1-31-2017

Effective Altruism and the Altruistic Repugnant Conclusion

Gianfranco Pellegrino
Libera Università Internazionale degli Studi Sociali Guido Carli, yorick612@gmail.com

Follow this and additional works at: http://commons.pacificu.edu/eip

Essays in Philosophy

Effective Altruism and the Altruistic Repugnant Conclusion

Gianfranco Pellegrino
Libera Università Internazionale degli Studi Sociali Guido Carli

Abstract

Effective altruism is committed to Altruistic Maximization – the claim that any impact of giving to charities ought to be maximized at the margins and counterfactually. This may lead to counterintuitive or contradictory conclusions in certain cases. For instance, when we can bring about a substantial benefit to few or a tiny benefit to a larger number at the same cost, spreading of benefits across a great number of recipients can compensate substantial loss for fewer people. However, sometimes the perspective of widely spreading tiny benefits instead than giving substantial benefits to small groups is counterintuitive, and repugnantly wrong. Call this the 'Altruistic Repugnant Conclusion.' Standard solutions to the Repugnant Conclusion do not work if applied to the Altruistic Repugnant Conclusion. The Altruistic Repugnant Conclusion can be rebutted by giving up on the idea that judgements on such cases should be precise. But imprecision undermines Altruistic Maximization. Hence a dilemma for effective altruists follows: either they accept some strongly counterintuitive judgments—weakening the attractiveness of their position—or they give up on the maximizing requirement, thus admitting sub-optimal contributions.

Introduction

Recently, Theron Pummer claimed that, once you decided to donate to charities, you don't have the option of making sub-optimal donations.[1] Defending this claim, he passingly noticed that "as many charitable causes are difficult to compare [...], there may be no best charities, but only an upper set of charities" (Pummer 2016, 85). In this paper, I shall explore this idea, and its consequences for effective altruism (EA). If Pummer is right, EA is in trouble, or so I argue.

According to Peter Singer, EA is "based on a very simple idea: we should do the most good we can" with our altruistic resources. (Singer 2015, vii) Doing the most good we can is a matter of differential impact: we should make "the most difference", thereby having "the biggest positive impact." (Singer 2015, 118) EA's core tenet, then, may be put as follows:

> *Effective Altruist Maximization (AM):* We ought to do the most good we can, maximizing the impact of donating to charities on the margin and counterfactually —which means that among the available charities, the one that is most effective on the margin should be chosen.[2]

Singer notices that "doing the most good is a vague idea that raises many questions" (Singer 2015, 7). Some of the questions are not empirical, but questions of value. However, Singer rejects an obvious conclusion that may be derived from vagueness— i.e., that there is no objective answer to the question: "Where can I do the most good?". (Singer 2015, 117) Singer suggests that there are many cases where we can reasonably agree, and disagreement on the few puzzling cases will not undermine agreement on clearer cases.[3] Moreover, as Singer points out, despite differences, effective altruists share some common principle. Effective altruists, Singer says,

> would all agree that a world with less suffering and more happiness in it is, other things being equal, better than one with more suffering and less happiness. Most would say that a world in which people live longer is, other things being equal, better than one in which people live shorter lives. (Singer 2015, 7)

---

[1] See (Pummer 2016).

[2] Effective altruist maximization is here framed in terms of marginal utility—in terms of the increase in utility brought about by a supplementary amount of money given to a charity. This is the meaning of the expression 'on the margin' in the main text.

[3] See (Singer 2015a, 107).

The principles Singer mentions in this passage concern features—such as happiness, or the length of lives—which ought to be maximized. Spelling out such principles is a way to make EA less vague.[4] But there are other features that any maximizing view of beneficent action should consider. For instance, the size of the benefits, but also the number of persons involved, or both.[5]

Indeed, AM comes in different versions, depending on the features to be maximized. An obvious contention is that effective altruists should decide which of these more specific versions of AM they endorse—this is part of making EA less vague. A less obvious claim is that none of the more easily predictable specific versions of AM are entirely satisfactory, and that further work needs to be done in trying to articulate more precisely EA's main core. In this paper, I shall try to substantiate this less obvious view.

The paper's main claims are the following. Sometimes, we can either spread tiny benefits across large numbers or give more substantial benefits to smaller groups at the same cost. If numbers are large enough, the first option seems obviously better. In virtue of the same logic, assuming that the cost remains the same, spreading minuscule benefits to immensely large groups should be better than giving a substantial benefit to very small numbers, or to individuals. However, often this conclusion seems counterintuitive, and repugnantly wrong. I call this the Altruistic Repugnant Conclusion (ARC)—a version of the standard Repugnant Conclusion (RC). Standard solutions to RC do not apply to ARC. Only one solution holds: giving up precision in comparisons. But this solution undermines AM. If we cannot precisely compare the impact of different charities, we cannot be required to maximize our donations.

The paper develops as follows. In §2, I lay out the specific principles upon which Singer and other effective altruists may suggest that certain charitable options are better than others. In §3, I show that the most convincing interpretation of AM leads to ARC.[6] In §4, I consider some ways to dismiss ARC, and rebut all of them. § 5 concludes.

---

[4] However, once these principles are endorsed, EA can be thicker than without them, thereby diminishing its appeal to a wide audience. For the difference between thick and thin versions of EA, see (Gabriel 2016, 2).

[5] Cp. (Stark 2008).

[6] On RC, see (Arrhenius, Ryberg, and Tannsjo 2014; Parfit 1984, 388; Sidgwick 1907, 418; Broad 1930, 249–50; McTaggart 1927, 452–3; Narveson 1967; Parfit 2016). There are many different versions of RC; see for instance (Arrhenius 2009). Here, I am not considering the differences, if any, between the original RC and these variations of it, and between them and ARC; nor do I consider the possible implications, if any, of these differences.

