Dominic Roser
Stefan Riedener
Markus Huppenbauer (Eds.)

Effective Altruism and Religion

Synergies, Tensions, Dialogue

PVER VALA ERNG LAGO | Nomos

Religion — Wirtschaft — Politik 32

ZRWP ZENTRUM FÜR RELIGION | WIRTSCHAFT | POLITIK

Effective Altruism and Religion

PVER
VALA
ERNG
LAGO

Religion – Wirtschaft – Politik

—

Schriftenreihe des Zentrums für Religion, Wirtschaft und Politik
Herausgegeben von
Prof. Dr. Martin Baumann, LU / Prof. Dr. Paul Dembinski, FR / Prof. Dr. Gerd Folkers, CH / Prof. Dr. Markus Huppenbauer †, UZH / Prof. Dr. Jens Köhrsen, BS / Prof. Dr. Antonius Liedhegener, LU / Prof. Dr. Jürgen Mohn, BS / Prof. Dr. Wolfgang Müller, LU / Prof. Dr. Georg Pfleiderer (Vorsitz), BS / Prof. Dr. Konrad Schmid, UZH / Prof. Dr. Jörg Stolz, LAU

—

Band 23

Dominic Roser | Stefan Riedener
Markus Huppenbauer (Eds.)

Effective Altruism and Religion

Synergies, Tensions, Dialogue

PVERVALAERNGLAGO | Nomos

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.d-nb.de

ISBN	978-3-290-22067-9:	Pano Verlag (Print)
	978-3-290-22068-6:	Pano Verlag (ePDF)
	978-3-8487-8119-5:	Nomos Verlag (Print)
	978-3-7489-2536-1:	Nomos Verlag (ePDF)

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

ISBN	978-3-290-22067-9:	Pano Verlag (Print)
	978-3-290-22068-6:	Pano Verlag (ePDF)
	978-3-8487-8119-5:	Nomos Verlag (Print)
	978-3-7489-2536-1:	Nomos Verlag (ePDF)

Library of Congress Cataloging-in-Publication Data
Roser, Dominic | Riedener, Stefan | Huppenbauer, Markus;
Effective Altruism and Religion
Synergies, Tensions, Dialogue
Dominic Roser | Stefan Riedener | Markus Huppenbauer (Eds.)
254 pp.
Includes bibliographic references.

ISBN	978-3-290-22067-9:	Pano Verlag (Print)
	978-3-290-22068-6:	Pano Verlag (ePDF)
	978-3-8487-8119-5:	Nomos Verlag (Print)
	978-3-7489-2536-1:	Nomos Verlag (ePDF)

1st Edition 2022

Published by
Nomos Verlagsgesellschaft mbH & Co. KG
Waldseestraße 3–5 | 76530 Baden-Baden
www.nomos.de

Production of the printed version:
Nomos Verlagsgesellschaft mbH & Co. KG
Waldseestraße 3–5 | 76530 Baden-Baden

ISBN	978-3-290-22067-9:	Pano Verlag (Print)
	978-3-290-22068-6:	Pano Verlag (ePDF)
	978-3-8487-8119-5:	Nomos Verlag (Print)
	978-3-7489-2536-1:	Nomos Verlag (ePDF)

DOI https://doi.org/10.5771/9783748925361

Onlineversion
Nomos eLibrary

Lara Buchak

Foreword

When I first heard about effective altruism, I assumed it was a Christian movement. Followers of effective altruism were trying to put into action the commandment to love their neighbour, or trying to abide by Jesus's words to the rich man: "If you wish to be perfect, go, sell your possessions, and give the money to the poor."[1] To my surprise, not only was effective altruism not primarily a Christian movement, but many Christians seemed suspicious of it.

As discussions of effective altruism came up, two worries were voiced most frequently among the Christians with whom I spoke. The first was that the focus on being "effective" – on saving the most total lives or on maximising the lives saved per dollar – reduces humanity to a mass to be weighed and measured, leaving no room to love one's neighbour as an individual made in the image of God: an individual who deserves our attention regardless of the cost of helping him. The second was that some of the more "fringe" elements of the movement, focused on extending human life indefinitely or colonising other planets, located the salvation of the world in human progress and a future utopia, rather than in something less bound in temporal existence. Not human enough, and too human.

I am a Christian, so I am particularly attuned to the reception of effective altruism among Christians. But I suspect that those from other religious traditions have had similar experiences. Effective altruism can initially seem like a movement that *embodies* their religious commitments, but their co-religionists turn out to be suspicious of it.

Effective altruists do not appear all that impressed with religion, either. The vocal majority of those involved in the effective altruist movement are non-religious, some having explicitly left the religious tradition of their youth. And they have worries about religious practitioners. Some simply worry that religious people are not particularly prone to thinking through things rationally—that they prefer tradition, authority, or plain old superstition to evidence-gathering. Others worry that the religious focus on spiritual things distracts from meeting the immediate and pressing needs

1 Matt. 19:21, New Revised Standard Version.

of food, shelter, and health; or that the focus on eternal things leads to complacency about temporal suffering.

Both religious commitment and effective altruism demand a singular focus. They both demand that one keep a particular aim at the forefront of one's mind, and make the bulk of one's life decisions with this aim in view. And they each can see the other as a competitor for that singular focus. As we know, you can only serve one master.

But, curiously, religious commitment and effective altruism are united in telling us we should not serve *mammon*. They are united in claiming that the ordinary, 21st-century American and Western European way of living has gone *drastically* wrong, and that we need to create a different way of living from the ground up. They are united in thinking that people who are not part of our everyday social group should occupy a *much larger* part of our concern. They are united in thinking that our focus should be on others rather than on ourselves, not just part of the time, but as a way of life.

So it seems that we ought to rethink the relationship between religious commitment and effective altruism; and that is just what the essays in this volume aim to do. While there have been some notable volumes addressed to religious audiences urging them to be both more altruistic (e.g., Ronald Sider's *Rich Christians in an Age of Hunger*) and more effective (e.g., Bruce Wydick's *Shrewd Samaritan*), nothing has been written directly on the relationship of religious commitment and effective altruism as a distinct movement that goes by that name.

The essay writers are commended not only for their insights, but for framing the questions and shaping the discussion, since they are writing against a background of very little that has come before. Dominic, Markus, and Stefan are especially commended for bringing together a volume on this topic. While volumes are often praised for moving the conversation forward, this one does something much more difficult, for it begins an entirely new conversation, one that I hope will continue.

Table of Contents

Dominic Roser and Stefan Riedener

Introduction

1. Effective Altruism and Religion: An Intriguing Encounter

The effective altruism (EA) movement matters. In the past decade, its adherents have put forth an ever increasing number of challenging ideas about how to improve the world the most. They have set about redefining our understanding of the most ethical life for individuals and the most urgent priorities for humanity. But EA is not just new fodder for academic debates or a further addition to a long line of ideologies offering intellectual entertainment. After only a few years of existence, and despite comprising only a couple of thousand people,[1] it has also already left a significant real-world footprint. One reason for this is that a number of highly influential actors have been influenced by its ideas. Bill Gates called William MacAskill, one of the movement's co-founders, "a data nerd after my own heart".[2] Sam Bankman-Fried – believed to be the richest person under thirty[3] – became wealthy precisely in order to promote effective altruist aims. And institutions like the World Health Organization, the World Bank or the UK Prime Minister's Office have been influenced by advice of Toby Ord, another co-founder of the movement. Another reason for EA's real-world effect is the simple fact that its distinguishing feature is a radical focus on impact. So even those of its adherents who are not global players have had remarkable leverage. The movement started out with a focus on channelling money towards poverty eradication. But it soon broadened into a more general project of improving the world as smartly and impartially as possible. And it did not fail to live up to its aim of "using evidence and reason to find the most promising causes to work on, and taking action […] to do the most good."[4]

The EA movement clearly has a secular character. When leaders of the movement state its core project, or articulate reasons for pursuing it, they rarely ever put forward explicitly religious claims. In a recent survey, 86

1 Moss, "EA Survey 2019."
2 Effective Altruism, "Doing Good Better."
3 Chan, "Hong Kong's 29-year-old crypto billionaire."
4 Effective Altruism, "Effective altruism is about doing good better."

percent of the members of the EA community reported being non-religious, agnostic, or outright atheist.[5] Indeed Peter Singer, the movement's father figure, has historically been met with profound opposition from religious quarters.[6]

At the same time, the world of charity – or more generally, of people for whom improving the world is a core part of their identity – has long been heavily populated by religious actors. Sometimes the religious background is very explicit. On other occasions it serves less visibly as the underlying motivation of individuals or the historical root of organisations. But it has often been, or still is, important for many. So with the rise of EA, a distinctively secular movement entered a territory that has long been, and still is, importantly shaped by religion. And this overlap has in some sense become even stronger when many effective altruists have become committed to longtermism: the idea that what matters most, today, is setting a good trajectory for the very long-term future. Themes like the end of humanity or a future radical utopia in particular have not always received intense attention outside of religious circles.

From the perspective of traditional, and often religiously influenced charitable initiatives, EA can thus have felt like the new kid on the block. But that kid has come of age rapidly, both intellectually and practically, and must now be taken seriously indeed. It calls for examination and discussion. First and foremost, there's a question about what people of different religions can learn from EA. EA is intellectually and sociologically rooted in very different soil than the charitable efforts of many people of faith. But many of its insights on questions of efficiency are independent from specific conceptions of the good: they should be relevant for a broad array of worldviews. It would thus seem natural that EA's fresh take has lessons to offer. At the same time, there's a question whether people of different religions should repudiate some EA stances as unimportant – or even as outright wrong and dangerous.

But religious perspectives on EA go far beyond the simple questions of what people of faith should accept or reject from the movement. They comprise a whole array of fascinating issues. In particular, there's the opposite question as well. Can EA learn some important lessons from religion? Can religious traditions offer interpretations of, or justifications for, doing good that EA has so far ignored? Have religions produced conceptual resources,

5 Dullaghan, "EA Survey 2019 Series."

6 This opposition primarily concerned Singer's views in bioethics, not his stance about global poverty.

or practical experience, that EA could helpfully adopt? Or should the adoption of EA in religious communities be seen as dangerous from EA's perspective? Finally, may EA in some sense be seen as a quasi-religious movement itself, considering how comprehensively life-orienting it is for some of its adherents?

At a practical level, the intersection between EA and religion is already a lively sphere: there has been an active community of Christian effective altruists for a number of years now, a Facebook group for Buddhists in EA, as well as an effective altruist initiative for Jews.[7] Astonishingly, however, at a theoretical level, hardly anything has been published on these wide-ranging questions. Aside from a small number of broader discussions about utilitarianism, Peter Singer, and religion,[8] the body of academic work on EA and faith appears to consist of no more than a handful of articles.[9] This, it seems to us, is not enough.

2. The Structure and Content of the Book

This book works toward filling this lacuna and getting the ball of discussion rolling. It consists of three blocks of chapters. The book opens with three contributions that provide an assessment of EA from a specific religion's perspective: Calvin Baker discusses Buddhism and EA, David Manheim considers an Orthodox Jewish perspective, and Dominic Roser elaborates on EA and Christianity. The middle block of four chapters then discusses EA in general but with a narrower perspective than a whole religion: Mara-Daria Cojocaru focuses on the type of love at stake, Jakub Synowiec on who counts as a neighbour, Stefan Höschele uses the lens of Relational Models Theory to compare EA and Christianity, and Kathryn Muyskens focuses on asceticism and activism. The book closes with three chapters that each look at a specific theme in EA from a religious perspective: Stefan Riedener examines existential risks from a Thomist perspective, Robert MacSwain explores the question of moral ambition and sainthood, and Markus Huppenbauer discusses donations.

7 https://www.eaforchristians.org; https://www.facebook.com/groups/buddhists.in.ea; https://eaforjews.org.

8 See in particular Camosy, *Peter Singer and Christian Ethics*, and Perry, *God, The Good, and Utilitarianism*.

9 See Liberman, "Effective Altruism and Christianity"; Miller, "80,000 Hours for the Common Good"; Gregory, "Charity, Justice, and the Ethics of Humanitarianism"; Chukwuma et al., "An Evaluation of the Concept of Effective Altruism."

Let us give a slightly more detailed summary of the chapters in this book. Baker begins his examination of Buddhism and EA by noting similarities between the two outlooks: they are akin to each other when it comes to such central commitments as impartially promoting the welfare of all sentient beings. He goes on to argue, however, that Buddhism would significantly diverge from EA on how to most effectively help others: it involves a radically different conception of ourselves and our place in reality – in particular, an idea of the ultimate good as quitting the cycle of rebirth. Nonetheless, Baker suggests that there can be a productive dialogue between EA and Buddhism, and he ends with a couple of constructive insights to this effect.

Manheim situates a number of EA tenets within the context of ancient and contemporary Orthodox Jewish debates: the moral obligation to help, consequentialist reasoning about altruism, cause prioritisation, and the use of reason and evidence to understand effectiveness. He suggests that Orthodox Judaism, with its unyielding emphasis on the Halacha norms, is not compatible with the complete framework of EA. Still, EA is not irrelevant to it. In particular, Halacha is engaged with complex questions about charitable giving, and thus EA insights matter – not to guide or change Halacha, but to inform it.

Roser characterises EA in terms of seven core commitments, and examines whether Christians can share these commitments. His verdict is very positive. The core EA commitments are not only compatible with Christianity, but novel and useful tools for living out Christian faith. So Christians ought to take up many of EA's insights. However, Roser also mentions a tension between the EA mindset and Christian faith: while EA is concerned with taking control and actively shaping the world in accordance with our values, a core thread in Christianity encourages us to renunciate control and to place ourselves trustingly in God's hands.

