

Essays in Philosophy

ISSN 1526-0569

Volume 18, Issue 1 (2017)

Effective Altruism: Introduction

William MacAskill
Oxford University

Essays Philos (2017)18:1 | DOI: 10.7710/1526-0569.1580

Correspondence: will@effectivealtruism.org



© 2017 MacAskill. This open access article is distributed under a Creative Commons Attribution 4.0 License (<https://creativecommons.org/licenses/by/4.0/>)

Effective altruism is a very new philosophical approach, and the present issue of *Essays in Philosophy* is the first academic volume entirely devoted to this idea.

As I and the Centre for Effective Altruism define it, effective altruism is the project of using evidence and reason to figure out how to benefit others as much as possible, and taking action on that basis.

On this definition, effective altruism is an intellectual and practical project rather than a normative claim, in the same way that science is an intellectual and practical project rather than a body of any particular normative and empirical claims. Its aims are welfarist, impartial, and maximising: effective altruists aim to maximise the wellbeing of all, where (on some interpretation) everyone counts for one, and no-one for more than one. But it is not a mere restatement of consequentialism: it does not claim that one is always obligated to maximise the good, impartially considered, with no room for one's personal projects; and it does not claim that one is permitted to violate side-constraints for the greater good.

Effective altruism is an idea with a community built around it. That community champions certain values that aren't part of the definition of effective altruism per se. These include serious commitment to benefiting others, with many members of the community pledging to donate at least 10% of their income to charity; scientific mindset, and willingness to change one's mind in light of new evidence or argument; openness to many different cause-areas, such as extreme poverty, farm animal welfare, and risks of human extinction; integrity, with a strong commitment to honesty and transparency; and a collaborative spirit, with an unusual level of cooperation between people with different moral projects.

I believe that effective altruism is an important idea for three reasons. First, it's an idea that should be supported by a very wide variety of moral views. Consequentialism, including utilitarianism, would of course support the idea of using one's resources to benefit others by as much as possible. But all plausible moral views care about making the world better, impartially speaking; and, given the radical inequality in the world today, they should support increased efforts from middle-class members of affluent countries to make the world better. As the late Derek Parfit noted in a talk about effective altruism, "I don't think the disagreement between utilitarianism, egalitarianism and prioritarianism makes much difference here, because the stakes are so high, and the difference that you can make is so obvious, that all those views are going to agree."¹

¹ Derek Parfit, 'Reasons, persons, and effective altruism', talk delivered at Harvard University on April 21, 2015, <https://youtu.be/q6glXJ7dVU0>

Second, it's an idea that is already doing a lot of good. Last year, GiveWell moved over \$100 million for the most effective charities; Giving What We Can members have between them made over \$1.4 billion in lifetime pledges to charity; GoodVentures is a \$10 billion foundation that describes itself as part of the effective altruism community. In addition, thousands of people around the world have made significant changes to their career plans on the basis of effective altruist ideas, and a number of new non-profit and social enterprises have been founded on effective altruist principles.

Finally, it's an area where philosophical work is of huge real-world importance: if we want to claim that one course of action is, as far as we know, the most effective way of increasing the welfare of all, we simply cannot avoid making philosophical assumptions. How should we value improving quality of life compared to saving lives? How should we value alleviating non-human animal suffering compared to alleviating human suffering? How should we value mitigating risks of human extinction, with the loss of hundreds of trillions of future lives that that would involve? When it comes to small chances of doing huge amounts of good, should we simply maximise expected value, or is some other decision theory correct? How should we act in light of deep uncertainty about what the morally right thing to do is?

Because philosophical assumptions are so important to the effective altruist project, I'm very happy to see new academic work on this area. The essays in this journal cover a variety of topics.

