# Effective Altruism and its Critics

IASON GABRIEL

ABSTRACT *Effective altruism is a philosophy and a social movement that aims to revolutionise the way we do philanthropy. It encourages individuals to do as much good as possible, typically by contributing money to the best-performing aid and development organisations. Surprisingly, this approach has met with considerable resistance among activists and aid providers who argue that effective altruism is insensitive to justice insofar as it overlooks the value of equality, urgency and rights. They also hold that the movement suffers from methodological bias, reaching mistaken conclusions about how best to act for that reason. Finally, concerns have been raised about the ability of effective altruism to achieve systemic change. This article weighs the force of each objection in turn, and looks at responses to the challenge they pose.*

The past ten years have seen the growth of a new social movement and approach to philanthropy called effective altruism. Effective altruism encourages individuals to make altruism a central part of their lives and do as much good as possible, typically by contributing a fixed percentage of one's income to the world's best-performing aid and development organisations. Beyond this, effective altruists are committed to the idea that scientific analysis and careful reasoning can help us identify which course of action is morally best.

To date, the movement has had a number of successes. It has led to the establishment of 'meta-charities', such as GiveWell and Giving What We Can, which provide the public with reliable information about how charities perform. It has also drawn attention to the vast difference in impact between charities, creating an incentive for them to demonstrate effectiveness in ways that can be tested. Finally, it has drawn public attention to what is perhaps the most important moral message of our time: individuals living in affluent societies have the power to do an incredible amount of good if they only stop and think for a moment about how best to achieve this aim. Nonetheless, there is reason to believe that effective altruism could turn out to have an even greater significance.

At present, private philanthropy in the United States exceeds $200 billion per year.[1] This figure is considerably *higher* than the total amount of global development aid combined.[2] At the same time, only a tiny percentage of the money donated by individuals makes its way to the world's poorest people for whom it would often do the most good.[3] Bad quality information, carelessness, and parochialism are major barriers to effective giving. Yet, considering how much is given away, there may be an opportunity for moral leverage: if effective altruists can persuade people to give more, or even just to give a bit more effectively, they could have a powerful and transformative effect on the world.

In spite of its promise, this approach to doing good has met with resistance among aid practitioners and activists committed to the cause of social justice at a global level. Taken at face value, their reaction is surprising. After all, both effective altruists and the humanitarian community work to improve the lives of the world's poorest people. I believe that the tension can be explained by competition within the charity sector and by questions about effective altruism that remain unresolved. To begin with, effective altruists have thrown down the gauntlet to nongovernmental organisations (NGOs) by asking them to demonstrate that their work is effective.[4] By challenging the reputation of these organisations, many of which have long track records of working to address poverty, the movement has provoked a negative reaction from traditional aid providers. Meanwhile, those working to combat global poverty also harbour deeper concerns about the normative and methodological assumptions of effective altruism. If effective altruists are mistaken about what it means to do good, or about how impact can be measured, then the movement risks doing more harm then good by diverting resources away from other important goals and objectives.

This article looks at these normative and methodological concerns in greater detail and asks how effective altruists can best respond to them. The first section considers the claim that effective altruism is insensitive to justice because it overlooks the value of equality, discriminates against the worst-off, and ignores rights. The second section focuses on objections to effective altruists' methodology for evaluating organisations. Here the concern is that observation bias, quantification bias, and an unduly instrumental approach to evaluation lead the movement astray. The final section looks at more specific concerns about the impact of giving to effective altruists' recommended charities and about the ability of the movement to achieve systemic change. These objections are so cogent that it makes sense to ask whether effective altruism is effective after all. However, I argue that many of these problems can be addressed and that the movement as a whole is worthy of our support.

## Is Effective Altruism Unjust?

There are thick and thin versions of effective altruism. The thin version of the doctrine holds that 'we should do the most good we can' and that this involves using a substantial amount of our spare resources to make the world a better place.[5] It is compatible with a wide range of moral theories and remains noncommittal both about the nature of the good and about the individual's relationship to it. This version is, however, committed to the methodological claim that through careful analysis of evidence it is possible to provide sound general advice about how individuals can have a positive impact.

The thick version of effective altruism makes a number of further assumptions. First, it adopts a largely welfarist understanding of value. According to this view, good states of affairs are those in which suffering is reduced and premature loss of life averted.[6] Second, it is broadly consequentialist, maintaining that we should do whatever maximises the sum of individual welfare at all times. As GiveWell cofounder Holden Karnofsky writes, effective altruism 'is a way of thinking about morality that insists on the maximisation of good accomplished and not just satisficing of rules and guidelines'.[7] Third, the movement takes 'a scientific approach to doing good', which

means using tools such as cost-effectiveness analysis and randomisation to help quantify and compare the impact of different interventions.[8]

This article focuses on the thick version of effective altruism and demonstrates its weaknesses. Not everyone who identifies with the movement shares each individual belief but they explain many of its judgments and capture much of what makes it unique. From a normative standpoint, the thick version of effective altruism remains vulnerable to the charge that it overlooks considerations of justice and therefore generates radically incomplete conclusions about how best to act. Indeed, it accords no intrinsic value to equality in the distribution of goods, to the contention that people with urgent claims deserve special attention, or to the idea that human beings are the bearers of moral rights that resist aggregation.

### Equality

Effective altruists recognise that equality is instrumentally important. They know that an unequal distribution of resources can have negative consequences including resentment, domination, and the erosion of public goods. They also recognise that inequality often gives rise to new opportunities to do good. Due to the fact that money tends to have declining marginal utility, there is strong theoretical reason to believe that more can be achieved by focusing on the global poor than by focusing on people who are already well off.[9] Given the extent of global inequality, William MacAskill, one of the founders of the effective altruist movement, calculates that those living in rich countries can reasonably expect to do *one hundred times* more for someone who is living in extreme poverty than they could by spending the money on themselves.[10] In an effort to avoid pernicious externalities and to make resources go further, effective altruists often support policies that are broadly egalitarian in outcome.