Going across cases and principles

Much of T*he Most Good You Can Do* is a discussion of possible alternative charitable actions.[7] The main argumentative strategy is to show that one of two alternative actions is strikingly more effective than the other, notwithstanding appearances or intuitions to the contrary. In this section, I shall focus on some of these cases, in order to elaborate the principles of distribution that Singer might presuppose.[8]

Consider
     Giving directly: You have $1,000 and you wish to give this amount to a poor family. You can either
          A. give it to an American poor family;
          or
          B. give it to an impoverished African family.[9]

Singer suggests to choose **B**. Here his reasoning: The American poor family is very likely to be already receiving aid from the government or domestic charities. Perhaps this aid is not "enough to raise [this family] out of poverty". However, the African family is very likely to get nothing from the government or fellow-citizens, and it remains in a much more impoverished context. Accordingly, "the cost of making a lasting, positive difference to the life of a poor American is far higher than the cost of making such a difference to the life of someone who is poor by global standards". (Singer 2015, 114, 115) Also, it seems highly plausible that the benefit that the American family will get from an additional $1000 is smaller than the benefit the African family might get. Hence, it is most effective to give money to the African family.[10]

---

[7] This seems to be a constant of the—still rare—scholarship on EA. Like Singer's *The Most Good You Can Do*, William MacAskill's *Doing Good Better* is replete with cases and discussions of them. See also (Pummer 2016; Gabriel 2016; Carey 2015).

[8] Singer's treatment of these cases is discursive and extended. I myself shall paraphrase it in a sketchier way, for the sake of shortness and in order to provide a comparative view of the various cases. References to Singer's passages will be given in the main text or in footnotes. Also, I am employing the figures about costs that Singer uses in *The Most Good You Can Do*. Nothing hinges on the choice of these figures.

[9] See (Singer 2015, 114, 115).

[10] This is an application of the law of diminishing marginal returns: the same sum of money yields different benefits depending on the well-being level of the recipient. Thus, *ceteris paribus*, the worse off the recipient, the greater the benefit. On EA and the law of diminishing marginal returns, see (MacAskill 2015).

In *Giving directly*, the comparison concerns *differently sized benefits* to be distributed to the s*ame number* of people. EA dictates choosing the option where the benefit has the greatest size—which is an obvious suggestion. The same suggestion might be given when *different numbers* of people are involved.

Consider:

> *Dog vs. preventing blindness:* You have $40,000 to donate to the best cause you can find. You can either
>> **A.** supply one blind person in the United States with a guide dog; or
>> **B.** preventing "anywhere between four hundred and two thousand causes of blindness in developing countries". (Singer 2015, 111)[11]

In *Dog vs. preventing blindness*, the benefits have different size (giving a guide dog to a US citizen is a smaller benefit than avoiding blindness to a developing countries citizen), and the recipients are different in number. However, option **B** is to be chosen for the same reason invoked in *Giving directly*, namely because it brings about larger benefits. It seems that difference in numbers is not relevant here. Indeed, the same logic applies to:

> *Wish vs. lives*: You have $7,500 to donate to the best cause you can find. You can either
>> **A.** make a child's wish come true. For instance, you can make Miles Scott's wish to be Batkid come true—Miles Scott "had been through three years of chemotherapy for leukemia"; or
>> **B.** save at least three children from malaria. (Singer 2015, 5–7)

Since **B** brings about a larger benefit, that's the option to choose. As Singer remarks,

> saving a child's life has to be better than fulfilling a child's wish to be Batkid. If Miles' parents had been offered that choice—Batkid for a day or a complete cure for her son's leukemia—they surely would have chosen the care. [...] Saving a life is better than making a wish come true [...]. (Singer 2015, 7)

In the cases considered above the right choice is determined by considering the size of the benefits produced: the reason to choose **B** is the larger benefit it could bring about. The principle underlying these choices might be put as follows:

---

[11] Cp. (Ord 2013, 1).

Size matters (from now onwards, *Size*): Doing the most good amounts to bringing about the *greatest* benefit.

This principle gives a specific articulation to AM.

Now, consider:

Art vs. sight: You have $100,000 to donate to the best cause you can find.

You can either

A. contribute to build a new museum wing, thereby giving an enhanced aesthetic experience to 100,000 visitors;

or

B. restore or preserve the sight of 1,000 poor people in developing countries.[12]

Plainly, restoring or preserving one's own sight is a greater benefit than seeing a new museum wing. Accordingly, **B** is the obvious choice here, and *Size* rightly dictates it. *Size* seems to apply both to same-number and different-number cases.[13]

But now consider:

Separation vs. vaccination: You have $1 million to donate to the best cause you can find. You can either

A. give it to a hospital in Palo Alto, California, where it will be used to fund a separation of two conjoined twins;

or

B. give it to a vaccination program in developing countries, whose impact will save hundreds of children by protecting them from measles.[14]

Here, the benefit at stake is escaping an early death: the benefit has the same size and significance in both options. However, the recipients of the benefit are *differently numerous*. *Size* would suggest that in this as well as in similar cases either option may be chosen. This may strike us as a counterintuitive suggestion: there should be a difference

---

[12] See (Singer 2015, 118–19).

[13] (Parfit 1984, 356) distinguishes same-number and different-number choices. *Size* applies to both.

[14] See (Singer 2015, 109–110).

between benefitting one or two individuals, and benefitting far many people. It seems that numbers cannot be completely irrelevant in cases like that. Numbers make the difference, here.

In *Separation vs. vaccination*, Singer suggests **B**: "it doesn't seem all that difficult to judge that it is better to use a million dollars to save the lives of hundreds of children by protecting them from measles than to use it to separate one pair of conjoined twins or save one extremely premature infant." (Singer 2015, 110)

Notice that Singer is not considering only the benefit's size; for him, numbers count. The principle he is assuming might be this:

> *Size and number matter* (from now onwards, *Size/number*): Doing the most good amounts to bringing about the greatest benefit to the greatest number.

