Cognitive Processing
https://doi.org/10.1007/s10339-024-01226-9

RESEARCH ARTICLE

Choosing between bad and worse: investigating choice in moral dilemmas through the lens of control

Revati Shivnekar[1,2] · **Narayanan Srinivasan**[1,2]

Received: 10 May 2024 / Accepted: 20 August 2024

Abstract
People's preferences for the utilitarian outcome in sacrificial moral dilemmas, where a larger group of individuals are saved at the cost of a few, have been argued to be influenced by various factors. Taking expected utility (EU) theory into consideration, we investigate whether the expected effectiveness of actions elucidate certain inconsistencies in moral judgments. Additionally, we also explore whether participants' role in the dilemma as the executor or a superior who merely makes a decision, which is carried out by a subordinate, influences judgments—a factor generally overlooked by classical EU models. We test these hypotheses using a modified moral dilemma paradigm with a choice between two actions, one highly successful and the other more likely to fail. Both actions are either expected to result in a favorable outcome of saving five individuals by sacrificing one or an unfavorable outcome of sacrificing five to save one. When the efficient action is anticipated to lead to a favorable outcome, in line with EU models, people almost invariably choose the efficient action. However, in conditions where the EUs associated with efficient and inefficient actions are close to each other, people's choice for favored outcome is above chance when they act as agents themselves. We discuss the implications of our results for existing theories of moral judgments.

Keywords Moral decision-making · Trolley problems · Control effectiveness · Utilitarianism

Introduction

Moral dilemmas that pit the utilitarian and deontological principles against each other have been widely used to study moral decision-making precisely because people are curiously inconsistent in judging them (Greene et al 2001; Koenigs et al. 2007; Bartels 2008; Suter and Hertwig 2011). A typical sacrificial moral dilemma poses a choice between an action that saves many by killing a few or letting the larger group die by choosing not to act. The former is called a utilitarian action because it maximizes lives saved, while the latter preserves the rule-based deontological ethical stance favoring means over the end. A famous example of such a moral dilemma is the trolley problem, first posed by Philippa Foot in 1967 (see Fig. 1 for illustration of trolley problems). Imagine you are the driver of a runaway trolley. After a bend, you see five repairmen working on the track you are on. There is no way to stop the trolley and avoid killing the repairmen unless you swerve the trolley onto another track by pressing a switch inside the trolley. However, this alternate track has one repairman working on it, whom you would end up killing if you switch tracks. The question is whether it is morally permissible to swerve the trolley.

In this dilemma, commonly called the Switch problem, the trade-off is saving five people by killing one. Interestingly, people's permissibility judgments do not solely depend on the trade-off but are also influenced by the action. To test this, philosophers and moral psychologists have created different versions of this dilemma by altering the action used to save the larger group while keeping the trade-off the same. In the Switch problem, pressing a switch diverts the trolley onto a sidetrack, killing the lone worker on that track. About 81% of participants choose the utilitarian action

Editors: Riccardo Brunetti (European University of Rome), Napoleon Mabaquiao (De la Salle University); Reviewers: Mark Lazara Dacela (De la Salle University) and a second researcher who prefers to remain anonymous.

✉ Narayanan Srinivasan
nsrini@iitk.ac.in

[1] Centre of Behavioural and Cognitive Sciences, University of Allahabad, Allahabad, India

[2] Department of Cognitive Science, Indian Institute of Technology Kanpur, Kanpur, India

