Against the no-difference argument

Adam Elga

August 23, 2023 *

Abstract

There are 1,000 of us and one victim. We each increase the level at which a "discomfort machine" operates on the victim—leading to great discomfort. Suppose that consecutive levels of the machine are so similar that the victim cannot distinguish them. Have we acted permissibly? According to the "no-difference argument" the answer is "yes" because each of our actions was guaranteed to make the victim no worse off. This argument is of interest because if it is sound, similar arguments threaten intuitive moral verdicts about many cases in which a large number of individual choices cumulatively make a great difference, while each choice seems to make no difference on its own. But the argument is not sound, as is shown by a simple objection based on a plausible dominance principle—an objection that avoids challenges that have been brought against previous criticisms of the no difference argument.

Keywords: No-difference argument, group harms, collective action problem, imperceptible benefits, imperceptible harms, collective harm, self-torturer, harmless torturers, thresholds.

1 Introduction

There are one thousand of us, and one victim. Each of us must independently decide whether to raise our hand. If i of us raise our hands, the victim will be zapped by a "discomfort machine" operating at power level i. At

*Forthcoming in *Analysis*. For insightful feedback and suggestions thanks to Zach Barnett, David Builes, Erik Carlson, Pietro Cibinel, David Elga, Alan Hájek, Brian Hedden, Martin Montminy, Jake Nebel, Julia Nefsky, and Gideon Rosen.

level zero the machine produces no discomfort, and at level one thousand it produces horrible discomfort. But the difference in sensation between any two consecutive levels of the machine is so small that the victim would not be able to distinguish between them (Glover and Scott-Taggart 1975, 174-5, Kagan 2011, 116, Nefsky 2019, 6, Parfit 1984, 76, Quinn 1990).

Nothing but the victim's well-being is at stake. Knowing all of the above, is it morally permissible for you to raise your hand?

Many of us answer: no. But some worry about the following *no-difference argument*: Raising your hand would increase the power level of the machine by just one unit, and hence not lead to any distinguishable difference to the victim. So raising your hand is guaranteed not to make the victim any worse off, and hence is permissible.

This argument is of interest because if it is sound, structurally identical arguments threaten intuitive moral verdicts about many cases in which a large number of individual choices cumulatively make a great difference, while each choice seems to make no difference on its own. For example, some altruistic acts benefitting a great number of people might be thought to have effects that are so dispersed as to be imperceptible (Parfit 1984, 76). The same might be said about the harms associated with some individual acts of pollution (Kagan 2011, Sinnott-Armstrong 2005). In each case one might worry that a "no-difference" argument shows that the relevant cumulative effects do not bear on whether each individual act is permissible.

Where does the no-difference argument go wrong? One might argue that it goes wrong by depending on an untruth:

> *No Threshold*. For all $i \leq 999$, being zapped at power level $i+1$ is not worse than being zapped at power level i.

Several authors have replied to the no-difference argument by criticizing No Threshold—or rather, analogs of it for similar cases (Arntzenius and McCarthy 1997, Barnett 2018, 8-9, Broome 1997, Broome 2019, 124–126,

Carlson et al. 2021, 6, Hedden 2020, §3, Kagan 2011, 132, Norcross 1997, Parfit 1984, 79, Voorhoeve and Binmore 2006, 107). These criticisms of No Threshold have significant force but are open to challenges: that they depend on questionable claims about the reliability of experience reports or controversial claims about vagueness or on how to resolve the sorites paradox (Nefsky 2011, Nefsky 2019; though see Regan 2000, 53–54); that they employ disputed semantic assumptions about the transitivity of relations of the form *exactly as F as*; or that they rely on contestable Pareto principles concerning the well-being of groups (Carlson et al. 2021, §4, Montminy 2022, §3). So it would be good to have a response to the no-difference argument that is not vulnerable to these challenges, and that is on as secure a dialectical footing as possible. My aim is to give such a response. My focus is narrow: I aim to show that the *no-difference argument* is unsound, but will not address other arguments that share its conclusion—such as "inefficacy" arguments which grant that each individual act might make a difference, but hold that the chance of making a difference is so small that it does not affect the act's permissibility (Budolfson 2019, 1714, Kagan 2011, 108).

2 *The no-difference argument*

The no-difference argument depends on two claims:

> (i) Raising your hand is guaranteed not to make the victim worse off.
> (ii) If raising your hand is guaranteed not to make a victim worse off, it is permissible.