At the same time, most effective altruists do not accord any intrinsic value to equality. As Peter Singer writes, 'They tend to view values like justice, freedom, equality, and knowledge not as good in themselves but good because of the positive effect they have on social welfare.'[11] This brings the approach into conflict with a central belief many people have about justice: it is bad, because unjust or unfair, that some are worse off than others through no fault or choice of their own.[12] Indeed, many people believe it's unfair for a morally arbitrary factor, such as skin colour or place of birth, to influence a person's life chances even if this state of affairs produces the greatest sum of overall wellbeing.

To illustrate how this disagreement plays out in practice, consider the following example:

> *Two villages*. There are two villages, each in a different country. Both stand in need of assistance but they are unaware of each other and never interact. As a donor, you must choose between financing one of two programs. The first allocates an equal amount of money to each community and achieves substantial overall benefit. The second allocates all of the money to one village and none to the other. By concentrating resources it achieves a marginally greater gain in overall welfare than the first project.

In this case, the unequal distribution of resources would not lead one village to resent or dominate the other, and since they do not maintain any public goods collectively,

there need be no fear of erosion. Therefore, exponents of the thick version of effective altruism see no reason to prefer the equitable program to the one that brings about the greatest overall welfare. However, those who have the egalitarian intuition about fairness will insist that there is a further reason to prefer the first program, namely that it reduces the influence of morally arbitrary factors on people's lives. According to this view of morality, it is unfair for one community to receive all the benefit and the other none at all, particularly since there is only a tiny difference in the overall welfare returns of each program.

### Priority

Although the declining marginal utility of money encourages effective altruists to focus on helping poor people, in practice, the most cost-effective interventions are unlikely to be those that focus on the poorest of the poor. This is because the ultra-poor suffer a composite of afflictions. Among other things, they tend to lack important capabilities and skills, to be victims of social marginalisation and geographical remoteness, and to suffer from chronic illness or disability.[13] This makes them some of the hardest people to help. Successful interventions need to be targeted and multidimensional, which increase their cost.[14] Furthermore, those who suffer from physical disabilities are often less efficient at converting resources into welfare. If our concern is only to achieve the largest overall gain in wellbeing, this means they will fare badly.

  Nonetheless, there is a strong case for holding that those at the bottom – those who have least or suffer most – should be prioritised because their level of need is the greatest. For those who endorse this priority principle, the worse off people are in absolute terms, the more important it is to offer them assistance.[15] The idea that urgent claims should be met first can be understood as a basic component of morality. Alternatively, it may arise from the need to justify coercion by the state. On the latter view, what matters for a state's legitimacy is that its political and economic intuitions can be justified to the worst-off people in society.[16] If this condition is not met, then those at the bottom have reason to prefer a different set of arrangements (that no one else has equal reason to object to) and are being dominated if their voices are not heard.

  To see how these considerations play out in the present context, consider the following example:

> *Ultra-poverty*. There are a large number of people living in extreme poverty. Within this group, some are worse off than others. As a donor, you must choose between financing one of two development interventions. The first program focuses on those who will benefit the most. It targets literate men in urban areas and has considerable success in sustainably lifting them out of poverty. The second program focuses on those who are most in need. It works primarily with illiterate widows and disabled people in rural areas. It also has some success in lifting these people out of poverty but is less successful at raising overall welfare.

Recent studies have found that donors face this dilemma when deciding how to target micro-finance initiatives.[17] But for effective altruists the choice is fairly straightforward. They favour projects that focus on fewer people when doing so delivers a greater gain

in overall welfare. But in *Ultra-poverty* this is not the case. Therefore, they would choose the program that supports literate men. When this pattern of reasoning is iterated many times, it leads to the systematic neglect of those at the very bottom, something that strikes many people as unjust.

### Rights

A right is a justified claim to some form of treatment by a person or institution that resists simple aggregation in moral reasoning. The strongest moral rights protect vital interests, functioning as trumps against the claim that a person's basic needs can be sacrificed for the sake of the greater good.[18] In this way, rights build upon the notion that it is wrong to use an individual to achieve some outcome in a way that 'does not sufficiently respect and take account of the fact that he is a separate person, that his is the only life he has'.[19]

In a world riven with uncertainty, the best way to promote total welfare is often to respect people's rights. This is because attempts to achieve the greatest good directly, by maximising expected value, are unreliable. As a result, it may be better for people and institutions to avoid imposing certain harms on others altogether. Yet effective altruists tend to take a different approach, as demonstrated by the case of sweatshop labour. Consider the following scenario:

> *Sweatshop.* The country you are working in has seen a rapid expansion of dangerous and poorly regulated factory work in recent years. This trend has helped lift a large number of people out of poverty but has also led to an increase in workplace fatalities. As a donor, you are approached by a group of NGOs campaigning for better working conditions. There is reason to believe that they can persuade the government to introduce new legislation if they have your financial backing. These laws would regulate the industry but reduce the number of opportunities for employment in the country as a whole.

Cast in these simple terms, I believe that most effective altruists would refuse to support the campaign even if there were no other available opportunities to do good. Following MacAskill's lead, they would argue that poor people are better off in a world where sweatshops exist than one in which they do not, and that a policy of non-interference is to be preferred for that reason.[20] But for those who believe that human beings have rights, including a right to adequate protection from dangerous working conditions, things are not so simple. This is because the case for sweatshop labour rests upon a morally impermissible trade-off: it improves the economic welfare of many by providing them with a source of income but at the expense of harming those who are severely injured, maimed, and killed in workplace accidents. It is not fair, one might think, that these people should pay the price in order to make others better off.[21]

Faced with these criticisms, effective altruists may respond in a number of ways. To begin with, they could hold the line and insist that equality, priority, and rights lack independent weight in moral evaluation. Alternatively, they could argue that these values come with scope-restrictions: they are important things for sovereign states to consider but individuals engaged in philanthropic action need not do so. Yet, these

replies are problematic for two reasons. First, they do not take seriously the possibility of moral error. Consequentialism may not be the correct account of personal morality, so conclusions that assume its truth are worryingly contingent. Second, they risk alienating potential supporters who want to do good effectively but also endorse fairness-based considerations. In the light of this, there is both epistemic and practical reason for effective altruists to look at other ways to proceed.

