



Should you save the more useful? The effect of generality on moral judgments about rescue and indirect effects

Lucius Caviola^{a,*}, Stefan Schubert^b, Andreas Mogensen^c

^a Department of Psychology, Harvard University, United States of America

^b Uehiro Centre for Practical Ethics, University of Oxford, United Kingdom

^c Global Priorities Institute, University of Oxford, United Kingdom

ARTICLE INFO

Keywords:

Moral judgment
Deontology
Utilitarianism
Rules
Construal level

ABSTRACT

Across eight experiments ($N = 2310$), we studied whether people would prioritize rescuing individuals who may be thought to contribute more to society. We found that participants were generally dismissive of general rules that prioritize more socially beneficial individuals, such as doctors instead of unemployed people. By contrast, participants were more supportive of one-off decisions to save the life of a more socially beneficial individual, even when such cases were the same as those covered by the rule. This *generality effect* occurred robustly even when controlling for various factors. It occurred when the decision-maker was the same in both cases, when the pairs of people differing in the extent of their indirect social utility was varied, when the scenarios were varied, when the participant samples came from different countries, and when the general rule only covered cases that are exactly the same as the situation described in the one-off condition. The effect occurred even when the general rule was introduced via a concrete precedent case. Participants' tendency to be more supportive of the one-off proposal than the general rule was significantly reduced when they evaluated the two proposals jointly as opposed to separately. Finally, the effect also occurred in sacrificial moral dilemmas, a general phenomenon occurring in multiple moral contexts. We discuss possible explanations of the effect, including concerns about negative consequences of the rule and a deontological aversion against making difficult trade-off decisions unless they are absolutely necessary.

"We must, however, not only make this general statement, but also apply it to the individual facts. For among statements about conduct, those which are general apply more widely, but those which are particular are more true".

Aristotle, *Nicomachean Ethics* 1107a31–2 (transl. Ross)

1. Introduction

People often face difficult moral choices. For example, a hospital board may have to decide which of two people's healthcare needs should be met when not everyone's needs can be met. In addition to having to decide on concrete cases, people also face decisions about what rules or policies to adopt for such cases. For instance, politicians in the health ministry may have to decide on a general principle for prioritizing among individuals when not everyone's healthcare needs

can be met.

Do people make the same judgments regarding one-off cases and the general rules that cover those cases? In this paper, we investigate whether level of generality affects moral judgments and show that case-specific judgments and general rules enjoining those case-specific judgments can differ in their level of support. Specifically, we consider situations in which people's judgments concerning general policies are more characteristic of a deontological approach to ethics, whereas their judgments about concrete cases are more characteristic of a utilitarian approach.

Our research is inspired by existing philosophical debates concerning the relevance of indirect benefits in priority setting. Some philosophers argue that medical need is the only proper distributive criterion for medical resources (Walzer, 1983; Williams, 1973). However, attending to the needs of some patients rather than others may indirectly produce greater (medical or non-medical) benefits for others, and it is not

* Corresponding author.

E-mail address: lucius.caviola@gmail.com (L. Caviola).

<https://doi.org/10.1016/j.cognition.2020.104501>

Received 13 September 2019; Received in revised form 15 October 2020; Accepted 22 October 2020

Available online 4 November 2020

0010-0277/© 2020 Elsevier B.V. All rights reserved.

obvious why these indirect benefits should be ignored (Brock, 2003; Du Toit & Millum, 2016; Kamm, 1993; Lippert-Rasmussen & Lauridsen, 2010; Mogensen, 2020). We focus on prioritization decisions that involve a choice between saving one of two individuals who may be said to differ in how much they contribute to society: for example, a choice between saving the life of a doctor specializing in a form of life-saving heart surgery that very few people in the world can perform and the life of a person who has been unemployed for a long time and is unlikely to find work within the foreseeable future.¹ Prioritizing the doctor in light of her abilities may seem unfair. On the other hand, doing so is likely to save more lives overall, given the doctor's unique skills. We thus have a moral dilemma, in which different moral criteria (fairness, saving lives) appear to be pitted against one another. We are interested in whether people weigh these criteria differently when considering general rules as opposed to one-off decisions.

1.1. Attitudes towards prioritization policies in the medical context

Surveys from health policy research suggest that people have strong objections to rules that prioritize individuals who contribute more to society. In an Australian survey, McKie and Richardson (2011) found that respondents strongly disagreed with a proposal to give higher priority to people who contribute more to society in the Australian public healthcare system. Similarly, Skitka and Tetlock (1992) showed that American students consider prioritizing people on the basis of their contributions to the community as inappropriate. Nord, Richardson, Street, Kuhse, and Singer (1995) found that 87% of Australian respondents believed that working people and non-working people should receive equal priority within the healthcare system. Similarly, only a minority accepted giving priority to people with special skills. Neuberger, Adams, MacMaster, Maidment, and Speed (1998) found that only 21% of the general public chose 'value to society' as a criterion which should be used in selecting patients for liver transplantation.

Thus, people are strongly opposed to government policies prioritizing more socially beneficial people, in spite of their potential societal benefits. But how would they themselves decide if faced with an emergency situation in which only one out of two people could be saved? Would their judgments about such one-off decisions in concrete cases be the same as their decisions about general rules or policies?

1.2. Moral judgments and levels of concreteness

In this paper we primarily focus on the issue of prioritizing more socially beneficial people in order to achieve greater indirect societal benefits (i.e., indirect benefit dilemmas). As far as we are aware, no previous research has explored such cases. While such cases are structurally different from sacrificial moral dilemmas (Greene, 2013), in which one person may be sacrificed to save many more (e.g., trolley problems), they may elicit similar psychological reactions. Just as people exhibit a deontological aversion to harming others for the sake of the greater good (violation of a non-harm principle), so they may exhibit a deontological aversion to prioritizing some people over others for the sake of the greater good (violation of an egalitarian principle). By a 'deontological aversion', we have in mind a form of moral aversion elicited by factors other than the anticipation of overall worse downstream effects, leading respondents to regard a certain act or rule as

objectionable in principle. As such, past research examining factors that increase or decrease utilitarian judgments in sacrificial dilemmas (Kawai, Kubo, & Kubo-Kawai, 2014; e.g., Petrinovich, O'Neill, & Jorgensen, 1993) may be informative for understanding people's judgments in indirect benefit cases.

In this section, we discuss some findings from previous research, primarily involving sacrificial dilemmas, which give us reason to expect that people's moral judgments about one-off cases and general rules may differ. While these studies are informative for our purposes, none have directly investigated moral judgments about rules and one-off cases.

Körner and Volk (2014) studied sacrificial moral dilemmas, varying the level of construal (Eyal & Liberman, 2012; Trope & Liberman, 2010). They found that when people have sufficient cognitive processing capacity, a more abstract level of construal led to more deontological decisions, whereas a concrete level of construal led to more utilitarian decisions (the pattern reversed when processing capacity was reduced). This is relevant for our purposes since general policies are arguably more abstractly construed than one-off decisions. Though general policies eventually lead to real-world outcomes, they only do so after some time, whereas one-off decisions immediately translate into outcomes. In addition, the causal path between a decision about a general rule and the relevant possible outcomes is more indirect than the causal path between a decision about a one-off case and the relevant possible outcomes associated with that case.

Another line of research demonstrates that decisions regarding hypothetical and real dilemmas come apart, and that people are more willing to break moral taboos when hypothetical dilemmas become reality. Bostyn, Sevenhant, and Roets (2018) found that participants were twice as likely to administer electric shocks to a mouse to prevent five other mice from being shocked when participants believed this decision would have real-world consequences than when the situation was merely hypothetical. Relatedly, Patil et al., (2014) found that people's behavior in virtual reality dilemmas was more utilitarian than their corresponding judgments about textual descriptions. Real dilemmas have a more concrete level of construal than hypothetical dilemmas (Eyal & Liberman, 2012; FeldmanHall et al., 2012; Trope & Liberman, 2010), and so these findings lend further credence to the hypothesis that people are more willing to break moral taboos under concrete levels of construal, such as in one-off decision situations.

Lastly, Kogut and Ritov (2015) studied discrepancies between moral decisions when considered from an abstract perspective as opposed to decisions made with respect to a specific individual. They found that people are more approving of 'bending the rules' when considering the situation from the perspective of a specific person. This may be because concrete cases tend to trigger more affective processes (Kogut, 2009).

1.3. The present research

The objectives of this project were to test the existence of a generality effect in moral judgments, probe a range of boundary conditions, and provide some insights into possible mechanisms. For example, would people be more supportive of proposals to prioritize more socially beneficial people over people whose indirect social contributions are smaller when such cases occur in one-off emergency situations compared to when they are codified as general rules? Based on the findings from previous research described above, we hypothesized that this was the case.

Our first seven studies establish the generality effect in indirect benefits dilemmas. We explore the robustness of this generality effect by controlling for various factors, such as the identity of the decision maker (Study 2), the type of people considered more socially beneficial (Study 3), the context (Study 3 and 6), the level of hypotheticality (Study 4), and the scope of the rule (Study 5 and 6). In Study 7, we explore the effect in a joint vs separate evaluation setup. In Study 8, we show that it also occurs in sacrificial moral dilemmas.

¹ The stimuli used in our studies are not confined to cases of medical decision-making, however. Some argue that medicine is (viewed as) a special kind of good, governed by unusually stringent egalitarian norms (Mogensen, 2020; Walzer, 1983). Because medicine may constitute a unique kind of good, it is valuable to consider people's resistance to prioritizing on the basis of indirect benefits beyond medical contexts. We also focus exclusively on indirect interpersonal benefits, as opposed to indirect intrapersonal benefits (like the effects of health-care provision on a patient's earnings).

1.3.1. Open science

Reports of all measures, manipulations, and exclusions, as well as all data, analysis code, and experimental materials are available for download at <https://osf.io/fz7q2/>

1.3.2. Ethics statement

For all studies, relevant ethical guidelines were followed and the research was approved through University of Oxford's Central University Research Ethics Committee, with the reference number MS-IDREC-R56657/RE002.

2. Study 1: personal one-off vs. government policy

In Study 1 we contrasted personal decisions involving indirect benefits with policy decisions. In both cases, the decision to be made concerned whether to prioritize people who may be thought to contribute more to society. More specifically, we compared people's willingness to prioritize such people when asked to imagine that they themselves must decide whom to save in a one-off scenario with people's willingness to endorse governmental policies that demand prioritizing socially beneficial people in scenarios of this kind. Based on previous research (e.g., McKie & Richardson, 2011), we hypothesized that people would be strongly opposed to a government policy that prioritizes saving those who contribute more to society. However, because this study used a context other than the public provision of healthcare, this hypothesis was not guaranteed to be confirmed. We hypothesized that people would be more willing to prioritize a more socially beneficial person over a less socially beneficial person in a one-off decision context in which they were asked to imagine themselves as forced to make the choice.

2.1. Methods

2.1.1. Participants

We recruited 162 British participants online via Prolific (111 female; $M_{age} = 34.57$, $SD_{age} = 11.54$). They received £0.26 in payment (in line with UK minimum wage) for their participation. Eight were excluded for failing at least one out of two simple attention checks. The attention check questions of all studies can be found in the online materials. As an example, one attention check question asked participants the following: "Please select the second option from the left to indicate that you're paying attention." We expected a medium effect size. Power analysis showed that 128 participants were required to detect an effect size of $d = 0.5$ with an alpha of 0.05, power of 0.80 and df of 1. We aimed to recruit 160 participants to account for any exclusions.

2.1.2. Materials and procedure

The study had two between-subjects conditions. In the personal condition, participants read a text that involved two people, Allen and Bob, who are trapped on an island where a volcano has just erupted. Participants were told that they could only save one of them, due to time constraints. They were also told that rescuing Allen would be more beneficial to other people because he is a doctor specializing in a form of life-saving heart surgery that very few people in the world can perform, whereas Bob is long-term unemployed and unlikely to find work for the foreseeable future. Participants then had to indicate on a 7-point Likert scale to what extent they would prioritize saving Allen ("I would prioritise saving Allen").

