Judgment and Decision Making, Vol. 15, No. 2, March 2020, pp. 173–181

Fewer but poorer: Benevolent partiality in prosocial preferences

Gabriele Paolacci* Gizem Yalcin†

Abstract

A prosocial action typically provides a more sizable benefit when directed at those who have less as opposed to those who have more. However, not all prosocial acts have a direct bearing on socioeconomic disadvantage, nor does disadvantage necessarily imply a greater need for the prosocial outcome. Of interest here, welfare impact may depend on the number of beneficiaries but not on their socioeconomic status. Across four preregistered studies of life-saving decisions, we demonstrate that when allocating resources, many people are benevolently partial. That is, they choose to help the disadvantaged even when this transparently implies sacrificing lives. We suggest that people construct prosocial aid as an opportunity to correct morally aversive inequalities, thus making relatively more disadvantaged recipients a more justifiable target of help. Benevolent partiality is reduced when people reflect beforehand on what aspects they will prioritize in their donation decision.

Keywords: prosocial preferences, altruism, consequentialism, distributive justice, inequality

1 Introduction

How should people decide whom to help? Consequentialist traditions in philosophy suggest that priority should be given to the causes that provide the largest benefit to the largest number of people. Most vocally, the Effective Altruism movement urges altruists to use evidence and reason to evaluate alternative charitable projects, and, all else being equal, to donate to charities that benefit more people as opposed to fewer (e.g., MacAskill, 2015; Pummer & MacAskill, 2019; Singer, 2015). Building on these propositions, many charity-rating organizations (e.g., GiveWell) develop high-level metrics that are valid across different countries or charity-specific goals, often expressing the efficiency of donations in terms of number of lives saved (e.g., cost to avert the death of an under-5 individual).

Underlying these efforts is the fundamental view that everyone counts as one. Consequentialists, as well the foundational documents of intergovernmental organizations, typically conceive human lives as having equal value, irrespective of characteristics such as gender, race, or socioeconomic status (e.g., Bentham, 1789; Mill, 1861; Parfit, 1978; Pummer & MacAskill, 2019; Singer, 2015; UN General Assembly, 1948). This does not negate that recipients' characteristics can be relevant to evaluate the impact of aid. However, from this perspective, recipients' characteristics are relevant to the extent that they translate into heightened need for help, and not because they grant a special moral status to recipients. Of particular relevance to the present paper, most consequentialist traditions concur that aid is more beneficial when directed at reducing socioeconomic disadvantage. However, this is contingent on aid being directed at reducing disadvantage, or on disadvantage generally increasing the need for aid.

In this paper, we study how people's preferences for prosocial aid respond to the socioeconomic disadvantage of recipients when the consequential impact of aid does not depend on recipients' socioeconomic conditions. Across four preregistered studies, we demonstrate that a substantial number of people have preferences for aid that are characterized by *benevolent partiality*. That is, when choosing where to direct prosocial aid between recipients with different background socioeconomic conditions, people help the more disadvantaged recipients more than what the disadvantage implies from a consequentialist viewpoint. In our studies, participants express their preference for seeing a number of lives saved from a disadvantaged country or a larger number of lives saved from a slightly less disadvantaged country. Contrary to the impartiality principle (e.g., Kahane et al., 2017; Pummer & MacAskill, 2019; Singer, 2015), a plurality of people prefers the former option, i.e., saving fewer lives than they could. We suggest that benevolent partiality is driven by distributive justice concerns. In particular, people prefer to help disadvantaged others not only because such recipients need more help, but also because they construct this action as a way to compensate for previously existing inequalities.

In the remainder of the article, we discuss how the impact of prosocial giving on welfare can be understood using the lenses of consequentialist ethics, and why people's preferences for aid might not align with consequentialist prescriptions. We then present four studies that show how people

*Department of Marketing Management, Rotterdam School of Management, Erasmus University. Email: gpaolacci@rsm.nl. (http://orcid.org/0000-0002-9938-7940)

†Department of Marketing Management, Rotterdam School of Management, Erasmus University (orcid.org/0000-0001-7671-4892).

https://doi.org/10.1017/S1930297500007324 Published online by Cambridge University Press

resolve the trade-off between *efficiency* (i.e., benefiting more people, in line with consequentialist ethics) and *compensation* (i.e., benefiting those who have relatively less, in line with distributive concerns).

2 Socioeconomic Conditions and the Efficiency of Help

As many activists warn, eliminating poverty is a global priority. Despite persistent economic improvement in the poorest regions of the world, about 1/10 of the world population lives below the poverty line of $1.90 per day (World Bank, 2018). Significantly, in 2015 the United Nations placed "No Poverty" as number 1 among the Sustainable Development Goals to achieve by 2030. Consequentialists, and Effective Altruists in particular, are no exception in arguing that alleviating disadvantage is one of the most impactful investment of prosocial resources (e.g., Singer, 2015).