Fortunately, there are a number of practical steps they could take to make the movement more inclusive. For example, they could treat equality and priority as tie-breaking principles in cases where the outcome of welfare-based analysis is indeterminate.[22] Given that their first-order estimations of impact invariably contain some margin for error, and that there is instrumental reason to prefer more equal distributions of resources, there is little reason not to go down this path. In a similar vein, effective altruists could conduct a rights audit of the interventions they recommend, to see if the preferred policies impose serious harms on individuals or groups. Here, too, there is general reason to avoid these outcomes. Welfare-based calculations that require these trade-offs should therefore be subject to additional scrutiny.

On a deeper level, the movement could address the problem posed by competing values by providing advice that is sensitive to different value systems, enabling advisees' values to bear upon personal cause selection. For those who adhere to a consequentialist worldview, existing recommendations would remain in place. While for others, who have egalitarian or prioritarian intuitions, analysis could be adjusted to include independent distribution weightings. Meanwhile, for those who accord special weight to rights, effective altruists could identify a subset of cases that do not entail objectionable trade-offs.[23] This would prevent the movement from getting bogged down in open-ended debates about value or from committing itself to substantive claims that are hard to justify in public life.[24] It would also allow effective altruism to accommodate the viewpoint of those who believe that an outcome may be better in one way because it is fairer. Finally, in cases where these different streams of analysis overlap, disjunctive analysis – applying effective altruism's methodology along diverse dimensions of value – would help to identify a set of priorities that are robust across different moral theories.[25]

## Is Effective Altruism Blind?

Effective altruists tend to rely upon a set of heuristics in order identify the areas in which they can have greatest impact. To begin with, they assess the scale of global problems by looking at major reports and publications in that area. If a problem is serious enough to warrant further investigation, they then ask whether there are proven interventions to remedy the situation. Finally, if a cause is tractable, they ask whether it is also neglected, something that determines whether new funds are likely to make a difference in that area.[26] With this background analysis in place, meta-charities such as GiveWell and Giving What We Can audit individual organisations working in priority areas to ensure that they deliver goods or services in an efficient, transparent, and competent way.

This approach has a number of advantages. It allows the movement to process a large volume of information in a systematic way and generates confidence in the

charities that they recommend. The more important methodological questions, though, concern which evidence should be admitted when evaluating interventions, the metric used when calculating impact, and how comparisons across different issue areas should be made. Given their choices in these areas, effective altruists are vulnerable to the charge of observational bias, quantification bias, and instrumentalism. Taken together, these effects add up to what some critics see as methodological blindness.[27]

### Observational Bias

Given that there are many areas in which we might hope to do good, it makes sense to focus first on fields like public health about which there is high-quality information: in these areas it will be easier to differentiate very good from very bad interventions. Nonetheless, this tendency to focus disproportionately on what is known, or readily verifiable, can lead to certain forms of bias and error. To begin with, those conducting initial assessments may become so deeply immersed in one area that they postpone inquiry into more rewarding areas about which less is known. More worryingly, they may grow accustomed to standards of proof that cannot be replicated in other domains, overlooking opportunities for significant impact.

In the case of effective altruism, a salient concern is that the preference for evidence garnered through randomised controlled trials (RCTs) has led the movement to ignore other less tangible opportunities to do good.

RCTs are important in the context of international development because they help isolate the impact of individual projects and activities from those of the environment. This has led them to be regarded as the 'gold-standard method of testing ideas in other sciences,' such as medicine.[28] However, RCTs also suffer from certain familiar limitations. In addition to being costly to run, Singer notes, 'They can be used only for certain kinds of interventions, in particular, those that can be done on a small scale with hundreds or thousands of individuals or villages, from which samples large enough to yield statistically significant results can be drawn.'[29] This scope-restriction means they are not suited to evaluating country-scale initiatives, nationwide advocacy programs, or projects that function over a longer time period. Indeed, only a small number of aid activities can meet this burden of proof.

In the present context, there is concern that overreliance on RCTs has led effective altruists to overlook other high-impact activities like campaigning. Indeed, all but one of the 'top charities' recommend by GiveWell and Giving What We Can focus on treating neglected tropical diseases, an area where evidence drawn from RCTs is plentiful. Having taken on the task of identifying effective interventions, and sharing this information with the public, effective altruists have a responsibility to ensure that their own standards can be met across a range of high-impact areas. Otherwise they risk causing unwarranted reputational damage to charities working outside their core areas and providing information about effectiveness that is misleading.

### Quantification Bias

While effective altruists sometimes value things other than subjective wellbeing at the level of axiology, critics argue that the metrics they use to evaluate interventions mean that it's the only thing that counts in practice. The charge of quantification bias

focuses primarily upon the movement's reliance, in its early years, on the disability adjusted life years (DALYs) metric. Healthcare professionals use this metric to compare the value of treating disease and prolonging life: by drawing on reports from people living with different diseases they are able to calculate conversion rates that make interpersonal comparisons of wellbeing possible. Yet, importantly, the DALY metric is an indirect measure of subjective wellbeing. Analysis conducted in these terms, therefore, tends to exclude other values.

When combined with further information about the monetary cost of medical interventions, the DALY metric can be also be used to determine the cost-effectiveness of a given activity. This kind of analysis underpins the effective altruist claim that it is possible to save a life for as little as $3,340.[30] It also is central to Giving What We Can founder Toby Ord's view that huge variations in impact create a 'moral imperative towards cost-effectiveness'.[31] However, attempts to maximise effectiveness simply by prioritising this kind of intervention also suffer from a number of shortcomings. In addition to those that have been mentioned already, they tend to ignore the significance of iteration effects that occur when an intervention is scaled-up or replicated across a community.