Cojocaru focuses on the "heart" in EA, or the concept of "love" that is central to it and to many religions. Building on Iris Murdoch, she distinguishes two spheres of morality: a public and a private one. In the former, agents operate on simple, uncontroversial ideas of the good, and utilitarian norms seem relevant. In the latter, however, much more complex conceptions of the good become pertinent, and those will only be detectable through really looking at the particulars of another person or a relationship. Different kinds of love operate at these different spheres. EA, Cojocaru claims, often ignores that humans need partial relations and perspectives in order to learn what is good.

Synowiec starts from the Christian imperative to love our neighbours. Along with other authors in this volume, he brings up the parable of the Good Samaritan and focuses on its core question: "who is my neighbour?" On standard interpretations of EA, the relevant neighbours would be all beings with interests – regardless of species, geographical distance, or temporal distance. Can Christianity share this understanding? Synowiec proposes a biblical interpretation according to which "neighbours" are persons that we can personally affect. He argues that, given the characteristics of our times, this includes all contemporary people. Animals are not exactly neighbours, but we have sufficient knowledge and power to treat them as such. Far future people are not neighbours either – and indeed, we have neither the knowledge nor the power to treat them as if they were.

Höschele discusses EA and Christian ethics through the framework of Relational Models Theory. According to this theory, there are four "elementary forms" of human sociality: four models of human interaction, governed by different norms. So Höschele asks which kinds of models, or human interaction, EA and Christian morality envision. He concludes that EA and New Testament ethics largely agree on the key element that characterises moral actions: the kind of love that values the other as much as the own person. Still, effective altruists and Christians can learn from each other, challenge each other on blind spots, and together steer philanthropy to appropriate levels of reflection and action.

Muyskens suggests that EA needs a kind of ascetisicm. She argues that charitable donation is not enough to stop systemic inequality or structural violence. Indeed, a focus on "charity" may even contribute to such injustice. And the cost-benefit analysis and randomised controlled trials favoured by the movement can produce distinctly biased perceptions of harms. So the traditional focus of EA on charity has problematic aspects. As a remedy, she argues, EA needs an ascetic type of action tackling systemic injustice, addressing the roots of the problem more directly.

Riedener examines EA's focus on reducing risks of human extinction, and asks whether such a focus can be justified within a Thomist moral framework. He argues that it can: Thomas's idea of the human end, his emphasis on the virtue of humility, and his conception of the place of humanity in the cosmos imply that anthropogenic extinction would be a tremendous moral disaster – a cosmologically important prideful failure to fulfil our God-given role. And this, Riedener suggests, should not only be relevant for Christians quite generally: similar thoughts also emerge on non-religious worldviews, based e.g. on the import of human dignity.

MacSwain discusses the relationship between EA, supererogation and sainthood. He begins with a discussion of supererogation in Singer, Urmson and Wolf – suggesting that according to Singer, effective altruists are not saints because their actions are often not supererogatory. He then argues for a reconsideration of Robert Merrihew Adams's notion of "real saints", as people who follow their own vocations. This suggests that some, but not all effective altruists are saints – just like some, but not all, non-EA people are saints.

Finally, Huppenbauer explores charitable giving. He looks at the motives for which people donate, endorses the moral obligation to give, and examines questions about when, how, and how much we should donate. He then considers two challenges to the present culture of donating: EA and a movement advocating social investment instead of donation. Concerning the former, he articulates a worry: if improving the world is such a dominant concern as it is for many effective altruists, and if there is no understanding of a good life outside of morality, people threaten to become "morality machines" – and to thus miss the meaning of life.

If there is anything like a common thread through these essays, then perhaps it is this: the relation between religions and EA is not without tensions, contradictions and differences. But in spite of this, or perhaps precisely because of it, a deepened dialogue will be mutually beneficial. Beyond this shared thread, we find that the essays also manifest a beautiful diversity: they are written from different backgrounds in religion, theology or philosophy, ask different questions and come to different conclusions. Thus they illustrate the richness of our topic. And they certainly still only cover a small part of the questions that emerge at the intersection between religion and EA. In particular, the majority of them still focus on Christianity. It is our firm hope that this collection is the beginning of a much larger story – the story of a more extensive and more serious engagement of religious people of all kinds with EA's ideas and practices.

3. The Book's Story

This book is not just the beginning of one story. It is also the end of another. This latter story had its bright and its sad moments. It started out in the late summer of 2019 with an inspiring workshop at the University of Fribourg entitled "Religious Perspectives on Effective Altruism". The present book is the outcome of this workshop: all the chapters, except for Calvin Baker's, were presented there.

On a warm summer day one year later, Markus Huppenbauer, the workshop's co-organiser and the book's co-editor, unexpectedly passed away. This was a tremendous shock. It is difficult to see how others could fill Markus's footsteps. While we, the remaining co-editors, occasionally disagreed profoundly with him on the promise of EA, our discussions about it were always delightful and informative. Markus had a uniquely generous and cheerful way of engaging with his interlocutors. And he always displayed an unflinching desire to push the debate forward, a strong commitment to making room for all kinds of viewpoints, and a remarkable boldness in taking up unpopular stances himself. Markus's plan was to revise the public talk he gave at the 2019 workshop for the purpose of this book. Sadly, he couldn't implement this plan anymore. Upon reflection, we decided to publish a translated transcript of his talk, and ask readers to keep in mind that he did not get the chance to edit and polish it anymore. We miss Markus dearly, and find that his talk – which seems characteristic of his style, attitude, and position – serves as a fitting memory.

This book has profited a lot from the generous efforts of many people. We are particularly grateful to Ludovico Conti, Véronique Dupont, and Joe Tulloch who got the manuscript in good shape. We would also like to thank Aryeh Englander, Elie Hassenfeld, Caleb Huffman, Frances Kissling, and Ben Schifman for their substantive inputs as well as Sarah Kirkby and Arianna Lanfranchi for their work on Markus Huppenbauer's chapter. We are indebted to Beate Bernstein from Nomos who skilfully led the book through the publication process. Without the funds from the Center for Religion, Economy and Politics and the University of Fribourg this whole project would not have been feasible in the first place. We would like to express our appreciation for their support.

We hope that the efforts of all these people and institutions – and particularly of the authors – will prove fruitful. May ensuing discussions not overemphasise differences between religions and EA. Rather, may all sides collaborate productively and take up insights from each other. The vision of a much better world inspires many effective altruists and people of faith alike. The beauty and importance of this goal, and the fact that at bottom it is shared, mean we should all listen to each other.

References

Camosy, Charles. *Peter Singer and Christian Ethics: Beyond Polarization*. Cambridge: Cambridge University Press, 2012.

Chan, Michelle. "Hong Kong's 29-year-old crypto billionaire: FTX's Sam Bankman-Fried." *Nikkei Asia*, 25 June 2021. https://asia.nikkei.com/Business/Business-Spotlight/Hong-Kong-s-29-year-old-crypto-billionaire-FTX-s-Sam-Bankman-Fried

Chukwuma, Joseph Nnaemeka, Tobias Chukwuemeka Ozioko, Obiora Anichebe, Gabriel Chukwuebuka Otegbulu, Charles Kenechukwu Okoro, George Ohabuenyi Abah, Collins Ikenna Ugwu, and Anthony Chukwudi Areji. "An Evaluation of the Concept of Effective Altruism in Peter Singer." *International Journal of Mechanical and Production Engineering Research and Development* 10, no. 3 (2020): 14803–14816.

Dullaghan, Neil. "EA Survey 2019 Series: Community Demographics & Characteristics." Rethink Priorities. 5 December 2019. https://www.rethinkpriorities.org/blog/2019/12/5/ea-survey-2019-series-community-demographics-amp-characteristics.

Effective Altruism. "Effective altruism is about doing good better." Accessed 30 September 2021. https://www.effectivealtruism.org/.

Effective Altruism. "Doing Good Better." Accessed 30 September 2021. https://www.effectivealtruism.org/doing-good-better/.

Gregory, Eric. "Charity, Justice, and the Ethics of Humanitarianism." In *Everyday Ethics: Moral Theology and The Practices of Ordinary Life*, edited by Michael Lamb and Brian A. Williams, 81–102. Washington, DC: Georgetown University Press, 2019.

Liberman, Alida. "Effective Altruism and Christianity: Possibilities for Productive Collaboration." *Essays in Philosophy* 18, no. 1 (2017): 6–29.

Miller, Ryan. "80,000 Hours for the Common Good: A Thomistic Appraisal of Effective Altruism." *Proceedings of the American Catholic Philosophical Association* (forthcoming).

Moss, David. "EA Survey 2019: How many people are there in the EA community?" Rethink Priorities. 26 June 2020. https://rethinkpriorities.org/publications/eas2019-how-many-people-are-there-in-the-ea-community.

Perry, John (ed.). *God, the Good, and Utilitarianism: Perspectives on Peter Singer*. Cambridge: Cambridge University Press, 2014.

Calvin Baker

Buddhism and effective altruism

Abstract

This article considers the contemporary effective altruism (EA) movement from a classical Indian Buddhist perspective. Following barebones introductions to EA and to Buddhism (sections one and two, respectively), section three argues that core EA efforts, such as those to improve global health, end factory farming, and safeguard the long-term future of humanity, are futile on the Buddhist worldview. For regardless of the short-term welfare improvements that effective altruists impart, Buddhism teaches that all unenlightened beings will simply be reborn upon their deaths back into the round of rebirth (*saṃsāra*), which is held to be undesirable due to the preponderance of *duḥkha* (unsatisfactoriness, dis-ease, suffering) over well-being that characterizes unenlightened existence. This is the *saṃsāric* futility problem. Although Buddhists and effective altruists disagree about what ultimately helps sentient beings, section four suggests that Buddhist-EA dialogue nonetheless generates mutually-instructive insights. Buddhists – including contemporaries, such as those involved in Socially Engaged Buddhism – might take from EA a greater focus on explicit prioritization research, which seeks knowledge of how to do the *most* good we can, given our finite resources. EA, for its part, has at least two lessons to learn. First, effective altruists have tended to assume that the competing accounts of welfare converge in their practical implications. The Buddhist conception of the pinnacle of welfare as a state free from *duḥkha* and, correspondingly, the Buddhist account of the path that leads to this state weigh against this assumption. Second, contrasting Buddhist with effective altruist priorities shows that descriptive matters of cosmology, ontology, and metaphysics can have decisive practical implications. If EA wants to give a comprehensive answer to its guiding question – "how can we do the most good?" – it must argue for, rather than merely assume, the truth of secular naturalism.

Introduction

This article addresses the following question: What perspective would Indian Buddhist philosophy take on effective altruism (EA)? EA is a young social movement that seeks to discover how we can maximise our altruistic impact and to put its discoveries into practice. In articulating an Indian Buddhist perspective on EA, we will focus on Indian Buddhist philosophy from approximately the first through eighth centuries CE, which corresponds to what Jan Westerhoff has recently described as the golden age of Indian Buddhist thought.[1]

1 Westerhoff, *The Golden Age*.

Neither Indian Buddhist philosophy (henceforth, "Buddhism")[2] nor EA is monolithic in its outlook. Both, however, are centred on core practices and commitments, which makes it possible to use the phrases "Buddhism" and "EA" meaningfully. When possible, I will conduct the discussion in terms that all schools of Buddhism would accept, and likewise for all branches of EA. When this is not possible, I will make it clear where the schools and branches diverge and what the implications of these divergences are.

Before outlining the article, it is worthwhile to motivate our guiding question. Why would we care what view an ancient philosophical tradition might take of a contemporary social movement? One reason is that inter-traditional philosophical dialogue can generate mutually-instructive insights. A second reason is that EA has tried to position itself as a movement whose aims are endorsable – and perhaps even required – by a wide range of ethical positions. Thinking carefully about how Buddhism would evaluate EA is one way to put this claim to the test. Third, there are several interesting *prima facie* similarities between Buddhism and EA. Each is centrally concerned with promoting the welfare of moral patients, which for both saliently includes, but is not necessarily limited to, alleviating suffering. The scope of welfare promotion is also similarly broad for each. Buddhism and most in EA agree that moral patienthood extends to all sentient beings. Regarding which sentient beings to benefit, EA is strongly impartial, and prominent strands of Buddhist thought point in this direction as well. Finally, some contemporary Buddhist practitioners (though, to be clear, not classical Indian Buddhists) believe that Buddhism and EA are kindred spirits when it comes to helping others. After a public conversation with Peter Singer, one of EA's major philosophical proponents, Matthieu Ricard, a Western-scientist-turned-Tibetan-Buddhist-monk and author of *Altruism* (2013), concluded that there is "no fundamental difference" between the stances he and Singer take on altruism.[3]

Despite these *prima facie* similarities, I will argue that Buddhism significantly diverges from EA in its practical and theoretical approach to altruism. The article proceeds as follows: sections one and two respectively give barebones introductions to EA and Buddhism. With this background in place, section three articulates a critical Buddhist perspective on EA.

2 I refer to Indian Buddhist philosophy as "Buddhism" only for the sake of brevity. I am not suggesting that Buddhism is reducible to philosophy or that non-Indian schools are ingenuine expressions of the tradition.

3 Matthieu Ricard, "Altruism Meets Effective Altruism."

Section four concludes with insights that Buddhism and EA might take from the dialogue.

1. Effective Altruism

I will follow William MacAskill, who co-founded EA with Toby Ord, in understanding the movement as devoted to a bipartite project.[4] The first part of the project is to make rigorous use of evidence and reason to discover how to maximise the good, given finite resources, without violating any side-constraints like human rights.[5] ("Resources" denotes anything that can be permissibly utilised to promote the good, such that the term refers not only to financial assets but also, e.g., to hours of research.) The good is provisionally equated with the welfare of moral patients, considered impartially. The second part of the project is to practically apply the conclusions of the first part. The cause areas on which EA has primarily focused so far include global health and poverty, nonhuman animal welfare (especially factory farming), the longterm future of humanity (especially existential risks), and global priorities research (research devoted to the first part of the EA project).