From a consequentialist point of view, the overall impact of a prosocial act can be understood as a function of the *size* of the benefit provided and the *number* of beneficiaries (Pellegrino, 2017). The size criterion to evaluate impact posits that one should invest towards causes that bring a larger benefit. For example, the same money could be raised to fulfill a child's wish to be Batkid for a day or to save the life of a child who may die from malaria. Because a child's life is a more sizable benefit than a child's wish coming true, donating towards malaria prevention is consequentially preferable (Singer, 2015). The number criterion posits that given a benefit of a certain size (e.g., saving a life), one should invest towards causes that benefit more people.

The socioeconomic condition of alternative recipients is often relevant to assess the consequential impact of aid. First and foremost, the size of the benefit that a monetary donation provides is often a direct function of the recipients' socioeconomic situation. In particular, donating to poorer recipients provides a larger benefit than donating to less poor recipients. In classic utilitarianism, this is true because the marginal utility of wealth is decreasing (e.g., Banerjee & Duflo, 2011; Diener & Biswas-Diener, 2002; Diener & Oishi, 2000). All else being equal, it is more beneficial to donate $100 to a poor family that would use it for survival than to a less poor family that would use it for recreational activities (MacAskill, 2015; Singer, 2015). Socioeconomic disadvantage is also important in prioritarianism, a variant of utilitarianism which ascribes a greater weight to increasing the welfare of those at lower welfare levels, independent of the marginal utility of wealth (e.g., Parfit, 2012).

Second, even if prosocial aid does not aim at alleviating disadvantage, socioeconomic conditions may be relevant to the extent that they correlate with access to the specific form of aid which is provided. Charities that distribute toys for children, for instance, do not aim at improving socioeconomic conditions; however, it is still sensible to donate toys to children of poor families, because poor families likely have less ability to acquire toys on their own. In sum, consequentialism provides a compelling argument for directing aid towards alleviating disadvantage, and generally emphasizes that the same outcome might provide a larger benefit to disadvantaged recipients.

However, not all prosocial acts have direct bearing on disadvantage, nor does disadvantage necessarily imply a greater need for the specific outcome that aid is trying to achieve. For example, several charitable projects (e.g., in malaria prevention) are explicitly trying to save lives. Number of lives saved is also a metric used by GiveWell in their efforts to quantify and equate the benefits of donating towards alternative heterogeneous causes. The rationale for relying on lives saved as an outcome, other than its obvious importance, is the presumably defendable principle of "impartiality": people's lives, or at least quality-adjusted life years (e.g., Singer et al., 1995), should be valued equally, irrespective of people's characteristics (e.g., gender, race, socioeconomic status), and in accordance with the tenet "everybody to count for one, nobody for more than one" (Bentham, 1789; Kahane et al., 2017; Singer, 1993).[1] That is, consequentialism suggests that one should not attribute increased priority to saving the life of any specific individual *per se*. A prosocial action, such as a donation towards aid, is assumed to be more beneficial when it allows saving more lives, irrespective of the socioeconomic situation of the people who are saved.

3 Benevolent Partiality

Several research streams suggest that people's preferences about which lives to save might not be always impartial. First, people do not always choose whom to help based on consequentialist considerations. Berman and colleagues (2018) demonstrated that people's donation decisions are less sensitive than investing decisions to cost-effectiveness arguments. When evaluating donation opportunities, people are often influenced by normatively irrelevant factors, such as whether victims are beautiful (Cryder, Botti & Simonyan, 2017) or identifiable (Small, Loewenstein & Slovic, 2007). People are also overly sensitive to statistics such as the number of dead people in disasters (Evangelidis & Van den Bergh, 2013) or overhead expenses (Gneezy, Keenan & Gneezy, 2014). Whereas previous research documented instances in which prosocial preferences respond to consequentially irrelevant factors, it does not directly show violations of impartiality. In fact, as explained by Berman and colleagues (2018), most documented failures to maximize

[1] In an interview with the authors, a research analyst at GiveWell confirmed our account.

https://doi.org/10.1017/S1930297500007324 Published online by Cambridge University Press

welfare tend to become weaker when people can directly compare alternative donation opportunities.

We believe, however, that socioeconomic differences between alternative recipients may exert a particularly strong influence on people's preferences for aid. A large amount of research shows that most people are averse to socioeconomic inequality. In the United States, for instance, many people appear to desire more redistribution, more egalitarian policies, and a generally more equal society (e.g., Engelmann & Strobel, 2004; Kiatpongsan & Norton, 2014; Norton & Ariely, 2011). This tendency is deeply rooted in people's moral intuitions about fairness. Psychologically, people often prefer equitable resource distributions, i.e., with people rewarded ("output") based on their contributions ("input"; e.g., Adams, 1965; Cook & Hegtvedt, 1993; Skitka & Tetlock, 1992). When inputs are perceived as equal, or there is little indication that inputs may differ, a distribution is deemed as just to the extent that outputs are equal (e.g., Mitchell et al., 1993). Importantly, people's aversion to inequitable allocations of resources can lead to sacrifice efficiency (i.e., total output) in order to benefit different parties equally (see Gordon-Hecker et al., 2017 for a review). For instance, beliefs that differences in income and wealth depend on external circumstances more than on individual effort translate into distress with existing inequalities, and preferences for redistributive policies (e.g., Alesina & Angeletos, 2005; Benabou & Tirole, 2006; Fong, 2001; Frank, Wertenbroch & Maddux, 2015).