This principle constitutes another specific articulation of AM.[15]

Consider:
> *Separation vs. preventing blindness*: You have $1 million to donate to the best cause you can find. You can either
> > **A.** give it to a hospital in Palo Alto, California, where it will be used to fund a separation of two conjoined twins;
> > or
> > **B.** give it to a program in developing countries, whose impact will prevent blindness by protecting 10,000 people from trachoma.[16]

The benefit brought about in **B** is smaller than the one produced by choosing **A**—death is obviously worse than blindness.[17] Then, *Size* would dictate choosing **A.** It is not clear which

---

[15] *Size/Number* applies only to different-number choices.

[16] I am assuming that the cost of curing trachoma is $100. I am relying on (Singer 2015, 109–11 and 194 n. 7),

[17] I am assuming that, without separation, the twins mentioned in B will face an early death. This upshot seems likely. In terms of DALY, "1 year when blind is equivalent to 0.8 years of healthy life". Accordingly, "curing a person of blindness for 5 years is equivalent to extending a healthy person's life by 1 year". (Singer 2015, 132) I am assuming that a $100 surgery is able to prevent one to thirty years of blindness, as per GiveWell's figure reported in (Singer 2015, 194 n. 7). Assuming the optimistic figure—30 years –, B amounts to prolonging 10,000 lives by 6 years. Anyway, it seems pretty obvious that if there is a discount rate of 0.2 for blindness, blindness is better than death. This is enough to make plausible my discussion of *Separation vs preventing blindness*.

verdict would be given by *Size/Number*, because this principle can be differently specified.

Consider a lexical specification of *Size/Number*:

> *Size first, number after* (from now onwards, *Lex*): Doing the most good amounts to **i.** bringing about the greatest benefit, and when **i.** is met, **ii.** spreading it over the greatest number.

Lex captures the idea that the significance of size overrides the significance of numbers. Therefore, no decrease in size can be compensated for by *any* increase in numbers. The best option is the option where the greatest benefit is brought about. If we can have same-sized benefits spread over differently numerous groups, we should choose the option benefiting the most crowded group. But, if benefiting a more crowded group requires diminishing the size of the benefit given, then this benefit should not be done to the many. Applied to *Separation vs. preventing blindness*, *Lex* dictates choosing **A**.[18]

Choosing **A** in *Separation vs. preventing blindness* can sound counterintuitive, though, for the same reason choosing **A** in *Separation vs. vaccination* is counterintuitive: numbers are relevant in these sorts of cases, and their relevance can outweigh the importance of benefit's size. Consider:

> *Rare disease vs. preventing blindness*: You have $1 million to donate to the best cause you can find. You can either
> > **A.** give it to a medical research centre, where it will be used to fund scientific research concerning very rare fatal diseases—such as Kuru or Creutzfeld-Jacob, for instance;
> > or
> > **B.** give it to a program in developing countries, whose impact will prevent blindness by protecting 10,000 people from trachoma.

Here, we can assume that, other things being equal, being cured from a rare fatal disease amounts to escaping an early death. Accordingly, the benefit produced by choosing **A** is larger than the one brought about by choosing **B**. Nevertheless, many effective altruists would choose **B** in these sorts of cases.

However, if **B** is a plausible choice in *Rare disease vs. preventing blindness*, why is it not in *Separation vs. preventing blindness*? The stakes are the same: **A** prevents an early

---

[18] On lexical views, see (Parfit 2016, 112–3, 115).

death for a small number of people, while **B** prevents blindness for a much larger number of people. If **B** is to be preferred to **A** in the latter case, then it should be preferred to **A** in the former too. But if one ought to choose **B** in both cases, we need a principled defence of this move—a principle allowing us to choose a smaller benefit, when it is given to a larger number of people. This principle can be the following:

> *Size and number mutually balance* (from now onwards, *Balance*): Doing the most good amounts to bringing about the **a.** greatest as well as **b.** maximally spread benefit, where gains in diffusion compensate for losses in size, and vice versa.[19]

As the benefit produced in **B** is more spread than the benefit brought about by **A**, and its larger diffusion compensates for its smaller size, then *Balance* dictates choosing **B** in both *Rare disease vs. preventing blindness* and *Separation vs. preventing blindness*. Losses in size can be compensated for by increases in the number of people benefited. And this seems very plausible. Indeed, Singer endorses this principle. He concedes that when considering cases such as *Rare disease vs. preventing blindness* one can be tempted to answer that "when we are faced with the needs of those who are […] severely burdened", the sum of the smaller benefits of the many have "no justificatory weight". But then he adds: "Intuitively appealing as this answer might be […], many effective altruists will be uncomfortable with the idea of setting a cutoff line for severe burdens against which a vast number of lesser harms simply do not matter."(Singer 2015, 120)

Lex is counterintuitive, then. It is also unfit to EA's rationale. In both GiveWell's and Giving What We Can's rankings the recommended top charities are engaged in bringing about benefits impacting over *large populations*—benefits such as providing anti-malaria bed-nets or deworming tablets. The relevance of the number of people involved is evident in these rankings. It seems that a principle such as *Balance* is the principle we should apply in order to give meaning and content to AM. But unfortunately, *Balance* can give counterintuitive results as well, as I shall show in the next section.

The altruistic repugnant conclusion

In *The Most Good You Can Do*, Singer takes on "difficult comparisons"—cases where either we have "no methods of making comparisons" between "a variety of causes", or such methods "exist, although they raise deep philosophical questions and so remain controversial". (Singer 2015, 129) Singer takes on a case discussed by Holden Karnofsky

---

[19] Cp. (Parfit 1984, 387).

(one of the founders of GiveWell), briefly sketched below.[20]

> *Three choices*: You have a sum that you would like to donate. With this sum, you can accomplish any of the following:
>
> > **A.** "prevent 100 deaths-in-infancy, knowing that in all likelihood these 100 people will grow up to have consistently low income and poor health for their ~40-year-long-lives";
> > **B.** "provide consistent, full nutrition and health care to 100 people, such as that instead of growing up malnourished (leading to lower height, lower weight, lower intelligence and other symptoms) they spend their lives relatively healthy", even though "this doesn't affect their actual life span—they still live about 40 years.
> > **C.** "prevent one case of relatively mild non-fatal malaria (say, a fever that last few days) for each of 10,000 people, without having a significant impact on the rest of their lives." (Singer 2015, 132–3).

