Review of Carol J. Adams, Alice Crary, and Lori Gruen (eds.) *The Good It Promises, The Harm It Does: Critical Essays on Effective Altruism,* 2023, Oxford: Oxford University Press.

Effective altruists (EAs) seek to persuade the globally wealthy to donate a proportion of their income to do good, and specifically to donate it to those charities that will do the most good, in expectation, with the given amount of money donated. Their movement began in the mid-2000s with the charity evaluators GiveWell and Giving What We Can, which focus on how to do good for humans who are currently alive. Animal Charity Evaluators was founded in 2012, and focusses on non-human animals, whether wild or domestic, bred for food or labour or entertainment. Despite its title and subtitle, the present volume is pitched as a critique only of this latter, animal-focussed part of the movement, though nearly all of the worries the contributors raise tell equally against the human-focussed part, if they tell against either. The volume also includes a coda by the editors where they address the community's recent turn to longtermism, which says that, as well as all humans and all non-human animals alive today, our moral circle should include those that will live in the future, including the very far future; and since, barring a catastrophe, these future beings will outnumber those currently living, longtermists say it is more urgent to do good for them. As well as the original charity evaluators, a series of further EA organisations have emerged, including grant-making organisations, research institutes, a vast discussion board known as the Effective Altruism Forum, and a variety of regular conferences around the world. So what began less than two decades ago as websites ranking charities by effectiveness has become both a large movement, with an estimated 7,400 active members, but also, importantly, an extremely well funded one, with an estimated $46billion pledged to EA causes.

It is only right that a movement of this size and power, backed by this much funding and seeking to make a substantial difference in the world, should be subject to scrutiny. This volume seeks to do that. The editors of the book are philosophers known for their work on animal ethics, among many other things; and the contributors include them alongside other scholars, activists, and scholar-activists in animal and vegan advocacy.

The book includes many criticisms of EA. Unfortunately, the editors chose not to focus each chapter on a specific criticism, but rather on the worries most prominent in the minds of the chapter's author. But, inevitably, many of the authors have similar lists of worries and only limited space to describe them, and so we read only brief versions of the same concerns again and again, and we rarely find the extended exploration they deserve. In what follows, I pick out what I take to be two of the book's central objections to the guiding principles of EA, though the book wisely takes both guiding principles and actual practices as legitimate targets for criticism. The objections are these: (i)

the moral theory that underpins much EA conflicts both with an ethics of care and with the ethics of self-determination espoused by many liberation movements; (ii) in the debate between reform and revolution, EAs pick reform, while the authors contend in favour of revolution. Of course, I too have only limited space and can't treat the concerns in sufficient detail. My aim is to persuade each side of the debate that the other has something to offer.

Besides the two issues I will cover, the criticisms of EA voiced in the volume include: its lack of diversity; its willingness to accept donations from the extremely wealthy; its willingness to work with morally compromised corporations; its lack of attention to local cultures in the countries it affects; its inattention to cultures of sexual harassment in the charities with which it works; and many, many more.

1. EA and a feminist ethics of care

A number of the chapters highlight the clash between the principles of effective altruism and the principles of a feminist or ecofeminist ethics of care.[1] In their Introduction, the editors write that one of the central commitments of effective altruism is impartiality: everyone with moral status should figure equally in your moral decision-making; you should give no preference to certain people over others. And this, they claim, is at odds with an ethics of care. I disagree: it is not a central commitment; and the commitment in the vicinity is not obviously in tension with much in care ethics.

Take two of the effective altruist's central claims: first, that the globally wealthy should contribute a certain proportion of their wealth to charity; second, that, having committed to donating that money, they should give it to a charity that distributes malarial nets in affected parts of the world, say, rather than to their local community centre. Neither of these requires absolute impartiality. Even if you give vastly greater weight to your own well-being and that of your family and friends, it will most likely still be better to donate a proportion of your salary than to keep it for yourself, because the good it will do for others is so much greater than the good it will do for you; and even if you give vastly greater weight to people in your local community than to those in a malarial region of the world, it will again most likely be better to use a certain amount you've committed to donating to help the latter, and for the same reason: a great good given a small weight can easily exceed a small good given a larger weight.