In the government condition, participants read a similar text which involved the government of a small island with an active volcano considering what to do in case everyone from the island would have to be quickly evacuated and not everyone could be saved. Participants were asked to consider a proposed policy according to which those who contribute more to society should be prioritized in such cases. As an example, it was explained that if a decision had to be made between saving a doctor specializing in a form of life-saving heart surgery that

very few people in the world can perform and a person who is otherwise similar in terms of age, health, family, and so on but long-term unemployed and unlikely to find work in the foreseeable future, then the policy suggests that the doctor should be saved, rather than the person who is unemployed. Participants had to indicate on a 7-point Likert scale to what extent they would endorse such a policy ("I would support the proposed policy").

Finally, participants responded to the Oxford Utilitarianism Scale (OUS; Kahane et al., 2017), the Empathic Concern Scale (Davis, 1980), and demographic questions.

2.2. Results

A *t*-test showed that participants were more willing to prioritize the more socially beneficial person in the personal condition ($M = 5.28$, $SD = 1.21$) than in the government condition ($M = 3.80$, $SD = 1.88$), $t(139) = 5.92$, $p < .001$, $d = 0.93$, 95% CI [0.66, 1.19] (Fig. 1).

Empathic concern correlated negatively with agreement in the government condition, $r = -0.31$, $p = .005$. However, empathic concern did not correlate with agreement in the personal condition, $r = 0.06$, $p = .62$. Endorsement of instrumental harm (OUS-IH) correlated with agreement in the government condition, $r = 0.49$, $p < .001$ as well as in the personal condition, $r = 0.23$, $p = .04$. Endorsement of impartial beneficence (OUS-IB) did not correlate with agreement in either case, $r = 0.05$, $p = .57$. No noteworthy correlations between responses and demographic variables were found.

2.3. Discussion

The results of our first study support the stated hypothesis. People are more willing to prioritize more socially beneficial people when asked to decide for themselves in a one-off decision context than when asked about their support for a government policy that demands that more socially beneficial people be prioritized in such situations.

Our correlational findings suggest that some people experience an empathetic aversion to the governmental general policy, but not to the decision in the one-off case. The fact that endorsement of instrumental harm correlated with agreement with the proposal suggests that moral judgments in sacrificial moral dilemmas may partly be driven by similar mechanisms to those that drive people's judgments in our indirect benefit dilemmas. In Study 8 we look more directly at sacrificial moral dilemmas.

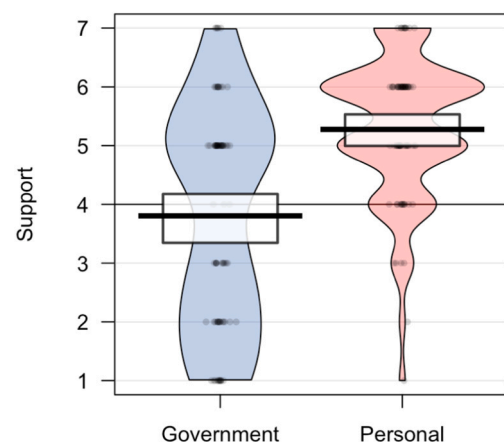


Fig. 1. Participants were more willing to prioritize a more socially beneficial person when they were asked to imagine that they themselves would make the decision in a one-off decision context than when they were asked about their support for a government policy that would prioritize more socially beneficial people in similar cases. 1 indicates 'Strongly disagree', 4 indicates 'Neither agree nor disagree', 7 indicates 'Strongly agree' (Study 1).

The two conditions in this study differ in multiple ways, which makes it difficult to draw a clear conclusion regarding the exact drivers of the effect. First, the personal condition involved an individual decision-maker, whereas the government condition involved group decision-making. Second, the individuals at stake were described in more concrete detail in the personal condition compared with the government condition (Kogut, 2009). Third, the proposed acts and/or policies may be felt to differ in the certainty they would attach to the death of the second person. The formulation of the question in the personal condition may have been felt to leave open the possibility of making some attempt to save both individuals (“I would prioritise saving Allen”), whereas this may not have been the case with respect to the content of the policy considered in the government condition (“the doctor should be saved, rather than the person who is unemployed”). In our subsequent studies, we control for these potential confounders.

3. Study 2: decision-maker

The previous study manipulated both the decision-maker (personal vs. government) as well as the level of generality (one-off vs. general rule). In Study 2, we attempted to disentangle these factors. Furthermore, we attempted to equalize another factor that differed across the vignettes used in Study 1: namely, that the personal condition focused on a single decision maker, whereas the government condition involved group decision making. Lastly, we used neutral labels, instead of names, for the people described in our vignettes, in order to reduce identifiability (Kogut, 2009).

Our hypothesis was that endorsement of prioritizing more socially beneficial people would be stronger in one-off situations compared to corresponding general rules. Furthermore, we hypothesized that there would be a main effect of the decision-maker factor. We expected that people would be more opposed to governmental decisions that prioritize more socially beneficial people compared with decisions of the same kind made by private individuals. We based this on the assumption that people would be more likely to view government decision-makers as constrained by egalitarian norms of fairness. We made no hypotheses about the sign and size of any interaction effects.

3.1. Methods

3.1.1. Participants

We recruited 250 British participants online via Prolific (164 female; $M_{age} = 35.12$, $SD_{age} = 12.65$). They received £0.18 in payment (in line with UK minimum wage) for their participation. None were excluded. We expected a small to medium effect size. Power analysis showed that 245 participants were required to detect an effect size of $\eta_p^2 = 0.03$ ($f = 0.18$) with an alpha of 0.05, power of 0.80, $df = 1$ and 4 groups. We aimed to recruit 250 participants to account for any exclusions.

3.1.2. Materials and procedure

The study had a 2 decision-maker (personal vs. government) \times 2 generality (one-off vs. rule) between-subjects design.

In the personal-one-off condition participants read a text very similar to the personal condition in Study 1. However, participants were asked to consider themselves as involved in a group decision. They were asked to imagine that they and some friends together owned a small boat and would have to decide together which of the two people should be rescued. They were asked to imagine that one member of the group of friends proposed rescuing Person A over Person B in light of the greater indirect utility of saving the former.

In the personal-rule condition, participants were asked to imagine that they and their friends, who together own a small boat, are considering what to do in case the volcano on the island erupts and there are several people whom they can choose to rescue, but not all can be rescued in time. One friend proposes that the group should adopt a rule that prioritizes saving more socially beneficial people in such cases.

The government conditions were analogous to the personal conditions. The government was imagined as either making a one-off decision about whom to rescue or as adopting a general rule that would enjoin rescuing the more socially beneficial person in cases like the one-off decision that participants were otherwise asked to consider.

In all conditions, participants indicated on a 7-point Likert scale to what extent they approved of the respective proposals they were asked to consider (“I would support the proposed policy [proposal to prioritize rescuing Person A]”). Finally, participants responded to demographic questions.

3.2. Results

Two-way ANOVA revealed a significant main effect of generality $F(1, 246) = 21.02$, $p < .001$, $\eta_p^2 = 0.08$, 95% CI [0.03, 0.15], but no significant main effect of decision-maker $F(1, 246) = 1.13$, $p = .29$, $\eta_p^2 = 0.002$, 95% CI [0.00, 0.03] (Fig. 2) Tukey’s HSD post-hoc analyses showed that ratings in the government-one-off condition ($M = 5.26$, $SD = 1.32$) were significantly higher than ratings in the government-rule condition ($M = 4.07$, $SD = 1.85$), $p < .001$, $d = 0.73$, 95% CI [0.36, 1.10]. Ratings between the personal-one-off ($M = 5.09$, $SD = 1.44$) and personal-rule ($M = 4.49$, $SD = 1.44$) conditions did not differ significantly, $p = .14$, $d = 0.41$, 95% CI [0.05, 0.79]. Note, however, that the interaction effect between generality and decision-maker was not significant, $F(1, 246) = 2.26$, $p = .13$, $\eta_p^2 = 0.009$, 95% CI [0.00, 0.05].

3.3. Discussion

These findings support one of the stated hypotheses, namely that the level of generality is an important driver of the effect we found in Study 1. People are more willing to endorse prioritizing more socially beneficial people in a one-off situation than when asked to consider a corresponding general rule. However, our other hypothesis was not confirmed. We found no significant main effect of decision-maker. We also did not observe any significant interaction effect between decision-maker and generality. We conclude that, of these factors, generality is likely the key driver of the effect. Given the absence of any effect related to decision-maker, our subsequent studies look only at cases that involve the government as an actor. This simplifies the construction and subsequent modification of new vignettes. Furthermore, it is more natural to imagine government actors as forced to consider general rules for life-or-death scenarios of the kind we consider.

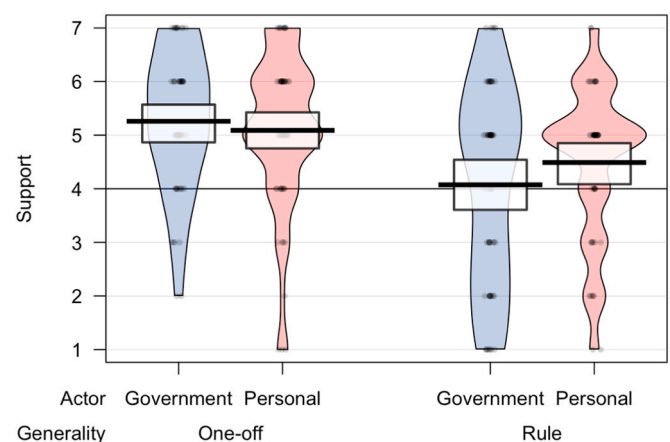


Fig. 2. Participants were more likely to agree to proposals that involve prioritizing more socially beneficial people in one-off cases as compared to adopting general rules enjoining prioritization decisions of this kind. There was no significant main effect for the decision-maker (personal vs. government). 1 indicates ‘Strongly disagree’, 4 indicates ‘Neither agree nor disagree’, 7 indicates ‘Strongly agree’ (Study 2).

One remaining limitation of this study was that the questions asked in the two conditions continued to differ in the third respect highlighted in our discussion of Study 1. Respondents in the one-off condition were asked about a proposal to “prioritize” rescuing one person rather than another, whereas the rule presented to respondents in the rule condition was said to suggest “that the doctor should be rescued, rather than the person who is unemployed.” As above, we may worry that the latter more clearly rules out any attempt to rescue the second person. In subsequent studies, we control for this potential confounder.

4. Study 3: types of useful people

In the previous two studies, we only considered a single, hypothetical context (rescue from an erupting volcano), one pair of potential victims differing in their indirect social utility (a doctor vs. an unemployed person), and one population (British citizens, sampled via Prolific). Before proceeding with further study of potential mechanisms, we sought to test the robustness of the generality effect by replicating it with 1) different pairs of individuals differing in their indirect social utility, 2) different contexts, and 3) a different population of respondents.

4.1. Methods

4.1.1. Pre-study

We first conducted a pre-study ($N = 48$) on MTurk, in which we tested eight different pairs of people whom we expected participants to perceive as differing in the extent of their social utility. More details about the materials and results are reported in the Supplementary Materials.

Participants were asked to indicate how much they believed the people whom they were asked to consider benefited others in society, using a scale ranging from 1 (*much less than average*), to 4 (*as much as the average person*), to 7 (*much more than the average person*). We found that in all eight pairs, the person we hypothesized would be considered to have greater social utility was viewed as such by participants. The pair of individuals used in the previous two studies (a doctor vs. a person who is long-term unemployed) resulted in the largest effect size. Based on the results, we chose as a second pair the following: an experienced firefighter vs. a telemarketer for a company selling vacuum cleaners. As a third pair, we created a new combination of people tested in the pre-study: a person who gives 25% of their income to charities that prevent the deaths of young children vs. a person who is homeless and must beg in order to support themselves. Participants’ verdicts regarding these three pairs resulted in large effect sizes of Cohen’s $d = 1.87$, 1.24 , and 1.40 , respectively.