According to Karnofsky,

> the answer is definitely B. I am very excited by the idea of changing someone's life in a lasting and significant way (B); I am much less excited by the idea of a temporary, less significant life change (C), and I don't think that the quality of a life equals the sum of the quality of the days in it. (A) excites me the least—I just don't put much value in "potential lives" (I think the death of a 20-year-old is more tragic than the death of an infant), and I especially don't put much value in saving "potential lives" riddled with health problems.[21]

Karnofsky seems to imply this: **A** will benefit merely possible people—the children mentioned in **A** are doomed to die, and their death is similar to their never being born. **B** will produce a larger benefit than **A** to the same number of people. **C** will yield a much smaller benefit, even though spread over a much larger number.

Singer does not take side on the issue, even though he quotes Karnofsky's preference for **B.** This assessment can be supported in terms of the principles stated in the previous section. If one endorses *Size*, then **B** is obviously better than either **A** or **C**, because

---

[20] Karnofsky aims at raising doubts about the use of DALY as a metric. But his point has general implications.

[21] http://blog.givewell.org/2008/07/28/significant-life-change/

it produces a larger and more substantial benefit. However, as we have seen above, *Size* can have counterintuitive implications.

Consider:

> *Three choices (no-death)*: You have a sum that you would like to donate. With this sum, you can accomplish any of the following:
>
> > **A.** provide consistent, full nutrition and health care to 100 people, such as that instead of growing up malnourished (leaning to lower height, lower weight, lower intelligence and other symptoms) they spend their lives relatively healthy;
> >
> > **B.** provide enhanced educational opportunities to 50 people living in a developing country, thereby empowering them to a more active and successful life;
> >
> > **C.** prevent one case of relatively mild non-fatal malaria (say, a fever that last a few days) for each of 10,000 people, without having a significant impact on the rest of their lives.[22]

The case above is highly stylized. To assess it, we should make a number of assumptions. First, I assume that all people involved live in developing countries. Giving greater educational opportunities to people in developed countries may be unrealistically costly. I don't consider these high costs, but only the relatively lower *per capita* costs needed to donors in order to guarantee greater educational opportunities in the context of developing countries.

Second, I grant that giving enhanced education opportunities to people in developing countries can have very good effects on third parties. Accordingly, under this perspective in *Three choices (no death)* **B** may be *overall* better than **A** and **C.** But here I put third-parties effects aside. I focus only on the good effects of more education *on the people getting it*. I assume that a boost in one's own educational opportunities is always better than more basic forms of aid.

---

22 *Three choices (no death)* differs from the original case, i.e. *Three choices*, in not including a death-or-life alternative. **A** in *Three choices* is about preventing deaths, while **A** in *Three choices* (*no death*) concerns providing goods not related to death-or-life alternatives. A case similar to *Three choices (no death)* is discussed in (Gabriel 2016, 4–5). However, in his discussion Gabriel focusses on EA's alleged blindness to the predicament of the worst-offs. Here, I am interested in the fact that EA can lead to Repugnant Conclusion-scenarios, in such sort of cases. Despite related, the two points are not coincident. Notice that according to some authors prioritarianism leads to the Repugnant Conclusion; see (Brown 2006; Holtug 2007; Persson 2012).

More needs to be said on that, though. Arguably, a seriously destitute or seriously burdened person would prefer any amount of basic intervention—in terms of food or medicines—to any amount of more sophisticated educational aid. In most cases, people starving or unhealthy prefer food or medicines to any opportunity to increase their education. Denying that satisfaction of these basic needs is more urgent than getting other sorts of goods would be nonsense. This may imply that in *Three choices (no death)*, **A** and **C** are much better than **B.**[23]

However, there is a sense in which improved educational opportunities are better than more basic aid. First of all, educational opportunities—or their effects—are more lasting. People can be fed and cured, and then they can be destitute again, if the flux of aid stops. By contrast, if people are educated, their life will be lastingly better. Second, greater educational opportunities can have an empowerment effect. Educated people acquire a control of their life able to keep them safer—to make them more resilient, so to say, to possible causes and circumstances of destitution. We may be facing increasing marginal returns, here. The best off is a person—the best off she is in a problematic context, such as that of living in a developing country and not belonging to the elites—, the larger is the benefits she got from sophisticated interventions on her education and similar sources of empowerment. Then, in *Three choices (no death)*, **B** is obviously better than **A** and **C**: it produces a substantially larger benefit.[24]

It might be argued that the figures displayed in *Three choices (no death)* are unsupported. Educational interventions are often difficult, and much more costly than other more basic interventions.[25] Likewise, impact on education very often comes as a side-effect of more basic interventions. It is unlikely that by paying the same cost 50 people can be educated, 100 people can be cured and fed, and 10,000 individuals can be cured of a mild non-fatal malaria.

To rebut this objection, here is my third (group of) assumption(s): first, the interventions considered in **B** in *Three choices (no death)* are *purely* educational, i.e. they are not a by-product of other interventions—rather, they can be achieved only through direct and specific interventions; second, the number of people involved in **B** amounts to *any* number of individuals whose educational opportunities can be enhanced at the same

---

[23] Notice that if this is the case, Karnofsky's judgment on *Three choices* can be challenged.

[24] Age and gender should also be taken into account. The younger the recipients, the greater and the more lasting the benefits of education. Educated women could acquire a greater capacity to control their lives as compared to men.

[25] See (MacAskill 2015, 8–9, 51) on deworming and its effects on improved school performances.

cost needed to fed and cure 100 people or to prevent a case of mild non-fatal malaria for 10,000 individuals.