To see how these factors can influence decision-making, consider the following choice:

> *Medicine.* According to recent estimates, condom distribution is a far more effective way of minimising the harm caused by HIV/AIDS than the provision of anti-retrovirals. Whereas anti-retrovirals help people who already have the virus, condoms help to prevent many more people from becoming infected. As a donor, you must choose between funding one of two national action plans. The first allocates the entire sum of money to condom distribution. The second allocates 90% to condom distribution and 10% to anti-retrovirals.

Assuming that the evidence in favour of condom distribution is reliable, effective altruists will tend to favour the first plan because it saves the greatest number of lives. Yet most governments and populations affected by the pandemic have rejected strategies that leave people with HIV/AIDS untreated.

One reason for this is recognition that it may be better to live in a society where one can *hope* to receive medical treatment if one is sick than to live in one where the largest numbers of people get treated overall.[32] Hope is valuable in this context either because it leads people to feel better about their lives or because it simply is feeling better about one's life. On either view, it contributes to wellbeing in a way that the original verdict overlooks. Hope is also an important resource for people who are undergoing serious hardship, encouraging them to tackle problems they might otherwise, in resignation, fail address. Taken together, these effects show why we cannot move directly from information about the most cost-effective intervention, on a DALY per capita basis, to reliable conclusions about the best overall policy or program.[33]

## Instrumental Bias

The final methodological charge, that effective altruism is unduly instrumental, holds that it fails to take seriously the effect of politics on outcomes and tends to favour

technocratic rather than democratic solutions to moral problems. Clearly a techno-cratic or data-driven approach to problem solving is sometimes appropriate. For example, it is unlikely that asking people to vote about the most effective aid and development organisations would provide a reliable guide to their impact. At the same time, reliance on these methods makes effective altruism vulnerable to certain forms of error.

At the micro-level, critics note that advocates of technical approaches to problem solving often develop a tacit preference for organisations that deploy the same kind of reasoning on site. They will tend to favour projects that focus on the priorities of experts and that can demonstrate their effectiveness in a way that is easily understood by evaluators.[34] What tends to be overlooked along the way is how different social processes influence outcomes and bear upon the good that is achieved. A technical approach to evaluation can also lead to the artificial narrowing of evaluative horizons, as the following example reveals:

> *Participation*. A group of villagers require help developing a water and sanitation system in order to tackle the problem of waterborne parasites. As a donor you face a choice between funding one of two projects. The first will hire national contractors to build the water and sanitation system, which they have done successfully in the past. The second works with members of the community to develop and build the new facilities. This approach has also worked in the past, but because villagers lack expertise, their systems tend to be less functional than the ones built by experts.

In this case, simple reliance upon cost-effectiveness analysis would lead us to support the view that the first project is better, especially if we conduct our evaluation at the point of completion. However, there are significant advantages to the second project, which cost-effectiveness analysis overlooks.

First, we might believe that it is valuable for people to choose the path their community takes and to participate in realising these goals, for reasons of autonomy and self-esteem. After all, there is an important moral difference between a receiving something as a gift and bringing it into existence through one's effort. The second approach also has important instrumental benefits. The resulting sanitation system is more likely to be valued by community members. It may therefore endure better despite its weaker functionality. Finally, the participatory approach may allow communities to develop new capacities and skills that can be deployed again in the future, creating a multiplier effect that spurs long-term development. Yet this too will tend to be overlooked if evaluators focus only on the technical problem of preventing waterborne disease. A more comprehensive assessment of the two projects would take these additional considerations into account.[35]

At the macro level, there is a related concern: namely, that the kind of strategic philanthropy which effective altruists favour has the potential to undermine the democratic process, particularly when it is directed overseas toward recipient nations.[36] Indeed, there is strong *pro tanto* reason to favour government leadership when it comes to public service provision or addressing social needs. Politicians are accountable to the electorate for the decisions they make and must provide a public justification of their policies for that reason. In contrast, private philanthropy and the work of foundations are not subject to democratic control in any direct

way. Nor are they under any pressure to provide public justification for their work. Decisions are usually based exclusively upon the comprehensive worldview of the agent in question.

This approach is problematic when it reduces popular control over social outcomes and when it weakens existing political institutions by circumventing standard mechanisms of accountability. Given the track record of many international organisations, these are dangers that effective altruists should be particularly cognizant of. Under certain circumstances there is evidence that service provision by non-state actors can diminish state capacity, impacting disproportionately on the worst-off people in society.[37] In addition to this, many of the most serious development failures have occurred at the hands of experts who, when freed from democratic oversight, enacted programs based on their own 'scientific' understanding of poverty that turned out to be deeply harmful.[38]

Recognising the force of these objections, effective altruists have responded in a number of ways. To begin with, they have sought to expand the evidence base by funding randomised trials and encouraging charities to better document the outcomes of their work.[39] Effective altruists have also tried to incorporate new forms of evidence and analysis into their thinking. This is clearest in the case of GiveWell, which has formally distanced itself from the DALY metric, reduced its reliance on cost-effectiveness analysis measured in this way, and expressed a preference for conclusions 'supported by multiple different lines of analysis, as unrelated to one another as possible'.[40] Finally, effective altruists have begun to explore new approaches to evaluating work done by advocacy organisations.[41] For example, the Open Philanthropy project, a new effective altruist research centre, has used natural experiments to conclude that campaigning for prison reform is the United States is cost effective.