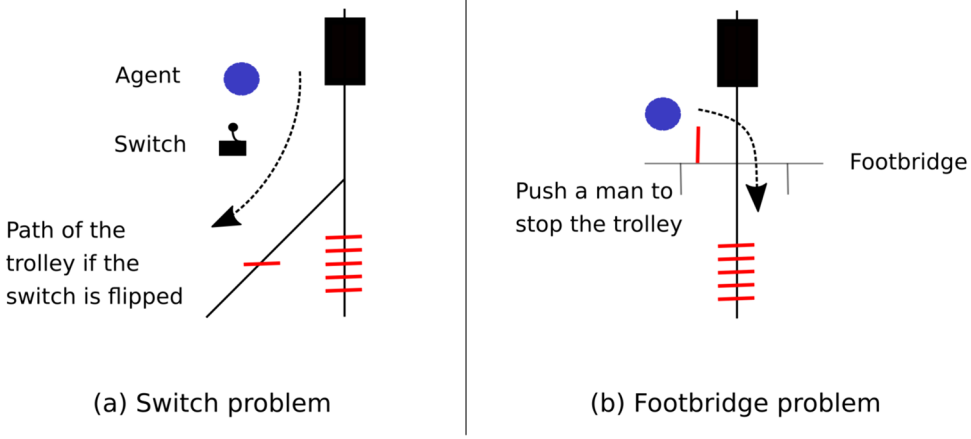

Fig. 1 Schematic depicting two trolley problems. The trade-off of lives saved and sacrificed is the same in both dilemmas, but the intervening action is different

in Switch. By contrast, in the Footbridge problem, when the utilitarian action is to push a heavy person off a footbridge to save the workers, the utilitarian response rate drops to a country-level average of 51% (Awad et al. 2020; see Fig. 1 for a schematic of Switch and Footbridge problems). Personal force or directness of action has been proposed as a potential predictor of such an inconsistency in moral judgments (Greene et al 2009). In personal dilemmas like Footbridge, the agent kills the victim by a direct application of muscle-force like in pushing and choking while in impersonal dilemmas like Switch, agents generate a muscle force that transforms into a mechanistic kind of force and ends up killing the victim, like by using switches and guns. People tend to be more utilitarian insofar as the dilemma is impersonal, and the action saves the larger group as a side effect, not as a means to an end.

Moral dilemmas that necessitate a choice between a utilitarian action and its deontological omission or inaction led us to consider *why* the personalness of an action contributes to the asymmetry in moral judgments. We propose that the inconsistency in utilitarian judgments may (at least in part) stem from the difference in the strength of the agent's belief in exerting control over situations through actions particular to Switch and Footbridge. In the former case, flipping a switch is perhaps perceived as less likely to fail than pushing a person, a larger person at that, off a footbridge. The person may fight back, or one may lack the necessary strength to push him. It is not hard to imagine that most would find it more challenging to execute pushing a man than flipping a switch successfully. Even if participants consider the utilitarian outcome a worthy goal, the action that affords them that goal may be ineffectual, leading to preference for a successful action at the cost of achieving the utilitarian outcome.

Indeed, previous research indicates that participants may not understand the success likelihood of actions across dilemmas systematically. Shou and Song (2017) demonstrated that participants' understanding of the outcome probabilities was far from deterministic as was intended in the task instruction. They asked participants to state the perceived probability of positive (people being saved) and negative (people dying) outcomes occurring after taking an action or resorting to inaction. Participants expected the positive outcome to be more probable in the Switch but not in the Footbridge variant. They were also less likely to choose the option if the negative consequence attached to it was perceived to be more probable. Similarly, a more direct manipulation of outcome probabilities by stating that the action is less successfully linked to the outcome made participants perceive the action as less appropriate and less moral (Kortenkamp and Moore 2014).

To further motivate our hypothesis, we draw upon the proposal by Higgins (2015) that behavior may be influenced by two independent factors: value effectiveness (the value inherent in achieving the goal) and control effectiveness (how efficiently a goal is achieved). When a highly worthwhile goal is achieved by employing a highly effective action (high control effectiveness), choosing that action means choosing the desired goal. However, if these two factors operate independently, then when the highly effective action is expected to lead to an unwanted outcome, participants' preferences should devalue the said outcome. We based our control effectiveness hypothesis on the possibility that people's judgments may be informed by how successful the actions given in the dilemma were in the past. Particularly, that the expected effectiveness of the action influences preferences for the utilitarian outcome. This hypothesis also aligns well with classical economic expected utility models, which propose that people's preferences are motivated not just by the costs and benefits realized by actions, but also the probability of those actions' success (see Shenhav and Greene 2010 for employing the expected utility model in moral dilemmas).