Nonetheless, a number of contributors find that it is permissible (and perhaps even mandatory) to give such great weighting to yourself and those connected to you that even when the benefit your money can bring elsewhere is so much greater, it is still allowed (or required) to favour the former. Carol J. Adams thinks it is obvious that Peter Singer was required to spend a large amount of

[1] E.g. the editors' Introduction and Chapters 2, 7, 8, 9, 12, 16.

money on the care of his mother when she was suffering from Alzheimer's disease rather than donating it to EA charities (Chapter 9); Krista Hiddema says it was good to spend the $600,000 needed to diagnose breast cancer in Esther, a domestic pig who had become a companion animal (Chapter 12); and Rachel McCrystal writes of a goat named Fred, "if [he] was able to be saved and come to Woodstock Sanctuary, where it would cost tens of thousands of dollars to care for him over his life, I would have done it in a second" (201, Chapter 14). What I have to say is the same for each case, so I will focus on Peter Singer's decision about his mother's treatment.

In an interview in 1999, Singer was asked how he could reconcile his decision to pay for his mother's care with his support for EA. He replied that it was "probably not the best use you could make of my money" (Specter 1999). Adams sees in this Singer's inability to recognise that he was moved by considerations that belong to care ethics and have no place in EA. He was moved by his obligation to care for a person, his mother, who had cared for him and with whose life his was so tightly entwined. And, Adams thinks, he was right to be so moved. Adams sees in Singer's later disavowal of his decision a failure, but I'm not so sure.

A claim that appears in a number of places throughout the book is that EA arises from and gives rise to an inappropriate emotional decoupling from the actions it takes and the effects it has. pattrice jones writes that "EA encourages a kind of calculating dispassion that can lead to callousness" (118). Animal lives, human and non-human, are treated as numbers on a ledger to be entered into a heartless mathematical formalism. Yet my experience with EAs suggests precisely the opposite understanding of their motivations: they have a deep visceral reaction to suffering regardless of its subject. Rather than feeling less profound dismay at the suffering of loved ones than most people feel, they feel more profound dismay at the suffering of people unknown to them than many people feel.[2] Where Carol J. Adams describes Singer as trying to make ethical decisions by taking the view from nowhere, I think many EAs might be more accurately described as trying to take the view from everywhere.

In her overview of care ethics (not from this volume), Stephanie Collins says that three key tenets of the approach are: (i) "deliberation should include sympathy and direct attendance to concrete particulars"; (ii) "agents should have caring attitudes"; and (iii) "agents should perform actions that are performed under the […] intention of fulfilling […] interests that the agent perceives some moral person […] to have" (2018). And I think EAs often live by exactly those maxims; and, whether a particular EA does or not, what the movement recommends can often be justified from the perspective of those maxims. For instance, EAs focus on the well-being of the individuals affected

[2] This is what Richard Yetter Chappell calls 'levelling-up impartiality'. https://forum.effectivealtruism.org/posts/xx4HBHd8bKnuNN6Hd/leveling-up-impartiality

by their interventions, and they collect evidence to discover how best to improve that well-being, which is surely what we would expect from an approach based on what sympathy demands when it is coupled with attention to the concrete particulars of those individuals' situations.[3]

From the view from everywhere, Singer's response to the interviewer's question about his mother makes sense. I don't know the details of his mother's illness or her care, but I watched my own grandmother suffer in a nursing home for the last ten years of her life, and I feel very immediately the benefits of better care that money can buy. Yet I also know how much money is required. On average, £800 per week in the UK; and I imagine a lot more for people with complex health conditions. That's £200,000 over five years, which is what effective altruists predict it costs to prevent the deaths of fifty children. So perhaps it was permissible for Singer to spend the money on better care for his mother, but it is surely understandable that those figures gave him pause.