Nonetheless, there is significant opportunity to further eliminate bias. The achievement of 'top charity' status continues to be heavily weighted toward organisations that appear cost-effective on the basis of the evidence discussed in this section. So long as this is the case, funding additional RCTs is not enough. Effective altruists should also to work closely with interested NGOs, providing them with the technical and financial support needed to demonstrate effectiveness over time.[42] Beyond that, effective altruism is still a long way from developing clear public standards of effectiveness for organisations engaged in advocacy work. In the absence of such standards, the movement needs to be more upfront about the bounded nature of the evaluations it provides or else risk causing reputational damage to campaigning organisations. Finally, as the movement scales up and has a larger social impact, questions of democratic accountability loom larger. There is already some evidence that global public health is becoming a crowded field.[43] As the low-hanging fruit in this sector begins to disappear, and donors to look for new high-impact opportunities, they will be pressed to consider more complicated interventions such as service provision or governance initiatives that have an overtly political dimension.

## Is Effective Altruism Effective?

Effective altruists aim to make the decisions that do the greatest amount of overall good. But how robust is their advice? I respond to that question in two parts. First, I

look at the impact of an individual's donation to one of their top charities. Second, I explore effective altruism's prospects for systemic change. In both cases, I argue, there are major obstacles to be overcome.

### Counterfactuals

A central claim of effective altruism is that for a person's action to have impact, it must also be the case that her contribution would not have been made by someone else if she had chosen to act differently. This type of counterfactual analysis explains their preference for neglected causes. It also explains their claim that it is usually better to choose a high-income career and donate a sizable portion of one's income to an effective charity than it is to work for a charity oneself.[44] Yet, this observation also raises doubts about the overall impact of donating to their recommended organisations.

To see why this is the case, we should first note that their favoured charities have limited room for more funding.[45] They can only absorb so much money while achieving the kind of outcome upon which the recommendations are based.[46] At the same time, the effective altruist community now contains a diverse set of actors, ranging from college students and young professionals to large philanthropic organisations – like GoodVentures – which have billions of dollars to spend. Bearing this in mind, we should ask what would happen if a smaller donor were to stop giving. Would larger donors simply fill the gap? This question is important because if they were to do so, then it follows that the impact of giving to a top charity is not in preventing malaria or ensuring that children are treated for intestinal worms, which would happen anyway. Rather, giving to these organisations simply makes additional funds available to larger foundations that are already flush with cash.[47]

To make progress on this front, we need to know more about why large donors don't fully fund the top charities already. There are three possibilities. First, they may have opportunities that are cost-effective and consume a large amount of resources but which are not available to individuals.[48] For example, large donors could establish new organisations to work on neglected causes. Second, large donors are aware that organisational pathologies occur when charities become too reliant on a single source of funding. In the absence of diverse funding streams, charities tend to become less accountable, focusing only on the wishes of a single donor. They are also less sustainable given that a donor's fortunes and priorities will often shift over time. Finally, donors who identify closely with effective altruism, such as GoodVentures, may be worried that by crowding out other potential supporters of top charities, they might jeopardise the long-term health of the movement. To begin with, the decision to fully fund the top charities would place considerable strain on existing meta-charities, leaving in place a demand for effective interventions that cannot be reliably identified. It would also create a precipitous rise in the *apparent* cost of the most effective interventions, making it harder for effective altruists to argue that individuals are in a position to do a massive amount of good for people living in extreme poverty and hence to recruit new supporters.[49]

Effective altruists should avoid these outcomes in order to preserve the promise of their movement. After all, if they manage to shift social norms around giving, then they could achieve gains on a scale that even the wealthiest foundations cannot now

bring about. At the same time, none of these considerations suggest the large donors would not pick up the slack if small-scale donations declined. To the contrary, it is very likely that at least one of these donors would do so.[50] This finding generates two important insights that effective altruists have tended to ignore. First, the willingness of large donors to act in this manner, though commendable, undermines the claim that ordinary people are having great impact by giving to the charities they recommend. Their individual contribution is quite different from the one that they think they are making: it consists largely in freeing up money for foundations and supporting a social movement that *may* do considerable good in the long run. Second, claims about what effective altruism may achieve in the future are quite speculative. This is not to say they are implausible, only that this potential cannot be tested or falsified in the rigorous way effective altruists prefer. They, too, appear to be acting on faith to some extent.

### Systemic Change

Many of the changes that would do most good for people living in extreme poverty are institutional or systemic in nature. According to recent estimates, trade protectionism costs people living in the world's poorest countries around $100 billion per year, and illicit financial flows cost them a further $25 billion per year.[51] Thus trade reform and better financial regulation could help to lift millions of people out of poverty on a sustainable basis.[52] In fact, the impact of these changes would be so great that campaigning for them may be the best thing to do, even when such efforts only have a small chance of success. Given their interest in expected value, we would expect effective altruists to be sensitive to this argument.[53] However, a number of critics worry that this is not the case in practice and that certain elements within the movement not only ignore long-term systemic change but also stand in its way.[54]

One argument to this end holds that by encouraging people living in rich countries to feel good about giving effective altruists trigger a set of ideas and associations that have a detrimental effect on public attitudes to poverty.[55] Proponents of this view argue that the language of charity is damaging in this context because it tacitly reinforces the idea that there is a moral hierarchy that separates giver and receiver, obscuring the fact that affluent people tend to benefit from economic and political structures that harm the global poor, while also diminishing the agency of those who stand to be assisted.[56] Psychological studies have shown that a focus on giving money also tends to enhance individualism and reduce communal motivation, dampening altruistic inclinations.[57] Finally, the idea of 'doing good' is itself problematic because it encourages people to believe that assistance is a matter of personal discretion rather than a moral responsibility, making collective action less likely. Taken together, these considerations suggest that those who embrace the effective altruist message are less likely to develop an accurate understanding of systemic sources of poverty or to put pressure on their governments to reform political institutions that exacerbate it.

A second reason for thinking that effective altruism is poorly equipped to bring about systemic change arises from its organisational logic and the model of reasoning it adopts. Crucially, effective altruists have the formal aim of doing as much good as possible. They therefore aspire to be *cause-neutral*, changing their practical focus as the evidence changes and as others move to address previously neglected causes. This

feature makes the movement particularly interesting. However, it may also reduce its ability to bring about systemic change.