What support is given for each claim? To get clear on what supports (i), notice that what happens to the victim is determined by the *state* of the world outside of your control (how many others raise their hands), together with your *act* (whether you raise your hand). So to say that an

act is "guaranteed not to make someone worse off" is to say that in each state, they are not worse off if you perform the act than if you don't.[1] So understood, (i) depends for its support on No Threshold, repeated here for convenience:

> *No Threshold*. For all $i \leq 999$, being zapped at power level $i+1$ is not worse than being zapped at power level i.

No Threshold in turn depends on the thought that zappings at consecutive levels of the machine are indistinguishable, together with an unstated principle that it is worse to be zapped at one level than another only if the two levels are distinguishable. Both the thought and the principle can be challenged (Graff 2001, Regan 2000, §1), but I set such challenges aside here in order to criticize the no-difference argument from another angle.

Turn now to (ii). What motivates it is a more general principle saying that in certain circumstances, acts which are guaranteed to make things no worse are permissible (Bales et al. 2014, 460, Doody 2019, 561, Hare 2010, 242, Schoenfield 2014, 267):

> *Permissible*. If (a) your choice is to perform an act or not, (b) nothing but a particular person's well-being is at stake, and (c) for each state that person is not worse off if you perform the act than if you don't, then the act is permissible.

The bottom line is that the no-difference argument crucially depends on both No Threshold and Permissible. So to answer the no-difference argument it is enough to show that their conjunction is false. That is the next order of business.

[1] I have given a gloss in the spirit of causal decision theory as opposed to evidential decision theory, but this is inessential: at the cost of some awkwardness (and some extra assumptions about the evidential independence of states on your action) the discussion could be made neutral between the two theories.

3 Notation and premisses

The argument against the conjunction of Permissible and No Threshold requires some notation and three premisses. Start with the notation: Let Z_i be the *simple prospect* of being zapped by the machine at power level i. Let $L[A, B, \dots]$ be the *chancy prospect* of being subjected to one of $A, B, \dots$, chosen at random by L, where $A, B, \dots$ are a finite number of simple prospects and L is a fair (uniform) lottery. For example, if F is a particular fair coin toss then $F[Z_{10}, Z_{500}]$ is the prospect of being zapped at level 10 if the coin lands heads and level 500 if it lands tails. (Within specifications of lotteries, order matters. So for example, $F[Z_{500}, Z_{10}]$ would be the "reversed" prospect of being zapped at level 500 if the coin lands heads and level 10 if it lands tails.) Let variables $A, B, \dots$ range over simple prospects, and variables $X, Y, \dots$ range over both simple and chancy prospects. Finally, let $X \prec Y$ be the claim that being subjected to X is worse than being subjected to Y. This claim will also sometimes be abbreviated "X is worse than Y", its negation will sometimes be written "$X \not\prec Y$", and the claim that X is *much* worse than Y will be written "$X \prec\prec Y$".

Now for the three premisses, which I introduce in turn. The first premiss is:

> *Harm*. If it is permissible to subject a victim to X rather than Y (given a choice between the two), then it is not worse to be subjected to X than Y.[2]

Since prospects are restricted to lotteries over various power levels of a discomfort machine, since only a single victim's well-being is at stake, and since (we may assume) no promises or special personal relationships or obligations are involved, this premiss is highly plausible.

[2]Choices are assumed to be fully informed in the sense that the options are explicitly specified lotteries and the chooser knows exactly how good or bad each lottery outcome is.

To motivate the second premiss, imagine that you have a ticket that gives you a one-in-a-million chance of winning $1 billion. Officials ask whether you wish to trade that ticket for one that instead offers a one-in-a-million chance of "winning" a year of extreme discomfort. Tempted to trade? Of course not—the discomfort ticket is obviously worse. This is true regardless of whether the two tickets depend on the same chance process. And it would still be true if the tickets also offered chances of additional prizes—provided that the additional prizes were the same for each ticket, and had the same chances. These ticket trades make one worse off because they replace the chance of one outcome with an equal chance of a much worse outcome, while leaving the chances of all other outcomes unchanged.

Notice that the above judgments do not depend on special assumptions about how good or bad the additional prizes are, or even whether the additional prizes can be assigned numerical levels of goodness. To illustrate this, suppose for the sake of argument that a particular finite set $\boldsymbol{Q}$ of prizes cannot be assigned numerical levels of goodness. For example, perhaps one of the prizes is so alien and bizarre that it is incomparable to winning $1 billion, and also incomparable to a year of extreme discomfort (Hare 2010, 239). (Two prizes are incomparable if they are not equally good, and neither is worse than the other.) Or perhaps three of the prizes are such that the first is not worse than the second, and the second is not worse than the third—but the first *is* worse than the third.