Carol Adams says Singer could have justified his decision as follows: "I care about my mother and have a responsibility to help her as she suffers this disabling illness" (138). I suspect he would agree these are important considerations, but he would point out that there are other important considerations, and he had to decide how to weigh these against each other; he had to determine how they traded off against each other. If he spent the money on his mother's care, he didn't spend it preventing the deaths of fifty children. And, in hindsight, he could not justify weighting his mother's comfort so greatly that it outweighed preventing those deaths. Where EAs make decisions using principles such as maximising4 the good in expectation, it need not be because they are shunning an approach based on sympathy, but because when your sympathy extends to everyone who is suffering, and when the resources at your disposal can't alleviate all of that suffering, you face trade-offs and need some method by which to make them.

Even if the official version of their approach to morality doesn't tolerate these sorts of trade-offs, most people who take an explicit approach in fact allow them. If I could save the lives of a billion people for sure by killing this one, I may; or perhaps even I should. However, while people do tend think trade-offs permissible, I think they imagine that the situations in which they create a tension with the demands of care ethics are not ones in which they will actually find themselves—surely only in the most tragic circumstances will I have to choose between a loved one and the lives of many others. Yet a key guiding insight of EA is that the world's wealth inequality is so vast and some of the life-saving treatments available so inexpensive from the point of view of the globally wealthy that the rich can often face these tensions repeatedly during their lives, just as Singer did.

---

[3] An alternative feminist ethics of care, specifically focused on the cases of animals, is provided by Donovan and Adams (2007).

Of course, Singer's decision makes Williams' integrity objection to utilitarianism vivid (Williams 1973). I follow Nikhil Venkatesh's reading of this objection on which Williams is concerned that a utilitarian can never make a commitment to any principle, other than utilitarianism itself; they can never say they will stick to that principle no matter the cost (2023, Chapter 2). Now, you might not think this such a terrible conclusion, but note that it applies also to personal relationships, as Singer's decision makes clear. The utilitarian cannot commit to caring unreservedly for a family member whenever they need it because they might face the choice between that commitment and the option of saving many children. Again, you might not think this upshot is so bad for utilitarianism, as long as such choices are rare. It would not diminish the quality of the relationship I enjoy with my family to know that they would choose to save fifty lives over helping me, as long as they were rarely to face that choice. But what EA makes clear is that globally wealthy people face that choice all the time because of the staggering, life-altering good that their money can do in the current world. Perhaps this is bad news for utilitarianism. But it's not good news for care ethics. Were there very little wealth inequality in the world, there would be little problem with an almost exclusive focus on caring for those with whom you have direct personal relationships and interlocking obligations of care. But when the circle to which a wealthy person belongs largely includes only other wealthy people, showing such partiality perpetuates inequality.

2. The ethics of autonomy and self-determination

The second ethical system with which the editors take EA to clash is the ethics that arises from some liberation movements. As they put it: "many [of these movements] are distinguished by their insistence on starting with the voices of the oppressed and taking simultaneously empathetic and critical engagement with these voices to guide the development of strategies for responding to suffering" (xxvii). Here, I think they make an important point: there is a paternalistic side to EA that might clash with an ethics that places importance on autonomy and self-determination—though I will note in passing that one of the charities that was long championed by GiveWell is GiveDirectly, which simply transfers donated funds directly and unconditionally to very poor families in Kenya, Uganda, and Rwanda, which seems both non-paternalist and redistributive.

If you want to do the most good with a given quantity of resource, you might think you need to find out two things: what do people want? What best obtains that for them? But a hedonistic utilitarian needn't stop to ask the first. Regardless of people want, what is good for them is greater pleasure and what is bad for them is greater pain. Desires whose satisfaction does not lead to greater pleasure will be thwarted by the utilitarian if doing so makes you happier. And it can often seem that EA assumes this account of well-being,

formalised in the concept of a QALY, or quality-adjusted life year, which is a unit of experiential well-being defined to be equivalent to a single year lived in perfect health. After all, while EAs spend much resource investigating how to maximise QALYs in expectation, they don't spend similar amounts asking people what they actually want.