In this work, we test how distributive justice concerns influence decisions about prosocial aid that do not affect, and are not consequentially affected, by recipients' background socioeconomic conditions. To clarify, we do not test whether people prefer resource allocations that are more equitable although less efficient: we test whether aid decisions prioritize those in worse background conditions (i.e., compensation) despite the inefficiency of the resulting allocation and the irrelevance of such conditions for the aid they provide. In other words, people choose between (1) *efficient* outcomes (i.e., more lives saved) and (2) *compensatory* outcomes that are less efficient (i.e., fewer lives saved) and not more equitable (because background conditions are unaffected).

Note that this trade-off may often apply to choices about aid, because of the geographical distance between those who provide help (frequently residing in Western societies) and those who need help. Given the type of help that is provided, people who are in worse socioeconomic conditions (e.g., in the developing world) tend to be those who are costlier to help (e.g., due to shipping costs; Li, Colby & Fernbach, 2018). It is thus important to understand how people reason about helping alternative recipients. In this paper, we are interested in the extent to which people's preference for giving to those who have the least are compensatory rather than consequentialist, and label this preference *benevolent partiality*: In our studies, participants who donate to those who are more disadvantaged behave as if lives had different value (a seeming violation of the impartiality principle); they do so in a benevolent attempt to compensate for distributions that they deem unjust.

4 Overview of Studies

We conducted four preregistered studies (plus three studies reported in the Supplementary Material) to investigate people's preferences for saving the lives of people in low but varying socioeconomic conditions.

In Study 1, participants chose which of two real charities they wanted the experimenters to donate $100 to, expressing the impact of such charities in terms of number of lives saved. We find that emphasizing socioeconomic status increases people's preferences for saving the lives of statistically more disadvantaged recipients, even if this implies saving fewer lives.

In the following studies, we investigated the presence benevolent partiality using scenarios that rule out consequentialist reasons to give to the more disadvantaged (e.g., socioeconomic status covarying with access to alternative forms of help), and confirm that people construct this situation as a compensation-efficiency trade-off. In Study 2A, we demonstrate that choices of aid allocation are more difficult to take when saving more lives implies saving those who are slightly more socioeconomically disadvantaged. In Study 2B, we find that saving the lives of the more disadvantaged is constructed as an opportunity to correct inequalities that are deem unjust.

Finally, in Study 3, we tested whether the presence of benevolent partiality depends on whether people reflected on the relative importance of different choice criteria prior to choosing between alternative charities. We find that this additional step, which is recommended in normative models of decision-making (e.g., Bazerman & Moore, 2008), results in choices that are more in line with consequentialist prescriptions.

All the four studies were preregistered. We report all the experimental conditions and the measures that we collected. Studies were conducted with participants recruited online, following best practices in online data collection (Goodman & Paolacci, 2017). We compensated participants with the equivalent of about $9/hour. All stimuli are presented in the Supplementary Material. Preregistrations, data, and analysis codes are available at: https://osf.io/js2fq.

5 Study 1: Benevolent Partiality

Study 1 provides an initial demonstration that people are more likely to prefer causes that benefit recipients that are in worse socioeconomic conditions, even if doing so reduces

https://doi.org/10.1017/S1930297500007324 Published online by Cambridge University Press

the consequential impact of their donation (i.e., the number of lives saved). This study was preregistered (https://aspredicted.org/j8tb4.pdf).

5.1 Method

5.1.1 Participants and Procedure

Participants were 304 US residents recruited from Amazon Mechanical Turk (MTurk) with 95%+ approval rate (45.7% female, mean age = 39.4). The study used a two-condition between-participants design.

At the beginning of the study, participants were told that the experimenters would donate $100 to a charity at the end of data collection. We explained to participants that they would indicate their preferred charity between two, the Schistosomiasis Control Initiative (SCI) and The END Fund (END), both delivering treatments for neglected tropical diseases. We clarified that the experimenters would donate the $100 to the charity receiving more preferences across all participants. To ensure participants knew that they were making a real decision, we provided them with a link to the webpage where we later published the receipt of our $100 donation.

Similar to what donors do while considering alternative projects on charity rating websites or Cause Marketing platforms (e.g., AmazonSmile), participants read some information about the two charities. Critically, we included cost-effectiveness information retrieved from http://www.givewell.org, i.e., the amount of money needed from the charity to avert the death of an individual under 5 years old. Given that this amount is lower for SCI ($1,100) than for END ($2,500), donating to SCI would allow for a larger impact in terms of lives saved than donating to END.