I will also discuss a set of normative principles that I take to motivate EA's bipartite project. I include the set for two reasons. First, it is plausible that social movements require guiding normative commitments to be distinctive *qua* movements[6] and, more fundamentally, to *be* social movements at all.[7] Second, I believe that most EAs would endorse the principles and that their conjunction justifies and explains characteristic EA behaviour. Since people participate in social movements and undertake substantive projects for (perceived) normative reasons, and since movements and projects are subject to normative assessment, including a set of motivating principles deepens our understanding of EA. I base the first three principles closely on those proposed by Berkey and by Crisp and Pummer;[8] the fourth is my own contribution. The principles are as follows:

> *Strong Welfare Promotion*: we have reason to promote the welfare of all moral patients, and this reason is sometimes, though not always, practically overriding.

4 MacAskill, "The Definition of Effective Altruism."
5 Pummer and MacAskill, "Effective Altruism."
6 Berg, "How Big Should the Tent Be?"
7 Berkey, "The Philosophical Core of Effective Altruism."
8 Berkey, "The Philosophical Core of Effective Altruism"; Crisp and Pummer, "Effective Justice."

Impartial Maximisation: all else equal, when we are acting on the reason to promote welfare, we should impartially maximise the amount of welfare we bring about per unit of resource input.

Methodological Rigour: a rigorous evaluation of the relevant evidence, broadly construed, should exclusively inform our attempts to promote welfare.

Weak Normative Uncertainty: in general, we should avoid basing our normative outlook exclusively on one ethical theory and instead be open to insights from multiple plausible theories. In particular, we should avoid behaviour that is seriously wrong according to common-sense morality, such as violating rights, even if such behaviour would impartially maximise welfare.

Since the inclusion of Weak Normative Uncertainty is the chief way in which my account differs from others in the literature, I would like to motivate the principle before moving on. There are at least two reasons for taking Weak Normative Uncertainty as a core principle of EA. First, on the descriptive level, there is widespread support within EA for taking normative uncertainty seriously. For instance, MacAskill and Ord have published extensively on normative uncertainty[9] and promulgated their views within EA,[10] with the result that the Centre for Effective Altruism includes moral uncertainty as a key concept in its primer on EA topics,[11] "moral uncertainty and moderation" is a guiding value of 80,000 Hours,[12] and "worldview diversification" is central to Open Philanthropy Project's grant-recommendation strategy.[13] Second, on the conceptual level, Weak Normative Uncertainty explains and justifies EA's respect for side-constraints (which may otherwise appear *ad hoc* on an impartial, welfare-maximising framework); agnosticism about what welfare consists in; openness to the possibility that goods other than welfare are worthy of promotion;[14] and interest in "moral circle expansion", i.e., in identifying entities that are not

9 For a comprehensive overview, see MacAskill et al., *Moral Uncertainty*.

10 See e.g. Wiblin and Harris, "Our descendants will probably see us as moral monsters"; and Ord, *The Precipice*, 213.

11 See https://concepts.effectivealtruism.org/concepts/moral-uncertainty/. The Centre for Effective Altruism is responsible for supporting and growing the movement.

12 See Todd and the 80,000 Hours team, "A guide to using your career." One of the most public-facing EA organisations, 80,000 Hours primarily advises early-career professionals on how to do the most good through their careers.

13 See e.g. Karnofsky, "Worldview Diversification" and "Update on Cause Prioritization." Open Philanthropy is an EA-aligned research and advisory organisation that *de facto* conducts the grant-making of Good Ventures, a philanthropic foundation with potential assets of $14 billion (MacAskill, "The Definition of Effective Altruism").

14 For these first three aspects of EA, see MacAskill, "The Definition of Effective Altruism."

popularly considered to be moral patients but in fact ought to be (EAs have discussed, e.g., insects and intelligent machines).[15] I believe these descriptive and conceptual considerations justify the inclusion of Weak Normative Uncertainty in the set of core EA principles.

2. The Buddhist Worldview

With the basics of EA on the table, we can proceed to a thumbnail sketch of the Buddhist worldview. On this view, all sentient beings exist within *saṃsāra*, the round of rebirth (lit. "wandering"). Beings are repeatedly reborn into the various realms of *saṃsāra*, such as the human, heaven, and hell realms, in accordance with their *karma*, which accrues to them in dependence on the ethical quality of their behaviour. There is, however, no purpose or meaning to *saṃsāra*. *Saṃsāra* is not progressing towards any goal and is not controlled by any creator deity or other intentional guiding force.

The fundamental problematic of life in *saṃsāra* is *duḥkha*. Numerous translations of "*duḥkha*" have been proposed, including dis-ease, unsatisfactoriness, and suffering, but since no English term captures its full sense, I will leave it untranslated. The problem of *duḥkha* is that it is predominant in *saṃsāra*. Taken on the whole, *saṃsāric* existence is characterised by a preponderance of *duḥkha* over whatever conventional goods may be found within *saṃsāra*, such as transient pleasures or successes.

The reason *duḥkha* predominates in *saṃsāra* is that beings are subject to a profound delusion (*avidyā*) about the ultimate nature of reality. Automatically and unconsciously, we perceive and think of ourselves and the objects in the world as substances, by which I mean real, independently-existing things that possess essences, bear properties, and endure diachronically. According to Buddhist ontology, in contrast, there are no substances. Instead, Buddhism holds that reality is constituted by impersonal, evanescent events and ever-becoming processes. The metaphysical picture that emerges is one of interdependence and thoroughgoing impermanence. The cause of *duḥkha*, Buddhism teaches, is the discrepancy between how we perceive and think about reality and how reality actually is – most importantly, the discrepancy between our delusion that we are substantial selves who are the subjects of our experiences and the agents of our actions and

15 See e.g. the work of Sentience Institute, an EA-aligned research organisation (https://www.sentienceinstitute.org/).

the truth that there is no self, but rather a causally interrelated sequence of impersonal physical and mental events.[16]

The final soteriological goal of Buddhism is to eliminate *duḥkha* by aligning our perception and cognition with the fundamental nature of reality, *viz.* the impermanent, substanceless process ontology just indicated. To accomplish this task, it is not sufficient to update our consciously-held beliefs via philosophical reflection. Rather, to transform our fundamental experience of ourselves and the world – which includes ceasing to perceive the world from the perspective of a substantial self – Buddhism holds that we must embark on its Ennobling Eightfold Path. Among other things, the Eightfold Path requires sustained meditative effort and cultivation of virtues such as loving-kindness (*maitrī*) and compassion (*karuṇā*) to effect the desired change in our perceptual and intentional orientation to the world. If followed to its end, the Buddhist path purportedly culminates in a complete awakening (*bodhi*, alt. trans. "enlightenment") to the ultimate nature of reality, which results in cessation (*nirvāṇa*, lit. "extinguishing") – specifically, the cessation of the volitions and actions that arise from delusion and cause *duḥkha* – and, thereby, liberation from the round of rebirth.

3. Buddhism and Effective Altruism

3.1. Altruism in Practice

With this background in place, I will argue that Buddhism would significantly diverge from EA on the question of how to most effectively help others.[17] I trace the divergence in part to an axiological disagreement. The ideal state of affairs, according to Buddhism, is, in principle, one in which all sentient beings attain awakening and thereby transcend *duḥkha* and *saṃsāra*. EA does not share with Buddhism the final end of liberating all sentient beings from a cycle of rebirth. This difference in final ends results in a corresponding difference in positions on (lowercase) effective altruism.

16 See Panaïoti, *Nietzsche and Buddhist Philosophy*, for an explanation of the psychological mechanics involved in the generation of *duḥkha*.

17 In this section, I am simply asking what perspective Buddhism would take on EA's efforts to help others, *qua* efforts to maximise expected total welfare. Part of thoroughly articulating this perspective is developing a Buddhist response to the question of how we can most effectively help others. I am not taking a stance in this section on whether Buddhism would find this question an interesting or practically important one. Consideration of whether Buddhists should find the topic of efficient welfare maximisation practically important, given their other commitments, will come in section 3.2.

The divergence also traces to differences in cosmology and ontology. Buddhism, with its literal belief in rebirth operative within *saṃsāra*, a complex of world-systems filled with heavens, hells, and supernatural beings, stands opposed to EA, which does not accept rebirth and opts instead, at least in large majority, for a secular, naturalist outlook.[18] These radically different conceptions of ourselves, reality, and our place in it result in radically different conceptions of how we can most effectively help others.

3.1.1. Saṃsāric Futility and the Bodhisattva Path

To see how differences in axiology and cosmology lead to divergences in altruistic practice, we can begin by contrasting EA's approaches with what is perhaps the most salient Buddhist approach: directly instructing others in Buddhist practice. To understand why this strategy is attractive from a Buddhist standpoint, imagine that we instead attempt to promote welfare by undertaking a humanitarian project, such as the distribution of anti-malarial bed nets (a long-time EA favourite).[19] On Buddhist assumptions, the limitation of such a project is that regardless of the extent to which it is successful at improving someone's present life – which indeed it might, for Buddhism accepts that there are relatively better and worse positions to inhabit within *saṃsāra* – the aid recipient will simply be reborn and again face the problem of *duḥkha* in their next life.

Similar points can be made about other "worldly" efforts to do good, EA-endorsed or not. Take, for example, activists' efforts to achieve economic or social justice by reforming basic social institutions. Even if these efforts were entirely successful from the perspective of the activists, from a Buddhist perspective, they would be of limited utility. A more just Earth would be a better world to inhabit for the beings who happen to be reborn on it as humans, and in this sense, the activists' efforts would have increased welfare. But from a macroscopic Buddhist viewpoint, greater justice on Earth is a relatively trivial improvement if the beings who temporarily inhabit the planet are simply reborn elsewhere in *saṃsāra* upon their deaths.

18 85.9% of respondents to the 2019 EA Survey identified as atheist, agnostic, or non-religious (Dullaghan, "EA Survey 2019 Series: Community Demographics & Characteristics").

19 NGOs focused on distributing long-lasting insecticide-treated nets to prevent malaria have consistently ranked at the top of the list of most effective charities maintained by GiveWell, a popular EA-aligned charity evaluator (see https://www.givewell.org/charities/top-charities).

At the limit, even if our descendants were to succeed in creating a worldly utopia, the problems of *duḥkha* and rebirth would remain. Our utopia would be analogous to the various heavenly realms of Buddhist cosmology, which are found lacking on account of the facts that the gods who inhabit them are eventually reborn (no longer as gods) into other realms and that the heavenly realms themselves, being parts of *saṃsāra*, are impermanent and ultimately subject to dissolution. We thus encounter what I will call the *saṃsāric* futility problem: all altruistic actions that address neither the fundamental problem of existence (*duḥkha*) nor its ultimate cause (delusion) and fail to advance the solution to that problem (Buddhist awakening) are, in the long run, ineffective. The conclusion we are forced to by the internal logic of Buddhism is that the best way to help beings in *saṃsāra* is to help them out of *saṃsāra*. That is, we can most help sentient beings by helping them to progress along the path to awakening. The practical question then becomes how we can help others to make this progress.

Perhaps the most salient answer to this question in the Buddhist tradition is that we can help others along the path directly, i.e., by offering them instruction in Buddhist practice. And, the thought goes, we can offer this instruction most effectively when we ourselves are advanced practitioners, who possess a deep, experience-based understanding of Buddhist praxis. That offering direct instruction is considered to be a highly effective mode of altruism is evidenced by two foundational elements of Mahāyāna Buddhism: first, the concept of *bodhicitta*, the altruistic aspiration to attain awakening for the sake of all sentient beings; and second, the closely-related ultimate ethical ideal of the *bodhisattva*,[20] a being who has attained (or nearly attained) awakening but, rather than passing out of *saṃsāra* upon the death of their physical body, elects to remain in *saṃsāra* and promulgate the *Dharma* – Buddhist teaching – until all sentient beings have themselves attained awakening.[21]

20 Theravāda Buddhism, the other main branch of Buddhism practised today alongside the Mahāyāna, also recognises the validity of the *bodhisattva* path. But whereas it is definitive of Mahāyāna practice to aim for the *bodhisattva* ideal, this ideal is seen in the Theravāda as a supererogatory option that, in practice, only the most zealous and promising candidates should pursue.

21 To avoid ambiguity, I will use "*bodhisattva*" to refer to an awakened or nearly awakened being on the *bodhisattva* path and "*bodhisattva*-in-training" to refer to someone who has taken the *bodhisattva* vows but is not yet advanced on the path.

One possible objection to the *saṃsāric* futility analysis stems from an observation about the *bodhisattva* path. I will argue that this observation provides an important qualification to the analysis, but does not constitute a counterexample. Here is the objection: canonical depictions of the *bodhisattva* path – e.g. the Jātaka tales, legends of the Buddha's rebirths prior to his awakening on Earth – often feature the *bodhisattva*-in-training performing decidedly worldly acts of altruism. These acts do not (in any clear way) advance the recipient's ultimate soteriological prospects, but rather improve their station within *saṃsāra*. In one famous scene, for instance, the Buddha-to-be offers himself as food for a starving tigress who is about to eat her own cubs out of desperation. But if offering worldly benefits to beings within *saṃsāra* were ultimately ineffective, the tradition would not applaud such acts in its canonical literature.