However, these models do not consider the participant's role within the moral judgments. It has also been shown

that people tend to be more permissive of utilitarian actions when participants are mere observers rather than executors themselves, often referred to as 'actor-observer bias' (Nadelhoffer and Feltz 2008; Avram et al. 2014; Tassy et al. 2013; Gold et al. 2015). People also report having lesser control over the outcome as executors than observers (Nadelhoffer and Feltz 2008). Nadelhoffer and Feltz (2008) argue that this may be so because people as executors are more heavily influenced by the negative emotions of killing someone and to justify not intervening, they assume that they have lesser control over the utilitarian action. When the payoffs associated with the actions and their likelihood is held constant, according to the classical EU models or Higgin's control effectiveness proposal, the participant's role should not make a difference in their judgments. We test whether manipulating the participant's role as the executor or a superior on whose behalf the action is carried out alters their preferences in sacrificial dilemmas.

To summarize, in this experiment, we manipulated the control afforded by actions in achieving a favorable outcome and the agent's role within a dilemma. By favorable outcome we mean an outcome that saves many by killing few, specifically, take one life to save five. On the contrary, an outcome is unfavored when it kills many to save a few. Two critical differences set our design apart from the traditional moral dilemma paradigm. Firstly, we presented participants with a choice between two actions of varying effectiveness (one highly successful and the other more likely to fail), rather than presenting a choice between action and omission. This was done to disentangle the effects of deontological inclination, removing the inaction option. This approach acknowledges that choosing the utilitarian action does not necessarily indicate a strong utilitarian influence and weak deontological one, and vice versa for the deontological choice (see Kahane et al. 2015; Conway and Gawronski 2013). Additionally, it allowed us to associate effective and ineffective actions separately with the favorable outcome. To achieve this, we adapted a Killer Shark dilemma (as described in Methods) such that favorable and unfavorable outcomes were obtained through two separate actions differing in their success rates. In the high control condition, the more successful action resulted in the favorable outcome, while in the low control condition, the less successful action was linked to it. In the high control conditions case, value and control effectiveness can be achieved through the same action, and thereby, participants could afford more control over the desired outcome. These conditions were set such that the expected utility (EU) of the highly efficient action was almost four times the inefficient action. Alternatively, in the low control condition, value and control effectiveness do not match up, limiting participant's control over the same outcome. The EUs associated with both actions were also almost equal. As a result, we expected fewer participants would choose the favorable outcome in the low control compared to high control condition. Secondly, before making a choice, we trained participants to learn the effectiveness of each of the two actions since our ability to obtain successful outcomes through actions is learned through experience and interactions with the world. In the past, it has been assumed that participants' experience with the action outside the lab affects their choices (Shou and Song 2017). We wanted to explore if training can induce such an effect within the lab.

Method

Participants

One hundred and ninety-two participants in the age group 18–30 years from the University or other educational institutions in the city (92 females; mean age = 23.3). The Institutional Ethics Review Board approved the study.

Experimental design

Our central aim in this experiment was to investigate how a participant's judgment in a moral dilemma is affected by the control effectiveness of the choice actions and who carries out the action. We employed a between-subjects 2×2 design with control (high and low) and agent-type (other and self) as independent variables. We modified the Killer Shark dilemma from Christensen et al. (2014) to manipulate these two variables across conditions. Unlike the usual moral dilemma settings, in our experiment participants had to make a choice between two actions that led to two different outcomes. An action was either highly successful or it frequently failed to achieve the associated outcome. The outcome was either favorable (killing 1 saves 5) or unfavorable (killing 5 to save 1). The control variable was operationalized based on the action-outcome contingency in a dilemma. In high-control settings, the action with the most success or effectiveness was associated with the favorable outcome. In contrast, in low-control situations, the action with the least success or effectiveness was linked to the unfavorable outcome. The variable agent type denoted whether the actions within the dilemma were executed by participants as themselves (self condition) or by participants ordering their subordinate (other condition). Effectively, we used a fully crossed design with four conditions: High-control + Other-agent (HO), High-control + Self-agent (HS), Low-control + Other-agent (LO), and Low-control + Self-agent (LS). Each participant was randomly assigned to one of these four conditions.