In a between-participants design, we manipulated the perceived socioeconomic differences between potential recipients of the donation. In particular, we mentioned that END collaborated with Ethiopia and that SCI collaborated with Nigeria, and varied how such countries were described. In the control condition, we described each country as "one of the countries with the lowest Human Development Index". In the disadvantage condition, we provided information about the annual GDP per capita and the literacy rate of both countries. Because these indicators were better for Nigeria ($1,990, 60%) than for Ethiopia ($870, 49%), we reasoned that participants in the disadvantage condition would perceive Ethiopia to be more disadvantaged (than Nigeria) compared to participants in the control condition.

After reading the information, participants selected the charity they wanted experimenters to donate to ("Which charity do you want us to donate the $100 to?"). As a manipulation check, participants indicated which country between Ethiopia and Nigeria was more disadvantaged on a 7-point scale (1= Ethiopia, 7 = Nigeria). Additionally, in a comprehension check, participants indicated whether SCI or END was more effective ("Based on the information you were given, which charity gets the bigger bang out of the buck? That is, which charity is more cost-effective?").

5.2 Results

The disadvantage manipulation was successful: Relative to Nigeria, Ethiopia was perceived as more disadvantaged in the disadvantage condition (M = 2.86, SD = 1.89) than in the control condition (M = 3.82, SD = 1.69; t(302) = 4.64, $p <$ .001, d = .54).

Emphasizing relative disadvantage made participants more likely to choose the charity that was less cost-effective but tied to more disadvantaged participants. That is, participants were more likely to choose END over SCI in the disadvantage condition (43.04%) compared to the control condition (25.49%; $\chi^2(1) = 9.64$, $p = .002$, $\varphi = .18$). Restricting the analysis to participants who correctly indicated that SCI was more effective than END (N = 243, 80% of the sample) yielded the same result (31.15% chose END in the disadvantage condition vs. 11.57% in the control condition; $\chi^2(1) = 12.70$, $p = .001$, $\varphi = .23$).

5.3 Discussion

Study 1 showed that tying a charity to a more disadvantaged country increased its attractiveness as a prosocial target, even if this implied forgoing impact in terms of lives saved. This study has the merit of considering a real choice of which charity to donate to, between one that was more cost-effective and one that was framed as focusing on disadvantaged recipients. However, giving to the latter charity was not a behavior necessarily inconsistent with consequentialist prescriptions. For example, participants may have reasoned that disadvantaged countries have scarcer access to alternative treatments for tropical disease. Donations towards such countries, then, could reflect a consequentialist assessment of benefit size, rather than partiality towards the disadvantaged. The following studies examine benevolent partiality using life-saving scenarios that facilitated controlling for consequentialist explanations and achieving higher internal validity.

6 Study 2A: Compensation-Efficiency Trade-off

Study 2A conceptually replicates Study 1 using a life-saving scenario which facilitates ruling out consequentialist explanations of benevolent partiality. That is, we took precautions to make observed partiality incompatible with a consequentialist approach to the donation problem. Most importantly, we emphasized that participants' donation decisions did not affect socioeconomic conditions and that socioeconomic

https://doi.org/10.1017/S1930297500007324 Published online by Cambridge University Press

conditions had no bearing on how critical donations were. Additionally, Study 2A investigates whether, consistent with our conceptualization, differences in recipients' socioeconomic status make the choice of how to allocate aid more difficult to make, even if such differences are, from a consequentialist point of view, irrelevant to the outcome that is produced. This study was preregistered (https://aspredicted.org/k6zi5.pdf).

6.1 Method

6.1.1 Participants and Procedure

Participants were 200 adults recruited from Prolific (62% female, mean age = 35.64, English as first language). The study used a two-condition between-participants design.

In the study scenario, a sudden outbreak of a new type of virus threatened the lives of thousands of people in two fictional third-world countries, Sangala and Naruba. A charity was collecting money to urgently ship antidotes unavailable in the developing world, and participants were faced with a choice of whether to ship antidotes to Sangala or to Naruba. Participants received information about the two donation projects that included their cost-effectiveness. Similar to Li et al. (2018), we explained that because of different shipping costs, the cost-effectiveness of donating to the two projects was different, and donating $10 towards the Naruba project would allow saving more people (i.e., 4–6) than donating towards the Sangala project (i.e., 2–4).

As in Study 1, we manipulated perceived differences in recipients' socioeconomic status by varying the description of the countries that the two projects were targeting. In a between-participants design we either described Sangala and Naruba as having a very low Human Development Index (control condition) or provided socioeconomic information such that Sangala was more disadvantaged than Naruba (disadvantage condition). Critically, donations would save lives but had no impact on socioeconomic conditions, i.e., the size of the benefit is not directly affected by disadvantage.