To see why this is the case, we should note that many of the most successful social movements from the past, such as abolitionism and the civil rights, were goal-oriented in a much stronger sense: they began with small groups of supporters who were committed to particular moral ends.[58] According to Martha Finnemore and Kathryn Sikkink, these groups then functioned as 'norm entrepreneurs', using advocacy and direct action to undermine the moral frames that made harmful practices appear legitimate, and creating the environment to which policymakers subsequently responded.[59] Against this backdrop, constancy of purpose was important in four ways. First, it allowed them to develop specialist knowledge and to constantly refine their message, factors that contributed to their long-term success. Second, it helped them win over supporters who admired their integrity. Third, it meant that they persevered even when the odds of success appeared to turn against them. Finally, it meant they continuously tested the status quo by applying pressure through actions and campaigns, a strategy that alerted them to significant opportunities that might otherwise have passed them by.[60]

In contrast to these historical movements, the fact that effective altruists are so willing to change their priorities can make the movement seem skittish to outsiders. Contingency of aim may also detract from its effectiveness over time in other ways, by leading supporters to give up too soon or miss opportunities that arise. Indeed, these insights suggest that trying to identify the most effective action by drawing on existing evidence may not be the most reliable way to maximise impact over time. Those who prioritise systemic change, may do better to take fair global institutions as their aim and 'reason backwards' from this outcome to appropriate forms of action.

There are no easy answers to the problems raised in this section. With regard to the overall impact of individual donations to top charities, effective altruists either need to revise current estimates downward or identify new high-impact interventions capable of absorbing large amounts of additional funding. Greater methodological pluralism, which was discussed in second part of this article, may help in this effort. Second, in the case of psychological framing effects, the problem is not with the movement's substantive commitments *per* se but rather with the language it uses. In fact, effective altruists are increasingly interested in systemic change and tend to believe (as consequentialists do) that promoting good outcomes is a moral duty. However, the ready invocation of charity obscures these facts and makes collective action around political issues harder to achieve. Therefore, the movement should work with social psychologists to help close this gap between the medium and the message. Finally, the importance of goal-orientation and specialisation speak to the need for division of labour when it comes to tackling different global problems. Despite the ideal of cause-neutrality, this truth is acknowledged within the movement. There are effective altruist organisations working on global poverty, animal welfare, and existential risk, among other things. However, it is also an insight that has external validity. Effective altruists should recognise the general utility of having highly committed partisan organisations working for progressive causes. Rather than being sceptical about the efficacy of these organisations, which approach the question of positive social change in a different way, they should recognise that there is deep uncertainty about how to do good,

partnering with these organisations when their goals overlap and trying to build the kind of political alliances that support large-scale institutional change.


## Conclusion

Effective altruism is a young movement committed to certain claims and ideals the details of which are still being worked out. Understood as broadly welfarist, consequentialist, and scientific in its outlook, the movement is vulnerable to the claim that it overlooks the importance of justice and rights, is methodologically rigoristic, and fails to isolate the activities likely to have the greatest impact overall. In most cases, I have shown that effective altruists are able respond to these objections, though sometimes this would mean changing their *modus operandi* in significant ways.

With regard to the problem posed by different values, I have suggested that the movement adopt a model of moral reasoning and advice that is disjunctive in character. Effective altruists would then allow advisees to choose the values they consider morally relevant (within certain parameters) and produce recommendations that optimise impact bearing these considerations in mind. By using independent equity and priority ratings, and identifying spheres of activity in which rights violations are unlikely to occur, effective altruists could make room for the view that outcomes are good when suffering is avoided, gains distributed fairly, and rights respected. With regard to the second charge of methodological blindness, effective altruists can respond by continuing to embrace new evidence and developing standards of effectiveness that are less reliant upon randomised trials. These changes would allow them to evaluate a wider range of causes, which could ease the funding bottleneck around their top charities and also make it easier to engage with others in the humanitarian and development sectors. Finally, effective altruists need to think more carefully about the language they use, to ensure that they don't reinforce negative associations with poverty. In this regard, they have much to learn from advocacy organisations. Stronger alliances with this sector are important when seeking to bring about long-term change and are wholly appropriate given the deep uncertainty about how best to do good.[61]

*Iason Gabriel, St John's College, The University of Oxford, Oxford, OX1 3JP, UK.*
*iason.gabriel@politics.ox.ac.uk*


## NOTES

1  Peter Singer, *The Most Good You Can Do: How Effective Altruism Is Changing Ideas About Living Ethically* (London: Yale Press, 2015a), p. viii.
2  Total global development aid in 2013 was $134.8 billion (OECD). Available at: http://www.oecd.org/dac/stats/documentupload/ODA%202013%20Tables%20and%20Charts%20En.pdf.
3  The Centre for Global Prosperity calculates that US citizens contributed $12 billion to organisations working in the developing world in 2009. Around one-third of this money went to Sub-Saharan Africa. See Patricia Miller, *The Index of Global Philanthropy and Remittances 2011* (Washington, DC: Centre for Global Prosperity: The Hudson Institute, 2011), pp. 9, 12.
4  Peter Singer, 'I want to shame charities into proving the worth of their spending', *The Guardian* 26 March, 2015b.

5 Singer 2015a op. cit., p. vii.

6 A large section of the effective altruist community is concerned with the suffering of non-human animals. Due to limitations of space, this article focuses on the way in which the movement aims to promote human wellbeing.

7 GiveWell, 'Effective altruism'. Available at: http://blog.givewell.org/2013/08/13/effective-altruism/.

8 William MacAskill, *Doing Good Better: Effective Altruism and A Radical New Way to Make a Difference* (London: Faber & Faber Ltd., 2015), p. 15.

