The ethical principles of effective altruism

Anthony Skelton

To cite this article: Anthony Skelton (2016) The ethical principles of effective altruism, Journal of Global Ethics, 12:2, 137-146, DOI: 10.1080/17449626.2016.1193552

To link to this article: http://dx.doi.org/10.1080/17449626.2016.1193552

Published online: 07 Aug 2016.
The ethical principles of effective altruism*

Anthony Skelton

Department of Philosophy, Western University, London, Ontario, Canada

Invited contribution received 4 May 2016

In The Methods of Ethics, Sidgwick remarks that

one who values conduct in proportion to its felicific consequences, will naturally set a higher estimate on effective beneficence in public affairs than on the purest manifestation of virtue in the details of private life … A sincere Utilitarian, therefore, is likely to be an eager politician, but on what principles his political action ought to be determined, it scarcely lies within the scope of this treatise to investigate. (1981 [1907], 495)

Sidgwick began the search for such principles in his The Elements of Politics (1919) and in his Practical Ethics (1898). He expressed plausible attitudes on how to do practical ethics in light of disagreement about ethical fundamentals (Skelton 2006). However, on other of the issues he discussed he put forward views that we now find quaint (or worse) (Schultz 2004) or, as in his discussion of the ethics of luxury, he barely scratched the surface, leaving much work undone. None of his immediate philosophical successors took up the challenge. Indeed, no one took it up in any serious way in philosophy until Peter Singer began writing in the 1970s. It is perhaps fair to say that he has done more to determine the principles on which the utilitarian’s social and political action, broadly speaking, ought to be based than any other philosopher since Sidgwick.

This is nowhere more obvious than in his work on the moral challenge of global poverty. Singer’s main focus in this domain has been the question of what the absolutely wealthy, those with a surfeit of what is necessary to meet their needs for food, shelter and basic medical care, owe the absolutely poor, those with too few resources to meet these same needs. Singer’s main contribution has been the articulation and defence of the philosophical conclusion that the absolutely wealthy ought to do more – nay, much more – for the absolutely impoverished (1972).

More recently, however, Singer has turned his attention away from strictly philosophical issues to more practical matters relating to the achievement of the aim that his argument obliges the absolutely wealthy to adopt (2009). In this capacity he has become a keen proponent of effective altruism, a new movement in the philanthropic community. This paper is an examination of the ethical principles of effective altruism as they are articulated in Singer’s The Most Good You Can Do. This paper discusses the nature and the plausibility of the principles that he thinks both guide and ought to guide effective altruists. It

*All unattributed page references are to Singer (2015).

© 2016 Informa UK Limited, trading as Taylor & Francis Group
argues in § II pace Singer that it is unclear that in charitable giving one ought always to aim to produce the most surplus benefit possible and in § III that there is a more attractive set of principles than the ones Singer outlines that ought to guide effective altruists in their philanthropic practices and in their lives more generally. These principles fit better with his practical ambitions and with plausible attitudes about the limits of beneficence.

Singer first discussed the issue of what the absolutely wealthy owe the absolutely poor in ‘Famine, Affluence, and Morality.’ He offers the following simple, but powerful argument:

P1. Death and suffering due to lack of food, shelter and basic medical care are very bad.

P2. If it is in one’s power to prevent something very bad from happening without thereby sacrificing anything of comparable moral significance, then one ought, morally, to do it.

P3. It is in one’s power to prevent death and suffering due to (among other things) lack of food, shelter, and basic medical care without thereby sacrificing anything of comparable moral significance by contributing to organizations working to alleviate absolute poverty.

C1. Therefore, one ought, morally, to contribute to organizations working to alleviate absolute poverty until doing so involves sacrificing something of comparable moral significance.