The scenario included several precautions to rule out consequentialist explanations for donations towards the disadvantaged, i.e., ensuring that the size of the benefit was also uncorrelated with disadvantage. We repeatedly emphasized that both countries would receive help in the future, but currently had no access to alternative forms of help (e.g., "no citizen has immediate access to antidotes independent of their geographic location or financial situation"). We also addressed a "strategic" consequentialist reason for why people might appear to show benevolent partiality, i.e., that people focus on lives saved and additionally consider how other donors might choose. If participants believed that the least cost-effective project is less likely to be supported, they may choose to donate to such project in an attempt to help where most help is needed. To rule this out, the scenario also reported that "so far, the two projects received about the same amount of money". The consequentialist prescription, therefore, is that participants donate to the project that allows saving more lives, i.e., Naruba.

After reading the scenario, participants were asked to choose the project that they wanted to donate to, further emphasizing that the money would not be used to alleviate disadvantage ("Which campaign do you donate your $10 to? That is, where do you want to ship antidotes to?"). To test whether people feel less certain about their choice when socioeconomic differences between beneficiaries is emphasized, we also measured participants' perceived conflict. Participants answered three statements on a 9-point scale (1 = not at all difficult/very certain/not at all conflicted to 9 = very difficult/not at all certain/very conflicted): "How difficult was it for you to decide between the two projects?", "How certain were you about which project to choose?" (reverse-coded), "How conflicted did you feel while choosing between the two projects?". Scores on these three items were averaged (α = .86).

6.2 Results

We found that emphasizing the relative disadvantage of Sangala made participants more likely to choose Sangala over Naruba, i.e., the project which would save fewer lives (42.42% in the disadvantage condition vs. 15.84% in the control condition; $\chi^2(1) = 15.89$, $p < .001$, $\varphi = .28$). Furthermore, people found it more difficult to choose between Sangala and Naruba when the relative socioeconomic disadvantage of the recipients was emphasized ($M_{control} = 4.70$, $SD_{control} = 2.00$, $M_{disadvantage} = 5.58$, $SD_{disadvantage} = 2.17$, $t(198) = -2.98$, $p = .003$, $d = .42$).

6.3 Discussion

The results of Study 2A replicate Study 1 under conditions that make observed partiality incompatible with consequentialist assessments. We conducted two conceptual replications of this study in which the prosocial endowment need not be donated to either project in its entirety. We found evidence of benevolent partiality (i.e., donations towards the disadvantaged-focused, less cost-effective option) both when participants could allocate $10 across projects (rather than choosing one project) and when participants were given an option not to donate. We report these studies in the Supplementary Material (Study S1 and Study S2).

Study 2A also illustrates that socioeconomic differences between alternative recipients makes it more difficult for choose based on effectiveness. The next study tests more directly whether distributive justice concerns can explain people's choices of which lives to save.

https://doi.org/10.1017/S1930297500007324 Published online by Cambridge University Press

7 Study 2B: Saving Lives as Compensation

Building on Study 2A, Study 2B tests whether distributive justice concerns play a role in benevolent partiality. In particular, we included a measure of whether people construct saving lives as a means to correct an undesirable distribution, and tested whether this measure mediates the effect of socioeconomic disadvantage on preferences for prosocial aid. This study was preregistered (https://aspredicted.org/et6hf.pdf).

7.1 Method

7.1.1 Participants and Procedure

Participants were 303 adults recruited from Prolific who did not participate in the previous study (63.0% female, mean age = 34.8, English as first language). The study used a two-condition between-participants design.

As in Study 2A, Study 2B asked participants to read about the outbreak of a disease in two fictional countries in the developing world, Sangala and Naruba. In a between-participants design, we again manipulated perceived socioeconomic differences between alternative recipients. In the disadvantage condition, Sangala was described as having worse socioeconomic conditions than Naruba, whereas in the control condition socioeconomic conditions were not spelled out. After participants read the scenario, they were asked where they preferred to allocate aid ("Which campaign do you donate your $10 to? That is, where do you want to ship antidotes to?"). Importantly, as in the previous study, donating towards Naruba allowed saving more lives (4–6) compared to donating to Sangala (2–4).

To test whether distributive justice concerns explain benevolent partiality, we also measured whether people constructed their donation as a means towards correcting an unfair resource distribution. Participants answered three statements on a scale from 1 (Sangala) to 7 (Naruba): "Donating to which campaign, if any, would make the third world a fairer place?", "Donating to which campaign, if any, would make the third world a more equal place?", "Donating to which campaign, if any, would most reduce inequality?". Scores on these three items were averaged (α = .83).

7.2 Results

Consistent with the results of previous studies, we found that emphasizing that Sangala had worse socioeconomic conditions than Naruba made participants more likely to choose Sangala over Naruba, thus saving fewer lives than they could (40.79% in the disadvantage condition vs. 20.53% in the control condition; $\chi^2(1) = 13.68$, $p < .001$, $\varphi = .21$). Moreover, a mediation analysis using 10,000 bootstrapped samples showed that this effect was explained by perceptions that donating towards Sangala would contribute to reduce unjust inequality (indirect effect = .44, CI 95% [.24, .73]).