Materials

The original Killer Shark dilemma read as follows:

> "You and ten divers are part of an U.N. team who is deactivating anti-ship mines from World War II. One team member has hurt himself and the blood in the water has attracted several sharks. You have an underwater rifle but only one harpoon and there are many sharks. The bleeding diver is swimming towards the last protective cage and will reach it before you and the others. The sharks, following the blood, are coming too close for you and the other divers to escape.
>
> If you shoot at the injured diver this will kill him and the sharks will stop to eat him, but you and the ten divers will be saved.
>
> Do you let the sharks eat the injured diver by shooting at him, so you and the other ten divers can reach the protective cage?"

We modified this dilemma so that the choice was between shooting with an underwater rifle or throwing and detonating an underwater grenade. These actions were explicitly tied to the favorable outcome of saving many by killing few or the reverse of it. In addition to manipulating the action-outcome contingencies, who executed the action was also manipulated across the conditions with either the participant as the executor (Self condition) or the subordinate whom participants could order (Other condition). Following is the dilemma that participants in HO condition saw (The trade-offs for all the conditions are shown in the Table 1):

> "You are a marine researcher studying sharks underwater. You and your guard are safe inside a protective cage where the sharks cannot reach you, but 6 of your colleagues are swimming outside the cage. One of your colleagues hurts himself and starts bleeding profusely. This attracts sharks' attention. The injured colleague has drifted away from the rest of the group.
> If you order your guard to shoot the injured colleague, sharks will eat him, and five colleagues will have enough time to reach the cage.
> If you order your guard to throw an underwater grenade at 5 colleagues and kill all of them, then the sharks will stop to eat them and your injured colleague will have time to reach the cage.
> The guard has time to use only one of the weapons. If he does not do anything, then all 6 colleagues will die.
> What do you do?
> ORDER TO SHOOT AND KILL 1 TO SAVE 5? Press A.
> Or
> ORDER TO THROW A GRENADE AND KILL 5 TO SAVE 1? Press L."

Before reading the dilemma, participants learned about the effectiveness of each action (see Procedure below). The rifle consistently achieved success, while the grenade often failed to hit its target accurately across all conditions. The rifle was associated with the favorable outcome in the high control conditions (HO, HS). It afforded higher control to participants by being the more successful weapon of the two. In contrast, in the low control conditions (LO, LS), the grenade linked with the utilitarian end was more likely to fail, leaving participants with less control over the favored outcome. Table 1 describes the trade-offs presented in each condition. Based on these trade-offs, we computed the expected.

utilities (EU) of actions in terms of number of lives lost. EUs of the rifle and grenade were calculated as follows under high-control conditions:

$$EU_{HO,HS}(rifle) = (p(success) \times (-killed)) + (p(failure) \times (killed))$$
$$= (0.9 \times -1) + (0.1 \times -6) = -1.5$$

$$EU_{HO,HS}(grenade) = (p(success) \times (saved - killed)) + (p(failure) \times (killed))$$
$$= (0.1 \times -5) + (0.9 \times -6) = -5.9$$

Table 1 Action-consequence trade offs within conditions

Condition	Description
High control + Other agent (HO)	Order a subordinate guard to shoot 1 to save 5 with the rifle or kill 5 to save 1 by throwing a grenade
High control + Self agent (HS)	Shoot 1 to save 5 with the rifle or kill 5 to save 1 by throwing a grenade
Low control + Other agent (LO)	Order a subordinate guard to shoot 5 to save 1 with the rifle or kill 1 to save 5 by throwing a grenade
Low control + Self agent (LS)	Shoot 5 to save 1 with the rifle or kill 1 to save 5 by throwing a grenade

Similarly, the EUs of the rifle and grenade in the low-control scenario were:

$$EU_{LO,LS}(rifle) = (0.9 \times -5) + (0.1 \times -6) = -5.1$$

$$EU_{LO,LS}(grenade) = (0.1 \times -1) + (0.9 \times -6) = -5.4$$

Procedure

Participants were brought into the lab where they completed the experiment alone. They were randomly assigned to one of the four conditions. After taking a written consent to participate, the experimenter instructed them to imagine themselves being in the story as much as they could. The experiment was split into training and testing phases. In the training phase, participants were introduced to the context for the dilemma to be introduced in the testing phase. For instance, participants in the Other condition (HO, LO) read the following text before beginning the training phase.