The case, then, can be restated as follows:

> *Three choices (no death)**: You have a sum that you would like to donate. With this sum, you can accomplish any of the following:
>
> > **A.** provide consistent, full nutrition and health care to 100 people, such as that instead of growing up malnourished (leading to lower height, lower weight, lower intelligence and other symptoms) they spend their lives relatively healthy;
> >
> > **B*.** provide enhanced education opportunities to *x* people living in a developing country, thereby empowering them to a more active and successful life;
> >
> > **C.** prevent one case of relatively mild non-fatal malaria (say, a fever that last a few days) for each of 10,000 people, without having a significant impact on the rest of their lives.

If only the size of benefit matters, and all the assumptions above hold, then **B*** is to be chosen, *no matter how few the people benefited are*. Following this principle, it would be better to fund separation for a pair of conjoined twins rather than preventing blindness for 10,000 people, or it would be better funding research about rare diseases rather than preventing blindness for 10,000 people. As said, this verdict does not fit EA's spirit—indeed, it does not fit the rankings provided by GiveWell, where charities having an impact on the most basic needs of large populations are preferred to charities focussing on difficult and more sophisticated interventions.

Now, recall how *Balance* emerged as the best interpretation of AM. It might be argued that, applied to *Three choices (no death)**, this principle gives the right answer. Promoting better education for *any number* of people cannot be better than feeding or curing 100 individuals or curing mild non-fatal malaria for 10,000 people. What is better depends on number and size, with mutual compensations. Then, while to provide consistent, full nutrition and health care to 100 people can be better than preventing one case of mild non-fatal malaria for 10,000 individuals, to provide better education to 50 people is not better than preventing one case of mild non-fatal malaria for 10,000 individuals. As increase in number can compensate losses in the size of benefit given, the relative difference in the size and the diffusion of the benefit compared are the relevant factors to establish which option to pursue. No option can have absolute priority over the others. Any judgment is contingent. This way of reasoning seems fit to the spirit and flexibility of EA.

However, consider:

> *Three choices (no death)***: You have a sum that you would like to donate. With this sum, you can accomplish any of the following:
>
>> **A.** prevent 100 deaths-in-infancy, knowing that in all likelihood these 100 people will grow up to have consistently low income and poor health for their ~40-year-long-lives;
>>
>> **B.** provide consistent, full nutrition and health care to 100 people, such as that instead of growing up malnourished they spend their 40-years long lives relatively healthy;.
>>
>> **C*.** preventing one case of relatively mild non-fatal malaria for 1 million people, without having a significant impact on the rest of their lives.

What about this new version of *Three choices (no death)*? Would the size of the benefit given in **B** compensate for the number of individuals benefited in **C***? If it did, then also in this case **B** would be better than **C*** or **A**, and this can be supported by invoking *Balance*.

But consider:

> *Three choices (no death)****: You have a sum that you would like to donate. With this sum, you can accomplish any of the following:
>
>> **A.** prevent 100 deaths-in-infancy, knowing that in all likelihood these 100 people will grow up to have consistently low income and poor health for their ~40-year-long-lives;
>>
>> **B.** provide consistent, full nutrition and health care to 100 people, such as that instead of growing up malnourished they spend their 40-years long lives relatively healthy;.
>>
>> **C**.** preventing one case of relatively mild non-fatal malaria for 1 billion people, without having a significant impact on the rest of their lives.

Is **B** still better than **C****? I think the confidence of most supporters of EA would shake here. It seems plausible to claim that sacrificing nutrition and health care for 100 individuals may be worthwhile, if the gain is preventing even mild non-fatal malaria for 1/7th of the world population.

If so, however, there are two problems. First, **C**** is surely better than **C***, and the latter is surely better than **C.** It is better to cure from malaria 1 billion people than curing only 1 million, and it is better to cure 1 million than curing only 10,000. However, we granted that **B** is better than **C** and **C***. This amounts to endorse the following rankings:

1. **C**** is better than **C***, which is better than **C**;
2. **C**** is better than **B**;
3. **B** is better than **C** and **C***.

But these three rankings are inconsistent, and one of them should go. To endorse all of them means to breach transitivity. Is EA committed to rejecting transitivity? This view is very controversial, and if EA required it, this would need serious inquiry and defence.[26]

Second, it seems that applied to *Three choices (no death)**** EA—and *Balance*—suggest that the increase in the number of people benefited in **C**** would compensate the loss of benefits for the people in **B.** This reasoning can lead to the following conclusion:

> *Altruistic Repugnant Conclusion (ARC)*: Compared to a possible charitable action **a**, whereby a given amount of benefit is given to a great number of people, there must be some alternative charitable action **b** that, at the same cost, can give a much greater amount of very tiny benefits to a much larger number of people and that, if other things are equal, should be preferred to **a**.[27]

I think that an 'altruistic' version of the Repugnant Conclusion (RC) is strongly counterintuitive. Bringing about very tiny, but immensely diffused, benefits instead than less diffused, but more substantial, benefits is seriously wrong.

However, this conclusion is a necessary consequence of *Balance*. If *Balance* articulates AM, and the latter expresses the core tenet of EA, then EA leads to RC.

My provisional conclusion, then, is this: Two principles can articulate the imperative to maximize the impact of our charitable actions. The first principle—*Size*—exclusively focuses on the size of benefits, and gives absolute priority to the greatest benefit that can be brought about. The second principle—*Balance*—establishes that losses in the size of benefits can be compensated for by gains in their diffusion, and vice versa. Now, the first principle runs against many effective altruists' views, and it contrasts the char-

---

[26] On the rejection of transitivity, see (Temkin 2012). (Norcross 1997, 139-158) mounts a detailed critique against arguments for intransitivity. Nothing in this paper hinges on settling the disputes between supporters and deniers of intransitivity, as my main point is that effective altruism can *either* deny transitivity (and this has the problems emphasised by Norcross) or accepting it (and this has problems, too—one of the main ones being the Altruistic Repugnant Conclusion.)

[27] Cp. (Parfit 2016, 110).

ity evaluations they provided. The second principle either breaches transitivity or leads to RC. In the next section I shall explore some avenues to rescue the second principle from its counterintuitive implications. My conclusions will be rather pessimistic.