7.3 Discussion

This study complements Study 2A in showing that socioeconomic differences, even when consequentially irrelevant in the decision problem, can present people with a compensation-efficiency trade-off. In particular, the results of Study 2B indicate that people might construct saving lives as a means to correct situations that they deem unfair, sacrificing efficiency in order to compensate disadvantaged recipients.

8 Study 3: Reflecting on How to Give Reduces Benevolent Partiality

Study 3 tests how the structure of the task employed in the donation decision affects donors' preferences for efficiency vs. compensation in prosocial aid allocation. Normative models of decision-making recommend approaching decisions by first evaluating the importance of each dimension of the problem and then evaluating how available solutions score on such dimensions (e.g., Bazerman & Moore, 2008). In this study, we asked how benevolent partiality varies depending on whether, before donating, people are prompted to reflect on the relative importance of the number of lives saved and the socioeconomic situation of the recipients. This study was preregistered (https://aspredicted.org/k4w7m.pdf).

8.1 Method

8.1.1 Participants and Procedure

Participants were 300 adults recruited from Prolific who did not participate in previous studies (60.7% female, mean age = 34.7, English as first language). The study used a two-condition between-participants design.

Participants went through the same material as the disadvantage condition in Study 2A, and made a choice between donating $10 to save 4–6 lives in Naruba or 2–4 lives in Sangala, with the latter being slightly more socioeconomically disadvantaged. In a between-participants design, we randomly assigned participants to either ranking the importance of the project attributes or not prior to making their choice. In particular, half of the participants were asked "if you had to decide where to ship the antidotes that your $10 donation would allow buying, how would you go about that decision? Please rank the five factors below from 1 (the most important factor in my decision) to 5 (the least important factor in my decision)." We listed five factors in an order randomized for each participant (i.e., lives saved

https://doi.org/10.1017/S1930297500007324 Published online by Cambridge University Press

per donation, population of the country, unemployment rate of the country, literacy rate of the country, average income in the country). All participants chose between the two donation projects (which were described along the same five attributes) using the usual wording.

As a further precaution, at the end of the study we asked participants "What will the collected donation be used for?" as a comprehension check. To pass the check, participants had to select only "For buying and shipping antidotes" among several options.

8.2 Results

Among participants who ranked the importance of the project attributes prior to their choice, 92% put "lives saved" as first in the ranking. Most importantly, participants who first ranked the attributes were less likely to display benevolent partiality than participants who did not rank the attributes. That is, participants were less likely to choose Sangala over Naruba in the ranking first condition (22.97%) compared to participants in the no ranking condition (46.05%; $\chi^2(1) = 16.63$, $p < .001$, $\varphi = .24$). This result did not change when restricting our analysis to participants who passed the comprehension check (21.97% selected Sangala in the ranking first condition vs. 43.88% in the no ranking condition; $\chi^2(1) = 13.69$, $p < .001$, $\varphi = .22$).

8.3 Discussion

The results of Study 3 revealed that reflecting beforehand on how to approach a donation decision makes preferences more aligned to the consequentialist prescription of saving more lives. This finding contributes to our understanding of when benevolent partiality is more and less likely to occur.

9 General Discussion

Across several studies of prosocial aid allocation, we investigated how people resolve the trade-off between efficiency (saving as many lives as possible) and compensation (saving the lives of those who are more socioeconomically disadvantaged). We elicited both consequential (Study 1) and hypothetical (Study 2A, Study 2B, and Study 3) preferences for alternative ways to donate to a charity — one allowing to save more people, and one allowing to save fewer people from a (slightly) more disadvantaged group. We took several precautions to rule out the consequentialist reasons why it often is more sensible to give to those who have the least. In our typical studies, prosocial aid did not address socioeconomic disadvantage, but was explicitly directed at preventing an outcome (death) that could be constructed as equally aversive irrespective of recipient characteristics. Further precautions, which we tested with comprehension checks, clarified that socioeconomic conditions did not correlate with access to help. For these reasons, within our paradigms, preference for saving fewer lives are difficult to reconcile with consequentialist assessments of impact. Contrary to the prescription of consequentialist philosophy and the Effective Altruism movement, we found that many people display what we termed benevolent partiality: They are partial, in that they prioritize some lives over others, contradicting the impartiality principle; they do so benevolently, in that they mean to prioritize those who belong to more disadvantaged groups.