The second premiss of this argument is its most contentious. The main worry concerns its demands (see, e.g. McGinn 1999; Timmerman 2015). On some readings, it requires the absolutely wealthy to make quite significant sacrifices. Premiss two of the argument is not, however, its only contentious one. Although it has been the target of less scrutiny, premiss three is subject to the worry that, even if one is convinced of some reasonably stringent version of premiss two, it is far from clear that one possesses the power to prevent very bad things from happening by means of philanthropic or other forms of giving. It is often not clear to what extent such giving achieves the goal of preventing suffering and death caused by, among other things, absolute poverty. It is often difficult to gauge the effectiveness of one’s philanthropic contributions. Even where it is possible to tell, it is hard still to assess the degree to which contributions are efficacious. Moreover, Singer’s argument tells us that the absolutely wealthy ought to do more but not how best to do it.

Effective altruism is designed to speak to all of these issues. On Singer’s reading, effective altruists endorse a package of views. First, they believe, in line with his argument, that the absolutely wealthy ought to do a lot more to prevent suffering and death (especially that caused by absolute poverty). At various points, Singer suggests that the above argument is driving the various choices of the effective altruists he showcases (3, 13, 17, 19, 27, 34, 56, 67). His effective altruists have a specific way of understanding the argument (especially its second premiss), and one aim of The Most Good You Can Do is to provide an account of it.

Second, effective altruists hold that, if one has decided to engage in philanthropy, then one ought to use ‘reason and evidence’ to determine which organizations or intervention to support (xi; also 4). The evidence on which effective altruists typically rely is taken from existing academic research (preferably using randomized controlled trials) regarding the effectiveness of a charity or (more likely) one of its interventions (e.g. deworming, vaccinating and distributing insecticide-treated bed nets). Effective altruists rely in this regard on meta-charities, for example, GiveWell, that perform the task of evaluating and recommending charities and their interventions (149–164).
Third, effective altruists hold that one ought to select, from those available, the philanthropic organization or intervention that is likely to produce the most surplus benefit possible in the situation (see, e.g. 65, 78, 79, 102, 123) in terms of reducing suffering and promoting happiness or in terms of preserving or saving lives (7, 9; see also MacAskill 2015). This is why effective altruists typically focus on charities or interventions operating in the developing world, where the expectation is that contributions will produce the greatest possible reduction of suffering and/or death (113), though some focus on the suffering of non-human animals (137–147) or the reduction of existential risk (165–178). Proponents of effective altruism, then, appear to address the nature and implications of Singer’s argument and the difficulties of realizing its conclusion in practice.

Singer contrasts effective altruists with so-called warm glow donors (5, 90), who give little thought to the efficacy of philanthropic giving. But this is only one way to deviate from Singer’s effective altruism. It is possible to endorse only some aspects of it. One might, for example, argue that while we have no obligation to contribute to charitable organizations, when we decide to do so we ought to aim to produce the most surplus good we can. One could, then, endorse this aspect of effective altruism, but reject Singer’s argument that we ought to do more for those in need. (I assume here that it is uncontroversial that we ought to rely on reasoning and evidence in directing philanthropy.)

One might, on the other hand, accept that the absolutely wealthy ought to do a lot more for the absolutely poor without accepting the claim that the only thing relevant to distributing philanthropy is quantity of surplus good produced. One might think this has some role to play in deciding what to do but that other considerations matter to figuring out how to contribute to charity. This will be my focus in the next section.

II

Singer takes it almost for granted that in thinking about which charity or intervention to support one ought to rely only on the quantity of suffering averted or the number of lives saved. It is, of course, relatively easy to agree that in the prominent cases that Singer notes, we ought, if we choose to give, do what is best quantitatively speaking. One ought to give to charities that work to encourage people to adopt vegetarianism – for example, Vegan Outreach – rather than to animal shelters (137ff.). One ought to donate to a charity that performs surgeries to restore the sight of people with trachoma in developing countries rather than to museums that seek only to improve their visitor’s aesthetic experiences (118ff.). One ought to donate to charities working to alleviate absolute poverty in the developing world rather than to those helping the poor in the developed world (107ff.). One might think that the reason it is hard to disagree about what we ought to do in these cases is that in each case the plausible choice yields significantly more surplus benefit.