> "Imagine you are a marine researcher. You are working underwater in an area that is infested with killer sharks. You have been given a subordinate guard whom you can order to either shoot an underwater rifle or throw an underwater grenade in case of danger. In the following session, you will get to learn how well the guard can shoot and throw a grenade underwater at a target. Press SPACE when you are ready."

Participants learned success rates for each weapon over the course of 40 trials by clicking a button that provided feedback indicating whether the target was hit or not. The rifle was always more effective than the grenade, with a 90% success rate vs. a 10% success rate for the grenade. During the testing phase, participants read a version of Christensen et al.'s (2014) Killer Shark dilemma altered according to the assigned condition, with each action and its corresponding consequence explicitly described. Participants were required to record their responses by pressing an appropriate key.

Participants were also informed that the action and its associated consequences represented the only possibilities in the dilemma. In low control conditions, for instance, where shooting with the rifle was linked to the scenario of sacrificing five to save one, participants were made aware that the rifle could not result in fewer casualties due to situational constraints. They were explicitly told that the action-outcome contingencies presented were the sole options available.

Finally, following the testing phase, participants were required to specify which of the two weapons they (or their subordinate) performed better. This question served as a qualifying criterion, ensuring that the participants had appropriately learned the efficacy of weapons in the training phase.

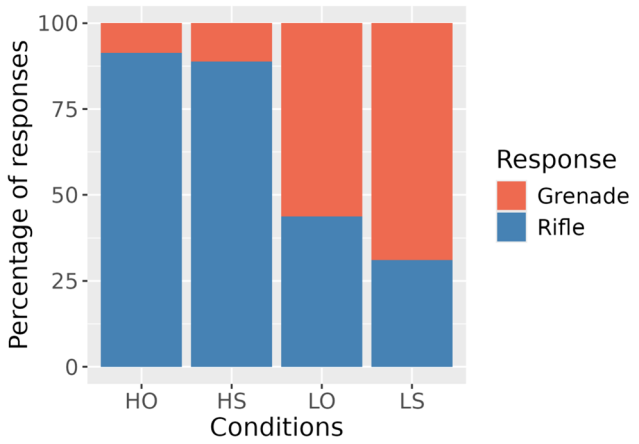

Fig. 2 Percentages of responses choosing the rifle within each condition. Note that favorable outcome in high control conditions (HO and HS) is linked to the high control affording rifle while in low control conditions (LO and LS), it is linked to the low control affording grenade

Table 2 Odds ratios with 95% CI

		Odds ratios with 95% CI bounds		
	Coefficient estimates with SE	Lower	Odds ratio	Upper
Intercept	0.25 (0.29)	0.73	1.29	2.3
Control	2.1 (0.6) *	2.76	8.17	30.36
Agent-type	0.54 (0.43)	0.74	1.72	4.09
Interaction	−0.82 (0.83)	0.0872	0.4424	2.25

* indicates $p < .001$

Results

Out of the total 192 participants, 184 successfully met the qualifying criterion by correctly identifying shooting with the rifle as the more effective action. Among these participants, 46 were randomly assigned to the High-Other condition (HO), 45 to the High-Self condition (HS), 48 to the Low-Other condition (LO), and 45 to the Low-Self condition (LS).

The rifle, which was the more successful of the two weapons, was consistently chosen over the grenade in high control (91% in HO and 89% in HS) but not in low control conditions (44% in LO and 31% in LS; see Fig. 2). Logistic regression analysis of participants' choice of weapon revealed a significant main effect of the control condition alone. Adding the control variable to the baseline logistic

regression significantly improved the prediction rather than without it ($\chi^2(1) = 59.42; p < .001$); however, the agent-type ($\chi^2(1) = 0.98; p = 32$). The odds of participants in the high control conditions choosing the rifle were 13.5 [4.5745, 50.2883] times more than those in low control conditions. The odds of choosing the rifle when participants were the executors themselves were 0.5806 [0.2443, 1.35] (see Table 2 for the full regression table).