How (not) to deal with the altruistic repugnant conclusion

There are many ways to deal with RC scenarios. I shall briefly consider some of them.[28] My main aim is to show that these coping strategies cannot be employed to face and dismiss ARC.

Many have challenged the claim that the RC is *genuinely* repugnant, i.e. wrong. Some of these *denial* arguments challenge our *intuition of repugnance* on various grounds. Some attempts to challenge the repugnance of the RC go as follows.

- *Large numbers skepticism*: RC scenarios concern very large numbers, long durations, and lives barely worth living. Our common sense, and the intuitions it produces, are unfit to assess these unusual circumstances. Our minds are unable to conceive a very large amount of barely worth living lives. Accordingly, our sense of repugnance is either spurious or biased by the presentation of the thought experiment.[29]

- *Relative preferability*: Compared to non-existence, even existing in Parfit's Z-world may be preferred behind a veil of ignorance. This shows that after all living in, and bringing about a, Z-world is not obviously wrong.[30]

- *Near-zero lives*: Contrary to intuitions, many actual lives, even in developed countries, are barely worth living, even though they are not less good than that. This can be easily viewed if one thinks about how much rare real bliss is, and how much suffering we experience in our actual lives. Then, if Parfit's Z-world is a world full of people living a barely worth living life, then it is not very different from ours. Then, if the former is repugnant, the latter is such as well. But our actual world is far from being repugnant, at least considering the fact that people keeps on living in it. Or better, our actual world is repugnant because the lives of the many living in dire poverty and destitution are *not* worth living. Then, it seems obvious that the best world we could create is to be very crowded—there

---

[28] See (Arrhenius, Ryberg, and Tannsjo 2014; Parfit 2016).

[29] See (Broome 2004, 55–9; Tannsjo 2002; Norcross 1997, 147-52). An attempt to challenge this strategy is in (Pummer 2013).

[30] See (Tannsjo 2002, 343–4).

are many possibilities to create more people living barely worth living lives. On the other hand, we cannot imagine, or at least cannot easily identify with, lives very much better than ours—the lives in Parfit's A-world. This makes difficult to us any comparison between A-worlds and Z-worlds—the very comparison on which the sense of repugnance rests.[31]

- *Debunking*: The intuition of repugnance underlying the RC can be debunked, by showing that **a**. it goes against much more entrenched and irresistible intuitions (such as Pareto principles or the preferability of egalitarian optimal distributions, or transitivity), and that **b**. it is the outcome of biasing factors—such as egoism (the preference for Parfit's A-world derives from our egoistic preference for supposedly luxurious worlds).[32]

- *Common-sense morality and RC*: In many ordinary life moral choices, we accept RC. For instance, many legal systems admit high speed limits that correlate with automobile accidents-related deaths, for the sake of the tiny, but widely spread, convenience of not imposing these limits—in terms of efficiency, ease, and so on. Not imposing these limits does not seem wrong, or morally repugnant. As there is no morally relevant differences in the structure of these RC scenarios and the RC scenarios discussed in scholarship, it seems that the repugnance we felt in considering the latter is only due to irrelevant features of the descriptions of certain cases discussed in scholarship.[33]

None of these denial strategies can be used to dissolve the repugnance of ARC. Consider **C****. Imagine what it could be for 1 billion people being cured of a mild non-fatal malaria, while continuing to live the rest of their rather destitute lives. Or better, if we can imagine how living in dire poverty would be, we can also imagine how living in dire poverty, but with few days of malaria spared, could be. I assume we can imagine how living in dire poverty would be, because this is a pre-requisite of any assessment of charitable actions. If imagining life in poverty were impossible, EA would not get off the ground.

It might be objected that the problem is imagining how the benefits of sparing few days of malaria to 1 billion people are aggregated, and that our feeling or repugnance derives from our failure to see that this aggregated benefit is immensely larger than the benefit of increased educational opportunities for few people. But this begs the question. The

---

[31] See (Tannsjo 2002, 345–6, 349, 350–1).

[32] See (Huemer 2008).

[33] See (Norcross 1997, 158-67).

problem with ARC is exactly that to many of us the benefit of giving better education to 100 people seems worthy of giving up the tiny aggregated benefit of sparing few days of non-fatal malaria to 1 billion people.

Second, it is commonly assumed that there is an asymmetry between stopping an existing life and failing to bring about a new life. On the basis of this asymmetry, we can claim that, compared to non-existence, a life in dire poverty is not preferable. Even though no one living in extreme destitution would commit suicide, nor it would be reasonable for her to do so, it might be argued that mere non-existence would be better than a short life conducted in dire straits. Likewise, if lives in Parfit's Z-world are similar to the lives conducted by most very needy people in the most extreme poor developing countries, as it turns out to be in ARC, then our lives in developed countries are much better than this, notwithstanding the rarity of real blissed moments. Then, if a Z-world is a world where the actual direst poverty would be much more diffused, because of a larger population, it seems obvious that this world is not an optimal world, and the increase in aggregated benefit cannot compensate for the badness of individual poverty. Some of us have the intuition that a sacrifice of empowerment opportunities to tiny, basic, and widespread benefits is unsound. This intuition resists, as well as the intuitions that outcomes where anyone is better or there is more equality and no losses of utility are to be preferred. We are facing a conflict of intuitions.

Finally, it seems to me that the intuition underlying ARC is not distorted by biasing factors, such as egoism. To prefer a life with enhanced educational opportunities to a life in poverty with few days of malaria spared is not particularly egoistic. Indeed, it would qualify as a properly reasonable attitude. If the price to get better education and empowerment were having two days of a non-fatal fever, I think most of us would accept the deal.

Another denial strategy is as follows. Z-world is not a realistic option, as in the current conditions of human nature to move from the actual world, or from A-world, to a Z-world would have prices, and those prices (things such as increased conflicts, suffering of very large families, memories and regrets of the past world, and so on) would make a Z-world patently worse than an A-world, even taking into account aggregated benefits. Accordingly, RC is an unrealistic thought experiment. If it is repugnant, its repugnance has no practical implications for actual policies. And perhaps it is impossible to say whether it would be repugnant or not, as we cannot imagine the changes in human nature needed to make a Z-world a real option at all.[34]

---

[34] See (Fotion 1997).

