full amount will be donated. You can choose to take a certain amount of the donation for yourself as an additional payment. You can choose any amount between 0 EUR and 20 EUR. This amount will then not be donated to the *Förderkreis*, but paid to you.
Thus, 20 EUR minus the amount you take for yourself will be donated.

E.3.3 Trust game sender

Information

In this section, you and another person can send money to each other. Both of you have a private account. Initially, you and the other person have an endowment of 5 EUR in the account. Like you, the other person is a participant in today's study.

Your decision

From your 5 EUR endowment, you can decide to send an amount to the other person. You can send any whole euro amount, that is 1,2,3,4 or 5 EUR to the other person. In this case, each euro you send to the other person will arrive tripled to the other person. So, for example, if you send 1 EUR, the other person will receive 3 EUR, if you send 2 EUR, 6 EUR and so on.

The decision of the other person

So after you make your decision, the other person has 5 EUR + triple the amount you sent. Now the other person decides on an amount of money to send back to you. The person can choose any amount that the account allows for it. After that, the process ends and you will receive your account balances as an additional payment.

Example

You have decided to send 2 EUR to the other person. Therefore, your account balance is 5 EUR - 2 EUR = 3 EUR. The person now has 5 EUR + 3 X 2 EUR = 11 EUR on the account and can accordingly send you back any amount between 0 EUR and 11 EUR. Suppose the person sends you back 4 EUR.

This will give you a total of 4 EUR + 3 EUR = 7 EUR as an additional payment. The other person will receive 11 EUR - 4 EUR = 7 EUR as a payment.

E.3.4 Trust game receiver

Information

In this section, you now take the other person's role in the previous section. That is, now the other person decides how much money to send you, and you decide how much to send back. Note that you are sending money to a different participant for this section than before. They are not the same person, and your decision in the previous section has no bearing on your decision in this section.

Both of you again have a private account with an endowment of 5 EUR. Now the other person can decide how much of the 5 EUR they want to send you.

Your decision

You can indicate what you would send back for different possible amounts the other person can send. Afterward, it will be checked what amount the other person actually sent, and your decision for exactly this case will be implemented.

E.3.5 Public goods game

Your group

In this section, you will form a group with 2 other people that also participate in today's study. Your group members have been randomly assigned to you. At no time will you or the other members of your group learn who the members of your group are. All group members will face the same decisions and receive the same information.

Your endowment

For this section, you and each other group member will receive an amount of 5 EUR which is called your endowment. You decide how to use your endowment. You can contribute all or part of it to a project. Any amount you do not contribute to the project will automatically be put into what is called a private account. For example, if you contribute 3 EUR to the project, 5 EUR- 3 EUR = 2 EUR will be put into your private account.

You will receive income from the project and your private account as described below.

Income from the private account

For every Euro that is put into your private account, you will receive exactly 1 EUR as income. For example, if you contribute 0 EUR of your endowment to the project and thus 5 EUR is put into your private account, you will receive 5 EUR as income from your private account. If 3 EUR is put into your private account (because you contributed 2 EUR to the project), you will receive 3 EUR as private income. Nobody but you will receive income from your private account.

Income from the project

For every EUR you or another group member contributes to the project, you and all other group members receive 0.50 EUR. So, each group member's income from the project is determined as follows:

Income from the project = 0.5 * total sum of contributions paid in by all group members.

Example: If the total sum of contributions to the project from all participants in your group is 9 EUR (for example, if you and the two other members each contribute EUR3), you and all other participants in your group will receive 9 EUR * 0.5 = 4.5 EUR from the project. If the total of your contributions to the project is 12 EUR, you and all other participants will receive 12 EUR * 0.5 = 6 EUR from the project.

Total income

Your total income from part 1 is simply the sum of your income from the private account and your income from the project.

Simulation

You can try out in the simulation below how your total income changes based on your contribution and the contributions of your group members to the project. To do this, enter whole numbers in the respective fields.