Next, we examined how participants' preferences for the weapons aligned with their expected utilities (EUs). In high control conditions, where the EU for the rifle was considerably higher than for the grenade, participants consistently chose the rifle, as expected. Conversely, in low control conditions, where the EUs for both the rifle and grenade were comparable, we expected that participants should not have a clear preference for either option in these conditions. But the frequency of choosing the rifle in these conditions was observed to be lower than chance, more so in LS. We wanted to see whether the alignment between participants' preferences and expected utilities was distorted in LS. We compared the ratios of expected and observed preference for choosing the rifle. This comparison revealed a significant difference in LS (Fisher's $p = 0.03$, odds CI = [0, 0.92]) but not in LO (Fisher's $p = 0.42$, odds CI = [0, 1.79]), suggesting that preferences aligned more closely with expected utilities when participants ordered their subordinate.

Discussion

Our experiment was motivated by the widespread observation that people prefer the utilitarian outcome disproportionately more in Switch than Footbridge. This asymmetry has been consistently attributed to how the action brings about the outcome. When the action is up-close and personal such as the pushing act in Footbridge, people are less likely to endorse it and less so to carry it out themselves. We propose that the personalness of an action matters to outcome selection at least partially because of what these actions entail about the effectiveness of the actions. Mikhail (2011) hypothesizes that the unrealistic expectation of actions to always have the expected consequences may potentially confound the interpretation of preferences in dilemmas. Shoving a man off a height is assumed to be successful as probable as being able to throw a switch. Certainly, this assumption is hard to justify and worth questioning (Shou and Song 2017; Kortenkamp and Moore 2014; Royzman et al. 2015).

We demonstrated that the preference for the favorable outcome (simply in terms of numbers of lives saved) is curtailed by the (in) efficacy of actions through which they are achieved. Participants in the high control condition chose the outcome that saved more lives more often than in the low control condition. In these dilemmas, the action-consequence contingency was matched the best way possible: the highly efficient action of shooting the rifle also saved a greater number of people, resulting in almost all participants choosing the rifle. In the low control condition, the preferences of more than half of participants demonstrate a devaluation of this favorable outcome (with only 56% and 69% choosing the favorable outcome in LO and LS, respectively). This shift away from the favorable outcome in low control conditions was a result of participants choosing the rifle, which led to the unfavorable outcome but afforded them more control instead. When faced with choosing between the bad and worse options in the low control condition, many preferred to have more control over the outcome, even if it meant selecting the worst option in terms of the number of lives saved.

Our results align well with Mikhail's (2011) model of choices based on expected utilities. Endorsing saving a larger number of lives by killing a few is easy in the high control conditions as the EU of the more successful action was much higher than the less successful ($|\delta(EU)| = 4.4$). Contrastingly, this difference was smaller in the low control conditions ($|\delta(EU)| = 0.4$), which is reflected in the more evenly split choice between the two actions especially in the low control LO condition.

However, participants' choices in our experiment cannot be fully explained within the purely deontological (people choosing to not intervene) nor within the expected utility maximizing model. Endorsing saving a larger number of lives by killing a few is easy in the high control conditions as the EU of the more successful action was much higher than the less successful ($|\delta(EU)| = 4.4$). Contrastingly, this difference was smaller and in favor of the rifle in the low control conditions ($|\delta(EU)| = 0.4$). In other words, if participants solely operated based on the EU maximizing model, then they should choose the rifle as often as or marginally more than grenade. This was, indeed, the case in LO which had an even split choice between the two actions. However, in LS, participants chose the grenade more than expected and the difference was significant. That is, when participants acted as executors themselves, they devalued the favorable outcome compared to the high control conditions but *not as much* according to the EU model. Here although we demonstrate the 'actor-observer bias', which is that people choose the utilitarian outcome more as observers than executors, it has the opposite pattern (Nadelhoffer and Feltz 2008; Avram et al. 2014; Tassy et al. 2013; Gold et al. 2015). One difference is that the utilitarian outcome in earlier studies is not the same as the favorable outcome in our experiment, since we separate an utilitarian outcome from a favorable one. Favorable outcome simply considers the lives saved, while the utilitarian outcome should consider the lives saved as well as the likelihood associated with it. Secondly, it could

also be argued that even in the *Other* conditions when the participant acts as a superior who orders the subordinate, the participant is still the executor with one degree of separation. Further experimentation is needed to understand this effect.