This strategy cannot be applied to ARC, for the latter is a very realistic scenario. It is perfectly possible that current aid strategies will lead to an overcrowded world of illiterate people, living very poor lives, even though being saved from some days of avoidable morbidity. If this scenario appears to us repugnant, this is not due to its being unrealistic.

Finally, even assuming that our common sense morality finds no repugnance in some RC scenarios such as allowing some automobile accidents-related deaths for the sake of convenience, this is not a rebuttal of RC. For it may be argued that lack of repugnance in the more common RC scenarios is biased by our getting used to them, while the intuition to be taken into account is the one voicing repugnance when we face the 'uncommon' scenarios discussed in philosophical literature. The less common scenarios can be signs of a moral mistake in our dealings with the more common scenarios.

Other authors have attempted to block RC through discarding the principles from which the conclusion derives. The following *blocking arguments* have been proposed in recent scholarship. First, one may return to focussing exclusively on the benefit's size, thereby claiming that no loss in benefit's size can be compensated for by gains in the diffusion of benefits. However, this view, expressed by *Size*, is equally problematic, as in certain cases it seems that very diffused benefits can compensate for less diffused losses. For instance, in *Separation vs. vaccination*, it seems pretty obvious that it would be better preventing the deaths of hundreds of children, rather than separating a couple of conjoined twins.

Parfit himself has proposed a second blocking strategy. Recently, he has defended the idea that some comparisons of value are *imprecise*: "in most important cases relative value does not depend only on any […] single, measurable property", and this holds especially when "the relative goodness of many different acts or outcomes, such as ones that would greatly benefit a few people, or give lesser benefits to many others" are at stake. "Such imprecision", Parfit pointed out, is due to the fact that many valuable things are "qualitatively very different", i.e. they have many different valuable features. (Parfit 2016, 113) Imprecise value relations are not transitive: "when there is imprecision, the relation *not worse than [or not better than]* does not imply *at least as good as*, so one of two things can be in one way better and in other ways not worse without being better all things considered". (Parfit 2016, 115) Accordingly, a more crowded world with less happy lives (let's say a N-world, along the continuum between an A-world and a Z-world) is neither better nor worse than a less populated (let's say a G-world along the continuum from A- to Z-world). The more crowded N-world is *imprecisely equally as good as* the less populated G-world. If this is true, then we can block RC, because we may

claim that a Z-world is *immensely more* crowded, with lives immensely less happy, even though worth living, and for this reason it is worse than an A-world, even though less crowded worlds, such as the N-world, are not worse than less populated worlds, such as a G-world.[35]

Is this strategy useful in blocking ARC? It might be argued that funding larger educational opportunities for few people is imprecisely equally good than funding more basic interventions for much larger groups, and that donors are free to choose between these two alternatives. As Pummer suggested, there might be an upper set of effective charities, among which donors are free to choose whichever they prefer, for whatever reason.

This strategy may be sound. However, I think effective altruists should avoid committing themselves to imprecision. Parfit says that his view appeals to "zones in which different things would be imprecisely equally good, bounded by zones in which it would be indeterminate whether one of two things would be better, or these things would be imprecisely equally good." (Parfit 2016, 116). Applied to charities, this involves indeterminate boundaries between the set of best charities and the lower sets of suboptimal charities. Now, if the borders between the upper and the lower sets are indeterminate, the reasons to act may be as follows. There is a reason to choose charities in the upper set, but there is neither a reason *to choose* nor a reason *not to choose* from the immediately inferior sets, even though as we go down through the sets of less and less effective charities there will be a moment when a reason *not* to choose certain very suboptimal charities will crop up.

Now, this may be a nuanced and attractive picture. It might be said that a sustainable choice between more and less effective charities cannot be dramatic, and Parfit's imprecision helps to make it non-dramatic. But the spirit, if not the letter, of EA is betrayed by this picture. EA suggests that choosing the best charity is what we ought to do—it is our *full* duty. Any suboptimal choice is good enough, but it is far from realizing a fully moral life. Indeed, according to Pummer, it is obviously wrong to choose a suboptimal charity, when we could have chosen an optimal one at the same cost. This seems to be the obvious meaning of AM. If so, imprecision contrasts with the spirit and the meaning of EA.

In order to avoid ARC, EA cannot avail himself of any of the most promising strategies employed to deny or to block Parfit's RC. Then, either effective altruists should

---

[35] See (Arrhenius 2016, 170, 173; Parfit 2016, 120).

accept ARC, thereby swallowing its repugnance, or they should give up their core tenet, i.e. AM.

Conclusion. Values and empirics

Recently J. MacMahan contended that EA has often been the target of criticisms historically pressed against standard Utilitarianism. This paper makes no exception, since many used RC to argue against Total Utilitarianism.[36]

MacMahan also maintains that the empirical worries coming from developmental economists are worthy of more consideration. The general thrust of these worries is that aid can be harming, as well as beneficial, due to unintended consequences.[37] The standard response by effective altruists is that "the positive benefits of [aid] vastly outweigh the cases in which it had negative consequences." (Hillebrandt 2015) But this kind of response turns the empirical issue into an axiological question. We need an axiology to compare harms and benefits to different persons and of different size. Angus Deaton levelled the following criticism to EA:

> In today's Rwanda, President Paul Kagame has discovered how to use Singer's utilitarian calculus against his own people. By providing health care for Rwandan mothers and children, he has become one of the darlings of the industry and a favorite recipient of aid. Essentially, he is "farming" Rwandan children, allowing more of them to live in exchange for support for his undemocratic and oppressive rule. Large aid flows to Africa sometimes help the intended beneficiaries, but they also help create dictators and provide them with the means to insulate themselves from the needs and wishes of their people. (Deaton 2015)

The only possible response to this *reduction ad Hitlerum* of EA is to argue that, *ceteris paribus*, the fact that more Rwandan children live, even under a dictatorship, is better than paying the price in lives necessary to bring about democracy in Rwanda. This is a compelling conclusion, at least for me. But actual Rwanda can be a repugnant Z-world for many others. Accordingly, EA-supporters must face ARC, and must either find a principled defence of their preference for widely spread basic improvements over promotion of democratic regimes and other superior forms of development or give up this preference. The second option, i.e. giving up the preference for basic aid

---

[36] See (McMahan 2016).