Singer is aware that not all agree that quantitative considerations are decisive. He notes, in passing, that John Taurek rejects the claim that we ought to do the best we can in trade-off situations, cases in which you can save from the same loss one or other, but not both, of two different but non-overlapping groups (196n6). Taurek would seem to deny that we ought always to select the charity that saves the most lives or that minimizes the greatest
suffering. He argues that in trade-off cases the number of individuals is not ethically relevant. In cases where you know and like the members of a smaller group or some single individual you are permitted to spare them rather than some larger group of strangers the same loss. In cases where all are strangers to you, you ought to flip a coin to determine which way to decide, for this expresses ‘equal concern and respect for each person’ (Taurek 1977, 303).

Taurek might, however, agree with Singer about the cases mentioned above. Taurek thinks that in cases where the loss to one is very small and the gain to another is quite big one ought to prevent the greater loss. He might agree with Singer that we ought to give to the charity that funds surgeries to restore the sight of those with trachoma and to the charity that alleviates poverty in the developing world. He might think this because he thinks that even those who lose out in these cases ‘should prefer’ that the one threatened with the greater loss should be ‘spared his loss’ at their expense (Taurek 1977, 302).

There is, then, some agreement between Taurek and Singer. Taurek would disagree, however, with Singer that if you are choosing between giving some significant sum of money to a charity that will save the lives of hundreds of children in the developing world by vaccinating them versus giving to a charity that will fund only an operation to separate conjoined twins who will otherwise die, you ought to give to the former (110). I agree with Singer that Taurek’s view is implausible in this case. But this is not the place to establish this. The important point is that one does not need to be Taurek to raise doubts about Singer’s claim that we ought always to do the best in every case of giving. One can appeal to Taurek-like ideas to do the same work.

Taurek appeals to knowing and liking someone in his account. It might be that knowing and liking is relevant to how we ought to donate. Singer himself has a relationship with Oxfam, to which he gives, even though it is not among the charities ranked highly in terms of effectiveness. One might argue that at least in cases where one is choosing between very effective charities one is permitted to let one’s relationship either with a charity or with its recipients play a role in which charity to donate to. It might, then, be permissible to donate to a slightly less effective charity because one knows and likes that charity than to a more effective one that one does not know and like. It does not seem wrong to forgo a small benefit for this reason. This might occur in cases where a charity works on a cause to which one is particularly attached.

Effective altruist William MacAskill says it would be unfair to let the fact that you have some relationship with a charity or with some potential recipients of charity to affect the pattern of your donations. He says

If I were to give to the Fistula Foundation rather than to the charities I thought were more effective, I would be privileging the needs of some people over others merely because I happened to know them. That would be unfair to those I could have helped more. If I’d visited some other shelter in Ethiopia, or in any other country, I would have had a different set of personal connections. It was arbitrary that I’d seen this problem close up rather than any of the other problems in the world. (MacAskill 2015, 41–42)

But it is far from clear that this is unfair. MacAskill says that such a decision is based on arbitrary considerations. But in some sense all close personal relationships are arbitrary. We might have loved someone else or pursued different friendships. It does not follow that it is impermissible to give them slightly more weight than strangers in one’s reasoning about
what to do, especially when the relationship involves a shared history. We might think in
addition that even if there is some unfairness here, that it is negligible and so outweighed.
In choosing between highly effective charities, it at least seems permissible to give to a
slightly less effective one to which you have some kind of connection, and some effective
altruists that Singer discusses actually do this (36).