4.1.2. Participants

We recruited 733 US participants online via MTurk. They received \$0.30 in payment (in line with US minimum wage) for their participation. Nine were excluded for failing at least one attention or comprehension check, leaving a final sample of 364 people (160 female; $M_{age} = 40.93$, $SD_{age} = 12.10$). We expected a small effect size. Power analysis showed that 351 participants were required to detect an effect size of $\eta_p^2 = 0.02$ ($f = 0.15$) with an alpha of 0.05, power of 0.80, $df = 1$ and 4 groups. We aimed to recruit 380 participants to account for any exclusions.

4.1.3. Materials and procedure

The study had a 2 generality (one-off vs. rule) \times 3 pairs of people (doctor/unemployed vs. firefighter/telemarketer vs. donor/homeless) \times 2 contexts (volcano vs. healthcare) design in which generality and contexts were manipulated between subjects and pairs of people within subjects.

The volcano vignette was almost identical to that used in Study 2. The healthcare vignette was analogous to the volcano vignette, but featured a situation in which the hospital board had to decide whether to

prioritize one person over the other in a one-off case (in the one-off condition) or whether to adopt a corresponding prioritization rule for such cases (in the rule condition).

The healthcare vignette read as follows in the rule condition: “Imagine that you are at a hospital run by the government. The hospital board has organs, such as kidneys, available, which they can use to save people’s lives. The hospital board is considering what to do if there are several people whose lives they can save through kidney transplants, but they do not have enough kidneys for everyone. One member of the hospital board proposes that people should be prioritized if giving them a kidney would be more beneficial to other people. Here is a typical example of how this policy would be used. Suppose that a choice will have to be made between giving a kidney to a doctor specializing in life-saving heart surgery that very few people in the world can perform, and a person who is otherwise similar in terms of age, health, family, and so on, but long-term unemployed and unlikely to find work in the foreseeable future. Then the policy suggests that the doctor should be given the kidney, rather than the person who is unemployed.”

The three pairs doctor/unemployed person vs. firefighter/telemarketer vs. donor/homeless person were presented in randomized order. This means each participant saw the same vignette three times with different pairs of people. As in the previous two studies, participants had to indicate on a 7-point Likert scale the extent to which they would support the proposal (“I would support the proposal”). Afterwards, participants responded to three follow-up questions, all with a 7-point Likert response scale. These questions were designed to guide further research on potential mechanisms. The questions concerned, firstly, whether the proposals could lead to negative consequences in the long-term, secondly, whether the proposals were unfair, and, thirdly, whether the proposals suggest that some people have higher moral worth than others. Finally, participants responded to demographic questions.

4.2. Results

The results showed that the generality effect replicated for all pairs in both contexts (Fig. 3). Three ANOVAs were conducted on the main dependent variable (Table 1). The analyses revealed that for all three pairs, participants were significantly more likely to agree with the proposal in the one-off condition than in the rule conditions (Table 2). For the doctor and donor pairs, there were no main effects of context. For the firefighter pair, participants considered the proposals less acceptable in the healthcare context than in the volcano context. For none of the three pairs were any significant interaction effects between

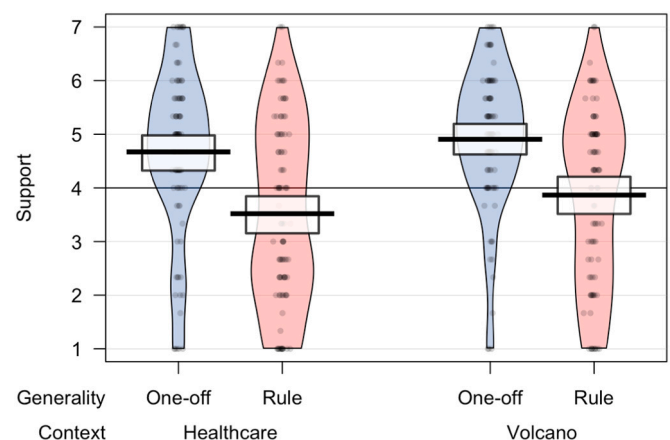


Fig. 3. Participants were more supportive of one-off proposals to prioritize socially more beneficial people than corresponding general rules. This was the case across different contexts and for different pairs of people. Figure shows aggregated results across the three types of pairs of people. 1 indicates ‘Strongly disagree’, 4 indicates ‘Neither agree nor disagree’, 7 indicates ‘Strongly agree’ (Study 3).

Table 1ANOVAs, $df = 360$. 95% CIs in brackets (Study 3).

	Generality			Context			Interaction		
	F	η_p^2	p	F	η_p^2	p	F	η_p^2	p
Doctor	43.58	0.11 [0.05, 0.17]	< 0.001	0.75	0.002 [0.00, 0.02]	0.39	0.06	< 0.001 [0.00, 0.01]	0.82
Firefighter	25.46	0.07 [0.02, 0.12]	< 0.001	8.65	0.02 [0.00, 0.06]	0.003	0.24	< 0.001 [0.00, 0.02]	0.62
Donor	43.84	0.11 [0.06, 0.17]	< 0.001	1.09	0.003 [0.00, 0.02]	0.30	0.46	0.001 [0.00, 0.02]	0.50

Table 2

Means and standard deviations. 1 indicates ‘Strongly disagree’, 4 indicates ‘Neither agree nor disagree’, 7 indicates ‘Strongly agree’ (Study 3).

	Volcano		Healthcare	
	One-off	Rule	One-off	Rule
Doctor	5.33 (1.44)	4.10 (1.84)	5.13 (1.60)	3.99 (1.90)
Firefighter	4.61 (1.53)	3.78 (1.85)	4.16 (1.68)	3.16 (1.85)
Donor	4.79 (1.51)	3.72 (1.74)	4.72 (1.75)	3.40 (1.88)

generality and context observed.

Participants considered the rule ($M = 5.23$, $SD = 1.60$) more unfair than the proposal in the one-off situation ($M = 4.67$, $SD = 1.66$), $t(362) = 3.25$, $p = .001$, $d = 0.34$, 95% CI [0.13, 0.55]. They considered it more likely that the adoption of the rule ($M = 5.36$, $SD = 1.42$) could lead to negative long-term consequences than the implementation of the proposal in the one-off situation ($M = 4.67$, $SD = 1.58$), $t(358) = 4.39$, $p < .001$, $d = 0.46$, 95% CI [0.25, 0.67]. For both perceptions of negative consequences and unfairness ratings there were no differences between contexts and also no interaction effects. Participants in all conditions believed that the proposal suggests that some people have higher moral worth than others. Here, there were no significant differences across any of the conditions. In other words, participants did not believe that the rule proposal ($M = 6.06$, $SD = 1.20$) suggests that the more socially beneficial people have a greater moral worth to a significantly greater extent than the one-off proposal does ($M = 5.87$, $SD = 1.19$), $t(362) = 1.49$, $p = .14$, $d = 0.16$, 95% CI [-0.05, 0.36].

Agreement rates for the proposals correlated negatively with perceptions of negative consequences ($r = -0.55$, $p < .001$), perception of unfairness ($r = -0.70$, $p < .001$), and perceptions of differential moral worth attribution ($r = -0.27$, $p < .001$). A linear regression analysis revealed that generality moderated the association between perceptions of negative consequences and agreement rates, i.e., interaction $b = -0.24$, $p = .01$. Thus, the negative association between agreement rates and perceptions of negative consequences was stronger in the rule condition ($r = -0.56$, $p < .001$) than in the one-off condition ($r = -0.47$, $p < .001$).

4.3. Discussion

This study demonstrates the robustness of the generality effect. It replicated with three different pairs in two different contexts and using a different population from that sampled in the previous studies.

This study also gives some insight into the potential mechanism of the effect. We found that people consider the rule more unfair and more likely to lead to negative consequences than the one-off decision. Perception of unfairness was strongly negatively correlated with agreement rates, more so than perceptions of negative consequences. This suggests that different perceptions of the fairness and long-term risks of the proposals could partly drive the generality effect. Perceptions of differential moral worth attribution also correlated negatively with agreement rates, but less strongly so.

Lastly, we note that in the healthcare vignettes used in this study, the rule condition did not ask respondents about their approval of a rule

characterized as entailing that “the doctor should be saved, rather than the person who is unemployed.” Instead, the rule is characterized as entailing that “the doctor should be given the kidney, rather than the person who is unemployed.” Earlier, we considered the worry that language stating that we should save A and not B, which characterized our rule conditions but not our one-off conditions, might be interpreted as ruling out that any additional effort would be made to save B. Since the one-off conditions used the weaker language of prioritization, this represented a potential confound. The rule used in the healthcare vignettes in this study is not characterized in terms of saving one person rather than another. For this reason, we do not think the wording used in the rule condition here can be suspected of implicating to a noticeably greater extent than the wording used in the one-off condition that we should ignore the possibility of making additional efforts to save the life of the second person if suitable means unexpectedly become available.

5. Study 4: precedent case

In the previous studies, the one-off cases always involved a concrete emergency. By contrast, the rule cases only mentioned a hypothetical catastrophe. In Study 4, we aimed to disentangle the effects of the rule vs. one-off framing and the hypothetical vs. concrete framing.

In this study, we examined whether there would be a difference in respondents’ levels of agreement with respect to two proposals, each involving a concrete emergency, in which an agent (the hospital board) either proposes a simple one-off decision (one-off) or adopts a general rule that applies to this specific scenario and other future cases of the same kind (precedent). The latter condition was, therefore, similar to a precedent case in which a general rule is established based on a concrete case. We also considered whether there would be any difference between the precedent case and a similar case involving the adoption of a rule in response to a merely hypothetical scenario (rule).

We hypothesized that people would be more willing to accept the proposal to prioritize more socially beneficial people if it applies just to this specific case (one-off) instead of representing a general rule (i.e., a precedent case), even if in both cases a concrete emergency is made salient. Further, we hypothesized that participants would be less willing to prioritize more socially beneficial people in the rule case than in the precedent case. Finally, we were interested in whether people would consider an agent that endorsed a rule to prioritize socially beneficial people as less trustworthy than an agent that decided to do so in a one-off case.

5.1. Methods

5.1.1. Participants

We recruited 233 US American participants online via MTurk. They received \$0.26 in payment (in line with US minimum wage) for their participation. Five were excluded for failing an attention check, leaving a final sample of 228 people (99 female; $M_{age} = 39.81$, $SD_{age} = 12.75$). We expected a small effect size. Power analysis showed that 204 participants were required to detect an effect size of $\eta_p^2 = 0.05$ ($f = 0.22$) with an alpha of 0.05, power of 0.80 and 3 groups. We aimed to recruit 230 participants to account for any exclusions.

5.1.2. Materials and procedure

The study had three between-subjects conditions: one-off, precedent, and rule. For all studies reported in this paper that include a precedent condition (Studies 4–6), the precedent condition always features a concrete case that is identical to the case described in the one-off condition, except that it is followed by the establishment of a general rule similar to that in the rule condition. By contrast, the rule condition does not feature a concrete case but introduces the general rule using hypothetical examples that could occur in the future.

Participants were presented with the healthcare vignettes of Study 3. The one-off and precedent conditions were almost identical to the vignette of the one-off healthcare condition in Study 3, with the only difference being the last sentence. In the one-off condition the last sentence was: “One member of the hospital board proposes that in **this particular emergency situation** the hospital should prioritize treating Person A over Person B, since helping Person A would be more beneficial to other people.” In the precedent condition the last sentence was: “One member of the hospital board proposes that **as a general policy**, the hospital should prioritize people who would be more beneficial to other people. In this particular emergency situation, the policy would suggest prioritizing treating Person A over Person B.” The rule condition was identical to the vignette of the rule condition in Study 3. In all conditions, participants indicated on a 7-point Likert scale to what extent they would support the proposal (“I would support the proposal”).