[37] See (Wenar 2011).

over more sophisticate forms of aiding can change the substance of effective altruism, by making less committal and precise the imperative to choose the best—i.e. the most efficient—charities. This is the big challenge left for Singer and the other effective altruists.[38]

References

Arrhenius, Gustaf. 2009. 'One More Axiological Impossibility Theorem'. In *Logic, Ethics, and All That Jazz. Vol. 57*. Uppsala Philosophical Studies. Uppsala: Department of Philosophy.

———. 2016. 'Population Ethics and Different-Number-Based Imprecision'. *Theoria 82* (2): 166–81. https://doi.org/10.1111/theo.12094

Arrhenius, Gustaf, Jesper Ryberg, and Torbjorn Tannsjo. 2014. 'The Repugnant Conclusion'. In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2014.

Broad, Charles Dunbar. 1930. *Five Types of Ethical Theory*. London: Routledge & Kegan Paul.

Broome, John. 2004. *Weighing Lives*. 1 edition. Oxford ; New York: Clarendon Press. https://doi.org/10.1093/019924376X.001.0001

Brown, Campbell. 2006. 'Prioritarianism for Variable Populations'. *Philosophical Studies 134* (3): 325–61. Carey, Ryan, ed. 2015. *The Effective Altruism Handbook*. Oxford: Centre for Effective Altruism. https://doi.org/10.1007/s11098-005-0897-5

Deaton, Angus. 2015. 'The Logic of Effective Altruism'. *Boston Review*, July 1. http://bostonreview.net/forum/logic-effective-altruism/angus-deaton-response-effective-altruism

Fotion, Nick. 1997. 'Repugnant Thoughts about the Repugnant Conclusion Argument'. In Contingent Future Persons, edited by Nick Fotion and Jan C. Heller, 85–97. *Theology and Medicine 9*. Springer Netherlands. https://doi.org/10.1007/978-94-011-5566-3_8

Gabriel, Iason. 2016. 'Effective Altruism and Its Critics'. *Journal of Applied Philosophy*, February, 1–17. https://doi.org/10.1111/japp.12176

Hillebrandt, Hauke. 2015. 'Effective Altruism Continued: On Measuring Impact'. *Boston Review*, July 31. http://bostonreview.net/blog/hauke-hillebrandt-giving-what-we-can-effective-altruism-impact

---

Holtug, Nils. 2007. 'On Giving Priority to Possible Future People'. In Hommage a Wlodek: Philosophical Papers Dedicated to Wlodek Rabinowicz, 1–21. www.fil.lu.se/HommageaWlodek

Huemer, Michael. 2008. 'In Defence of Repugnance'. *Mind 117* (468): 899–933. https://doi.org/10.1093/mind/fzn079

MacAskill, William. 2015. *Doing Good Better: How Effective Altruism Can Help You Make a Difference*. New York, N.Y: Avery.

McMahan, Jeff. 2016. 'Philosophical Critiques of Effective Altruism'. *The Philosopher's Magazine 73*.

McTaggart, J.M.E. 1927. T*he Nature of Existenc*e. Cambridge: Cambridge University Press.

Narveson, J. 1967. 'Utilitarianism and New Generations'. *Mind 76*: 67–72. https://doi.org/10.1093/mind/lxxvi.301.62

Norcross, Alastair. 1997. 'Comparing Harms: Headaches and Human Lives'. *Philosophy & Public Affairs 26* (2): 135–67. https://doi.org/10.1111/j.1088-4963.1997.tb00079.x

Ord, Toby. 2013. 'The Moral Imperative toward Cost-Effectiveness in Global Health'. http://www.cgdev.org/publication/moral-imperative-toward-cost-effectiveness-global-health

Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Clarendon Press.

———. 2016. 'Can We Avoid the Repugnant Conclusion?' *Theoria 82* (2): 110–27. https://doi.org/10.1111/theo.12097

Persson, Ingmar. 2012. 'Prioritarianism and Welfare Reductions'. *Journal of Applied Philosophy 29* (4): 289–301. https://doi.org/10.1111/j.1468-5930.2012.00570.x

Pummer, Theron. 2013. 'Intuitions about Large Number Cases'. *Analysis 73* (1): 37–46. https://doi.org/10.1093/analys/ans134

———. 2016. 'Whether and Where to Give'. *Philosophy & Public Affairs 44* (1): 77–95. https://doi.org/10.1111/papa.12065

Sidgwick, Henry. 1907. *The Methods of Ethics*. 7th ed. London: MacMillan & Co.

Singer, Peter. 2015. *The Most Good You Can Do. How Effective Altruism Is Changing Ideas About Living Ethically*. New Haven: Yale University Press.

Stark, Andrew. 2008. 'Benefit versus Numbers versus Helping the Worst-off: An Alternative to the Prevalent Approach to the Just Distribution of Resources'. *Utilitas 20* (03): 356–82. https://doi.org/10.1017/S0953820808003208

Tannsjo, Torbjorn. 2002. 'Why We Ought to Accept the Repugnant Conclusion'. *Utilitas 14* (03): 339–59. https://doi.org/10.1017/S0953820800003642

Taurek, John M. 1977. 'Should the Numbers Count?' *Philosophy & Public Affairs* 6 (4): 293–316.

Wenar, Leif. 2011. 'Poverty Is No Pond: Challenges for the Affluent'. In *Giving Well. The Ethics of Philantropy*, 104–32. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199739073.003.0007