Now suppose that you have a ticket that gives you $1 billion with a one-in-a-million chance, and otherwise gives you a prize chosen uniformly at random from $\boldsymbol{Q}$. Officials ask whether you wish to trade that ticket for one that gives you a year of extreme discomfort with a one-in-a-million chance, and otherwise gives you a prize chosen uniformly at random from $\boldsymbol{Q}$. In this case, too, trading would be unwise.

The second premiss, *Worsen*, is a special case of the idea that trades of

the above sort always make one worse off:[3]

Worsen. If $A' \prec\prec A$ then
$L[B_1, B_2, \ldots, B_n,\ A'] \ \prec\ L[A,\ B_1, B_2, \ldots, B_n]$.

The third and final premiss, *Endpoints*, is a stipulation from the initial case: that being zapped at level one thousand is much worse than being zapped at level zero:

Endpoints. $Z_{1000} \prec\prec Z_0$.

4 *An argument against the conjunction of Permissible and No Threshold*

To show that Permissible and No Threshold are not both true, it is enough to show that if Permissible is true, then No Threshold is false. So assume for the sake of argument that Permissible is true. I will now argue that given Harm, Worsen, and Endpoints, it follows that No Threshold is false:[4]

(1) If $A_i \not\prec B_i$ for all $i \leq n$, then it is permissible to choose $L[A_1, A_2, \ldots, A_n]$ over $L[B_1, B_2, \ldots, B_n]$ (given a choice between the two).

(2) If $A_i \not\prec B_i$ for all $i \leq n$, then $L[A_1, \ldots, A_n] \not\prec L[B_1, \ldots, B_n]$.

(3) $L[Z_1, Z_2, \ldots, Z_{999},\ Z_{1000}] \prec L[Z_0,\ Z_1, Z_2, \ldots, Z_{999}]$.

(4) Not: ($Z_{i+1} \not\prec Z_i$ for all $i \leq 999$).

(1) is an application of Permissible to chancy prospects.[5] (Permissible

[3]Worsen follows from the conjunction of a finite statewise dominance principle (Williamson forthcoming, 9, Doody 2021, 250) and the principle that prospects are equivalent when they offer the same chances of the same outcomes (Hare 2010, 240-1, Meacham 2019, 999 Doody 2021, 249, Williamson forthcoming, 8). That said, I mean to use the ticket comparisons in the main text to motivate Worsen directly.

[4]This argument is a modal analog of the "staircase argument" (Barnett 2018, 8-9), where modal variation (different possible outcomes of a lottery) replaces interpersonal variation (different actual outcomes to members of a population).

[5]Throughout I leave implicit that the stated constraint is claimed to hold for all simple prospects $A_1, A_2, \ldots, B_1, B_2, \ldots$ and all uniform lotteries L.

applies in this way because replacing each B_i with A_i ensures that for each state—result of lottery L—the victim is not worse off if the replacement is performed than if it is not.) (2) follows from (1) and Harm by modus ponens. (3) follows from Endpoints by Worsen. (4) follows from (2) and (3). But (4) is the negation of No Threshold, so the argument is complete.

The overall conclusion is that Permissible and No Threshold are not both true, and hence the no-difference argument is unsound.

Notice that this argument avoids challenges that might be raised against previous responses to the no-difference argument: It relies on no claims about the reliability of experience reports. It does not depend on claims about the transitivity of relations of the form *exactly as F as*, or on Pareto principles concerning the well-being of groups. It employs no sorites-style reasoning and assumes no particular account of vagueness. (Admittedly, the argument does concern a series of cases whose successive elements differ at most to a small degree. But the argument does not appeal to a tolerance premiss. It does not, for example, assume—even for *reductio*—that if Z_i is painless then so is Z_{i+1}.) Nor does the argument beg questions about vagueness by ruling out from the start that the "worse than" relation is vague. And the argument does not illicitly settle controversial questions about sharp cutoffs. (For example, according to some views of vagueness: (4) is true, but no cutoff claim of the form "$Z_{i+1} \prec Z_i$" is true.)

So however strong the case against the no-difference argument was before, it is now even stronger.

References

Frank Arntzenius and David McCarthy. Self torture and group beneficence. *Erkenntnis*, 47(1):129–144, 1997. doi: 10.1023/a:1005376607563.

Adam Bales, Daniel Cohen, and Toby Handfield. Decision theory for

agents with incomplete preferences. *Australasian Journal of Philosophy*, 92 (3):453–470, July 2014. doi: 10.1080/00048402.2013.843576.