Next, participants were asked how trustworthy they would consider the hospital board to be in case it were to agree with the proposal and decided to adopt it, using a 7-point scale ranging from 1 (*extremely untrustworthy*) to 7 (*extremely trustworthy*). Afterwards, participants indicated on a 7-point Likert scale whether they believed that the proposal could lead to negative consequences in the long-term. Participants who believed this to be the case were asked to briefly explain in an open text field how the adoption of the proposal could lead to negative consequences. Finally, participants responded to demographic questions.

5.2. Results

A first one-way ANOVA revealed a significant main effect of condition on support for the proposal $F(2, 225) = 20.66, p < .001, \eta_p^2 = 0.16$, 95% CI [0.07, 0.24] (Fig. 4, Table 3). Tukey’s HSD post-hoc tests clarified that all conditions differed significantly from each other. Support in the one-off condition was higher than in the precedent condition ($p =$

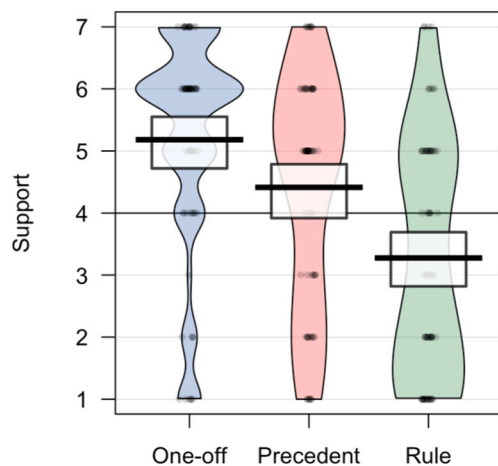


Fig. 4. Support for a precedent case in which a general rule to prioritize more socially beneficial people was set up as a result of a concrete one-off case was between support for the corresponding one-off proposal and support for a corresponding general rule that was set up without a concrete one-off case. 1 indicates ‘Strongly disagree’, 4 indicates ‘Neither agree nor disagree’, 7 indicates ‘Strongly agree’ (Study 4).

Table 3

Means and standard deviations. 1 indicates ‘Strongly disagree’, 4 indicates ‘Neither agree nor disagree’, 7 indicates ‘Strongly agree’ (Study 4).

	One-off	Precedent	Rule
Support	5.18 (1.75)	4.41 (1.87)	3.28 (1.91)
Trustworthiness	4.61 (1.71)	4.35 (1.62)	3.09 (1.58)
Negative consequences	4.64 (1.62)	5.07 (1.38)	5.47 (1.54)

.03, $d = 0.42$, 95% CI [0.10, 0.75]) and support in the precedent condition was higher than in the rule condition ($p < .001, d = 0.60$, 95% CI [0.27, 0.93]).

A second one-way ANOVA revealed a significant main effect of condition on perceived trustworthiness of the hospital board after having agreed to the proposal $F(2, 225) = 18.69, p < .001, \eta_p^2 = 0.14$, 95% CI [0.06, 0.22] (Table 4). Tukey’s HSD post-hoc tests clarified that the one-off and precedent conditions did not differ from each other ($p = .58, d = 0.16$, 95% CI [−0.16, 0.48]). However, trustworthiness ratings in the precedent (as well as the one-off) condition were significantly higher than in the rule condition ($p < .001, d = 0.78$, 95% CI [0.44, 1.12]).

A third one-way ANOVA revealed a significant main effect of condition on expected negative consequences of adopting the proposal $F(2, 225) = 5.83, p = .003, \eta_p^2 = 0.05$, 95% CI [0.006, 0.11] (Table 3). Tukey’s HSD post-hoc tests clarified that the one-off and precedent conditions did not differ from each other ($p = .19, d = 0.29$, 95% CI [−0.04, 0.61]). Similarly, there were no differences between the precedent and rule conditions ($p = .22, d = 0.28$, 95% CI [−0.04, 0.60]). However, expectations of negative consequences were significantly higher in the rule condition than the one-off condition ($p = .002, d = 0.53$, 95% CI [0.20, 0.86]).

There were no significant correlations with demographic variables.

5.3. Discussion

The results confirm our first hypothesis that people are more likely to agree to prioritizing more socially beneficial people in a one-off case as opposed to adopting a general rule for such cases, even when a concrete emergency is made salient. We also found that hypotheticality made a difference to people’s willingness to support a general rule. Participants were more likely to agree to a general rule saying that more socially beneficial people should be prioritized if this rule is suggested in the context of a concrete emergency.

One natural interpretation of these results is that there are at least two independent drivers of the observed generality effect. First, people have a preference for saving the more socially beneficial person in concrete one-off cases. Second, people have an aversion to general rules that prioritize more socially beneficial people. In the one-off case only the former is present, in the rule case only the latter. However, in the precedent case both drivers are present and act in opposing directions, meaning that people have conflicting intuitions.

6. Study 5: repealed rule

In Study 5, we tested whether people would still be opposed to a general rule prioritizing more socially beneficial people even if there was no reason to be concerned about possible knock-on effects resulting from corrupt use of the rule or other forms of misapplication that might occur in contexts outside the vignette. We tested this by

Table 4

Means and standard deviations. 1 indicates ‘Strongly disagree’, 4 indicates ‘Neither agree nor disagree’, 7 indicates ‘Strongly agree’ (Study 5).

	One-off	Precedent	Repeal
Support	4.97 (1.76)	4.22 (1.96)	4.09 (1.99)
Negative consequences	4.68 (1.68)	5.28 (1.62)	4.83 (1.76)

introducing a condition in which a general rule is proposed in response to a concrete emergency case (precedent case) but where this general rule will be repealed by outside forces after one day and therefore has practical application only to the case at hand.

Although we had found that people tended to cite the possibility of negative indirect consequences as justification for their opposition to a general principle saying that more socially beneficial people should be prioritized, we suspected that this might represent a form of *post hoc* rationalization. Our hypothesis was that people's greater opposition to a general rule prioritizing the more socially beneficial person was not driven by concerns about potential negative knock-on effects of this kind, but rather by a deontological aversion to principles of this kind. Thus, we hypothesized that people would be similarly dismissive of a precedent case resulting in establishment of a long-lasting general rule as of a precedent case that results in adoption of a general rule whose applicability is restricted to a single day.

6.1. Methods

6.1.1. Participants

We recruited 373 US American participants online via MTurk. They received \$0.2 in payment for their participation (in line with US minimum wage). Sixteen were excluded for failing an attention check, leaving a final sample of 357 people (169 female; $M_{age} = 39.72$, $SD_{age} = 11.39$). We expected a small effect size. Power analysis showed that 339 participants were required to detect an effect size of $\eta_p^2 = 0.03$ ($f = 0.17$) with an alpha of 0.05, power of 0.80, and 3 groups. We aimed to recruit 370 participants to account for any exclusions.

6.1.2. Materials and procedure

The study had three between-subjects conditions: one-off, precedent, repeal. The vignettes were based on the vignettes in Study 4. The one-off and precedent conditions were identical to the corresponding cases used in Study 4. The repeal condition was identical to the precedent condition, but with the following additional sentence at the end: "*Unbeknownst to the board, any general policies they adopt will not be applied beyond today, because of new changes to hospital administration that will be implemented by the government tomorrow.*"

In all conditions, participants indicated on a 7-point Likert scale to what extent they would vote for the proposal ("I would vote for this proposal [policy]"). Next, participants indicated on a 7-point Likert scale whether they believed that the proposal could lead to negative consequences in the long-term. Finally, participants responded to demographic questions.

6.2. Results

A one-way ANOVA revealed a significant main effect of condition on support for the proposal, $F(2, 354) = 7.22$, $p < .001$, $\eta_p^2 = 0.04$, 95% CI [0.007, 0.08] (Fig. 5, Table 4). Tukey's HSD post-hoc tests clarified that support in the one-off condition was significantly higher than in both the precedent condition ($p = .008$, $d = 0.40$, 95% CI [0.14, 0.66]) and the repeal condition ($p = .001$, $d = 0.47$, 95% CI [0.20, 0.73]). ByBy contrast, support in the precedent condition and repeal condition did not differ significantly from each other ($p = .85$, $d = 0.07$, 95% CI [-0.18, 0.32]).

A second one-way ANOVA revealed a significant main effect of condition on expected negative consequences of adopting the proposal $F(2, 354) = 4.07$, $p = .02$, $\eta_p^2 = 0.02$, 95% CI [0, 0.06] (Table 4). Tukey's HSD post-hoc tests clarified that expected negative consequences were perceived to be significantly higher in the precedent condition than in the one-off condition ($p = .02$, $d = 0.36$, 95% CI [0.10, 0.62]). Expected negative consequences in the repeal condition lay between those of the precedent condition ($p = .10$, $d = 0.27$, 95% CI [0.01, 0.52]) and those in the one-off condition ($p = .77$, $d = 0.08$, 95% CI [-0.17, 0.34]).

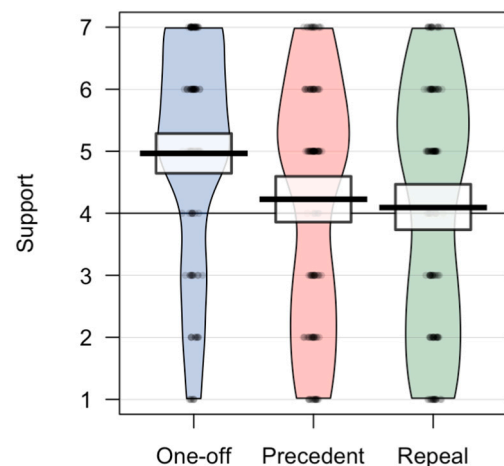


Fig. 5. Support for a general rule that was adopted via a precedent case was the same when the rule repealed after one day. This casts some doubt on the possibility that people's dismissal of the general rule is driven by consequentialist concerns about potential negative long-term consequences that the rule could have. 1 indicates 'Strongly disagree', 4 indicates 'Neither agree nor disagree', 7 indicates 'Strongly agree' 1 indicates 'Strongly disagree', 4 indicates 'Neither agree nor disagree', 7 indicates 'Strongly agree' (Study 5).

6.3. Discussion

The findings of this study support our hypothesis that people continue to be opposed to a general rule prioritizing more socially beneficial people even if the rule only applies to a single concrete emergency situation. This casts some doubt on the possibility that people's dismissal of the general rule is driven by consequentialist concerns about potential negative long-term consequences that the rule could have. A plausible conclusion to be drawn from this study is that people's dismissal of the general rule is instead driven by a deontological aversion to rules of this kind.

Even though many concerns about negative future consequences are ruled out in the repeal condition, there may still be some such concerns that persist. For example, people might object to the rule because they worry that it would send a signal to the new hospital administration, which in turn could create negative consequences. Or, people may worry that the rule could cause risk to patients in the short-term, even if it did not cause harm in the future. Therefore, while this study provides some support for the hypothesis that people's weaker support for rules compared to one-off cases is not driven by concerns about consequences, we also cannot fully rule it out.

7. Study 6: narrow rules

It is possible that people were more opposed to rules than one-off decisions in our previous studies because the rules had a wider scope than the one-off decision. Obviously, there is a sense in which any rule must of necessity have broader scope than a corresponding one-off decision. What we have in mind is the idea that the rules considered were relatively *open-ended*. Participants may have had concerns about how these rules would be applied beyond the specific example given. In particular, they may have worried that these rules could be used to support verdicts about other cases with which they would not agree. There are at least two dimensions of scope that we control for in this study.

First, there is the context to which the rule applies. Participants may interpret the rule to prioritize more beneficial people as a general principle of care that would apply to various contexts, including contexts that differ significantly from the example given. In this study, we control for context scope by explicitly specifying that the rule would

only apply to particular kinds of emergency cases.