Now, while you might agree that this sort of autonomy and self-determination are important when effective altruism focusses on alleviating human suffering, it is less relevant for its animal advocacy work, since it is less possible to involve non-human animals in decisions about their own lives, and it is difficult to assess what would be good or bad for them by their own lights. But, as many contributors to this book point out, it is not only animals who are affected by much animal advocacy: for one thing, there are those employed in the factory farming industry, such as those in Brazil who sustained frostbite during their work and were denied proper medical treatment for it (18); for another, there are the humans who are the target of various veganism campaigns. However, in both cases, it is difficult for animal advocates to avoid imposing their own values to some degree. While factory farming workers will likely want improved working conditions, they may well not want factory farms to disappear, as animal advocates typically do; and while some meat-eaters might want to learn how to switch to a vegan diet, others will not. Animal advocacy has always been about changing attitudes rather than listening to the attitudes people actually have.

3. The epistemology of Effective Altruism

Nonetheless, I do think there are two ways in which effective altruism could learn from the importance liberation movements place on listening to the people whose lives their interventions will affect. Both are epistemic. The first is the familiar insight of standpoint epistemology that people from a community with which a charity hopes to work will often have useful knowledge about what's likely to work for that community and what isn't. Brenda Sanders' and Simone de Lima's illuminating chapters in this volume provide a variety of cases (Chapters 1 and 2): both note that a certain sort of veganism campaign that promotes 'fake meat' alternatives, either through supermarkets or through fast food chains, will likely not make any difference to the diets of people living outside cities in Brazil or low-income communities of colour in the US: for both, these alternatives will be too expensive, but also geographically inaccessible, since they live in food deserts, with very poor access to the sorts of supermarket that will sell it.

But there is also a more extreme version of the claim from standpoint epistemology that is sometimes suggested by some of the book's contributors, namely, that being a member of the affected community gives such good epistemic access to what will work for that community that the sorts of evidence of effectiveness that EAs seek becomes less necessary. This seems implausible.

The social world is remarkably complex, and its causal structure very opaque even to those who are embedded in it and harmed by it.

The book holds out another advantage of listening to those involved in liberation movements: they have often thought hard and creatively about alternative ways societies might be organised. And this connects to a serious problem for EAs that has become particularly acute as it has shifted its focus to the long-term future. EAs tend to decide what to recommend by calculating expected utilities. So they begin by listing the different things you might do with your donation. Then for each option available to you, they look at all the different ways the world might turn out, estimate how much goodness the world would contain if it were to turn out that way, weight that by the probability the world would turn out that way were you to donate in the way in question, and then add up those probability-weighted amounts of goodness to give the expected goodness of that way of using your money. Then they recommend donating in whichever way maximises goodness in expectation. You might question that way of proceeding, but even if you find it compelling, there are many ingredients that must be fixed before the comparison can take place: critics often focus on the problem of measuring goodness and the problem of setting probabilities; but there is also the problem of specifying the different ways the world might go and the problem of specifying the different actions that are available to you. And I take it that this is one of the things Amia Srinivasan means when she says that liberation movements are valuable sources of "radical political imagination" (x, Foreword). We'll come back to the question of reform vs revolution in a moment, but one strength of movements that aim at revolution is that they really think about ways in which the world and its social structures might be significantly different. And this, it seems to me, is an area where EA has much to learn. It has been most visible in recent times in debates about longtermism, where EAs have had to envisage different ways the future might go and then assign probabilities and utilities to those. The scenarios that are most widely discussed are those that particular high-profile or particularly loud voices have imagined, or ones drawn from science fiction. But there's no reason to think that such a method samples the space of possible futures with any sort of evenness; and indeed we don't yet have a good account of how to conceive of that space or how to think of uniform sampling from it. So EAs could attend better to the different ways the world might be that they take to be possible, which they include, and how they come up with them. Liberation movements are good sources of imagination; artists, novelists, and the like are as well. In general, this is an area in which you want a diversity of approach: if everyone in your movement has a similar outlook, they will likely imagine similar possibilities.