There may be other ways in which fairness more clearly matters. Taurek, as noted,
appeals to fairness in his account. But we do not need to hold his view of fairness to
think that it matters in philanthropy. There are other ways that fairness might figure
into one’s reasoning about philanthropy that conflicts with the view that Singer (and pre-
sumably MacAskill) advocates. First, it might, for example, be unfair to give to a charity that
will perform operations to correct a thousand obstetric fistulas rather than to a charity that
will prevent a smaller number of people from dying from starvation even though giving to
the former produces slightly more surplus benefit. It might be unfair to contribute to a
charity that would provide robust educational opportunities to some very large number
of children living in the developing world who are not in need of food and medicine
rather than to a charity that gives food and medicine to a smaller number of very
poorly off children who are malnourished and who are living in a different part of the
developing world, even though giving to the former would produce on balance slightly
more surplus benefit. One might argue here that it would be unfair to choose the
better off in these cases, and that it would not be wrong on this ground to donate to
the latter charities. It might be unfair not to give priority to the least well off, on the
assumption that the charities in question are highly but not equally effective. Indeed, it
might be patently unfair to rule out helping the least well off in these kinds of cases at
least some of the time out of the concern for fairness. It might be fairer to give them in
cases of this sort at least some chance of being saved, though perhaps not the chance
that Taurek advocates. These are views that ought to appeal to some of Singer’s effective
altruists, since even on his reckoning some hold that equality matters for its own sake (8–
9). Equality or fairness seems to matter here to what we ought to do.

These are some considerations that are meant to count against the claim that only
quantitative considerations matter in considering how to distribute one’s charity. These
considerations are not meant to impugn appeal to quantitative concerns entirely. The
idea is that these are not the only considerations. It might be better to hold, then, the
view that, other things being equal, we ought to promote in giving the greatest quantity
of good. This is more difficult to deny and certainly captures many of Singer’s high
profile examples.

III

According to Singer, effective altruism involves not only subscription to the claim that if we
decide to contribute to charity, we ought to do the most good we can, but to the claim
that we ought to promote surplus aggregate good more generally in our lives. A sizable
portion of The Most Good You Can Do is devoted to showcasing the various different
styles of life that Singer thinks people can pursue while functioning as this kind of effective
altruist (23–72). One might choose to live on very little of one’s existing income giving a
sizable proportion of one’s surplus away to charity, or one might take a job in which one
can earn a large salary, for example, by working on Wall Street, in order to live well and
contribute a lot to charity, or one might work in an organization in which one can have a greater than ordinary impact on the fate of the most needy or one can, together with donating to charity, give parts of one’s body away (e.g. one’s kidney). These are distinct ways to live as an effective altruist, though all share the aim of contributing to preventing death and suffering.

According to Singer, effective altruism of this kind is based on a set of abstract principles that he finds in Sidgwick, on which

The good of any one individual is of no more importance, from the point of view (if I may say so) of the Universe, than the good of any other; unless, that is, there are special grounds for believing that more good is likely to be realized in the one case than in the other.

And it is evident to me that as a rational being I am bound to aim at good generally, – so far as it is attainable by my efforts, – not merely at a particular part of it. (1981 [1907], 382)

Singer thinks these are the rational principles explicating why effective altruists adopt the various strategies and lifestyles they do (80, 81–82, 87). These imply for Singer that in reasoning about what to do we are required to produce the impartial best (80, 102). He does not claim that to be an effective altruist you must be a utilitarian (53, 79). Effective altruists might, then, arrive at their conclusions by means of other principles. He says only that these are the principles that ‘would’ guide effective altruists (80). But he clearly favours them over others. (For a defence of these principles, see de Lazari-Radek and Singer [2014]. As Singer does here, they construe them as all things considered principles.)

However, it is unclear, first, that these are the principles that explain best the behaviour of effective altruists and, second, that these principles are normatively the most desirable. Singer thinks that effective altruists are motivated by the idea that we should maximize value in our lives more generally (65, 94, 102). That is, they embrace a specifically utilitarian version of his argument’s second premiss, implying that the absolutely wealthy donate until the loss to them of doing so is as great as the benefit to the absolutely poor.

One difficulty with advocating these principles as normatively most desirable is that doing so fits uncomfortably with Singer’s previous work on the issue of the demands of his argument’s premiss two in which he appears to back away from the more stringent versions of the premiss. The utilitarian version, as he acknowledges, is one among others. In his discussions of the practical implications of his argument in both Practical Ethics (1993) and The Life You Can Save he advocates for more modest readings of the argument’s second premiss and so for more modest practical implications. In Practical Ethics he argues that the absolutely wealthy ought to give 10% of their income to the absolutely poor. In The Life You Can Save he argues for a sliding scale so that the more you make the more you are required to give. He imposes this on the top 10% of earners. For one earning less than US$105,000 the suggestion is that one contribute 1% of one’s pre-tax income to charity.