Second, there are the people to whom the rule would apply. Participants may worry that the rule would be applied in cases where they disapprove of the classification of people as more and less beneficial. We control for people scope by specifying that the rule would only apply to cases where one patient is a doctor and the other patient is an unemployed person.

A further possibility is that participants were worried about how rigidly the rule would be applied, and whether it might also be applied in cases where there were strong reasons to prioritize helping a person of lower indirect social utility: for example, because their medical needs are more urgent. To control for rigidity, we specify that the rule would only apply as long as no considerations speak against it and as long as the patients have the same level of medical need and are otherwise similar in terms of age, health, family, and so on.

7.1. Methods

7.1.1. Participants

We recruited 360 US American participants online via MTurk. They received \$0.33 in payment (in line with US minimum wage) for their participation. 27 were excluded for failing an attention check, leaving a final sample of 333 people (128 female; $M_{age} = 38.37$, $SD_{age} = 11.74$). We expected a small effect size. Power analysis showed that 318 participants were required to detect an effect size of $\eta_p^2 = 0.03$ ($f = 0.175$) with an alpha of 0.05, power of 0.80 and 3 groups. We aimed to recruit 340 participants to account for any exclusions.

7.1.2. Materials and procedure

The study had the same design as Study 4. There were three between-subjects conditions: one-off, precedent, rule. Participants were presented with two scenarios in randomized order: an adjusted healthcare scenario (from Study 4) and a new firefighter scenario. The firefighter scenario described a case in which only one out of two people could be saved from a burning building. In all conditions the following background was given to make the proposals more plausible: *"The city suffers from a lack of doctors, since many doctors have left for other cities. At the same time, it has an unusually large number of unemployed people."*

In both scenarios, and in both the rule and precedent condition, the scope of the rule was defined in narrow terms, such that it was clear that it would only apply to a very narrow type of context and only to specific types of people: *"Therefore, a member of the hospital board proposes the following. From now on, in all emergency cases where both a doctor and an unemployed person urgently need a kidney to survive but only one kidney is available, the hospital should save the doctor. That is because saving the doctor would be more beneficial to other people."* To control for rigidity, participants were told the following: *"The member of the hospital board clarifies that this holds as long as no other considerations speak against it and as long as the patients have the same level of medical need and are otherwise similar in terms of age, health, family, and so on."*

Participants indicated on a 7-point Likert scale to what extent they would support the proposal ("I would support the proposal"). Participants also indicated on a 7-point scale to what extent they found the proposal morally wrong. This was included to test whether we would find the same effect with perceived wrongness as with support. Next, participants indicated on two 7-point Likert scales whether they believed that the proposals were unfair and whether they could lead to negative consequences in the long-term. Participants were then asked to note down their reasons for whether they believed the proposals to be unfair or capable of leading to negative consequences in a text field. Finally, participants responded to demographic questions.

7.2. Results

We conducted two separate one-way ANOVAs for the healthcare and firefighters scenarios. The patterns were the same in both scenarios.

There was a significant main effect of condition on support for the proposal in the healthcare scenario, $F(2, 330) = 13.20$, $p < .001$, $\eta_p^2 = 0.07$, 95% CI [0.03, 0.13] (Fig. 6, Table 5). Tukey's HSD post-hoc tests showed that support in the one-off condition was higher than in the precedent condition ($p = .001$, $d = 0.49$, 95% CI [0.22, 0.76]) and than in the rule condition ($p < .001$, $d = 0.68$, 95% CI [0.40, 0.96]). Support in the precedent condition did not differ significantly from support in the rule condition ($p = .38$, $d = 0.17$, 95% CI [-0.09, 0.43]). Similarly, there was a significant main effect of condition on support for the proposal in the firefighters scenario, $F(2, 330) = 19.12$, $p < .001$, $\eta_p^2 = 0.10$, 95% CI [0.05, 0.17]. Tukey's HSD post-hoc tests showed that support in the one-off condition was higher than in the precedent condition ($p < .001$, $d = 0.63$, 95% CI [0.35, 0.90]) and than in the rule condition ($p < .001$, $d = 0.81$, 95% CI [0.53, 1.10]). Support in the precedent condition did not differ significantly from support in the rule condition ($p = .37$, $d = 0.17$, 95% CI [-0.09, 0.44]).

Next, we looked at the perceived moral wrongness of the proposals. There was a significant main effect of condition on perceived moral wrongness of the proposal in the healthcare scenario, $F(2, 330) = 5.21$, $p = .006$, $\eta_p^2 = 0.03$, 95% CI [0.003, 0.07] (Table 5). Tukey's HSD post-hoc tests showed that perceived moral wrongness was lower in the one-off condition than in the precedent condition ($p = .02$, $d = 0.38$, 95% CI [0.11, 0.65]) and than in the rule condition ($p = .02$, $d = 0.37$, 95% CI [0.10, 0.63]). Perceived moral wrongness did not differ significantly between the precedent condition and the rule condition ($p = 1.0$, $d = 0.001$, 95% CI [-0.26, 0.27]). Perceived wrongness of the proposal were significantly above the midpoint in both the precedent ($t(108) = 2.68$, $p = .009$, $d = 0.26$) and rule conditions ($t(111) = 2.51$, $p = .01$, $d = 0.24$). Similarly, there was a significant main effect of condition on perceived moral wrongness of the proposal in the firefighters scenario, $F(2, 330) = 3.90$, $p = .02$, $\eta_p^2 = 0.02$, 95% CI [0.00, 0.06]. Tukey's HSD post-hoc tests showed that perceived moral wrongness was lower in the one-off condition than in the precedent condition ($p < .05$, $d = 0.33$, 95% CI [0.06, 0.60]) and than in the rule condition ($p = .04$, $d = 0.32$, 95% CI [0.05, 0.58]). Perceived moral wrongness did not differ significantly between the precedent condition and the rule condition ($p = 1.00$, $d = 0.006$, 95% CI [-0.26, 0.27]). Perceived wrongness of the proposal were significantly above the midpoint in both the precedent ($t(108) = 3.68$, $p < .001$, $d = 0.35$) and rule conditions ($t(111) = 3.37$, $p = .001$, $d = 0.32$).

There were significant differences in perceived unfairness across the three conditions, $F(2, 330) = 4.49$, $p = .01$, $\eta_p^2 = 0.03$, 95% CI [0.00, 0.07]. Tukey's HSD post-hoc tests showed that concerns about unfairness were higher in the rule condition compared to the one-off condition ($p = .01$, $d = 0.40$, 95% CI [0.13, 0.66]). But there were neither any significant differences in fairness ratings between the one-off condition and the precedent condition ($p = .12$, $d = 0.26$, 95% CI [-0.01, 0.52]) nor between the rule condition and the precedent condition ($p = .62$, $d = 0.13$, 95% CI [-0.14, 0.39]).

There were also significant differences in perceived negative consequences across the three conditions, $F(2, 330) = 4.21$, $p = .02$, $\eta_p^2 = 0.02$, 95% CI [0.00, 0.06]. Tukey's HSD post-hoc tests showed that participants were more concerned about negative consequences in the rule condition compared to the one-off condition ($p = .01$, $d = 0.39$, 95% CI [0.13, 0.66]). But there were neither any significant differences in perceived negative consequences between the one-off condition and the precedent condition ($p = .30$, $d = 0.19$, 95% CI [-0.07, 0.46]) nor between the rule condition and the precedent condition ($p = .34$, $d = 0.19$, 95% CI [-0.07, 0.46]).

A linear regression analysis revealed that both perceived unfairness ($b = -0.53$, $p < .001$) and perceived negative consequences ($b = -0.22$, $p < .001$) significantly predicted lower support of both the healthcare and firefighter proposals (aggregated scores across both scenarios), even when both factors were controlling for the other.

Support ratings and perceived moral wrongness correlated negatively in both the healthcare scenario, $r = -0.71$, $p < .001$, and the firefighters scenario, $r = -0.71$, $p < .001$. Support ratings correlated

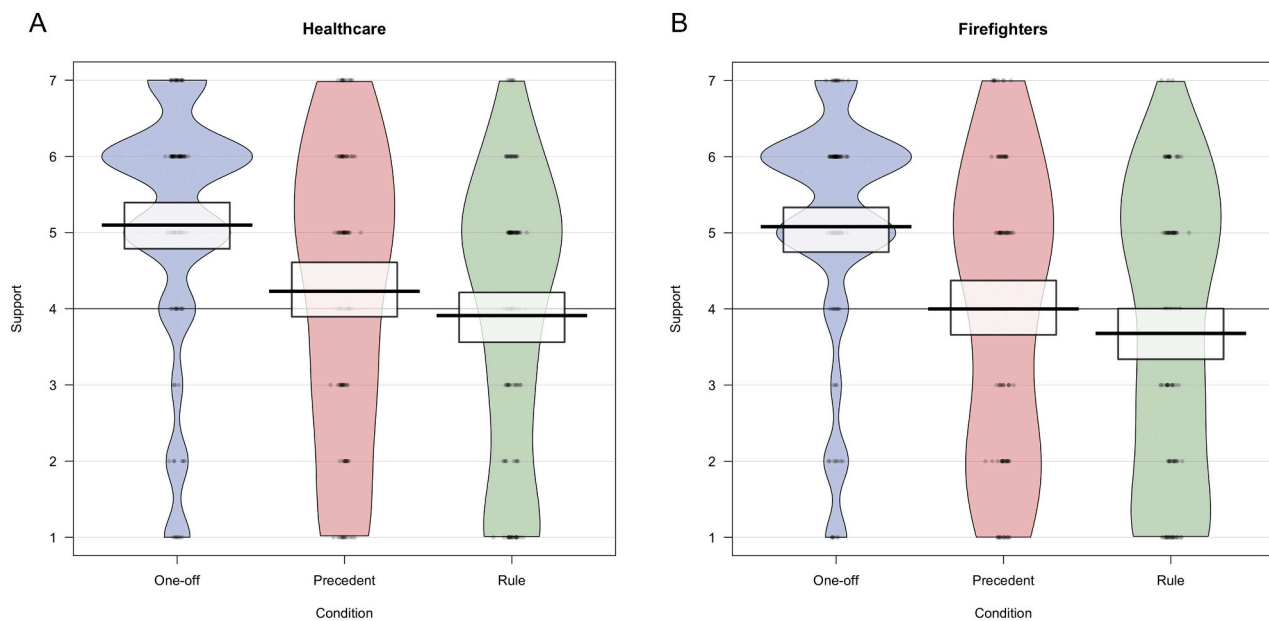


Fig. 6. Participants continued to be more supportive of proposals to prioritize more socially beneficial people in one-off cases than of corresponding general rules even when the rules were so narrow that they only covered cases that are exactly the same as the situation described in the one-off case. 1 indicates 'Strongly disagree', 4 indicates 'Neither agree nor disagree', 7 indicates 'Strongly agree' (Study 6).

Table 5

Means and standard deviations. 1 indicates 'Strongly disagree', 4 indicates 'Neither agree nor disagree', 7 indicates 'Strongly agree'. (Study 6).

	One-off	Precedent	Rule
Support healthcare	5.10 (1.62)	4.23 (1.89)	3.91 (1.85)
Support firefighters	5.08 (1.55)	4.00 (1.88)	3.68 (1.88)
Wrongness healthcare	3.78 (1.74)	4.42 (1.65)	4.42 (1.77)
Wrongness firefighters	4.02 (1.74)	4.58 (1.64)	4.59 (1.85)
Unfairness	4.38 (1.81)	4.84 (1.75)	5.06 (1.62)
Negative consequences	4.54 (1.79)	4.88 (1.79)	5.21 (1.61)

negatively with concerns about negative consequences in both the healthcare scenario, $r = -0.55$, $p < .001$, and the firefighters scenario, $r = -0.56$, $p < .001$. Similarly, support ratings correlated negatively with perceived fairness in both the healthcare scenario, $r = -0.63$, $p < .001$, and the firefighters scenario, $r = -0.66$, $p < .001$.