Zach Barnett. No free lunch: The significance of tiny contributions. *Analysis*, 78(1):3–13, 2018. doi: 10.1093/analys/anx112.

John Broome. Is incommensurability vagueness? In Ruth Chang, editor, *Incommensurability, Incomparability and Practical Reason*, pages 67–89. Harvard University Press, Cambridge, MA, 1997.

John Broome. Against denialism. *The Monist*, 102(1):110–129, January 2019. doi: 10.1093/monist/ony024.

Mark Bryant Budolfson. The inefficacy objection to consequentialism and the problem with the expected consequences response. *Philosophical Studies*, 176(7):1711–1724, July 2019. doi: 10.1007/s11098-018-1087-6.

Erik Carlson, Magnus Jedenheim-Edling, and Jens Johansson. The significance of tiny contributions: Barnett and beyond. *Utilitas*, 33(4):488–496, 2021.

Ryan Doody. Opaque sweetening and transitivity. *Australasian Journal of Philosophy*, 97(3):559–571, July 2019. doi: 10.1080/00048402.2018.1520269.

Ryan Doody. Hard choices made harder. In Henrik Andersson and Anders Herlitz, editors, *Value Incommensurability*, pages 247–266. Routledge, New York, 2021. doi: 10.4324/9781003148012-17.

Jonathan Glover and M. Scott-Taggart. "It makes no difference whether or not I do it". *Aristotelian Society Supplementary Volume*, 49(1):171–210, July 1975. doi: 10.1093/aristoteliansupp/49.1.171.

Delia Graff. Phenomenal Continua and the Sorites. *Mind*, 110(440):905–936, October 2001. doi: 10.1093/mind/110.440.905.

Caspar Hare. Take the sugar. *Analysis*, 70(2):237–247, April 2010. doi: 10.1093/analys/anp174.

Brian Hedden. Consequentialism and collective action. *Ethics*, 130(4): 530–554, 2020. doi: 10.1086/708535.

Shelly Kagan. Do I make a difference? *Philosophy & Public Affairs*, 39(2): 105–141, 2011. doi: 10.1111/j.1088-4963.2011.01203.x.

Christopher J. G. Meacham. Difference Minimizing Theory. *Ergo, an Open Access Journal of Philosophy*, 6(35), 2019. ISSN 2330-4014. doi: 10.3998/ergo.12405314.0006.035.

Martin Montminy. Do tiny contributions make a difference? Reply to Barnett. *Analysis*, 82(4), October 2022. doi: 10.1093/analys/anac033.

Julia Nefsky. Consequentialism and the problem of collective harm: A reply to Kagan. *Philosophy & Public Affairs*, 39(4):364–395, 2011. doi: 10.1111/j.1088-4963.2012.01209.x.

Julia Nefsky. Collective harm and the inefficacy problem. *Philosophy Compass*, 14(4), 2019. doi: 10.1111/phc3.12587.

Alastair Norcross. Comparing harms: Headaches and human lives. *Philosophy & Public Affairs*, 26(2):135–167, April 1997. doi: 10.1111/j.1088-4963.1997.tb00079.x.

Derek Parfit. *Reasons and Persons*. Oxford University Press, Oxford, 1984.

Warren S. Quinn. The puzzle of the self-torturer. *Philosophical Studies*, 59(1): 79–90, 1990. doi: 10.1007/bf00368392.

Donald H. Regan. Perceiving imperceptible harms (with other thoughts on transitivity, cumulative effects, and consequentialism). In Michael J. Almeida, editor, *Imperceptible Harms and Benefits*, Library of Ethics and

Applied Philosophy, pages 49–73. Springer Netherlands, Dordrecht, 2000. doi: 10.1007/978-94-011-4144-4_3.

Miriam Schoenfield. Decision making in the face of parity. *Philosophical Perspectives*, 28(1):263–277, 2014. doi: 10.1111/phpe.12044.

Walter Sinnott-Armstrong. It's not my fault: Global warming and individual moral obligations. In Walter Sinnott-Armstrong and Richard B. Howarth, editors, *Perspectives on Climate Change: Science, Economics, Politics, Ethics*, volume 5 of *Advances in the Economics of Environmental Resources*, pages 285–307. Emerald Group Publishing Limited, Bingley, January 2005. doi: 10.1016/S1569-3740(05)05013-3.

Alex Voorhoeve and Ken Binmore. Transitivity, the sorites paradox, and similarity-based decision-making. *Erkenntnis*, 64(1):101–114, 2006.

Timothy Luke Williamson. A risky challenge for intransitive preferences. *Noûs*, forthcoming. doi: 10.1111/nous.12455.