The reason that he does this is quite plausible: advocating more modest norms for giving makes it on balance more likely that more people will contribute to philanthropic organizations or interventions aimed at the prevention of death and suffering due to extreme poverty, among other ills. The idea that effective altruism requires subscription to the Sidgwickian principles (as Singer construes them) may strike many as running
counter to this ambition and will strike many as an unattractive ideal for the purposes of practice and of attracting individuals to the cause. If you have to buy into Sidgwick, many might reason, that is enough to dampen enthusiasm, for doing so could commit you to serious (uncompensated) sacrifices.

In reply, Singer might suggest that he has changed his mind about the effects of advocating this (more utilitarian) ideal over other, more modest ones. It might be that this is better than rivals in terms of motivating people to prevent loss of life and suffering. If he is right, it has moved some individuals to do some highly effective things in the service of combating suffering and death caused by crushing poverty, among other things.

The difficulty is that to establish this we would need evidence of the sort that effective altruists seek about the various interventions or charities that they support. This is not, however, in our possession.

Of course, Singer might argue that the examples of effective altruists that he discusses will lead us to think the principles more attractive because, as he describes them, their practical manifestations are attractive. He is keen to show that ‘typical effective altruists leave themselves time and resources to relax and do what they want’ (8). The idea is, then, that we can accept the principles and avoid the demandingness worry. If this is so, we should relax and learn to love the Sidgwickian principles. The difficulty with this is that while it may be true that this is what effective altruists do, it is not clear that this is consistent with the Sidgwickian principles. Singer has not, at any rate, shown that these principles justify these permissions. He might have shown that the altruists promote surplus good but he is far from showing that what they do is produce the most surplus good they can.

There is, anyway, a better set of norms to both explain why certain effective altruists modify their behaviour in the ways that they do and to recommend as normatively desirable. These seem to provide a more productive image for the purpose of attracting individuals to the effective altruist fold and serve as a more attractive normative ideal.

It is pretty clear from at least some of the descriptions that many of the effective altruists gave more weight to their own interests than the interests of others. True, they did not seem to adopt the more conventional balance between one’s own and others’ interests, but they did nonetheless give some greater weight to their own interests.

Consider as evidence the following activities by effective altruists. One was ‘writing his thesis, studying a language, practicing his guitar, and writing children’s books, a novel for adults, short stories, and a translation of Plato’s Cratylus that preserves the puns most translators consider to be untranslatable’ (34). Another wrote novels and followed various political projects to which she felt a special connection (35–36). Yet another ‘plays soccer with friends, enjoys listening to music, and goes cycling most weekends’ (45). Others placed limits on what they were willing to do to promote the aims and ambitions of effective altruism and so they had children and gave greater weight to the interests of their friends and family and kept their spare organs.

Let me be clear: this is not meant as a criticism of these individuals. The point is that like most of us, effective altruists accord more weight to their own interests than to the interests of strangers in their deliberations about what to do. This does not mean that the approach is in some way defective. The issue is whether the most plausible account of their behaviour involves reference to Sidgwick’s or cognate principles.
It does not. Indeed, the behaviour of the effective altruists suggests that they accept in addition to the principle of beneficence a principle of prudence, a principle to which Sidgwick himself has an affinity. He accepts that there are principles other than the ones that Singer mentions. Sidgwick thinks that the fact that you and I are distinct individuals gives each of us a reason to be more concerned with the ‘quality’ of our own existence than with the ‘quality of the existence of other individuals’ (1981 [1907], 498). What separates effective altruists from most of us is that they have a particularly strong view of beneficence that sets limits on the pursuit of one’s prudential aims. But all the same they think that beneficence is at least to some extent limited by prudence, and in particular by what contributes to life satisfaction and/or to a meaningful life (viii, 94; 97–104).