To respond to this objection, I suggest that we begin by reflecting on the goal of the *bodhisattva* path. To reemphasise, the goal is to reach awakening oneself and then to lead other beings to awakening. It is neither to persist indefinitely with worldly altruism nor to live longer in *saṃsāra* for its own sake. Upon his awakening, the Buddha did not work in the ancient Indian equivalent of a soup kitchen, feed himself to another hungry animal, or spend the rest of his days basking in his enlightenment, but rather embarked on a decades-long teaching career in which he introduced the *Dharma* and established the Buddhist monastic order on Earth. Likewise, Śāntideva, whose work is plausibly the most important primary source for Mahāyāna ethics, writes that the goal of the *bodhisattva* is to bring about the end of *saṃsāra* by leading all sentient beings to awakening.[22] But helping sentient beings to achieve comparatively better short-run *saṃsāric* results is not going to bring about the end of *saṃsāra*. Even if the tigress avoids the anguish in this life and the descent into the hell realms in the next that would have resulted from eating her cubs, she will eventually end up in hell regardless (as she will in heaven and all realms in between, just as we all will as we wander aimlessly through *saṃsāra*). That worldly altruism plays some as yet unspecified role in the *bodhisattva* path does not, therefore, constitute a counterexample to the *saṃsāric* futility analysis. Despite the real proximate benefits that beneficiaries of worldly altruism enjoy, until their fundamental delusion is addressed, they will continue in their indefinite journeys through the cyclic existence, which is, again, considered to be undesirable.

22 Goodman, *Consequences of Compassion*, 100–01.

What, then, is the role of worldly altruism in the *bodhisattva* path? In light of facts about ultimate *saṃsāric* futility, one plausible answer is that practicing worldly altruism is necessary for making progress on the *bodhisattva* path, particularly in its earlier stages. Buddhism teaches that before one is able to effectively instruct others, one must cultivate compassion (alongside wisdom (*prajñā*)). And according to the tradition, one cultivates compassion in part by engaging in conventional, i.e. worldly, altruism. We can therefore see the practice of worldly altruism as an earlier stage on the longterm path of personal transformation that culminates in *bodhisattva*-hood. Although individual instances of worldly altruism are not maximally effective *qua* discrete actions, for the Buddhist practitioner their performance constitutes an important (if preliminary) section of the *bodhisattva* path, the cultivation of which *is* maximally effective when considered holistically.[23]

We have therefore reached an important qualification to the *saṃsāric* futility analysis: although it remains that any altruistic action that does not bring the patient closer to awakening is, in the long run, ineffective for the patient, a Buddhist altruist may nonetheless have sufficient reason to perform such an action when doing so is (i) partially constitutive of the *bodhisattva* path and (ii) anyway better for the patient (relative to the altruist doing nothing instead). Given this qualification, however, it is natural to think that insofar as the *bodhisattva*-in-training engages in worldly altruism, they should do so most effectively (in worldly terms). We might, then, conjecture that the classical Indian Buddhist *bodhisattva*-in-training would join forces with EA, were they transported to the present.

23 Note that the longterm futility of worldly altruism does not imply that in conferring worldly benefits, the *bodhisattva*-in-training is treating sentient beings as mere means to the end of cultivating compassion. Since the *bodhisattva*-in-training is not yet advanced enough to teach the *Dharma*, they are helping sentient beings in the best way they possibly can. In such a scenario, everyone wins: the aid recipient ascends to a better position in *saṃsāra*, which, while worse than awakening, is better than whatever *saṃsāric* position they would have inhabited without the aid. Meanwhile, the *bodhisattva*-in-training makes progress on the *bodhisattva* path, which is good for them and for the future recipients of their altruism.
See also Lazar and Lee-Stronach, "Axiological Absolutism and Risk," for a distinction between individual acts and "campaigns" (diachronically coordinated sequences of individual acts collectively aimed at some end) and a discussion of the relevance of this distinction in moral philosophy and decision theory. Adopting this parlance in the Buddhist setting, an individual act A of worldly altruism may not be maximally effective in itself, and yet the campaign of walking the *bodhisattva* path – in which A features as a part – may be.

Despite the initial appeal of this conjecture, I think such an alliance would be unlikely. The type of worldly altruism we encounter in the Jātaka tales is often local and even somewhat *ad hoc* in character. (Think again of the tigress.) EA, in contrast, focuses most of its efforts on "faraway" moral patients like persons living in extreme poverty in the developing world, nonhuman animals in factory farms, and future generations on the ground that research reveals this focus to be maximally effective. Underlying this difference in orientation is an epistemological disagreement. Buddhism cautions against attempting to help others before addressing one's own fundamental delusion and the pervasive, though often undetected, egocentrism it engenders.[24] Such attempts, the tradition warns, may be ineffective or even result in unintended negative consequences.

This epistemic worry is most pronounced regarding attempts to teach others about the good life and thereby to change fundamental aspects of their psychologies, such as their core values or characters. But the worry also pertains to interventions that are not explicitly ethical or soteriological – particularly those that rely on complex altruistic reasoning. As the route to impact becomes increasingly removed from oneself, Buddhism warns that the doors to self-deception and to causal misunderstanding open more widely. I believe the tradition would consequently be suspicious of many EA endeavours, which are distinctive in their reliance on complex and often speculative reasoning (e.g. the calculations of GiveWell, a charity evaluator, as to which NGOs maximise lives saved or quality-adjusted life years (QALYs) generated per dollar of marginal donations and EA arguments that advanced artificial intelligence is among the most likely causes of human extinction in the twenty-first century[25]). Of course, the existence of this suspicion does not imply that the contemporary *bodhisattva*-in-training could not consistently support *any* EA-endorsed intervention. But it does provide substantial evidence against the conjecture that such an agent should simply join EA. For them, any EA-endorsed intervention must meet the epistemic worry about unawakened altruism and also perform better, in expectation, than the various non-teaching-related interventions that have precedent in the Buddhist tradition, to which we will now turn.

24 Gold, "More Things in Heaven and Earth," 174.

25 See Bostrom, *Superintelligence*; and Ord, *The Precipice*.

3.1.2. Indirect Altruism

Though the path of direct teaching holds a uniquely revered position in the Buddhist world, it does not follow from Buddhist assumptions that this is the only effective mode of altruism. For in addition to helping others progress along the path to awakening via direct instruction, one can – at least in principle – help them along the path indirectly, by maintaining, augmenting, or creating the conditions that are conducive to successful Buddhist practice. Here we will explore some interventions that would, assuming Buddhist soteriological values, plausibly rank as effective modes of indirect altruism.

Before discussing these interventions, three caveats are in order. First, the epistemic worry about unawakened altruism is still in play. Second, the grounds in favour of the interventions' Buddhist credentials are more speculative than the evidence cited in the previous section. Whereas we have significant canonical evidence regarding the nature of the *bodhisattva* path, traditional Buddhist texts have not explicitly asked questions in the genre of "given finite resources, how can we maximise value (i.e., *nirvāṇa* attainments)?". I will attempt to address, or at least to mitigate, both of these concerns by basing the interventions on precedents drawn from Buddhist political and monastic culture. These precedents offer at least suggestive support for the view that Buddhism would neither regard these interventions as entirely novel suggestions nor regard us as clueless about their efficacy.

Finally, third, all unawakened acts of altruism fall short relative to one Buddhist ideal of altruistic activity, namely the mode of altruism of awakened beings. On this ideal, the awakened being acts for the benefit of the unawakened in an entirely spontaneous, non-intentional manner, in contrast to the deliberately goal-oriented mode of engagement of the unawakened altruist.[26] In Buddhist terminology, the actions of both the awakened being and the unawakened altruist are wholesome or skilful (*kuśala*). However, whereas the action of the awakened being does not produce *karma* in virtue of being non-intentional, the action of the unawakened altruist does produce *karma* (albeit good *karma*), which keeps them tethered to *saṃsāra* (albeit in the form of relatively better rebirths). But the fact that the unawakened altruist's action is in this way worse for them neither

26 This caveat therefore also applies to *bodhisattvas*-in-training who are not yet fully awakened and so are not acting out of spontaneous compassion, but it is particularly apposite in the context of unawakened beings who are consciously engaging in indirect altruism.

detracts from the action's status as altruistic (or wholesome) nor diminishes its efficacy.

With these caveats in mind, imagine that a Buddhist altruist amasses worldly resources such as wealth or power. (If relevant, we can assume that she is genuinely motivated by compassion and amasses the resources in a permissible manner, e.g., by establishing a Buddhist nonprofit and gathering donations, which are given freely in the Buddhist spirit of generosity (*dāna*).) The altruist could then use the resources to support established Buddhist monasteries and teachers, create new Buddhist practise centres specifically designed for long-run durability, or spread Buddhism to people whom it otherwise would not have reached.

It seems possible for these interventions to be more effective, in Buddhist terms, than direct instruction. For if the altruist were to facilitate others' (progress towards) awakening via patronage, and especially if these others went on to become *bodhisattvas*, she could enjoy a multiplier effect on her altruistic impact. Rather than becoming a *bodhisattva* herself – the best-case outcome of devoting her life to Buddhist practice – the altruist would, say, have played a central causal role in the arising of two *bodhisattvas*. And it is of more benefit to all sentient beings in *saṃsāra* to have an additional two *bodhisattvas* than it is to have an additional one. Moreover, each of these interventions has precedent in the Buddhist world. Historically (and, in certain states, continuing through the present), kings have supported monasteries and often been considered praiseworthy for doing so.[27] Likewise, Buddhists have historically engaged in missionary activity, helping to spread the tradition from present-day Nepal and India (the locus of its origination) to Tibet, East Asia, and, in a process that is currently ongoing, to the West. That some of these missionaries have ascended to legendary status[28] within the tradition suggests at least in-principle support for missionary activity.

More creative resource allocations are also available to the altruist. She could, for example, fund pertinent research, such as research into whether certain forms of Buddhist practice are, on average, more conducive to the

27 The third century BCE king Aśoka of the Mauryan Empire (in present-day India) is the touchstone example.

28 The hagiography of Padmasambhava, who (according to tradition) spread Buddhism to Tibet, is an excellent example.

attainment of awakening than others, and if so, which forms these are.[29] Again, there is a plausible route to substantial impact here. If certain claims about efficacy were verified or even substantiated to any nontrivial degree, and if Buddhist practitioners then increasingly adopted the most effective techniques, the rate at which people attained awakening could be significantly increased. Such an increase would be welcome from all Buddhist perspectives, but perhaps particularly so from that of the Mahāyāna, for it would mean an increase in the rate of new *bodhisattvas* joining the ranks of already-active *bodhisattvas* in their enlightened efforts to lead all sentient beings to awakening. And again, we find historical precedent for resource allocations in at least the spirit of this suggestion. In addition to supporting monasteries focused on practice and ritual, kings have supported Buddhist monastic universities in India and Tibet, the idea perhaps being that a better understanding of the *Dharma*, or at least a better understanding of how to teach it, will lead to greater welfare.

3.1.3. Politics, Society, and the Longterm Future

So far, our discussion has focused on altruistic actions that are available to individuals and to relatively small groups, such as nonprofits and research teams. But we can also ask what a larger Buddhist community (*saṃgha*) working together might accomplish. After all, the conditions that are maximally conducive to the ideal Buddhist outcome on Earth – that as many sentient beings as possible attain awakening, while those who do not awaken progress on the path (or at least generate sufficient merit to secure a decent rebirth) – obtain at the widest social level. That is, it would appear that the ideal Buddhist society is optimised for progress on the Buddhist path for those who are willing and able to walk it and for wholesome

29 Of course, this research would be difficult to pursue in the actual world, in which it is taboo in most Buddhist circles to disclose one's attainments. Yet, the taboo is due to a leeriness of monastics abusing power gained via (false) claims of advanced attainment, as opposed to any in-principle opposition to empirical claims about the efficacy of practice. Vajrayāna Buddhism (the "Diamond" or "Thunderbolt" Vehicle), a subset of the Mahāyāna, claims for instance that its distinctive tantric methods are the most efficient means of attaining awakening by a significant margin. Or, to take another example, there is debate within contemporary Theravāda Buddhism as to whether solely practising insight meditation (*vipaśyanā*) is generally as effective as combining insight meditation with calming concentration meditation (*śamatha*).

conventional welfare for those who are uninterested or unable.[30] Such a society would feature, among other things, a sufficient level of material and physical security, such that everyone's subsistence, health, and safety needs were met; a robust monastic system, able to fully support everyone who desired to commit their lives to Buddhist practise; and cultural norms that, in conjunction with political, economic, and educational institutions, maximally inclined people to wholesome living, e.g. by instilling loving-kindness and compassion as foundational values and ensuring that no one had to resort to an unwholesome occupation out of financial necessity.

I am sceptical that for a Buddhist community with finite resources and the goal of maximising the number of sentient beings who attain awakening, attempting to actualise the ideal Buddhist society is the most instrumentally rational course of action. One issue is the question of political tractability. Another issue, more pressing in my estimation, is that many of the targets that must be hit to achieve the ideal Buddhist society, such as eliminating poverty and improving public health, are not neglected. Billions of dollars and countless research hours, among other resources, are already devoted to these causes every year, so it is unlikely that a Buddhist community could significantly improve these efforts via additional marginal contributions. In my view, even a larger Buddhist community would accomplish more good, in Buddhist terms, by allocating most of its resources to a more tractable and neglected set of narrower interventions such as augmenting Buddhist institutions, which possess the additional virtue of more directly addressing the core problems of *duḥkha* and rebirth.

It is nonetheless productive to consider the ideal Buddhist society, for it serves as a further counterpoint to EA. In the short and medium run, Buddhists working to establish the ideal Buddhist society and EAs might share and even collaborate on goals such as improving global health and putting an end to factory farming. Such Buddhists might even adopt the EA priority of minimizing existential risk, not to prolong existence in *saṃsāra* for its own sake, but to preserve a set of conditions that was unusually conducive to awakening. After all, on Buddhist assumptions, an existential catastrophe on Earth would simply result in Earth's inhabitants being reborn elsewhere in *saṃsāra*, so there is no reason Buddhists would welcome such an event. (Extinction, whether voluntary or not, also suffers from the *saṃsāric* futility problem.) Rather, although Buddhists hold that all societal conditions are ultimately impermanent, they accept that certain

30 For work on Buddhist political thought see e.g. Moore, "Political Theory in Canonical Buddhism"; and Bodhi, *Buddha's Teachings on Social and Communal Harmony*.

conditions are more conducive to successful Buddhist practice than others. Relative to the end of awakening, Buddhists therefore have instrumental reason to maintain propitious societal conditions to whatever extent possible.