4. Reform or revolution?

One reason to be pessimistic that EAs will be interested in exploring a wider range of possible radically alternative futures for human society is that they

tend not to recommend any interventions that aim to significantly change society in the more radical ways that are well known, such as socialist approaches, and so on. This leads to their opponents to criticise EAs much as Rosa Luxembourg criticised the socialist reformer Eduard Bernstein for abandoning the goal of revolution: the interventions that EA recommends aim to make improvements within the existing social system; they do not aim to change it; and yet that system is the cause of much of the suffering EA seeks to alleviate; what's more, through their interventions, they prolong the system's existence.[4]

In response, EAs argue they are not in a position to change the social system, even if they thought that would improve the world—and many, I suspect, don't think this, believing instead that all systems will cause much suffering, but that the current one with substantial voluntary philanthropy will alleviate it best. But of course, by their own lights, they don't have to be able to change the social system for sure in order to recommend trying. It is enough that the expected goodness of aiming at the change outweighs the expected goodness of other possible interventions. A comparison with longtermism is telling. After all, it recommends interventions that increase the probability of a good outcome—a long happy future for sentient beings—from an extraordinarily small amount to a very slightly higher but still extraordinarily small amount. But that's what the revolutionaries recommend as well! Of course, EAs could point to the ways in which revolutions can go very badly wrong and make things dramatically worse for the worse off. But this is true also for their longtermist interventions: a very long future for humanity might be very good, but it might also be terribly bad. So the case for longtermism and the case for revolutionary change seem analogous. In the end, everything turns on the probabilities you use to calculate the expected utilities. EAs often talk of 'evidential probabilities', as if for any body of evidence and any way the world might be, there is a unique precise probability that gives how likely the world is to be that way given that evidence. I'm sceptical that such exist. And it seems to me that, given the evidence we have, both the reformers and the revolutionaries assign probabilities that are rational responses to that evidence. But of course we need not, and certainly should not, rest content with the evidence we currently have. And again, it seems to me that this is an area where again EA could improve. As well as imagining alternative ways the world might be, many liberation movements build small-scale versions of the alternative societies they would like to see: off-grid communities, socialist communes, etc. And these provide a rich source of information about what their alternative systems would be like. The evidence will be limited, of course. These societies are, after all, still based in the larger context of the current social system. But the evidence is still helpful, and EAs should assess it.

---

[4] Cf. Bernstein 1899 [1970]; Luxembourg 1899/1908 [1919]. This framing of one component of the so-called 'institutional critique' of Effective Altruism goes back at least to Broi (2019), and Lori Gruen adopts it in her excellent contribution to the book (Chapter 17).

So, in the end, this book suggests some important critiques of Effective Altruism. However, with a few exceptions, the authors aren't given the space to explore them in as much depth as they deserve. But I hope this will not lead EAs and others interested in this movement to ignore what is valuable in the book.

Declaration of Interest

I read about charity evaluators in the mid-2000s, and while I've never had any involvement in the EA community, I've since donated a portion of my income to charities recommended by GiveWell. Later, when EA pivoted towards longtermism, I stuck with my original charities and tried to explore what I found so wrong-headed about that move, eventually writing up my argument against it. On the basis of this, I was invited to spend a month at the Global Priorities Institute in Oxford, which is partly funded by EA charities.

Acknowledgments: I'm very grateful to Jacob Barrett, Heather Browning, David Thorstad, Jack Sagar, Katie Steele, Nikhil Venkatesh, and comments from anonymous referees for helpful discussions of this material.

References

Bernstein, Eduard (1899 [1978]) *Die Voraussetzungen des Sozialismus und die Aufgaben der Sozialdemokratie* (Stuttgart: Dietz Nachfolger)

Broi, Antonin (2019) 'Effective Altruism and Systemic Change.' *Utilitas*, 31(3): 262-276.

Collins, Stephanie (2018) 'Care Ethics: The Four Key Claims'. In David Morrow (ed.) *Moral Reasoning,* 193-204 (Oxford: Oxford University Press)

Donovan, Josephine and Carol J. Adams (eds.) (2007) The Feminist Care Tradition in Animal Ethics (New York: Columbia University Press)

Luxembourg, Rosa (1898/1908 [1919]) *Sozialreform oder Revolution?* (Leipzig: Vulkan-Verlag)

Specter, Michael (1999). 'The Dangerous Philosopher.' *New Yorker,* September 6, 46– 55.

Venkatesh, Nikhil (2023) *Utilitarianism and the Social Nature of Persons* PhD Thesis, University College London

Williams, Bernard (1973)' A Critique of Utilitarianism'. In *Utilitarianism: For and Against*, 77–150 (Cambridge: Cambridge University Press)