7.3. Discussion

The findings of this study confirm that the generality effect exists even when the scope of the rule is held narrow. In contrast to Study 5, we did not find that participants were more approving of the precedent case than the general rule. However, they were still less approving of the precedent case than the one-off decision. In all our other studies, we asked participants to indicate their level of support for the proposals but we did not directly ask about their perceived wrongness. In this study, we explicitly asked about the moral wrongness of the proposals. We found that participants were not just less supportive of the general rule than the one-off proposal, but also considered the general rule to be morally wrong to a greater extent than the one-off proposal. In line with previous studies, we found that participants were more concerned about negative consequences when asked to assess general rules and considered them more unfair than the one-off proposal.

Participants were asked to give free text responses explaining why they believed that the proposals could lead to negative consequences, or why they found them unfair. Typical responses included "it is simply unfair", "it puts the value of one human life over another", or "it can lead to a slippery slope". We did not find any types of reasons that

participants only gave in the general rule condition, and not in the one-off condition.

8. Study 7: joint vs separate

In Study 6 we found that people are less supportive of a general rule to prioritize more socially beneficial people than a corresponding one-off decision even if that rule has such a narrow scope that it only covers exactly those cases described in the one-off scenario. This suggests that people may be inconsistent. If they approve of every single case that the rule covers, why then would they be less approving of the general rule? In Study 7, we investigated whether people would continue to be more supportive of the one-off proposal than the general rule even when evaluating both scenarios jointly. Previous studies have shown that moral judgments in joint evaluation vs separate evaluation setups can differ (e.g., Barak-Corren, Tsay, Cushman, & Bazerman, 2018; Caviola, Faulmüller, Everett, Savulescu, & Kahane, 2014).

If participants' judgments remained the same irrespective of whether they were evaluating the two types of proposals separately or jointly, that would suggest that they believe themselves to have good reasons for viewing the one-off proposal as more appropriate than the general rule. For example, if the primary explanation for why people are less supportive of the general rule than the one-off proposal involves explicit consideration of reasons related to negative consequences or a principled distinction among different degrees of (un)fairness, we would expect judgments to remain mostly unaffected by a joint vs separate evaluation manipulation. If, however, people's judgments about the general rule and the one-off proposal are more similar under joint evaluation, this would suggest that they are not confident that they have adequate reasons for being more supportive of the one-off proposal than the general rule. Thus, if judgments are more similar under joint evaluation compared to under separate evaluation, that would provide evidence that the effect under separate evaluation is at least partly driven by factors that people are not able to explicitly justify when confronted with the apparent inconsistency.

8.1. Methods

8.1.1. Participants

We recruited 309 US American participants online via MTurk. They received \$0.30 in payment (in line with US minimum wage) for their participation. 11 were excluded for failing an attention check, leaving a final sample of 298 people (134 female; $M_{age} = 38.23$, $SD_{age} = 10.88$). We expected a small to medium effect size. Power analysis for the separate evaluation tests showed that 158 participants were required to detect an effect size of $d = 0.45$ with an alpha of 0.05, power of 0.80 and df of 1. We aimed to recruit 300 participants to account for the additional joint evaluation condition as well as for exclusions.

8.1.2. Materials and procedure

The study had a joint evaluation vs separate evaluation design. There were three between-subjects conditions: joint evaluation ($N = 92$), separate evaluation one-off ($N = 101$), and separate evaluation rule ($N = 105$). In the joint evaluation condition participants saw both firefighter scenarios on the same page (in randomized order). The scenarios were similar to the ones used in Study 5. One crucial difference was that the two cities were described as being two separate cities: city A and city B. At the end of the page, participants had to indicate on two separate 7-point Likert scales to what extent they would endorse the proposed general rule and the one-off proposal respectively ("I would support the proposal [proposed general policy] in city A [B]"). In the separate evaluation conditions, participants only saw one of the respective scenarios and had to indicate their endorsement of the proposal. In the joint evaluation condition, participants were asked on the next page to what extent they consider their two responses to be consistent, using a 7-point scale ranging from 1 (*completely inconsistent*) to 7 (*completely consistent*).

8.2. Results

An independent t -test revealed that in the separate evaluation conditions support for the one-off proposal ($M = 5.15$, $SD = 1.28$) was stronger than for the general rule ($M = 4.06$, $SD = 1.86$), $t(184) = 4.92$, $p < .001$, $d = 0.68$, 95% CI [0.39, 0.97] (see Fig. 7). A paired sample t -test revealed that in the joint evaluation condition support for the one-off proposal ($M = 4.57$, $SD = 1.88$) was also stronger than for the general rule ($M = 4.33$, $SD = 1.95$), $t(95) = 2.16$, $p = .03$, $d = 0.22$, 95% CI [0.02, 0.43]. The differences in support of the one-off proposal and the general policy was significantly stronger² in the separate evaluations conditions than in the joint evaluation condition, $t(305) = 3.04$, $p = .003$, $d = 0.36$, 95% CI [0.09, 0.58]. Support for the one-off proposal in the separate evaluation condition was stronger than support in the one-off condition in the joint evaluation condition, $t(158) = 2.50$, $p = .01$, $d = 0.37$, 95% CI [0.08, 0.65]. By contrast, support of the general rule did not differ across the two types of conditions, $t(189) = -0.99$, $p = .33$, $d = 0.14$, 95% CI [-0.14, 0.42].

Participants in the joint evaluation condition believed that their two responses were consistent ($M = 6.17$, $SD = 1.18$). 75% of the participants in the joint evaluation condition were equally supportive of the one-off proposal and the general rule, 17% were more supportive of the one-off proposal, and 8% were more supportive of the general rule.

8.3. Discussion

The findings of this study demonstrate that people's judgments about the one-off proposal and the general rule tend to converge when they evaluate these jointly. When directly comparing the one-off proposal and the general rule, 75% of participants were equally supportive of both. Our results show that when evaluating both proposals jointly,

participants in particular become less supportive of the one-off proposal, without significantly altering their views about the general rule. However, our results also show that even under joint evaluation 17% of participants are less supportive of the general rule than the one-off proposal. Thus, on average, support for the one-off proposal remains stronger than for the general rule.

One way to interpret these findings is that when evaluating the two proposals jointly, many people feel that it would be inconsistent of them to be more supportive of the proposal in the one-off emergency case than of the corresponding general rule, if that general rule prescribes behaving exactly as they would approve of behaving in the one-off case. As a consequence, they equalize their judgments. If they instead had access to explicit reasons for viewing the general rule as less acceptable, we would expect them to retain the same relative judgments about the one-off proposal and the general rule, irrespective of whether evaluating them jointly or separately. Indeed, that may be exactly why a small percentage of participants continue to be more supportive of the one-off proposal than the general rule when evaluating them jointly.

In sum, our findings show that the generality effect we have been exploring in this paper may partly be driven by factors that a majority of people are unable to explicitly justify when asked to directly compare a one-off proposal with a general rule.

9. Study 8: sacrificial dilemma

All our previous studies focused on indirect benefit dilemmas that involved prioritizing a socially more beneficial person over another person. In Study 8, we examined whether the generality effect also shows in other types of moral dilemmas. In particular, we looked at a sacrificial dilemma, a type of dilemma widely studied in moral psychology. The most well known sacrificial dilemma is the trolley problem (Foot, 1967; Thomson, 1985). In this study, we relied on a modified version of the rescue boat scenario (Greene et al., 2008), in which one person has to be thrown overboard in order to prevent the boat from sinking, which would cause all passengers to drown.

9.1. Methods

9.1.1. Participants

We recruited 202 US American participants online via MTurk. They received \$0.30 in payment (in line with US minimum wage) for their participation. 14 were excluded for failing an attention check, leaving a final sample of 188 people (70 female; $M_{age} = 36.39$, $SD_{age} = 10.52$). We expected a small to medium effect size. Power analysis showed that 158 participants were required to detect an effect size of $d = 0.45$ with an alpha of 0.05, power of 0.80 and df of 1. We aimed to recruit 200 participants to account for exclusions.

9.1.2. Materials and procedure

The study had two conditions. Participants in the one-off condition read the following: "(...) Due to an accident the cruise ship is on fire and has to be abandoned. Since there are not enough lifeboats available, there are presently too many people on one of the lifeboats, which causes the boat to sink and everyone on it to drown. The captain proposes to throw one person overboard because this would be the only way to save all remaining people." Participants in the rule condition read the following: "(...) The management of the cruise ship is considering what to do if, in case of an accident, the ship has to be abandoned but there are not enough lifeboats to save everyone. The captain proposes that as a general policy, from now on in all emergency situations where the cruise ship has to be abandoned and where there are too many people on a lifeboat causing it to sink and everyone to drown, one person should be thrown overboard if this is the only way to save all remaining people."

Participants indicated on a 7-point Likert scale to what extent they would support the proposal ("I would support the proposal [proposed general policy]"). Next, participants indicated on two 7-point Likert

² According to the formula to assess preference reversals by Hsee (1996): $t = (M_{JA} - M_{JB}) - (M_{SA} - M_{SB}) / (S_J^2/N_J + S_{SA}^2/N_{SA} + S_{SB}^2/N_{SB})^{1/2}$

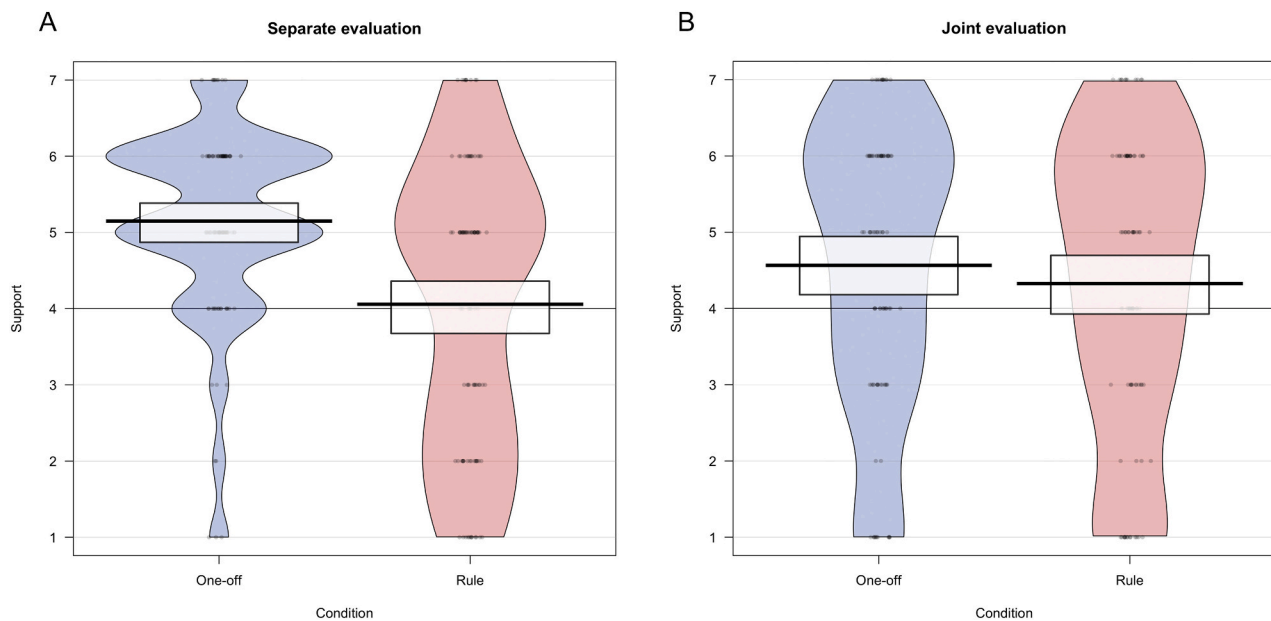


Fig. 7. When participants directly compared and evaluated the one-off proposal and the general rule side-by-side, the generality effect became weaker, suggesting that many participants could not access normative reasons for viewing the general rule as less acceptable than the one-off proposal. 1 indicates 'Strongly disagree', 4 indicates 'Neither agree nor disagree', 7 indicates 'Strongly agree' (Study 7).

scales whether they believed that the proposal was unfair and whether it could lead to negative consequences in the long-term. Finally, participants responded to demographic questions.