We suggest that benevolent partiality is driven by people's distributive justice concerns, and in particular by people's discomfort with resource distributions that are perceived as inequitable. In support of this interpretation, we found that socioeconomic differences make the decision harder to take when compensation and efficiency arguments conflict (Study 2A), and that people construct saving lives as an opportunity to reduce inequalities that they deem unjust (Study 2B). As a further validation of this interpretation, we conducted a separate study (reported in the Supplementary Material as Study S3) that used the same scenario as Study 2A and Study 2B, and in which we framed disadvantaged recipients as either responsible for their socioeconomic circumstances (i.e., they overharvested crops) or not (i.e., their soil was of poor quality). Previous research suggests that distributive justice concerns are weaker when disadvantaged people are perceived as responsible for their situation (e.g., Fong, 2001, Skitka & Tetlock, 1992). Consistent with benevolent partiality depending on the strength of distributive concerns, we found that people were less likely to sacrifice cost-effectiveness when disadvantaged recipients were perceived as responsible for their disadvantage.

We also found that the prevalence of benevolent partiality depends on how the aid allocation task is structured. In particular, we prompted people to reflect upon the relative importance of alternative dimensions of the problem, a recommended step in prescriptive models of decision-making (e.g., Bazerman & Moore, 2008). We found that this additional step translated into choices that were more aligned with consequentialist prescriptions to save the largest number of lives, irrespective of socioeconomic status.

This work contributes to our understanding of the of moral psychology of altruism, and in particular of the role of distributive justice concerns in prosocial preferences. Discomfort with inequalities can often motivate behaviors that align with consequentialist prescriptions, such as reducing disadvantage (e.g., Alesina & Angeletos, 2005; Frank et al., 2015). We found that distributive justice concerns also play a role in decisions that do not directly affect socioeconomic conditions. In particular, distributive justice concerns can make people more sensitive to the needs of disadvantaged recipients, even if such needs are equal to alternative recipients in less disadvantaged conditions. As a result, people display

https://doi.org/10.1017/S1930297500007324 Published online by Cambridge University Press

preferences and behaviors that deviate from the prescriptions of consequentialist ethics and Effective Altruism.

Our work also informs the psychology of impartiality, which is one key tenet of utilitarian and, more generally, consequentialist philosophy (Kahane et al., 2017). Previous research focused on instances of partiality that depended on the psychological distance between the actor and the victim. For example, people are less willing to help unidentified victims that are more geographically distant (Kogut et al., 2018), or that are perceived as out-group rather than in-group members (Levine et al., 2002). One exception is the work by Goodwin and Landy (2014) showing that the young are often prioritized over the old in life-and-death decision-making contexts (which is consistent with attention to life years rather than lives). Similarly, our research documents partiality among actors that evaluate equally distant targets that differ in one moralized characteristic. We hope that our paper will stimulate more descriptive research on the thorny issue of how people value the lives of others (e.g., Awad et al., 2018). Whereas we focused on the role of socioeconomic disadvantage in the light of its relevance for charity, future research may investigate other determinants of partiality, and their consequences for how aid is ultimately allocated.

Finally, we believe these results are important for governmental agencies and rating organizations that are concerned with quantifying the impact of alternative prosocial initiatives. Rating models are built upon precise philosophical prescriptions (e.g., impartiality), and their normative appropriateness is a long-standing topic in philosophy. Importantly, however, the assumptions in these models may be more or less aligned with the prosocial preferences of the users of such models (e.g., potential donors), also depending on the situation (e.g., on whether the donation task prompts people to reflect on their preferences). The potential for disagreement suggests the need for rating agencies, such as GiveWell, to increase even further the transparency of their consequentialist assumptions. This would result in choices that are more informed and better aligned with people's philosophical orientations.

10 References

Adams, J. S. (1965). Inequity in social exchange. *Advances in Experimental Social Psychology*, *2*, 267–299.

Alesina, A., & Angeletos, G. (2005). Fairness and redistribution. *American Economic Review, 95*, 960–980.

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J.F., & Rahwan, I. (2018). The moral machine experiment. *Nature*, *563*, 59–64.

Banerjee, A. V., & Duflo, E. (2011). *Poor economics: Rethinking poverty & the ways to end it*. India: Random House.

Bazerman, M. H., & Moore, D. A. (2008). *Judgment in managerial decision making*. New York, NY: Wiley.

Bénabou, R., & Tirole, J. (2006). Incentives and prosocial behavior. *American Economic Review*, *96*, 652–1678.

Bentham, J. (1789). *An introduction to the principles of morals and legislation*. England, London: Athlone.

Berman, J. Z., Barasch, A., Levine, E. E., & Small, D. A. (2018). Impediments to effective altruism: The role of subjective preferences in charitable giving. *Psychological Science*, *29*, 834–844.

Cook, K. S., & Hegtvedt, K. A. (1983). Distributive justice, equity, and equality. *Annual Review of Sociology*, *9*, 217–241.

Cryder, C., Botti, S., & Simonyan, Y. (2017). The charity beauty premium: Satisfying donors' "want" versus "should" desires. *Journal of Marketing Research*, *54*, 605–618.

Diener, E., & Biswas-Diener, R. (2002). Will money increase subjective well-being?. *Social Indicators Research*, *57*, 119–169.