In the long run, however, even society-focused Buddhist altruists and EAs would practically diverge. Given the range of axiological disagreement and the role of normative uncertainty in EA, it is difficult to say what the ideal state of affairs is from an EA perspective. What we can say is that it is not a world in which human activity is devoted, more-or-less, to the large-scale pursuit of Buddhist enlightenment. To offer a concrete contrast, Ord offers one positive vision for the future that attracts many EAs (though certainly not all).[31] In this vision (which is presented as an ideal, rather than likely, outcome), our descendants solve all the world's problems, including poverty, injustice, and nonhuman animal suffering; reach the zeniths of science, technology, philosophy, art, and (post)human flourishing; and, over the long run, spread to all accessible portions of the universe via interstellar colonisation. To pursue Ord's vision would be to pursue continued existence in *saṃsāra* for its own sake,[32] a project that Buddhism can only view as profoundly misguided. Indeed, on one foundational Buddhist analysis, craving for existence is one of three subspecies of craving (*tṛ́ṣṇā*)[33], which is one of the two proximate causes of *duḥkha* (the other being aversion (*dveṣa*)).

In this section, we have explored a number of altruistic interventions that plausibly rank as highly effective on Buddhist assumptions. These have included the direct route of becoming an advanced Buddhist teacher, which archetypically begins with personal cultivation in the form of local, conventional altruism and meditation and culminates in *bodhisattva*-hood. They have also included several indirect options, including building and supporting monasteries, spreading Buddhism, funding pertinent research, and working towards macroscopic societal changes. For the most part, these approaches to altruism stand in contrast to the approaches that (currently) hold sway within EA, such as improving global health, fighting factory farming, and mitigating existential risks. In the Buddhist analysis, these approaches on their own all fall prey to the *saṃsāric* futility problem.

31 Ord, *The Precipice*.
32 Cf. the *bodhisattva* rationale for remaining in *saṃsāra*, *viz.* to save other sentient beings.
33 Ṭhānissaro Bhikkhu, "Dhammacakkappavattana Sutta."

As we saw, society-oriented Buddhists *may* align their priorities with some of EA's core cause areas in the short-to-medium run, although I raised doubts about whether doing so would maximise altruistic impact in Buddhist terms. But even in a world where such an alignment occurred and was successful – i.e., in a world where humanity solved the most pressing worldly problems and achieved existential security – Buddhism and EA would again go their separate ways. For Buddhists, the *raison d'être* of civilization would remain as it always was: to lead the greatest possible number of beings to awakening, i.e., to effect the greatest possible escape from *saṃsāra*. (It is only in the service of this end that Buddhist priorities would have aligned with those of EA in the first place.) Buddhists would therefore advocate measures such as interstellar colonization only insofar as, and only in the manner that, they were expected to further this aim. For many EAs, in contrast, the aim of a well-off, existentially secure civilization would be to make life as wonderful and as expansive as it could possibly be until the heat death of the universe, i.e., to dive with the greatest possible zeal into *saṃsāra*.[34]

Before moving on, we may note, as a point of contrast to this majority EA view, that a relatively small minority of EAs primarily or entirely value the reduction of suffering.[35] This focus may seem to be more in line with the preoccupation with *duḥkha* in Buddhist thought. However, the suffering-focused branch of EA prioritises causes like factory farming and access to pain relief in the short term and preventing risks of suffering posed by advanced artificial intelligence in the long run. Moreover, it tends to emphasise uncertainty and cooperation: "all [suffering-risk] reducers should aim to compromise with those who want to ensure that humanity has a cosmic future…Rather than fighting other people's efforts to ensure humanity's survival and the chance to develop into an intergalactic, long-lasting and flourishing civilisation, we should complement these efforts by taking care of the things that could go wrong. Cooperation between

34 So, to note just two of many possible contrasts, the spacefaring Buddhist society, motivated solely by compassion for aliens, would have fewer people than Ord's massive interstellar civilization, likely by orders of magnitude. Its members would also attain full awakening and pass out of *saṃsāra* as quickly as possible upon learning that there were no aliens in reach, whereas those in Ord's civilization would, again, perpetuate (post)humanity for as long as possible until the heat death.

35 Dullaghan, "EA Survey 2019 Series: Cause Prioritization." See e.g. the Center on Long-Term Risk (https://longtermrisk.org), the Center for Reducing Suffering (https://centerforreducingsuffering.org), and the Organisation for the Prevention of Intense Suffering (https://www.preventsuffering.org).

future optimists and future pessimists will be best for everyone."[36] All these efforts differ significantly from the various Buddhist interventions we have explored in this section. Again, there would not seem to be much practical convergence in the altruistic activities of Buddhism and EA, even on this minority EA view – though I admit the topic could be fruitfully explored in greater depth than I am able to here.

3.2. Altruism in Theory

We can now turn to offering a Buddhist assessment of the four normative principles that I suggest underpin EA. Before doing so, however, it is important to acknowledge that the theoretical structure of Buddhist ethics – if there is one at all – is a matter of significant controversy in the contemporary literature. If Buddhist ethics is amenable to thinking in terms of universal normative principles, then claims of the form "Buddhism would accept (reject) principle P" can be taken at face value. If it is not, such claims must be taken *mutatis mutandis*. For example, if Buddhist ethics is best interpreted as a form of moral phenomenology,[37] then we might read the claim "Buddhism would accept Impartial Maximisation" as stating that across a wide range of cases, the action that would impartially maximise welfare would appear to the Buddhist adept as the most skilful or wholesome.

With this methodological consideration in mind, let us begin by assessing Strong Welfare Promotion. According to this principle, altruism should be a significant commitment in our lives, but it need not be all-consuming. Relative to the *bodhisattva* ideal of Mahāyāna ethics, Strong Welfare Promotion is actually not strong enough, for the sole end of the *bodhisattva* is the awakening of all sentient beings. Said differently, the paradigmatic *bodhisattva* is entirely, rather than partially, committed to altruism. In contrast to Mahāyāna Buddhism, Theravāda Buddhism does not hold that all practitioners should strive to become *bodhisattvas*. However, loving-kindness and compassion are two of its principal virtues, respectively aimed at promoting the positive component(s) of welfare and alleviating suffering. What exactly Theravāda ethics would say about Strong Welfare Promotion depends on which ethical theory it is implicitly committed to, if it is committed to one at all. Still, the centrality of loving-kind-

36 Althaus and Gloor, "Reducing Risks of Astronomical Suffering."
37 See e.g. Garfield, "What is it Like to be a Bodhisattva?"

ness and compassion to the ethical outlook of the Theravāda suggests a nontrivial place for altruism in the tradition.

According to Impartial Maximisation, insofar as we are trying to behave altruistically, we should seek to impartially maximise the amount of welfare we bring about (other things equal). To begin our examination of this principle, recall that Buddhism is centrally concerned with the welfare of all sentient beings. The precise moral standing of nonhuman animals relative to humans in Buddhist thought is somewhat unclear, however.[38] One plausible generalisation is that although nonhuman animals do have moral status, they are less important than humans, who uniquely enjoy the capacity to attain awakening during their current rebirth.[39] Buddhism would therefore agree with the majority[40] of EAs who consider all sentient beings to be moral patients, disagree with the minority who do not hold this view, and find general accord with the strong current of uncertainty within EA[41] about how, precisely, to perform inter-species welfare comparisons and aggregations.

Buddhism also shares with EA a strong emphasis on impartiality. Impartiality (*upekṣā*) is another of the principle Theravāda virtues. It involves (among other things) a neutral orientation with respect to welfare, which means not assigning more intrinsic importance to the welfare of anyone (including oneself) over that of anyone else (*Visuddhimagga* 9.96, 9.108–09, 9.124).[42] Impartiality in this sense remains important in the Mahāyāna as well; it is evident, for instance, in the work of Śāntideva (e.g. *Bodhicaryāvatāra* 8.90–103).[43]

Impartial concern for the welfare of all sentient beings seems to commit Buddhism to the maximisation of altruistic impact, at least across an interestingly broad range of cases in which one's goal is to benefit others. If one intends to allocate a unit of resources altruistically, and one expects allocation A to generate more welfare than all other possible allocations, then, if one is impartial with respect to welfare and other things are equal, one

38 Finnigan, "Buddhism and Animal Ethics."

39 Waldau, "Buddhism and Animal Rights"; Barstow, "On the Moral Standing of Animals."

40 71% of respondents to the 2019 EA Survey voted that the cause area of "animal welfare/rights" should receive "at least significant resources" (Dullaghan, "EA Survey 2019 Series: Cause Prioritization").

41 See e.g. Muehlhauser, "2017 Report on Consciousness and Moral Patienthood"; and Schukraft, "Comparisons of Capacity for Welfare."

42 Buddhaghosa, *Path of Purification*, 312, 315, 319.

43 Śāntideva, *Guide to the Bodhisattva Way of Life*, 96–97.

ought to select allocation A. To select any allocation other than A would be to value the welfare of some more than the welfare of others, violating impartiality.[44] Śāntideva appears to support this stance in maintaining that the *bodhisattva* should strive to eliminate the *duḥkha* of all sentient beings, "bring about all present and future pleasure and happiness, … [and] *abandon a small benefit in order to accomplish a greater benefit.*"[45] There is, therefore, strong conceptual and suggestive canonical evidence to support the conclusion that Buddhism would accept Impartial Maximisation in a variety of practically-important decision contexts.

Note, however, that Impartial Maximisation enjoins us to be welfare maximisers *when we are acting on whatever reason we have to promote welfare*, not to be welfare maximisers *tout court*. The scope of the principle in our life, therefore, varies with the strength of our reason to promote welfare. Moreover, Impartial Maximisation is qualified by a *ceteris paribus* clause. If other things are not equal – if, for instance, certain side-constraint violations are impermissible, even when their expected impacts are welfare-positive (see below for relevant Buddhist positions) – then, even when one is acting altruistically, one may not be required to impartially maximise welfare. The *ceteris paribus* clause would also come into play if Buddhism accepted some form of egalitarianism, prioritarianism, or Rawlsian lexicality. For if Buddhism accepted such a view, in at least some cases, it would recommend resource allocations that promoted a relevant form of equality or benefitted the worst off at the cost of failing to maximise total welfare.

For reasons of space, I cannot fully address whether Buddhism would accept such a view. To offer a preliminary response, though, the case for Rawlsian lexicality rests importantly on the separability of persons. But such a distinction becomes difficult to sustain if one accepts Buddhist ontology, according to which the self does not exist and the person exists merely as a matter of convention. In this vein, Śāntideva holds that "[w]ithout

44 See Gowans, *Buddhist Moral Philosophy,* 134–37 for a similar argument.

45 Śāntideva, *Training Anthology*, 17 (quoted in Goodman, "Ethics in Indian and Tibetan Buddhism," italics added.) Although the possibility of inferring ultimate objectives (e.g. leading beings to awakening) and general values (e.g. impartiality) from Śāntideva's work is not disputed, not all interpreters believe we should read Śāntideva as attempting to set out precise, universal ethical principles (for discussion see Gowans, "Buddhist Moral Thought"; Finnigan, "Madhyamaka Ethics"; McRae, "Psychology of Moral Judgment"; and Gold, "More Things in Heaven and Earth"). Given this hermeneutical uncertainty, the canonical evidence here is only suggestive, as I say presently in the main text.

exception, no sufferings belong to anyone. They must be warded off simply because they are suffering. Why is any limitation put on this?" (*Bodhicaryāvatāra* 8.102).[46] For this reason, I think it is unlikely that Buddhism would accept so strong a view as Rawlsian lexicality. But it is conceivable (though I will remain neutral on how likely it is) that Buddhism would accept a milder form of egalitarianism or prioritarianism. This need not imply any disagreement with EA, however, for at least on the characterisations offered by MacAskill and by Pummer and MacAskill, EA is compatible with (though not committed to) egalitarianism and prioritarianism.[47]

Methodological Rigour requires us to base our altruistic efforts on a rigorous evaluation of the relevant evidence. As it pertains to the nature of welfare, Methodological Rigour is, from a Buddhist standpoint, redundant in practice, if commendable in spirit. Buddhism disagrees with EA on what the relevant evidence is, for it regards the word of the Buddha (*buddhavacana*) as authoritative in a way that EA (alongside all other intellectual traditions) does not. Buddhism would hold that a rigorous evaluation of the relevant evidence is already complete in the form of the Buddha's teaching and its subsequent elaboration by accomplished Buddhists. The state of being awakened is the pinnacle of human welfare, and the attainment of *nirvāṇa* is the best possible outcome for all sentient beings. Regarding the ultimate goal of existence, it is on the Buddha's teaching and its canonical exegesis that we should base our understanding, not on philosophy produced by our own minds, shrouded as they are in delusion.

Yet despite the end being fixed, there is much we can learn about effective means. The production of novel scriptures and meditative techniques throughout Indian Buddhist history attests to the tradition's amenability to innovation in how, concretely, the *Dharma* is to be taught and awakening is to be realised. And as we have discussed, there are a multitude of indirect altruistic avenues one might pursue, from maintaining a particular monastery to striving for broad trajectory changes in society. From our current epistemic position, it is uncertain which of the direct modes of instruction and practice, and which indirect avenues, are most effective. Given this uncertainty and the objective of impartially maximising welfare, basing our altruistic efforts on a rigorous evaluation of the relevant evidence appears to be a requirement of instrumental rationality. Buddhism would therefore accept a modified version of Methodological Rigour that

46 Śāntideva, *Guide to the Bodhisattva Way of Life*, 97.