9 GiveWell, 'Your dollar goes further overseas'. Available at: http://www.givewell.org/giving101/Your-dollar-goes-further-overseas.

10 MacAskill 2015 op. cit., p. 28.

11 Singer 2015a op. cit., p. 146.

12 Larry Temkin, *Inequality* (Oxford: Oxford University Press, 1993), p. 13.

13 Akhter Ahmed, Ruth Vargas Hill, Lisa Smith, Tim Frankenberger, Kajal Gulati, Doris Wiesmann, Wahi-dand Quabili & Yisehac Yohannes, *The World's Most Deprived: Characteristics and Causes of Extreme Poverty and Hunger*, 43 (Washington, DC: International Food Policy Research Institute, 2007).

14 Oriana Bandiera, Robin Burgess, Narayan Das, Selim Gulesci, Imran Rasul & Munshi Sulaiman, 'Can basic entrepreneurship transform the economic lives of the poor?' *IZA Discussion Paper No. 7386* (2013): 1–38; Narayan Das & Farzana Misha, 'Addressing extreme poverty in a sustainable manner: Evidence from CFPR programme', *BRAC CFPR Working Paper* 19 (2010): 1–36.

15 Derek Parfit, 'Equality and priority', *Ratio* 10,3 (1997): 202–221, at p. 213.

16 John Rawls, *A Theory of Justice* revd. edn. (Cambridge: Belknap Press, 1999), p. 75.

17 Paul Mosley & David Hulme, 'Microenterprise finance: Is there a conflict between growth and poverty alleviation?', *World Development* 26,5 (1998): 783–790.

18 Ronald Dworkin, 'Rights as trumps', in Jeremy Waldron, *Theories of Rights* (Oxford: Oxford University Press, 1984), pp. 153–167.

19 Robert Nozick, *Anarchy, State and Utopia* (Oxford: Basil Blackwell, 1974), p. 33.

20 MacAskill 2015 op. cit., pp. 159–163.

21 This may be true even if they *choose* to work in sweatshops either because the right to adequate workplace protection cannot be forfeited or because the choice is made against a backdrop of acute vulnerability. Speaking about the plight of child workers, Debra Satz writes that, 'if the background circumstances and options poor children and their parents face are unjust, the option chosen does not by some mysterious process suddenly become just ... whether a choice confers legitimacy depends on other conditions besides its being voluntary'. See her *Why Some Things Should Not Be for Sale: The Moral Limits of the Market* (Oxford: Oxford University Press, 2010), p. 144.

22 Note that this revision would not change the final verdict for either *Two Villages* or *Ultra-poverty*.

23 Rights can also be accommodated in other ways. For example, Amartya Sen argues that rather than seeing them as constraints on action, we should view rights as an important stratum of non-welfarist goods, which bear upon our evaluation of outcomes and function as goals to be pursued. See his 'Rights and agency,' *Philosophy and Public Affairs* 11,1 (1982): 3–39, at pp. 15–16.

24 This is particularly important given Rawls's observation that 'a plurality of reasonable yet incompatible comprehensive doctrines is the normal result of the exercise of human reason within the framework of free institutions of a constitutional regime'. See his *Political Liberalism* (New York: Columbia University Press, 1993), p. xvi.

25 Nick Beckstead, a researcher at the Open Philanthropy Project, has suggested, in personal correspon-dence, that the distribution of bed nets to prevent malaria is an intervention of this kind. It helps some of the world's poorest people and is an effective way of providing a normal lifespan for children who would otherwise die in infancy. It therefore scores highly according to utilitarian and prioritarian values without infringing rights.

26 MacAskill 2015 op. cit., pp. 17, 224.

27 See, for example, comments made by Angus Deaton & Daron Acemoglu in 'The logic of effective altru-ism', *Boston Review* July/August (2015), pp. 19, 21.

28 Against the view that this status is deserved, Angus Deaton argues, 'Evidence from randomized controlled trials can have no special priority' and that it does 'not occupy any special place in some hierarchy of evi-dence'. See his 'Instruments, randomization, and learning about development', *Journal of Economic Litera-ture* 48 (2010): 424–455. at p. 426. For the original citation, see MacAskill 2015 op. cit., p. 9.

29 Singer 2015 op. cit., p. 157.

30 Where this means giving someone sixty years of life at full health. See MacAskill 2015 op. cit., p. 63 and GiveWell, 'Against Malaria Foundation (AMF)'. Available at http://www.givewell.org/international/top-charities/amf#footnoteref81_9yixouy.

31 Toby Ord, 'The moral imperative toward cost-effectiveness in global health', *Center for Global Development* March (2013): 1–8.

32 Dan W. Brock & Daniel Wikler, 'Ethical issues in resource allocation, research and new product development', in Dean T. Jamison *et al.* (eds) *Disease Control Priorities in Developing Countries* (Washington, DC: International Bank for Reconstruction and Development/The World Bank, 2006), p. 264.

33 Another argument appeals to the priority principle and concludes that – because it is worse to have HIV/AIDS than to have some chance of contracting the disease – those who already have it deserve a chance to be treated.

34 James Ferguson, *The Anti-Politics Machine: 'Development', Depoliticization, and Bureaucratic Power in Lesotho* (Minneapolis, MN: University of Minnesota Press, 1994).

35 In the words of Amartya Sen, they will focus on the 'comprehensive' outcomes of projects rather than their 'culmination' outcomes. See his *The Idea of Justice* (Harmondsworth: Penguin, 2010), p. 215.

36 Rob Reich, 'What are foundations for?', *Boston Review* (1 March 2013), available at: http://bostonreview.net/forum/foundations-philanthropy-democracy and Emily Clough, 'Effective altruism's political blind spot', *Boston Review* (14 July 2015), available at: http://bostonreview.net/world/emily-clough-effective-altruism-ngos.