Diener, E., & Oishi, S. (2000). Money and happiness: Income and subjective well-being across nations. *Culture and Subjective Well-being*, 185-218.

Engelmann, D., & Strobel, M. (2004). Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. *American Economic Review*, *94*, 857–869.

Evangelidis, I., & Van den Bergh, B. (2013). The number of fatalities drives disaster aid: Increasing sensitivity to people in need. *Psychological Science*, *24*, 2226–2234.

Fong, C. (2001). Social preferences, self-interest, and the demand for redistribution. *Journal of Public Economics, 82*, 225–246.

Frank, D., Wertenbroch, K., & Maddux, W. W. (2015). Performance pay or redistribution? Cultural differences in just-world beliefs and preferences for wage inequality. *Organizational Behavior and Human Decision Processes, 130*, 160–170.

Gneezy, U., Keenan, E. A., & Gneezy, A. (2014). Avoiding overhead aversion in charity. *Science*, *346*, 632–635.

Goodman, J. K., & Paolacci, G. (2017). Crowdsourcing consumer research. *Journal of Consumer Research, 44*, 196–210.

Goodwin, G. P., & Landy, J. F. (2014). Valuing different human lives. J*ournal of Experimental Psychology: General, 143*, 778–803.

Gordon-Hecker, T., Rosensaft-Eshel, D., Pittarello, A., Shalvi, S., & Bereby-Meyer, Y. (2017). Not taking responsibility: Equity trumps efficiency in allocation decisions. *Journal of Experimental Psychology: General*, *146*, 771–775.

Kahane, G., Everett, J. A., Earp, B. D., Caviola, L., Faber, N. S., Crockett, M. J., & Savulescu, J. (2017). Beyond sacrificial harm: A two-dimensional model of utilitarian psychology. *Psychological Review*, *125*, 131–164.

https://doi.org/10.1017/S1930297500007324 Published online by Cambridge University Press

Kiatpongsan, S., & Norton, M. I. (2014). How much (more) should CEOs make? A universal desire for more equal pay. *Perspectives on Psychological Science*, *9*, 587–593.

Kogut, T., Ritov, I., Rubaltelli, E., & Liberman, N. (2018). How far is the suffering? The role of psychological distance and victims' identifiability in donation decisions. *Judgment and Decision Making, 13*, 458–466.

Levine, M., Cassidy, C., Brazier, G., & Reicher, S. (2002). Self-Categorization and bystander non-intervention: Two experimental studies 1. *Journal of Applied Social Psychology, 32*, 1452–1463.

Li, M., Colby, H. A., & Fernbach, P. (2018). Efficiency for lives, equality for everything else: How allocation preference shifts across domains. *Social Psychological and Personality Science, 10,* 697-707.

MacAskill, W. (2015). *Doing good better: Effective altruism and a radical new way to make a difference*. London, England: Guardian Faber Publishing.

Mill, J. S. (1861). *Utilitarianism*, ed. Roger Crisp. Oxford, England: Oxford University Press, *5*, 3–4.

Mitchell, G., Tetlock, P. E., Mellers, B. A., & Ordonez, L. D. (1993). Judgments of social justice: Compromises between equality and efficiency. *Journal of Personality and Social Psychology*, *65*, 629-639.

Norton, M. I., & Ariely, D. (2011). Building a better America—One wealth quintile at a time. *Perspectives on Psychological Science*, *6*, 9–12.

Parfit, D. (1978). Innumerate ethics. *Philosophy & Public Affairs*, *7*, 285–301.

Parfit, D. (2012). *What makes someone's life go best*. Ethical Theory: An Anthology, *13*, 294- 298.

Pellegrino, G. (2017). Effective altruism and the altruistic repugnant conclusion. *Essays in Philosophy*, *18*, 1–24.

Pummer, T., & MacAskill, W. Effective altruism (2019). *International Encyclopedia of Ethics.*

Singer, P. (1993). *Practical ethics*. Cambridge, England: Cambridge University Press.

Singer, P., McKie, J., Kuhse, H., & Richardson, J. (1995). Double jeopardy and the use of QALYs in health care allocation. *Journal of Medical Ethics*, *21*, 144–150.

Singer, P. (2015). *The most good you can do: How effective altruism is changing ideas about living ethically.* New Haven, CT: Yale University Press.

Skitka, L. J., & Tetlock, P. E. (1992). Allocating scarce resources: A contingency model of distributive justice. *Journal of Experimental Social Psychology*, *28*, 491–522.

Small, D. A., Loewenstein, G., & Slovic, P. (2007). Sympathy and callousness: The impact of deliberative thought on donations to identifiable and statistical victims. *Organizational Behavior and Human Decision Processes*, *102*, 143–153.

UN General Assembly (1948). *Universal declaration of human rights*. Paris, France.

World Bank (2018). *Poverty and shared prosperity 2018: Piecing together the poverty puzzle*. Washington, DC: World Bank. License: Creative Commons Attribution CC BY 3.0 IGO