Our experimental design was unlike the classic or typical dilemma problem to serve three broad objectives. One, we wanted to investigate the utilitarian choice isolated and deconfounded from the deontological influences (see Conway and Gawronski 2013). Two, we expected that the expected likelihood of an action's success (its efficacy) may be informed by the decision maker's previous experience with it (see Greene et al 2009; Shou and Song 2017; Kortenkamp and Moore 2014). And lastly, our experimental manipulation required actions with different efficacies to facilitate comparison. We modified a sacrificial moral dilemma to satisfy these objectives such that participants learned the success rates of two actions that were starkly different in their efficacies. Ultimately, they made a choice between the two actions informed by their likelihood of success. However, one may argue that such a dilemma does not remain an ethical dilemma anymore. Participants may be driven by wanting to minimize harm alone, without considering whether they 'should' take an action. In other words, rule-based moral considerations may not factor into decision-making at all. However, if participants were influenced solely by outcomes, then we would not see above-chance preference for the maximized outcome in low control conditions. Further, ethical dilemmas are not only those that contrast action with inaction. When a decision-maker decides to intervene, the ethical choice about which course of action to choose still remains. The choice becomes especially tricky in low control conditions when the value effectiveness of the utilitarian outcome is compromised. Earlier proposals such as the Subjective Utilitarian Thoery (SUT) by Cohen and Ahn (2016) also argue that instead of calling inaction as "deontological" and action as "utilitarian", the moral decision-making can be characterized as a unitary process that is aimed at maximizing perceived psychological utility within the choice (see Hennig and Hütter 2020). Hence, we believe that our alterations to the classic dilemma setting allowed us to isolate and investigate the preference for utilitarianism without losing the moral aspects of the dilemma.

While our control conditions were comparable to Switch and Footbridge, our control effectiveness proposal extends beyond these two trolley problems. The foundation of our hypothesis is the possibility that actions may not always succeed once initiated. For instance, in the Footbridge scenario, the victim might fight back, increasing the likelihood of the action failing. In our experimental manipulation, we anticipated that learned action-effect contingencies would inform moral decisions. Predictions about the success of actions influence moral judgments; if people expect the action to fail, their endorsement of that action decreases (Greene et al. 2009).

Reasoning about moral permissibility involves not just the intended goal of an action but also the likelihood of the action achieving that goal. This expectation may be based on prior experience with the specific action. It would be interesting to examine how these expectations extend to similar actions. For example, if participants in the test phase of the experiment were given similar weapons but not exactly the same one they had prior experience with, we could observe whether their learning about the action is specific to that weapon or extends to similar ones. We also demonstrate that theorizing about moral permissibility should also consider the role of the executor within the dilemma. A decision-maker's role as the direct or indirect executor can influence the choice made. Further experimentation is needed to understand why and under what conditions the role of the executor affects the choice. A possible explanation for this effect can be found within the dimension of control effectiveness, as we suggested earlier.

Additionally, the control effectiveness might be inherently low in certain actions. These may include personal actions which operate through unmediated muscle force directly applied to the victim (Greene et al. 2009; Moore et al. 2008), physical contact (Cushman et al. 2006), direct actions without personal force (Greene et al. 2009; Royzman et al. 2015) etc. The possibility of our moral judgments being constrained by embodied cognition has been alluded to in earlier studies in terms of" representations such as goals-within-the-reach-of-muscle-force" (Greene et al. 2009). They speculate that any action plan within moral judgments may reject harm as a goal state when attempted by a direct application of muscle force. Our results indicate that such an embodied representation would also be constrained by the control effectiveness of the actions.

Conclusion

We propose an explanation of an action's control over the utilitarian outcome to describe the inconsistent utilitarian judgment tendencies in personal and impersonal dilemmas. We show that participants are less utilitarian when the utilitarian action is perceived to be ineffective. In such cases, people choose the action that affords more control over the outcome, even if it would cause more harm than save. We also suggest that the personal force explanation of moral judgments may be informed by including considerations from control affordances.

Author contribution Both authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Revati Shivnekar. The first draft of the manuscript was

written by Revati Shivnekar and edited by both authors. Both authors wrote and approved the final manuscript.

Data Availability All the stimuli and analyses are available online (osf.io/w9cqp).