47 MacAskill, "The Definition of Effective Altruism"; Pummer and MacAskill, "Effective Altruism."

defers to Buddhist teaching on the ultimate good (*nirvāṇa*) but requires rigorous research on all other practical questions and subgoals that are part of the altruistic project.

We now arrive at the final EA principle, Weak Normative Uncertainty. According to this principle, we should avoid basing our normative outlook exclusively on one ethical theory and instead be open – epistemically and practically – to insights from multiple plausible theories. Buddhism would not accept that we should hedge across different normative theories. Although there is room in Buddhism for epistemic uncertainty about certain doctrinal disputes, such as how precisely to understand the non-self (*anātman*) teaching, fundamental evaluative uncertainty, e.g. about what our final end should be, is not a feature of classical Buddhist thought. Rather, the Buddhist diagnosis of anyone who does not accept, or who is uncertain about, the core components of the Buddhist evaluative position – that *duḥkha* is the fundamental problem of existence and that *nirvāṇa* is the best possible outcome for all sentient beings – is that the person is, like most of us, deluded.[48] This contrast between EA and Buddhism highlights an essential difference between the two: whereas the former begins from an assumption of uncertainty – that we may be mistaken about our ultimate values, the true ethical theory, and our prioritisation of causes and interventions – the latter is committed to a definite soteriological program.

EA draws from Weak Normative Uncertainty the corollary that we should avoid behaviour which seriously violates common-sense morality. There is some evidence that if presented with this position, Śāntideva would demur. In certain passages, he appears to endorse the violation of (what contemporary philosophers would call) side-constraints for the sake of welfare maximisation, arguing in one passage that the compassionate person ought to cause one person to suffer if doing so is necessary to alleviate the suffering of many (*Bodhicaryāvatāra* 8.105)[49] and in another that the *bodhisattva* should cause a small amount of suffering if doing so is necessary to prevent a large amount of suffering.[50] However, it is unclear whether the purpose of these passages is to introduce a face-value moral principle, and if so, whether the scope of the principle includes everyone who is trying to act compassionately or, in contrast, only advanced *bodhisattvas*, who have the requisite wisdom to foresee the consequences of their actions. In my view, it

48 See e.g. Williams and Tribe, *Buddhist Thought*, 7–8.

49 Śāntideva, *Guide to the Bodhisattva Way of Life*, 97.

50 Śāntideva, *Training Anthology*, 17 (quoted in Goodman, "Ethics in Indian and Tibetan Buddhism").

is not maximally charitable to read Śāntideva as endorsing the position that whenever we (unawakened beings) think we can maximise expected utility by violating a side-constraint, we should go ahead and do so.

Moreover, it is not the case that all or most Buddhist philosophers would accept that one should prevent larger quantities of suffering by inflicting smaller quantities of suffering, even when one is certain about the relevant consequences. For instance, as Gethin argues,[51] killing a sentient being cannot, according to the Theravāda Abhidhamma, be entirely an act of compassion, and hence cannot be entirely wholesome or skilful.[52] Killing is also forbidden by the five lay precepts (rules that committed Buddhist laypersons undertake to follow) and by all forms of the Buddhist monastic code (*Vinaya*), both of which additionally forbid other commonly-condemned actions such as lying and stealing. With a possible exception made for *bodhisattvas* acting in special cases, then, Buddhism would broadly concur that we should avoid behaviour that is seriously objectionable in the eyes of common-sense morality. The reason that Buddhism concurs with EA on this point does not derive, though, from any fundamental normative uncertainty internal to Buddhist thought.

4. Conclusion

If the analysis of the preceding section is accurate, Buddhism is unlikely to agree with EA on how we should allocate the finite resources we have available for altruistic purposes. Among other things, this result constitutes one negative datum point against EA's hope to be appealing to, and consistent with, a wide range of evaluative positions. However, I suggested above that assessing EA from a Buddhist perspective could result in a productive dialogue between the two. To close, I would like to suggest some positive insights that Buddhism and EA might take from our discussion.

I have argued that Buddhists should accept Impartial Maximisation and Methodological Rigour, though in modified forms so as to fit within the Buddhist soteriological framework. Accepting these principles gives Buddhists reason to engage in prioritisation research: research into which causes are most pressing and which means within them are most effective.

51 Gethin, "Can Killing a Living Being Ever Be an Act of Compassion?"

52 It does not follow that the Theravāda is categorically opposed to killing, regardless of circumstance. But its general disapproval of killing does provide strong evidence against the view that the Theravāda Abhidhamma would approve the rule of thumb "kill whenever you expect killing to result in more total welfare than not killing."

Yet, as I noted above, Buddhists have not traditionally asked questions in the explicit form of "given finite resources, how can we save the greatest possible number of sentient beings?". Part of the reason for this lacuna is, I believe, that the path of practice, attainment, and direct instruction has been extremely compelling to many Buddhists and has therefore overshadowed the route of indirect altruism, to which prioritisation research is more germane. Another part of the reason is likely the epistemic worry about unawakened altruism: because Buddhists maintain that we are caught in the grips of a thoroughgoing delusion, they are more pessimistic about our ability to successfully engage in complex altruistic reasoning. And this does admittedly weaken (from a Buddhist perspective) the reason in favour of engaging in prioritisation research.

Still, it would be a mistake to attribute to Buddhism the implausibly conservative position that Buddhists already know everything that can be known about effectively helping others. To reemphasise points made in the main text and in footnotes above, the production of new scriptures and intra-Buddhist debates about the relative potencies of different meditative regimens attest to the tradition's openness to inquiry into effectiveness and implicit affirmation of the possibility of progress in this domain. Even if it differs from that of EA in assuming Buddhist soteriological values and in possessing more modest ambitions, a new prioritisation research program may therefore be a practical takeaway for Buddhists from the dialogue with EA. I should also stress for the sake of contemporary relevance that this takeaway is not merely of interest to those who accept the classical Indian Buddhist worldview, as we have been using the term "Buddhist". It is also of practical import to all contemporary Buddhists who share with their classical forebears an impartial concern for the welfare of all sentient beings. Indeed, it seems that many contemporary Buddhists would find the case for prioritisation research even more compelling than would their classical counterparts, given the increasing popularity of movements like Socially Engaged Buddhism[53] and the novel social and technological conditions of the modern world – conditions that the Dalai Lama, for one, has interestingly acknowledged in recommending that everyone take an "enlightenment pill" were such an item ever to be invented.

53 Socially Engaged Buddhism seeks to address contemporary issues such as climate change and injustice and to avoid what it perceives as an objectionable withdrawal from the world by other, more conservative Buddhist traditions.

In terms of insights that EA might leave with, one point our discussion has highlighted is that one's choice of foundational evaluative perspective can carry significant practical implications. Although this point may appear to be rather obvious, it warrants careful reflection in the present context. Buddhists take life in *saṃsāra* – i.e., life in our world – to be suffused, and thus, in some sense, irredeemably marred, by *duḥkha*. It is for this reason that ceaseless rebirth in *saṃsāra* is taken to be a problem in the first place. (If life in *saṃsāra* were on-balance good, rebirth would not be a problem and might even be a blessing.) Since the Buddhist analysis of *duḥkha* and its origins in certain psychological and perceptual delusions is plausible (or at least, not obviously false), and, if correct, profound in its welfare-relevant implications, any complete answer to the EA question of how to maximise welfare must eventually take a stance on *duḥkha*. This problem and, conversely, the Buddhist conception of the pinnacle of welfare as a state free from *duḥkha*[54] therefore weigh against the view that the competing accounts of welfare tend to converge in their practical implications, despite their theoretical differences – a view that appears to hold nontrivial sway within EA (and analytic philosophy).[55] While this view may be accurate if we restrict our attention to the theories of welfare that have dominated the Western philosophical literature (*viz.* mental state, preference satisfaction, objective list, and nature fulfilment theories), it is inaccurate if we transcend this parochial focus on the West and consider other visions of welfare, such as that of Buddhism, which have been articulated in world philosophy. In rejecting many of our intuitions about what is good for us as products of unenlightened delusion and recommending a theory of welfare, and a path to it, that is at times remarkably counterintuitive, Buddhism forces EAs (and everyone else interested in promoting welfare) to confront the possibility that our thinking on the matter may be less clear than we initially supposed. This point takes on even greater force when we recognise that the problem of *duḥkha* is logically distinct from that of rebirth, or, put differently, that it is consistent with naturalism to hold that *duḥkha* exists and detracts significantly from our welfare.

54 This is not to say that the Buddhist vision of ultimate welfare is solely one of a state free from *duḥkha*. Arguably, Buddhists would also understand wisdom (*prajñā*) and virtues such as compassion (*karuṇā*) to be necessarily present in the state of highest welfare. But whether or not wisdom and virtue are welfare components, the absence of *duḥkha* is unambiguously a welfare component on the Buddhist view.

55 See e.g. MacAskill and Ord, "Why Maximize Expected Choice-Worthiness?," 341, who refer to the purportedly "much greater agreement between people on what constitutes a good life than on how to act morally."

Our discussion has also brought into sharp relief the fact that like foundational evaluative positions, descriptive matters of cosmology, ontology, and metaphysics can have decisive practical implications. To make the point bluntly: if Buddhism is true, quintessential EA activities like funding the distribution of antimalarial bed nets are a relative waste of resources, suffering as they do from the *saṃsāric* futility problem, whereas if secular naturalism is true, building Buddhist monasteries is a relative waste of resources. In getting the cosmology, ontology, or metaphysics wrong, we risk getting the ethics wrong. Since the answers to these cosmological, ontological, and metaphysical questions are in this way crucial considerations, EA must eventually come to explicit answers on the questions rather than merely assuming naturalism. At minimum, EA must seriously consider these questions before undertaking actions that are both irreversible and optimal only on the assumption of naturalism, such as, arguably, allocating significant resources to interstellar colonisation. To fail to give the questions serious consideration prior to taking irreversible action would be, in EA parlance, to risk negative value lock-in, i.e., to risk cementing values and goals that are either suboptimal or dependent on false beliefs. To what extent it should be a current priority within EA to work on cosmology, ontology, and metaphysics – in particular, to examine non-naturalist soteriologies, and perhaps more importantly, to reflect on how to handle the uncertainty that would inevitably remain even after a thorough examination – is a question that must be left to further research.[56]

References

Althaus, David and Lukas Gloor. "Reducing Risks of Astronomical Suffering: A Neglected Priority." Center on Long-Term Risk. 14 September 2016. https://longtermrisk.org/reducing-risks-of-astronomical-suffering-a-neglected-priority.

Barstow, Geoffrey. "On the Moral Standing of Animals in Tibetan Buddhism." *Études Mongoles et Sibériennes, Centrasiatiques et Tibétaines*, no. 50 (2019). https://doi.org/10.4000/emscat.3865.

Berg, Amy. "Effective Altruism: How Big Should the Tent Be?" *Public Affairs Quarterly* 32, no. 4 (2018): 269–87.

56 I am indebted to Rupert Gethin, Jonathan Gold, Barry Grimes, David Leech, Stefan Riedener, and Dominic Roser for extremely helpful comments on earlier drafts of this article.

Berkey, Brian. "The Philosophical Core of Effective Altruism." *Journal of Social Philosophy*: 1–23 (2020). https://doi.org/10.1111/josp.12347.

Ṭhānissaro Bhikkhu (translator). "Dhammacakkappavattana Sutta" (SN 56.11). Access to Insight. 1993. https://www.accesstoinsight.org/tipitaka/sn/sn56/sn56.011.than.html.

Bodhi. *The Buddha's Teachings on Social and Communal Harmony: An Anthology of Discourses from the Pāli Canon*. Somerville: Wisdom Publications, 2016.

Bostrom, Nick. *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press, 2014.

Buddhaghosa. *The Path of Purification* (Visuddhimagga). Translated by Ñāṇamoli Bhikkhu. 4th ed. Kandy: Buddhist Publication Society, 2010.

Crisp, Roger and Theron Pummer. "Effective Justice." *Journal of Moral Philosophy* (2020): 1–19.

Dullaghan, Neil. "EA Survey 2019 Series: Community Demographics & Characteristics." Rethink Priorities. 5 December 2019. https://www.rethinkpriorities.org/blog/2019/12/5/ea-survey-2019-series-community-demographics-amp-characteristics.

———. "EA Survey 2019 Series: Cause Prioritization." Rethink Priorities. 2 January 2020. https://www.rethinkpriorities.org/blog/2020/1/7/ea-survey-2019-series-cause-prioritization.

Finnigan, Bronwyn. "Buddhism and Animal Ethics." *Philosophy Compass* 12, no. 7 (2017): e12424.

———. "Madhyamaka Ethics." In *Oxford Handbook of Buddhist Ethics*, ed. by Daniel Cozort and James Mark Shields, 162–82. Oxford: Oxford University Press, 2018.

Garfield, Jay L. "What is it Like to be a Bodhisattva? Moral Phenomenology in Śāntideva's *Bodhicaryāvatāra*." *Journal of the International Association of Buddhist Studies* 33 (2010): 333–357.

Gethin, Rupert. "Can Killing a Living Being Ever Be an Act of Compassion? The Analysis of the Act of Killing in the Abhidhamma and Pali Commentaries." *Journal of Buddhist Ethics* 11 (2004): 167–202.

Gold, Jonathan C. "More Things in Heaven and Earth: The Path to Nirvana, Naturalized." In *Naturalism, Human Flourishing, and Asian Philosophy: Owen Flanagan and Beyond*, ed. by Bongrae Seok, 166–85. Abingdon: Routledge, 2019.

Goodman, Charles. *Consequences of Compassion: An Interpretation and Defense of Buddhist Ethics*. New York: Oxford University Press, 2009.