[Calculator]

[Comprehension questions]

E.3.6 Moral luck game

Information

The following is about the decisions of another person who, like you, is a participant in this study. However, this person faced a different decision than you will be presented. At an earlier time, the person had a choice between the following two lotteries S and G:

Lottery S – S for Donation [Spende in German].

With a probability of 70%, a donation of 15 EUR to the *Förderkreis für krebskranke Kinder und Jugendliche Bonn e.V.* is triggered. With a probability of 30%, the person receives 10 EUR as an additional payment and no donation is triggered.

Lottery G – G for Money [Geld in German].

With a probability of 30%, a donation of 15 EUR to the *Förderkreis für krebskranke Kinder und Jugendliche Bonn e.V.* is triggered. With a probability of 70%, the person receives 10 EUR as an additional payment and no donation is triggered.

Determination of the consequences

At the end of the trial, the computer will randomly draw a number for the person from numbers 1 to 10. Depending on which lottery the person chose, the consequences of the number drawn will be different.

If the person chose lottery S, the donation will be triggered with a probability of 70%, so if either a 1, 2, 3, 4, 5, 6, or 7 is drawn. If the number drawn is an 8, 9, or 10, the person will receive 10 EUR as an additional payment.

If the person chose lottery G, the donation is triggered with a probability of 30%, so if either a 1, 2, or 3 is drawn. If the number drawn is a 4, 5, 6, 7, 8, 9, or 10, the person will receive 10 EUR as an additional payment.

Thus, the two lotteries differ in their probabilities of triggering the donation and the amount of additional payment to the other person. For Lottery S, a donation is much more likely than for Lottery G. Conversely, the probability of receiving an additional payment is higher for Lottery G than for Lottery S.

The consequences of the lotteries have no influence on your own payments.

Your decisions

In this section you will make decisions for different scenarios. For your decisions, you will be given 5 EUR for each scenario. You have the choice to divide the 5 EUR between you and the person who decided between the lotteries. You will receive every Euro you allocate to yourself as an additional payment. For every Euro you allocate to the other person, the person will subsequently receive 3 EUR as an additional payment.

For example, you can divide the 5 EUR so that you receive 4 EUR. Accordingly, the other person will then receive 1 EUR * 3 = 3 EUR as an additional payment. If you allocate yourself the entire 5 EUR, the other person will receive 0 EUR. If you allocate the entire amount to the other person you will receive 0 EUR and the other person will receive 5 EUR * 3 = 15 EUR and so on.

You will be asked how you want to distribute the amount of money from 5 EUR for each of four different scenarios.

Scenarios

The scenarios differ in which lotteries the person chose in each case, as well as which consequence happened to materialize. The following four scenarios exist:

Scenario 1:
The other person chose Lottery S. The donation was triggered and the person does not receive the additional payment.

Scenario 2:
The other person chose Lottery S. The donation was not triggered and the person receives the additional payment.

Scenario 3:
The other person chose Lottery G. The donation was triggered and the person does not receive the additional payment.

Scenario 4:
The other person chose Lottery G. The donation was not triggered and the person receives the additional payment.

For each of the four scenarios, you can decide individually how to distribute the 5 EUR based on the person's decision and the realized consequences. At no point does the other person receive any information about you.

At the end of the trial, you will be informed which scenario occurred by the matching of a person's decision to the consequences the computer has drawn for them.

Afterwards, you and the person then receive the amounts of money you selected for the scenario that occurred.

On the next screen you will see an example.

Example

You will see an example decision for Scenario 1. You can make the distribution of 5 EUR between you and the person using a scale and a slider. The associated slider for the scale appears when you click on the scale.

[Example]

If scenario 1 occurs, you would receive [XX] EUR by your decision and the other person would receive [XX] EUR. The other person will receive a total of [XX] EUR, because in this scenario the realized

consequence is that the other person will not receive an additional amount of money through the lottery, but will have triggered a donation.