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

References

Avram M, Hennig-Fast K, Bao Y, Pöppel E, Reiser M, Blautzik J, Gutyrchik E (2014) Neural correlates of moral judgments in first- and third-person perspectives: implications for neuroethics and beyond. BMC Neurosci 15(1):1–11

Awad E, Dsouza S, Shariff A, Rahwan I, Bonnefon JF (2020) Universals and variations in moral decisions made in 42 countries by 70,000 participants. Proc Natl Acad Sci 117(5):2332–2337

Bartels DM (2008) Principled moral sentiment and the flexibility of moral judgment and decision making. Cognition 108(2):381–417

Christensen JF, Flexas A, Calabrese M, Gut NK, Gomila A (2014) Moral judgment reloaded: a moral dilemma validation study. Front Psychol 5:607. https://doi.org/10.3389/fpsyg.2014.00607

Cohen DJ, Ahn M (2016) A subjective utilitarian theory of moral judgment. J Exp Psychol Gen 145(10):1359

Conway P, Gawronski B (2013) Deontological and utilitarian inclinations in moral decision making: a process dissociation approach. J Pers Soc Psychol 104(2):216

Cushman F, Young L, Hauser M (2006) The role of conscious reasoning and intuition in moral judgment: testing three principles of harm. Psychol Sci 17(12):1082–1089

Foot P (1967) The problem of abortion and the doctrine of double effect. Oxford 5:5–15

Gold N, Pulford BD, Colman AM (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. J Econ Psychol 47:50–61

Greene JD, Cushman FA, Stewart LE, Lowenberg K, Nystrom LE, Cohen JD (2009) Pushing moral buttons: the interaction between personal force and intention in moral judgment. Cognition 111(3):364–371

Greene JD, Sommerville RB, Nystrom LE, Darley JM, Cohen JD (2001) An fMRI investigation of emotional engagement in moral judgment. Science 293(5537):2105–2108

Hennig M, Hütter M (2020) Revisiting the divide between deontology and utilitarianism in moral dilemma judgment: a multinomial modeling approach. J Pers Soc Psychol 118(1):22

Higgins, E. T. (2015). What is value? where does it come from? a psychological perspective. *Handbook of value: Perspectives from economics, neuroscience, philosophy, psychology and sociology*, 43–63.

Kahane G, Everett JA, Earp BD, Farias M, Savulescu J (2015) 'Utilitarian' judgments in sacrificial moral dilemmas do not reflect impartial concern for the greater good. Cognition 134:193–209

Koenigs M, Young L, Adolphs R, Tranel D, Cushman F, Hauser M, Damasio A (2007) Damage to the prefrontal cortex increases utilitarian moral judgements. Nature 446(7138):908–911

Kortenkamp KV, Moore CF (2014) Ethics under uncertainty: the morality and appropriateness of utilitarianism when outcomes are uncertain. Am J Psychol 127(3):367–382

Moore AB, Clark BA, Kane MJ (2008) Who shalt not kill? Individual differences in working memory capacity, executive control, and moral judgment. Psychol Sci 19(6):549–557. https://doi.org/10.1111/j.1467-9280.2008.02122.x

Mikhail J (2011) Elements of moral cognition: Rawls' linguistic analogy and the cognitive science of moral and legal judgment. Cambridge University Press

Nadelhoffer T, Feltz A (2008) The actor–observer bias and moral intuitions: adding fuel to Sinnott-Armstrong's fire. Neuroethics 1:133–144

Royzman EB, Kim K, Leeman RF (2015) The curious tale of Julie and Mark: unraveling the moral dumbfounding effect. Judgm Decis Mak 10(4):296–313

Shenhav A, Greene JD (2010) Moral judgments recruit domain-general valuation mechanisms to integrate representations of probability and magnitude. Neuron 67(4):667–677

Shou Y, Song F (2017) Decisions in moral dilemmas: the influence of subjective beliefs in outcome probabilities. Judgm Decis Mak 12(5):481–490

Suter RS, Hertwig R (2011) Time and moral judgment. Cognition 119(3):454–458

Tassy S, Deruelle C, Mancini J, Leistedt S, Wicker B (2013) High levels of psychopathic traits alters moral choice but not moral judgment. Front Hum Neurosci 7:229