———. "Ethics in Indian and Tibetan Buddhism." In *Stanford Encyclopedia of Philosophy,* ed. by Zalta Edward, Spring 2017 Edition. https://plato.stanford.edu/entries/ethics-indian-buddhism/.

Gowans, Christopher W. *Buddhist Moral Philosophy: An Introduction*. New York: Routledge, 2015.

———. "Buddhist Moral Thought and Western Moral Philosophy." In *A Mirror is for Reflection: Understanding Buddhist Ethics*, ed. by Jake H. Davis, 53–70. Oxford: Oxford University Press, 2017.

Karnofsky, Holden. "Worldview Diversification." Open Philanthropy. 13 December 2016. https://www.openphilanthropy.org/blog/worldview-diversification.

———. "Update on Cause Prioritization at Open Philanthropy." Open Philanthropy. 26 January 2018. https://www.openphilanthropy.org/blog/update-cause-prioritization-open-philanthropy.

Lazar, Seth and Chad Lee-Stronach. "Axiological Absolutism and Risk." *Noûs* 53, no. 1 (2019): 97–113.

MacAskill, William. "The Definition of Effective Altruism." In *Effective Altruism: Philosophical Issues*, ed. by Hilary Greaves and Theron Pummer, 10–28. Oxford: Oxford University Press, 2019.

MacAskill, William, Krister Bykvist, and Toby Ord. *Moral Uncertainty*. New York: Oxford University Press, 2020.

MacAskill, William and Toby Ord. "Why Maximize Expected Choice-Worthiness?" *Noûs* 54, no. 2 (2020): 327–53.

McRae, Emily. "The Psychology of Moral Judgment and Perception in Indo-Tibetan Buddhist Ethics." In *The Oxford Handbook of Buddhist Ethics*, ed. by Daniel Cozort and James Mark Shields, 335–58. Oxford: Oxford University Press, 2018.

Moore, Matthew J. "Political Theory in Canonical Buddhism." *Philosophy East and West* 65, no. 1 (2015): 36–64.

Muehlhauser, Luke. "2017 Report on Consciousness and Moral Patienthood." Open Philanthropy Project. June 2017. https://www.openphilanthropy.org/2017-report-consciousness-and-moral-patienthood.

Ord, Toby. *The Precipice: Existential Risk and the Future of Humanity*. London: Bloomsbury, 2020.

Panaïoti, Antoine. *Nietzsche and Buddhist Philosophy*. Cambridge: Cambridge University Press, 2012.

Pummer, Theron and MacAskill, William. "Effective Altruism." In *International Encyclopedia of Ethics*, ed. by Hugh LaFollette, 2020. doi:10.1002/9781444367072.wbiee883.

Ricard, Matthieu. "Altruism Meets Effective Altruism: Thoughts on Conversations with Peter Singer." Accessed 12 January 2016. https://www.matthieuricard.org/en/blog/posts/altruism-meets-effective-altruism-thoughts-on-conversations-with-peter-singer.

Śāntideva. *A Guide to the Bodhisattva Way of Life* (Bodhicaryāvatāra). Translated by Kate Crosby and Andrew Skilton. Oxford: Oxford University Press, 2008.

———. *The Training Anthology* (Śikṣā-Samuccaya). Translated by Charles Goodman. New York: Oxford University Press, 2016.

Schukraft, Jason. "Comparisons of Capacity for Welfare and Moral Status Across Species." Rethink Priorities. 17 May 2020. https://www.rethinkpriorities.org/blog/2020/5/16/comparisons-of-capacity-for-welfare-and-moral-status-across-species.

Todd, Benjamin and the 80,000 Hours team. "A guide to using your career to help solve the world's most pressing problems." 80,000 hours. March 2019. https://80000hours.org/key-ideas/#moral-uncertainty-and-moderation.

Waldau, Paul. "Buddhism and Animal Rights." In *The Oxford Handbook of Buddhist Ethics*, ed. by Daniel Cozort and James Mark Shields, 650–74. Oxford: Oxford University Press, 2018.

Westerhoff, Jan. *The Golden Age of Indian Buddhist Philosophy.* Oxford: Oxford University Press, 2018.

Wiblin, Robert and Keiran Harris. "Our descendants will probably see us as moral monsters. What should we do about that?" The 80,000 Hours Podcast. 19 January 2018. https://80000hours.org/podcast/episodes/will-macaskill-moral-philosophy/.

Williams, Paul and Anthony Tribe. *Buddhist Thought: A Complete Introduction to the Indian Tradition*. London: Routledge, 2000.

Dominic Roser

Effective Altruism as Egyptian Gold for Christians

"This is my prayer:
that your love may abound more and more
in knowledge and depth of insight,
so that you may be able to discern
what is best."
(Philippians 1)

Abstract

Despite primarily emerging in secular circles, effective altruism is not merely compatible with Christianity but is of significant value to it. Its insights offer much support to Christians aiming at serving their neighbour well. The chapter characterises effective altruism by way of seven commitments. Not all of these commitments are embraced by Christianity to their maximal extent. But they all point in the right direction if we compare the actual practice of Christians with the ideal. Christians are called to be more altruistic, and their altruism should put more emphasis on effectively achieving good consequences. In particular, the impartially assessed welfare consequences for those in need should receive more attention. Such a focus would benefit from more careful belief formation about what works in line with the epistemic practices advocated by effective altruism. The article also mentions one tension between the underlying mindset of effective altruism and Christianity. While effective altruism is driven by the aim of intentionally taking responsibility for results into one's own hands, Christianity includes an affirmation of trustfully letting go of control.

1. Introduction

Effective altruism (EA) started out as a very secular movement. Its roots are among Bay Area rationalists, Oxford philosophers, and East Coast hedge fund managers – not circles known for a religious slant. 86 % of the movement's members profess to be atheist, agnostic, or non-religious.[1] This article makes the case that a large part of EA's message is not merely compatible with Christianity but even of significant importance to it.

In the tradition of accepting prophetic words from outside the community of the faithful – think of Balaam in the Old Testament or Westphal's suggestion that Christians make use of the critical questions asked by Marx, Nietzsche, and Freud[2] – Christians should humbly take up many insights that EA has developed. At the very least they should examine them with an open mind on the basis of the attitude propagated in 1 Thess. 5:21: "Do

1 Dullaghan, "EA Survey 2019 Series."
2 Westphal, *Suspicion and Faith*.

not treat prophecies with contempt but test them all; hold on to what is good."

The African theologian Augustine of Hippo was particularly forceful in his endorsement of integrating elements from non-religious culture into the Christian faith. While I presented it as humbly taking up insights he relies on the imagery of plundering the Egyptians (Exod. 12:35–36):

> "If the [pagans] have said things which are indeed true and are well accommodated to our faith, they should not be feared; rather, what they have said should be taken from them as from unjust possessors and converted to our use. Just as the Egyptians had not only idols and grave burdens which the people of Israel detested and avoided, so also they had vases and ornaments of gold and silver and clothing which the Israelites took with them secretly when they fled, as if to put them to a better use."[3]

Just like the escaping Israelites took the gold of the Egyptians with them and later had it available to build the sanctuary (Exod. 25), so Christians, too, can recognise secular ideas and "baptise" these ideas for their purposes. In the context of this article, these purposes include in particular loving their neighbour well.

The article is structured as follows. Section 2 characterises EA by way of seven commitments. Section 3 examines the core question by running through the commitments one by one: in what ways is EA a useful tool for Christians in living out their faith? Section 4 discusses a big-picture tension between EA and faith. Section 5 concludes and also briefly comments on the reverse question: is the Christian faith a helpful tool for practicing EA?

3 Augustine and D. W. Robertson, *On Christian Doctrine*, section 40.60.

2. Characterising Effective Altruism

EA has been understood in a number of different ways.[4] For the purpose of examining its usefulness to Christians, I characterise it by way of seven commitments typically exhibited by its adherents:

(1) Altruism. Effective altruists dedicate a significant amount of resources to benefiting others.
(2) Consequentialism. Effective altruists embrace an altruism that is much focused on outcomes. This focus contrasts with other features of the moral situation that might guide one's altruism such as good intentions, virtues, respect for rights, etc.
(3) Welfarism. Effective altruists focus much on one specific aspect of outcomes: the welfare of individuals. This includes in particular its blunter elements such as physical and mental wellbeing. This focus contrasts with other aspects of outcomes that might guide one's actions, such

4 Examples can be found in MacAskill, "Definition," 12–14, including his own definition:
"(i) the use of evidence and careful reasoning to work out how to maximise the good with a given unit of resources, tentatively understanding 'the good' in impartial welfarist terms, and
(ii) the use of the findings from (i) to try to improve the world."
My characterisation deviates from MacAskill's definition mainly for practical reasons: comparing EA to other sets of commitments (in my case: Christian commitments) makes it helpful to split it up into distinct elements.
In terms of substance, my characterisation primarily deviates from his definition by making altruism an integral part of EA. While I agree that one could hold commitments 2–7 regardless of one's level of altruism, it seems somewhat contrived to exclude altruism from the definition of a movement called effective *altruism* (on this, see also the helpful discussion in Berkey, "Philosophical Core," 103–106). First, altruism plays an important role in the motivation of most real live effective altruists. Secondly, some parts of the movement have explicitly been built around altruism and not merely around practicing altruism in a specific way (including the *Giving What We Can* pledge and the concept of earning to give). Thirdly, altruism is the most natural explanation for why anyone would be interested in pursuing commitments 2–7 (cf. the oddness mentioned in MacAskill and Pummer, "Effective Altruism," 4). I agree however that EA can be seen as a project to which one is committed rather than a normative claim one holds: while 70 % of effective altruists in fact do identify with utilitarianism (Dullaghan, "EA Survey 2019 Series") one could exhibit the seven commitments, i.e. pursue the EA project, for non-moral reasons such as excitement, a quest for meaning in life, "compulsively" being drawn to altruism by our evolutionary heritage, or selfish reasons in terms of hedonically profiting from the commitments.

as the value that relationships, knowledge, or dignity have apart from whether they make anyone better off.[5]

(4) Impartiality. Effective altruists embrace a very impartial form of altruism. Individuals do not count more or less merely due to their sex, geographical location, time of birth, species, etc.

(5) Effectiveness.[6] Effective altruists focus much on maximising the ratio at which a given amount of resources yields good outcomes.

(6) Truthseeking. Effective altruists oppose thoughtlessness and put much effort into forming beliefs that track reality well, in particular beliefs about the ratio at which resources yield good outcomes.

(7) Rationalism.[7] Effective altruists' method for forming beliefs that track reality well emphasises empirical evidence, science, Bayesian thinking, skepticism, open-mindedness, quantitative tools, forecasting research, awareness of cognitive biases, etc. rather than intuition, conventional wisdom, superstition, etc.

In order to see the forest for the trees, it helps to summarise these seven elements into three steps and to highlight the logic behind grouping these three steps into one social movement.

Step I: Effective altruists are do-gooders (commitment 1) whose concern for others takes a specific form: impartial, welfarist consequentialism (commitments 2–4).

5 This non-welfare value could either be impersonal value or it could be personal value in the sense that it makes individuals' lives better without making it better *for them*, i.e. without increasing their welfare (see Wall, "Perfectionism in Moral and Political Philosophy," section 1). Note that things such as knowledge, friendship, or dignity could additionally be valuable, first, as elements of welfare in objective list theories and, secondly, instrumentally valuable for promoting welfare according to any theory of welfare. Further, these things could be morally relevant apart from their value.

6 Many would consider "efficiency" or "cost-effectiveness" to be more accurate terminology. They would use "effectiveness" more specifically for how much of a good outcome is achieved (rather than for how much of it is achieved *per unit of resources*). Unfortunately, this convention for a more precise use of language seems not to be followed widely.

7 It is hard to capture the eclectic bundle of views that make up the typical effective altruist's epistemology in a single word. The *Centre for Effective Altruism* captures it by the expression "evidence and careful reasoning". In order to be more specific but still broad, I crudely use the label "rationalism" in reference to the rationalist community, including the *Center for Applied Rationality*, as one paradigmatic exemplification of the endorsed style of thinking. This is admittedly a source of confusion since rationalism has many further meanings in other contexts (including as an opposite to empiricism).

Step II: Consequentialism implies that one should focus on effectiveness (commitment 5).

Step III: The focus on effectiveness implies that one should focus on carefully assessing which efforts yield the best outcomes and this is best done by relying on a "rationalist" style of thinking (commitments 6 and 7).

None of these seven commitments need be pure or maximal.[8] Rather, the lives of effective altruists are shaped by these commitments to a greater or lesser extent.[9] Many people outside the movement embrace combinations of these commitments as well. Hence, EA could be reproached for lacking novelty. In response to this worry it should first be noted that even if EA lacked novelty it would not thereby automatically be irrelevant: the value of EA could be seen in publicising how the developments of the last, say, three decades in terms of the growth of technology, data, scientific understanding, and wealth have radically increased the impact of attending to the seven commitments compared to the previous history of humanity. But, in fact, EA *is* novel in at least three ways anyhow. First, EA is often more radical about the seven commitments than others. For example, while EA is not unique in emphasising that effectiveness matters, it stands out in emphasising just *how much* it matters. The best ways of benefiting others are often not just a bit better than average ways but by orders of magnitude better.[10] Secondly, EA has been more forceful in teasing out underappreciated implications of the seven commitments. For example, EA is fairly unique in highlighting and exploring impartiality between people living in the present and the *far* future. Another example: EA points out that a commitment to effectiveness does not merely imply careful attention to prioritising between different interventions within a cause area but also between cause areas. There are few other cause-neutral movements united

8 Cf. MacAskill and Pummer, "Effective Altruism," 4.

9 Since adherents of the movement often emphasise different subsets of the seven commitments, different understandings of EA can sometimes look as if they were connected by family resemblance rather than by a core of individually necessary and jointly sufficient conditions (cf. Moss, "Maximising Altruism," 4).