Alida Liberman's 'Effective altruism and Christianity: possibilities for productive collaboration' considers potential areas of overlap—as well as possible sources of tension—between effective altruists and Christians. Liberman points out that one of the two tenets that she believes are at the core of effective altruism, namely that we should give away a significant portion of our income to assist the world's neediest people, is shared by the long-standing Christian tradition that recognizes a strong moral obligation to help the poor. (A similar point has been made by Toby Ord in his article 'Global poverty and the demands of morality'.²) The other tenet, that we should use the best available scientific tools to maximize the good done with the resources we give, seems more in tension with Christianity. Some of this tension is practical: as a matter of fact, Christian giving is not very effective, but this could change as a result of effective altruist outreach. But the tension is also partly doctrinal: unlike (most) effective altruists, Christians attach high value to giving aimed at supporting local churches

² In J Perry (ed.) *God, The Good, and Utilitarianism: Perspectives on Peter Singer*, Cambridge: Cambridge University Press, 2014, pp. 177-191.

and faith communities, at enabling communal worship of God, and at supporting evangelism.

Andrew Fisher's 'Theory-neutral arguments for "effective animal advocacy"' defends the extension of the effective altruist approach to the specific issue of animal advocacy. Fisher's proposal, specifically, is to rely on the best available evidence to identify the best ways of helping animals. According to Fisher, such an approach has been resisted by some in the animal advocacy movement because it is assumed that it involves a commitment to consequentialism, which many animal advocates reject. Fisher tries to defuse this source of resistance by noting, correctly, that effective altruism, and hence effective animal advocacy, does not presuppose a consequentialist moral theory.

Gianfranco Pellegrino's 'Effective altruism and the altruistic repugnant conclusion' argues that a component of effective altruism—which he calls altruistic maximization—has implications that are as repugnant as the well-known repugnant conclusion in population ethics. Such cases arise in situations where agents can bring about either a substantial benefit to few individuals or a tiny benefit to a much larger number of people. Pellegrino also claims that this "altruistic repugnant conclusion" can only be avoided by abandoning the idea that our judgments in these cases are precise. This, according to Pellegrino, creates a dilemma: since imprecision undermines altruistic maximization, effective altruists face the choice between embracing a view that has deeply counterintuitive implications or rejecting a central tenet of effective altruism.

Joshua Kissel's 'Effective altruism and anti-capitalism: an attempt at reconciliation' attempts to find common ground between left-wing critiques of effective altruism and effective altruists' distrust of anti-capitalism strategy. Kissel argues that effective altruism and anti-capitalism are compatible both in theory and in practice, and that increased dialogue between these two approaches would result in outcomes that are preferable from each of these perspectives.

Kathryn Muyskens's 'The other half of effective altruism: selective asceticism' agrees with the ideas, central to effective altruism, that we have a duty to make the world a better place and that we should act on this principle by applying our powers of reason so as to make our efforts as effective as possible. Muyskens, however, seeks to emphasize an aspect of effective altruism that she believes has been unduly neglected: effective altruists should consider not only action, but also inaction, as an effective method of achieving social change.

Max Elder and Bob Fischer's 'Focus on fish: a call to effective altruists' argues that the three-factor framework that effective altruists often rely upon to assess rival causes justifies making fish a focus of animal-focused effective altruists to at least the same degree that hens and chickens currently are. The authors claim that the fish killed for food outnumber chickens by about one order of magnitude; that fish have been largely neglected by animal advocates in general and effective animal advocates in particular; and that the comparatively low tractability of fish suffering is more than offset by its much higher scale and neglectedness.

Peter Murphy's 'But does it hurt' explores the relationship between altruism and sacrifice. Murphy argues that, while it is not the case that the morally best acts necessarily involve personal sacrifice, such sacrifice is involved in acts deserving of the highest praise. He thus concludes that self-sacrifice bears on the moral evaluation of agents, but not on the moral evaluation of acts, with which effective altruists are primarily concerned.

Rhys Southan's 'Peter Singer, R.M. Hare, and the trouble with logical consistency' examines Richard Hare's metaethical theory and its influence on Peter Singer. Southan argues that Singer's strategy, inspired by Hare, of trying to persuade others by describing the principles underlying their prescriptions often fails because it attributes to them principles which they do not in fact endorse. Southan thus concludes that effective altruists, who like Singer often rely on consistency arguments, face a more daunting task than they generally believe.