37 Emily Clough, 'Effective altruism's political blind spot', *The Boston Review* (14 July 2015), available at: http://bostonreview.net/world/emily-clough-effective-altruism-ngos

38 Timothy Mitchell, *Rule of Experts: Egypt, Techno-politics, Modernity* (Berkeley, CA: University of California Press, 2002); William Easterly, *A Tyranny of Experts: Economists, Dictators and the Forgotten Rights of the Poor* (New York: Basic Books, 2014).

39 For example, GiveWell recently made a grant of $300,000 to GiveDirectly for this purpose.

40 See GiveWell, 'Some more considerations against investment in cost-effectiveness' available at: http://blog.givewell.org/2011/11/04/some-considerations-against-more-investment-in-cost-effectiveness-estimates/ and GiveWell, 'Why we can't take expected value estimates literally (even when they are unbiased)', available at: http://blog.givewell.org/2011/08/18/why-we-cant-take-expected-value-estimates-literally-even-when-theyre-unbiased/.

41 Singer 2015 op. cit., pp. 158–164.

42 Funds like the Edna McConnell Clark Foundation already provide this kind of support to grantees, ensuring that recipients are able to demonstrate impact within three to six years.

43 GiveWell, 'Trying (and failing) to find more funding gaps for delivering proven cost-effective interventions'. Available at: http://blog.givewell.org/2013/03/21/trying-and-failing-to-find-more-funding-gaps-for-delivering-proven-cost-effective-interventions/.

44 Their reasoning is that if you don't work for the charity then someone else who is similarly qualified will probably do so. However, if you turn down the job in financial services it is very unlikely that the person who replaces you would give a large portion of his or her income to the world's best charitable organisations. See William MacAskill, 'Replaceability, career choice, and making a difference', *Ethical Theory and Moral Practice* 17, (2014): 269–283.

45 Of the top charities that GiveWell recommends, GiveDirectly has the greatest capacity to absorb additional funds. In December 2014, GiveWell estimated it could absorb a further $25–30 million in the coming year. In March 2015, GiveDirectly received a grant of $25 million from GoodVentures, which suggests their current funding gap is now much smaller than originally anticipated.

46 The major constraints in this regard are organisational capacity, skilled staff, ability to identify appropriate targets, and political support for an organisation's work.

47 These organisations often struggle to spend the money they have. See, for example, David Callahan, 'Why isn't Bill Gates giving away his money faster?', *Inside Philanthropy* (May 16th 2014). Available at http://www.insidephilanthropy.com/home/2014/5/16/why-isnt-bill-gates-giving-away-more-money-faster.html.

48 Against this, Elie Hassenfeld, one of GiveWell's founders, states they are yet to identify large-scale opportunities for impact that outperform their top recommended charities. See Dylan Matthews, 'You have $8 billion. You want to do as much good as possible. How do you spend it?', *Vox* 24 April (2015). Available at: http://www.vox.com/2015/4/24/8457895/givewell-open-philanthropy-charity.

49 This follows from the claim that the most-cost effective organisations are vastly more effective than the median intervention (Ord op. cit.; MacAskill 2015 op. cit., pp. 59–62).

50 Many of them already have a surfeit of funds and it would go against their core principles to ignore high-impact opportunities to avert death and suffering.

51 Thomas Pogge, *Politics as Usual: What Lies Behind the Pro-Poor Rhetoric* (Cambridge: Polity Press, 2010), p. 20; Dev Kar, 'Illicit financial flows from the least developed countries: 1990–2008', *United Nations Development Programme Discussion Paper* (2011), p. 5.

52 So too could other reforms such as greater labour mobility. Michael Clemens, 'Economics and emigration: Trillion-dollar bills on the sidewalk?', *The Journal of Economic Perspectives* 25,3 (2011): 83–106.

53 Reasoning in this way number of effective altruists have now turned their attention to existential risk, studying low-probability events that would have a catastrophic effect on human life if they occurred. Open Philanthropy has also conducted shallow or medium-depth investigations into areas like labour mobility and macroeconomic policy.

54 Amia Srinivasan, 'Stop the robot apocalypse, review of *Doing Good Better: Effective Altruism and a Radical New Way to Make a Difference*, by William MacAskill', *London Review of Books* 37,18 (2015): 3–6; and Matthew Snow, 'Against charity', *Jacobin* 25 August (2015). Available at: https://www.jacobinmag.com/2015/08/peter-singer-charity-effective-altruism/

55 Martin Kirk, 'Beyond charity: Helping NGOs lead a transformative new public discourse on global poverty and social justice', *Ethics & International Affairs* 26,2 (2012): 245–263.

56 Andrew Darnton & Martin Kirk, *Finding Frames: New Ways to Engage the UK Public in Global Poverty* (London: Bond, 2011), p. 90; Tomohisa Hattori, 'The moral politics of foreign aid', *Review of International Studies* 29,2 (2003): 229–247.

57 Kathleen Vohs, Nicole Mead & Miranda Goode, 'The psychological consequences of money', *Science* 314,5802 (2006): 1154–1156.

58 Neta Crawford, *Argument and Change in World Politics: Ethics, Decolonization, and Humanitarian Intervention* (Cambridge: Cambridge University Press, 2002).

59 Martha Finnemore & Kathryn Sikkink, 'International norm dynamics and political change', *International Organization* 52,4 (1998): 887–917.

60 As Steven Teles and Mark Schmitt argue, it is the nature of political advocacy that 'events evolve rapidly and in a non-linear fashion so that an effort that doesn't appear to be working might suddenly bear fruit'. As a result, advocates 'may require a long period of trial and error', before they succeed. See 'The elusive craft of evaluating advocacy', *Stanford Social Innovation Review* (2011), pp. 39, 40.

61 I would like to thank William MacAskill, Toby Ord, Rob Reich, Nick Beckstead, Daniel Brock and the referees and editor of this journal for their feedback. I would also like to thank all the participants at the Stanford Centre on Philanthropy and Civil Society workshop, 9 April 2015.