What effective altruists appear to have in common is a particular reading of the second premiss of Singer’s argument. They are committed to preventing suffering and loss of life, and they have a distinct view of what is comparable in terms of moral importance to it. They believe that we ought to give or contribute until doing so compromises what makes life satisfying or meaningful. This seems indicated by the fact that all of the effective altruists that Singer discusses found their lives satisfying or meaningful in some respect (97–104).

These are, it seems, the main principles underlying effective altruism. The mixture and balance of principles is normatively attractive. This is a more plausible ideal on which to rest effective altruism, giving weight to self- and other-regarding reasons. It involves accepting pro tanto rather than all things considered versions of Sidgwick’s principles. These principles are combined with a set of other virtues that effective altruists rightly exhibit. First, a heightened receptivity to evidence and facts in thinking about how best to achieve philanthropic ends. Second, effective altruists have, perhaps unwittingly, hit on a sensible way of to some extent reconciling prudence and beneficence; they appear to have stumbled on what psychologists have been reporting, that above a certain threshold increases in income and wealth produce fewer and less long-lasting increases in life satisfaction. Effective altruists have discovered that for happiness, for life satisfaction, one needs much less materially than our cultural norms suggest.

Singer could reply that the tendency to give oneself greater weight might just be instrumental to happiness. A person could be more effective if they are happy or living meaningfully. But this is not clear since Singer allows that the effective altruists have values other than happiness (and suffering) (7; though cf. 50, 146). These could be doing the work and there is nothing against saying that they do and ought to.

Finally, finding a place for prudence or self-regarding reasons makes it easier for Singer to reply to an argument that he considers from Williams (1981). Williams contends that the trouble with utilitarianism and with Sidgwick’s principles in particular is their request that we detach from our personal point of view, ‘from more personal considerations that otherwise dominate the way in which we live’, and evaluate all our loyalties, loves and dispositions to determine how felicific they are (85). He argues that this is impossible. Singer argues otherwise.

But it is not clear that Singer’s argument is entirely successful. He concedes that the detachment is not ‘total’ (85). There is good reason for this: it is not clear what it would mean to evaluate one’s dispositions and loyalties from some completely detached point of view. If you accept the above view about what does and ought to undergird the effective altruist programme, you have a reply to Williams: the personal point of view of
prudence is the thing from which one views what matters to them. But effective altruists do not leave things here. They limit what they do in pursuit of what they care about by means of a principle of beneficence or the version of premiss two of Singer’s argument that they accept. To do this there is no need to detach completely from one’s dispositions, loyalties and other ground projects and to then figure out what to do. One just has to limit the extent to which one pursues one’s own ends by means of other, compelling other-regarding principles.

IV

This paper discussed the nature and the plausibility of the ethical principles that Singer thinks both guide and ought to guide the behaviour of effective altruists. It argued in § II that it is unclear that in charitable giving one ought always to aim to produce the most surplus benefit possible and in § III that there is a more attractive set of principles than the ones Singer outlines that ought to guide effective altruists in their philanthropic practices and in their lives more generally. These principles fit more comfortably with some of Singer’s practical ambitions and with plausible attitudes about the limits of beneficence.2

Notes

1. It is unclear that Taurek would deny this in the case of non-human animals. It is in fact unclear what his view entails in these cases. Perhaps here the numbers count, in which case he may agree with Singer.
2. Thanks to audiences at The University of Western Ontario, Stockholm University and Oakland University and to Lisa Forsberg, Isra Black, Anne Skelton and (especially) Henrik Ahlenius for helpful feedback on previous versions of this paper.

Disclosure statement

No potential conflict of interest was reported by the author.

Notes on contributor

Anthony Skelton is Associate Professor in the Department of Philosophy at the University of Western Ontario and Associate Director of the Rotman Institute of Philosophy. He researches in the areas of normative ethics and the history of ethics. He has published articles in, among other journals, Utilitas, Canadian Journal of Philosophy, Journal of the History of Philosophy and Ethics. He is the co-editor of Bioethics in Canada (OUP, 2013). He is currently working on a book entitled Henry Sidgwick and the Conflicts of Ethics. This is his first publication in the Journal of Global Ethics.

References