9.2. Results

A *t*-test showed that participants were more supportive of the proposal in the one-off condition ($M = 4.55$, $SD = 1.83$) than of the general rule ($M = 3.69$, $SD = 1.89$), $t(183) = 3.17$, $p = .002$, $d = 0.46$, 95% CI [0.17, 0.76] (Fig. 8). Ratings of unfairness in the one-off condition ($M = 4.95$, $SD = 1.68$) and the rule condition ($M = 5.27$, $SD = 1.52$) did not statistically differ from each other, $t(186) = -1.36$, $p = .17$, $d = 0.20$, 95% CI [-0.09, 0.49]. Participants were more worried about the long-term negative consequences of the general policy ($M = 5.60$, $SD = 1.29$) than the one-off decision ($M = 4.93$, $SD = 1.50$), $t(185) = -3.30$, $p = .001$, $d = 0.48$, 95% CI [0.18, 0.77].

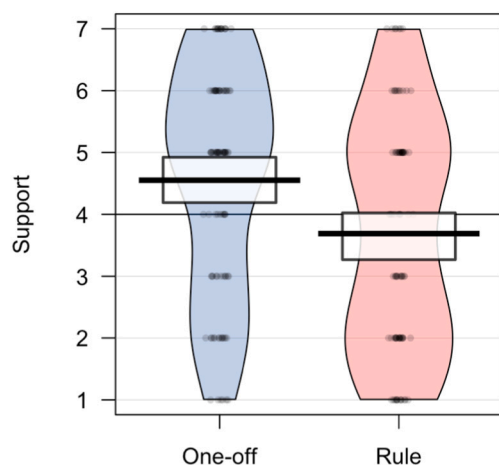


Fig. 8. Participants were more supportive of sacrificing one to save many in a one-off decision context than if it was stated as a general rule. 1 indicates 'Strongly disagree', 4 indicates 'Neither agree nor disagree', 7 indicates 'Strongly agree' (Study 8).

9.3. Discussion

The findings of this study suggest that the generality effect does not just apply to the indirect benefit cases we have studied so far in this paper. Instead, the effect also applies to sacrificial dilemma cases. This suggests that people in general are more approving of characteristically utilitarian responses to one-off decisions in moral dilemmas than of corresponding general rules.

These findings also rule out some hypotheses about the effect that are specific to indirect benefit cases. For example, they make it less likely that our effect was driven by people's concerns about the classification of people into beneficial and less beneficial people, since this aspect was not present in the sacrificial dilemma we used.

Previous research has found that moral judgments differ depending on the type of sacrificial dilemma. In this study, we have looked at a "personal" sacrificial moral dilemma that involves active personal harm. Future research could investigate if the generality effect also exists in other types of sacrificial moral dilemmas.

10. General discussion

Across our studies we found evidence for a generality effect: participants were more supportive of a proposal to prioritize people who are more beneficial to society than others if this applies to a concrete one-off situation than if it describes a general rule. The effect showed robustly even when controlling for various factors. It occurred even when the decision-maker was the same in both cases (Study 2), when the pairs of people differing in the extent of their indirect social utility was varied (Study 3), when the scenarios were varied (Study 3, Study 6), when the participant samples came from different countries (Study 3), and when the rule only entails cases that are exactly the same as the one-off case (Study 6). The effect also occurred when the general rule was introduced via a concrete precedent case (Study 4 and 6). The tendency to be more supportive of the one-off proposal than the general rule was significantly reduced when participants evaluated the two proposals jointly as opposed to separately (Study 7). Finally, we found that the effect also occurred in sacrificial moral dilemmas (Study 8), suggesting it is a general phenomenon occurring in multiple moral contexts.

Despite their structural differences, indirect benefit dilemmas and

sacrificial dilemmas likely trigger similar psychological reactions. In both types of dilemmas, deontological and utilitarian moral concerns can be invoked. People may be more opposed to general rules than corresponding one-off decisions either because they expect that such a rule *leads to worse outcomes* (the consequentialist interpretation) or because they think such a rule is *intrinsically wrong* (the deontological interpretation). We will now discuss reasons speaking for and against these two interpretations in turn.

10.1. Concerns about negative consequences

In Studies 3–6 and 8, participants considered it more likely that the adoption of the rule could lead to negative long-term consequences than the one-off decision. People may worry that in many cases it will be difficult to decide whether one person is more socially beneficial than another. They might fear that in such cases the scope for subjective judgment on behalf of decision-makers would lead to misuse of the rule, involving nepotism or other forms of corruption. It is also possible that any concerns people may already have about the one-off case could be heightened for a respective general rule, since the rule would lead to multiple decisions of the kind.

While we consider it plausible that concerns about negative consequences play a role in explaining the generality effect, there are reasons that speak against treating it as its only driver. First, concerns about corruption may seem apt in respect of government decisions, but arguably have little merit when applied to the decisions of private individuals. In Study 2, we found no main effect of decision-maker on people's judgments, nor any interaction effect involving decision-maker and generality. Second, concerns about corruption could not easily explain why the generality effect occurred in the sacrificial dilemma of Study 8 because that dilemma did not involve a distinction between socially more and less beneficial individuals.

Therefore, one possibility is that worries about adverse consequences represent *post hoc* rationalizations as opposed to true drivers of the effect. One reason to adopt this hypothesis is that in Study 5, we found that participants were more resistant to adopting a rule than endorsing a corresponding one-off verdict even in cases where the rule will not be used to guide behavior beyond one particular case. However, we also acknowledge the limitations of that study, which allow for alternative interpretations. Another reason to assume that worries about adverse consequences represent *post hoc* rationalizations is our finding that in joint evaluation, participants rated the two types of proposals more similarly than in separate evaluation (Study 7). The fact that people's level of support for the one-off decision differs across the joint evaluation and separate evaluation conditions suggests that they lack principled reasons (e.g., concerns about negative consequences) for being more supportive of saving the more socially beneficial individual in one-off cases. If in fact the differences observed in separate evaluation were explained by concerns about negative consequences, then we would expect participants to continue to be more supportive of the one-off proposals to the same extent in joint evaluation. Given that this is not the case, it is unlikely that such principled reasons are the full driver of the effect. But concerns about negative consequences might play a partial role in explaining the effect.

10.2. Deontological aversion

10.2.1. Unfairness

In Studies 3 and 6 we found that participants considered the rule more unfair than the one-off proposal. Thus, people may have a deontological aversion to general rules that they consider to be unfair or as violating an egalitarian principle. They might experience such an aversion to a lesser extent for analogous one-off decisions.

While we believe that concerns about unfairness likely play a role in explaining the generality effect, we also believe that the explanation is not especially informative. It may be thought to come close to a mere

restatement of the fact that they judged the rule as more objectionable. The appeal to perceived unfairness as a driver naturally raises the question of why participants viewed the rule as more unfair.

10.2.2. Trade-off aversion

One possibility is that the generality effect can be explained by trade-off aversion (Tetlock, Kristel, Elson, Green, & Lerner, 2000): the fact that people dislike making difficult trade-off decisions unless absolutely necessary. This is particularly the case for decisions that involve weighing different sacred values, such as human lives, against each other, i.e. *taboo trade-offs*. In a concrete one-off emergency situation people may realize that such a decision must be made and thus they choose the more pragmatic, utilitarian solution. In the general rule case, however, people might consider it unnecessary to make such a trade-off decision and thus prefer to avoid it.

Furthermore, people might consider it inappropriate to establish a general rule about difficult trade-off cases because they perceive preparing for such cases as expressing complacency in the face of potential emergencies. Instead of considering how to decide in such potential future emergency situations, people might consider it more appropriate to avoid these emergencies in the first place.

It is worth noting that the explanation highlighted here can arguably be given a consequentialist gloss and so may illustrate a way of rendering trade-off aversion consistent with a consequentialist approach. In other words, people might believe that the adoption of the rule, by expressing complacency in the face of potential emergencies, will in fact lead actors to be less concerned about preventing potential emergency cases from arising, leading to worse outcomes in aggregate.

10.2.3. Algorithm aversion

A related possibility is that people believe such decisions should be made on a case-by-case basis by the individuals responsible, rather than by a formal general rule. It has been demonstrated that people show *algorithm aversion*, e.g., they more quickly lose confidence in statistical algorithms as forecasters than in human forecasters (Dietvorst, Simmons, & Massey, 2015). It is possible that similar mechanisms that explain algorithm aversion explain people's greater aversion to general rules over one-off decisions that have to be made by humans on a case-by-case basis.

People may also oppose the general rules because they perceive them as disrespectful to the individuals who are being submitted to them. They may think that such general rules process all individuals at stake in a "mechanistic" way without appropriately taking into account everyone's special circumstances and interests. By contrast, people may perceive proposals in a one-off case as more appropriate because they think that in such cases everyone's special circumstances and interests could be taken into account.

10.3. Abstract vs. concrete levels of construal

Another theory that could be thought to explain why people are more opposed to a rule saying that we should prioritize more socially beneficial people than a corresponding one-off decision is construal level theory (CLT) (Eyal & Liberman, 2012; Trope & Liberman, 2010). CLT relies on the familiar notion that people mentally represent objects and events in different ways. In particular, a distinction is drawn between high-level construals that are abstract and decontextualized, and low-level construals that are concrete and context-specific. Psychological distance is found to facilitate evaluations based on high-level construals (Trope & Liberman, 2010). Eyal and Liberman (2012) further hypothesize that more central values are conceptualized as higher-level constructs, given that representing an object more abstractly involves retaining more central features while omitting secondary features. As a result, more central values should be expected to take priority when evaluating stimuli that are psychologically distant, with secondary values gaining in importance at greater psychological proximity.

In order to apply CLT to our findings, we would adopt the hypothesis that a deontological egalitarian principle mandating equal treatment irrespective of further indirect benefits represents a more central value for most participants than the utilitarian principle of maximizing aggregate welfare. As noted previously, although general policies eventually lead to real-world outcomes, they do so only after some time, and the causal path between a decision about a general rule and its application in any given instance is more indirect than the causal path between a decision about a one-off case and the relevant possible outcomes associated with that case. As a result, CLT would lead us to expect that subjects' evaluation of general rules are more likely to be based primarily on deontological egalitarian considerations, whereas their evaluations of descriptions of concrete scenarios are more likely to evoke utilitarian considerations.

Setting aside the need to adopt the stated hypothesis about value centrality, for which we know of no independent evidence, the explanatory power of CLT should be considered limited in light of our results in Studies 4 and 6. Here we find that people continue to be more opposed to the rule than the one-off decision even when the rule was defined by a concrete one-off situation, i.e., a precedent case. However, it may be thought possible that the establishment of the general rule at the end of the vignette triggered a more abstract way of thinking in participants.

10.4. Limitations and future research

Our studies demonstrate that there is a robust generality effect with regard to decisions involving prioritizing among people who may be thought to differ in their indirect social utility, as well as decisions involving sacrificing one to save many. We were also able to rule out several possible confounding factors. Although we have suggested potential explanations for the effect we observe, more research is needed to fully understand the psychological mechanisms driving the effect. It is likely that the effect is driven by multiple factors. Future research could therefore investigate more thoroughly whether concerns about negative consequences are causally driving the effect or whether they are rather *post hoc* justifications, as we have suggested above. More research is also needed to directly test the hypothesis that people have a deontological aversion, perhaps driven by trade-off aversion, against general moral rules of the kind studied here.

Another avenue for future research is to explore the scope of the generality effect. We have shown that the effect exists in indirect benefit dilemmas and in a sacrificial moral dilemma. Future research could explore whether the effect also shows up in other types of moral dilemmas and even certain non-moral cases. Research could test whether the effect only applies to cases that involve difficult trade-off decisions or whether it applies more generally to decision-making.

10.5. Rules and one-off decisions in philosophical utilitarianism

The findings of our studies are relevant for debates concerning the place of rules in utilitarianism. Many have argued that utilitarians should not attempt in every case to calculate the value of the consequences of the acts available to them, but should instead adopt simpler, concrete rules as guides to practical decision making (Frey, 2000; Richard M. Hare, 1981; Mill, 1861; Railton, 1984; Sidgwick, 1874). This saves time and helps to guard against our tendencies to bias calculations of aggregate utility to get the results we want in the heat of the moment. These concrete rules are often viewed as likely to be similar in character to deontological principles, facilitating convergence between consequentialism and deontology (Frey, 2000). For instance, Hare (1993), argued that medical professionals should adopt an absolute ban on participating in torture, while also believing firmly that there are possible cases in which it is for the best to use torture, arguing that things will go better overall if an absolute ban is adopted than if a more flexible rule is put in place. Insofar as our findings can be taken to suggest that people exhibit a deontological aversion to certain kinds of

general rules that they are disposed to rationalize by appeal to characteristically consequentialist considerations, a view like Hare's may well have been influenced by factors of this kind.

11. Implications

Our findings indicate that people's willingness to endorse one-off moral decisions and corresponding general moral rules come apart. Though they may endorse proposals to prioritize more over less socially beneficial people in one-off cases, they refrain from endorsing the corresponding general rule. This has important implications for moral psychology.

As discussed, Bostyn et al. (2018) have shown that people's behavior in hypothetical moral dilemmas does not predict their behavior in corresponding real dilemmas. (Plunkett and Greene (2019) have challenged that claim.) Based on this finding, they argue that psychological research about hypothetical moral dilemmas does not predict behavior in the real world. However, as discussed, decisions in hypothetical cases and decisions about general rules both typically have an abstract level of construal. And indeed, our findings suggest that people may be more deontological about general rules than about concrete one-off decisions, just like they are more deontological in hypothetical moral dilemmas than in concrete moral dilemmas. Therefore, it may be that studies of hypothetical decisions *do* predict real-world decisions after all, albeit decisions concerning general rules (rather than concrete one-off decisions).

The fact that people take a different view on deontological constraints with respect to general rules and one-off decisions is all the more important considering that much of our behavior is governed by rules in modern society. This is especially true of government staff: e.g., those employed in the public provision of health-care. They typically do not have discretion to act according to their own conscience, being instead required to follow existing laws and guidelines. Hence, the finding that people tend to be more deontological when deciding on general laws and guidelines points to the potential for an important mismatch between the moral judgments of those who set the rules and those who must act in accordance with them in concrete scenarios.

Another implication of our findings is that the framing of a general rule or policy can affect people's intuitions. Discussing precedents or other concrete real-world examples in the context of general rules could sway people in a consequentialist direction. Such precedents make people focus more on outcomes and less on deontological moral principles. In the same vein, our findings suggest that legal systems that rely to a greater extent on precedent and case law, as in the common law tradition characteristic of the English legal system, should be more consequentialist than legal systems that emphasize codified statutes, as in the civil law tradition common in continental Europe.

Our findings thus relate to an on-going discussion among moral philosophers concerning the place of rules in moral thinking and the methodological significance to be attached to verdicts about particular cases versus abstract principles. Philosophers in the utilitarian tradition have tended to place significant weight on the apparent self-evidence of abstract consequentialist principles, downplaying intuitive verdicts about particular cases as likely to be tainted by biases (Greene, 2008; Huemer, 2008; Sidgwick, 1874; Singer, 2005). Greene (2008) writes: "I have often observed that consequentialism strikes students as appealing, even as tautologically true, when presented in the abstract, but that its appeal is easily undermined by specific counterexamples." (65) Horne, Powell, and Hummel (2015) provide empirical evidence for this phenomenon. The results of our studies are striking in that they seem to illustrate the opposite pattern.

Author notes

Reports of all measures, manipulations, and exclusions, as well as all data, analysis code, and experimental materials are available for

download at https://osf.io/fz7q2/?view_only=a64f3a7c02424bb89cedbb0831f382f6.

Author credit statement

LC, SS, and AM planned the studies, interpreted the data, and wrote the paper. LC collected and analyzed the data and wrote the first draft of the paper.

Acknowledgements

Our research was funded by Jesus College at the University of Oxford and the Centre for Effective Altruism.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cognition.2020.104501>.

References

- Barak-Corren, N., Tsay, C.-J., Cushman, F., & Bazerman, M. H. (2018). If you're going to do wrong, at least do it right: Considering two moral dilemmas at the same time promotes moral consistency. *Management Science*, 64(4), 1528–1540.
- Bostyn, D. H., Sevenhant, S., & Roets, A. (2018). Of mice, men, and trolleys: Hypothetical judgment versus real-life behavior in trolley-style moral dilemmas. *Psychological Science*, 29(7), 1084–1093.
- Brock, D. W. (2003). Separate spheres and indirect benefits. *Cost Effectiveness and Resource Allocation*, 1.
- Caviola, L., Faulmüller, N., Everett, J. A. C., Savulescu, J., & Kahane, G. (2014). The evaluability bias in charitable giving: Saving administration costs or saving lives? *Judgment and Decision making*, 9(4), 303–316.
- Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. <https://www.uv.es/friasnav/Davis.1980.pdf>.
- Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology-General*, 144(1), 114–126.
- Du Toit, J., & Millum, J. (2016). Are indirect benefits relevant to health care allocation decisions? *The Journal of Medicine and Philosophy*, 41(5), 540–557.
- Eyal, T., & Liberman, N. (2012). Morality and psychological distance: A construal level theory perspective. In Mikulincer, M. & Shaver, P. R. (eds.), *Herzliya series on personality and social psychology. The social psychology of morality: Exploring the causes of good and evil* (p. 185–202). American Psychological Association.
- FeldmanHall, O., Mobbs, D., Evans, D., Hiscox, L., Navrady, L., & Dalgleish, T. (2012). What we say and what we do: The relationship between real and hypothetical moral choices. *Cognition*, 123(3), 434–441.
- Foot, P. (1967). The problem of abortion and the doctrine of the double effect. *The Oxford Review*, 5, 5–15.
- Frey, R. G. (2000). Act-utilitarianism in Blackwell guide to ethical theory. In H. Follette (Ed.), *Blackwell guide to ethical theory* (pp. 165–182). Oxford: Blackwell Publishers.
- Greene, J. D. (2008). The secret joke of Kant's soul. In W. Sinnott-Armstrong (ed.), *Moral psychology, volume 3: The neuroscience of morality: Emotion, brain disorders, and development* (pp. 35–80). Cambridge, MA: MIT Press.
- Greene, J. D. (2013). *Moral tribes: Emotion, reason, and the gap between us and them*. Penguin.
- Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, 107(3), 1144–1154. <https://doi.org/10.1016/B978-0-444-51115-7.00023-1>.
- Hare, R. M. (1981). The ethics of medical involvement in torture: commentary. *Journal of Medical Ethics*, 19(3), 138–141.
- Hare, R. M. (1993). Could Kant have been a utilitarian? *Utilitas*, 5(1), 1–16.
- Horne, Z., Powell, D., & Hummel, J. (2015). A single counterexample leads to moral belief revision. *Cognitive Science*, 39(8), 1950–1964.
- Hsee, C. K. (1996). The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational Behavior and Human Decision Processes*, 67(3).
- Huemer, M. (2008). Revisionary intuitionism. *Social Philosophy & Policy*, 25(1), 368–392.
- Kahane, G., Everett, J. A. C., Earp, B. D., Caviola, L., Faber, N., Crockett, M. J., & Savulescu, J. (2017). Beyond sacrificial harm: a two dimensional model of utilitarian decision-making. <https://ora.ox.ac.uk/objects/uuid:57a7b50e-bbb7-4ec6-b8dc-9be8e69fa0e2>.
- Kamm, F. M. (1993). *Morality, mortality, volume 1: Death and whom to save from it*. Oxford University Press.
- Kawai, N., Kubo, K., & Kubo-Kawai, N. (2014). “Granny dumping”: Acceptability of sacrificing the elderly in a simulated moral dilemma. *Japanese Psychological Research*.
- Kogut, T. (2009). Public decisions or private decisions? When the specific case guides public decisions. *Journal of Behavioral Decision Making*, 22(1), 91–100.
- Kogut, T., & Ritov, I. (2015). Target dependent ethics: Discrepancies between ethical decisions toward specific and general targets. *Current Opinion in Psychology*, 6, 145–149.
- Körner, A., & Volk, S. (2014). Concrete and abstract ways to deontology: Cognitive capacity moderates construal level effects on moral judgments. *Journal of Experimental Social Psychology*, 55, 139–145.
- Lippert-Rasmussen, K., & Lauridsen, S. (2010). Justice and the allocation of healthcare resources: Should indirect, non-health effects count? *Medicine, Health Care, and Philosophy*, 13(3), 237–246.
- McKie, J., & Richardson, J. (2011). Social preferences for the inclusion of indirect benefits in the evaluation of publicly funded health services: Results from an Australian survey. *Health Economics, Policy, and Law*, 6(4), 449–468.
- Mill, J. S. (1863). *Utilitarianism*. London: Parker, Son, and Bourn.
- Mogensen, A. L. (2020). Meaning, medicine, and merit. *Utilitas*, 32(1), 90–107.
- Neuberger, J., Adams, D., MacMaster, P., Maidment, A., & Speed, M. (1998). Assessing priorities for allocation of donor liver grafts: Survey of public and clinicians. *British Medical Journal*, 317(7152), 172–175.
- Nord, E., Richardson, J., Street, A., Kuhse, H., & Singer, P. (1995). Who cares about cost? Does economic analysis impose or reflect social values? *Health Policy*, 34(2), 79–94.
- Patil, I., Cogoni, C., Zangrando, N., Chittaro, L., & Silani, G. (2014). Affective basis of judgment-behavior discrepancy in virtual experiences of moral dilemmas. *Social neuroscience*, 9(1), 94–107.
- Petrinovich, L., O'Neill, P., & Jorgensen, M. (1993). An empirical study of moral intuitions: Toward an evolutionary ethics. *Journal of Personality and Social Psychology*, 64(3), 467.
- Plunkett, D., & Greene, J. D. (2019). Overlooked evidence and a misunderstanding of what trolley dilemmas do best: Commentary on Bostyn, Sevenhant, and Roets (2018). *Psychological Science*, 30(9), 1389–1391.
- Railton, P. (1984). Alienation, consequentialism, and the demands of morality. *Philosophy & Public Affairs*, 13(2), 134–171.
- Sidgwick, H. (1874). *The methods of ethics*. London: Macmillan.
- Singer, P. (2005). Ethics and intuitions. *The Journal of Ethics*, 9(3–4), 331–352.
- Skitka, L. J., & Tetlock, P. E. (1992). Allocating scarce resources: A contingency model of distributive justice. *Journal of Experimental Social Psychology*, 28(6), 491–522.
- Tetlock, P. E., Kristel, O. V., Elson, S. B., Green, M. C., & Lerner, J. S. (2000). The psychology of the unthinkable: Taboo trade-offs, forbidden base rates, and heretical counterfactuals. *Journal of Personality and Social Psychology*, 78(5), 853–870.
- Thomson, J. J. (1985). The trolley problem. *The Yale Law Journal*, 94(6), 1395–1415.
- Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, 117(2), 440–463.
- Walzer, M. (1983). *Spheres of justice. A defence of pluralism and equality*. Oxford: Robertson.
- Williams, B. (1973). *The Idea of Equality in his Problems of the Self: Philosophical Papers* (pp. 230–249). Cambridge: Cambridge University Press.