10 See Caviola et al., "Donors Vastly Underestimate."

by such an extremely general goal as doing the most good.[11] Thirdly, EA has brought this combination of messages to new audiences. While some of the seven commitments taken on their own are thoroughly familiar to professional philanthropists, scientists, philosophers, business people, animal activists, policymakers, etc. bundling the seven commitments together as a common agenda for donors, activists, and voters is something new.

3. Effective Altruism as a Plug-In for Christians

I now examine whether EA helps Christians to practice their faith better. In other words: if Christians use a broadly rationalist epistemology in order to carefully assess which measures most effectively improve the welfare of all beings impartially considered, and if they put much resources into improving this welfare, does this make them live out their faith more fully? Or, more broadly: are the insights, the practices, and the style of the movement that sprung up around these seven commitments a helpful tool for Christians? In the words of Prov. 27:17: is EA an iron that sharps the iron of faith? Answering these questions includes – but is not limited to – the minimal test of whether the seven commitments are compatible with the doctrines of Christianity.

(a) Altruism

When it comes to emphasising concern for others, EA preaches to the converted in addressing Christians. The call to care for the widow, the orphan, the stranger, and the poor is a constant chorus from the Mosaic law to the Psalms, the prophets, the gospels, and the epistles. This has often remained a theory in the history of Christianity but time and again it has also been

11 There is a distinctive challenge for EA movement building with respect to the first and second point, i.e. with respect to being radical about the seven commitments and teasing out underappreciated implications. The EA movement would already be an interesting and valuable movement if it were, say, twice as radical as the broad population about these commitments and if it teased out a small number of underappreciated implications. However, the EA movement is also interesting if it is ten times more radical and teased out utterly unfamiliar implications. How far should EA take its own ideas? The challenge is that the movement becomes overly heterogeneous if it groups those who are only at some distance from the broad population – which might well include quite some Christians in the movement – together into one movement with those who are at much further distance (though in the same direction) from the broad population.

put into practice. Many social reforms and charitable interventions have ultimately been driven by Christian conviction even if these motivational roots were not visible at the surface. One of the signature features of *Giving What We Can* – donating at least 10 % of one's income – is a common reference point for giving among believers, going back to the Old Testament concept of tithing.[12] Even if only a minority lives up to the 10 % benchmark, religious people give away a larger fraction of their income than secular people do.[13] Thus, as far as the generosity message is concerned, Christianity can mostly profit from EA by finding new momentum through an ally. It is of course an ally that does not come out of nowhere but whose intuitions have been in many ways shaped by centuries of Christian concern for the poor and marginalised. The earliest builders of utilitarian theory were even explicitly Christian in their motivation.

EA is an ally not only in terms of endorsing altruism but also in terms of normalising or even celebrating altruism, including in its demanding forms. Rather than coyly acting for the benefit of others in secret out of fear of deviating from social norms, both Christianity and EA are blunt about viewing altruism as a key part of a life well-lived. The only difference is that the worry about character corruption of public giving looms larger in Christianity.[14]

Some Christians might worry about the specific terminology of altruism. The term was popularised by Auguste Comte who had a firmly non-Christian agenda. Nothing hinges on the term, however. It is simply a useful concept for effective altruists – and also for Christians – due to its extreme generality which distinguishes it from other concepts such as love or justice. According to one understanding of the concept, any behaviour that benefits others is altruistic, regardless of its motives.[15] This suits both Christianity

12 This is despite the fact that the rules for the various tithes in the Old Testament are difficult to understand and that tithing is hardly mentioned in the New Testament at all. When it comes to giving, the New Testament mentions various ideas – including giving half of one's possessions (Luke 19:8), selling all one's possessions (Mark 10:21), having everything in common (Acts 2:45), giving according to one's means (Acts 11:29; 1 Cor. 16:2), or giving beyond one's means (2 Cor. 8:3).

13 See for example Austin, *Giving to Religion*. Even those who are critical of there being a positive link between religion and generosity acknowledge that it is clearly the mainstream view (cf. Sablosky, "Does Religion Foster Generosity?"). A recent literature review even suggested that religion "most often positively relates with giving to *outgroup and secular organizations*." (Yasin et al., "How Does Religion." Emphasis mine).

14 Jesus' injunction not to announce with trumpets when giving to the needy is discussed specifically with relation to EA in a blogpost by Synowiec, "Should we sound."

15 Kraut, "Altruism."

and EA as they want to make space for motivational roots of altruism that are not directly about wanting to benefit others or give them what they are owed – such as excitement, selfishness, personal character development, or an urge to express gratitude for overflowing blessings by passing them on.

(b) Consequentialism

Christian ethics features strong consequentialist elements. Making good things happen is part of the Christian vision – obviously so. Camosy observes:

> "[M]any traditional Christians display a broadly consequentialist structure to their moral reasoning. Singer wants to produce the consequence of maximal preference satisfaction, whereas many Christians (especially Thomists) understand the moral life teleologically – with the proximate goal, some say, of the flourishing of all creation (with a special emphasis on vulnerable persons), and an ultimate goal of achieving union with God."[16]

The much more difficult question is how dominant we should take the consequentialist element of Christian ethics to be. There is significant terminological confusion – which leads to substantive confusion – in the debate about consequentialism. The idea of a moral view which pays no attention to consequences at all is a strawman. The debate is rather about the question of whether we should pay attention to *nothing but* consequences (consequentialism) or, in case considerations other than consequences – such as rights, rules, virtues, or intentions – matter as well (non-consequentialism), *how much* attention these other considerations should receive.

I welcome the explicit acknowledgement of leading effective altruists that EA does not imply a commitment to consequentialism in the sense that *only* consequences matter. MacAskill and Pummer even define EA as "the project of using evidence and reason to try to find out how to do the most good, and on this basis trying to do the most good, *without violating constraints*."[17] In this crude form, the principle is not spelled out enough to do any serious work. (Possibly, the lack of nuance reveals that this qualifying clause is more of a deferentially accommodating – rather than fully authentic – attempt at

16 Camosy, "Engaging with Peter Singer," 69.

17 MacAskill and Pummer, "Effective Altruism," 5. The emphasis is mine. Another representative example is Todd and MacAskill's statement on the 80,000 hours website that their *default* position – which allows for exceptions – is not to "take a career for the greater good if that career directly causes significant harm." (Todd and MacAskill, "Is it ever okay.")

ecumenism?) Unless constraints were understood such that the acceptable exceptions are already baked into the constraint ("Don't lie unless x, y, or z is the case"), this crude wording rules out all the interesting cases where we should violate constraints in order to promote good results. It is hardly a controversial position that we should sometimes do so, both for the deontological sensibilities of common sense as well as for Christians. While the Bible obviously gives center stage to moral constraints, there are numerous instances where it is distinctively non-pharisaical about breaking these rules. An example is David eating the consecrated bread (1 Sam. 21) – an occurrence which is approvingly cited by Jesus in Luke 6 where the reason for breaking the recognised rules notably consists in down-to-earth benefits such as eating, healing, and saving lives. And it is not just ritual rules. In Exod. 1 the Egyptian midwives are rewarded by God for helping the Israelites even if doing so necessitated lying to the Pharaoh.

Leaving these preliminary remarks on consequentialism aside, the question then is: is the effective altruist call to focus much on consequences a helpful exhortation for Christians? The answer is: yes, at least as long as the call is limited. The call is useful but it should stop short of asking Christians to embrace full-blown consequentialism. Let me discuss the yes first and then turn to its limits. My argument for the yes is indirect: I portray consequentialism's main competitor – deontology – as a temptation. There are three ways in which humans are seduced to follow an excessively deontological stance and to give too little emphasis to consequences.

First, there is the temptation to seek an easy route to feeling assured of being on the morally right side by following simple rules. In Luke 11:42 Jesus exclaims "Woe to you Pharisees, because you give God a tenth of your mint, rue and all other kinds of garden herbs, but you neglect justice and the love of God." He does not dismiss rules but shifts the focus of his audience by telling them to "be generous to the poor, and everything will be clean for you." This is reminiscent of a parallel skepticism in the Old Testament towards focusing on ritual laws, for example in Mic. 6:6–8: "Shall I come to Him with burnt offerings, with yearling calves? (...) He has told you, O man, what is good; and what does the Lord require of you but to do justice, to love kindness, and to walk humbly with your God?" Following clear rules allows for peace of mind about having clean hands. In comparison, the open-ended task of taking responsibility for every way in which one's choices impact others is harder to bear.

Secondly, there is the temptation to dial down demandingness by relying on an excessively deontological morality. Admittedly, deontology can also be more demanding than consequentialism because it demands respect

for rights and rules even if doing so comes at great cost to oneself and others. But the bigger theme is how deontology can be less demanding than consequentialism because it makes space for supererogation. According to deontological morality, we do not have the *duty* to go for the morally best option. This allows for pegging the dividing line between duty and supererogation at a non-ambitious level. The temptation to exploit this is particularly strong given that deontology's complexity allows for hiding our motivation behind elaborate distinctions. Jesus criticises the complacency involved in merely jumping through deontological hoops in the Sermon on the Mount (Matt. 5:21–22): "You have heard that it was said to the people long ago, 'You shall not murder, and anyone who murders will be subject to judgment.' But I tell you that anyone who is angry with a brother or sister will be subject to judgment." In Matt. 19:8 Jesus presents the possibility for divorce explicitly as an accommodation of the human struggle with a more demanding morality. A particularly relevant example of implausibly toning down morality's demands by exploiting deontological distinctions happens in the case of positive duties. These duties – to actively do something rather than merely refrain from doing something – are often seen as significantly weaker than negative duties in a deontological perspective. In contrast, the Bible mentions positive and negative duties naturally together (cf. Prov. 14:31 and Luke 19:8). Charles Camosy – who remarks that Jesus rarely speaks of hell except in the context of duties to the poor – notes that in the famous passage of Matt. 25 the sins that separate the hell-bound from the heaven-bound are all sins of omission.[18] Key figures in church history equated the failure to share one's wealth with theft. And regular penance in many churches explicitly involves confession of both acts and omissions. The consequentialist call to take responsibility for all the consequences our actions and omissions is a very helpful *corrective* to the natural biases of sinful creatures who are tempted to hide selfishness in the complex edifices of deontological theories and to categorise too many actions as supererogatory nice-to-haves.

Thirdly, a further temptation of deontology is vengefulness. As consequentialism is forward-looking it only sees a point in retribution if this creates positive incentives for the future. Christianity, too, is skeptical of vengefulness: dealing with wrongful behaviour is God's business (and it can include forgiving it). While the Christian skepticism towards retribution has different roots than consequentialism's, Christians can still value the

18 Camosy, "Engaging with Peter Singer," 72, 74.

"letting the past be past" attitude as a helpful crutch to counteract our sinfully vengeful nature.[19]

My claim is thus that the consequentialist elements of any, including a Christian morality, are in danger of being crowded out by an overblown attention to deontological considerations. The temptation to do so exists because deontology allows us a simpler way to ascertain that we have paid our dues, makes space for toning down demandingness through motivated reasoning, and gives fuel to our feelings of vengefulness.[20]

There is a further – and key – reason why Christians typically give too little emphasis to consequentialist considerations in their practical reasoning. This reason has nothing to do with deep moral theory but rather with changing empirical circumstances. Most sensible deontological theories are stakes-sensitive: rights and rules can be overridden if the stakes – in terms of axiological value – are high enough (and this is so even if deontological theories struggle to spell out the precise conditions under which such overriding is justifiable). The stakes of many decisions have massively risen in the past few years and decades. Due to the growth of technology, wealth, and data, the resources and knowledge of ordinary people allow for much bigger impact than they used to. We should expect consequences to override deontological considerations much more often than in the past. Thus, *regardless* of whether our overall moral theory gives much or little space to consequentialist considerations, these considerations have become practically much more relevant. EA opens our eyes to this implication of almost any sensible moral theory.

Thus, EA is helpful in pushing Christians to be more consequentialist. However, Christians can only go so far in embracing consequentialism. Its scope must be limited in at least two ways for Christians.

First and straightforwardly, a number of key non-consequentialist considerations are forcefully present in scripture. The decalogue features a number of rules for specific kinds of actions rather than merely injunctions to promote certain values. The call to forgive only makes sense under the assumption that we can wrong each other whereas consequentialism struggles to explain how actions cannot just be wrong but wrong *someone*. The Bible

19 This is especially relevant in applying EA to the political realm where a lot of conventional action seems optimised for confronting the bad guys rather than serving their victims.

20 Of course, deontology is not the only competitor to consequentialism. In the common philosophical classification, virtue ethics is the most prominent third option. Here, too, there are temptations, in particular to shape the focus on one's own character development in a self-centered way.

exhibits much respect for the choices of agents even when these decisions wreak havoc. Paul says in 1. Cor. 13 that if we gave all we possess to the poor but did not have love, something essential would be missing. The list is endless.[21] First impressions could of course be deceptive and it might, for all we know, be possible to uncover a consequentialist rationale behind all the superficially non-consequentialist considerations. This would not be hard for some instances of non-consequentialism in scripture. For example, the command to honour the Sabbath or not to kill could reasonably well be consequentialised. However, the number and extent of non-consequentialist stances makes it a steep uphill battle to *fully* consequentialise the Bible.

Secondly, even if true morality were ultimately consequentialist, it is not clear that humanity is tasked with promoting *any* good consequence or whether we have particular responsibility for only a subset of all good consequences. It could be argued that there is a sort of division of labour between the Heavens and the Earth: God gives less responsibility to humans for certain areas of life since he has particular plans for these areas and since he has great epistemic and motivational advantages over humans. One possible example is this: there is a clear scriptural mandate addressed to humans to relieve suffering and oppression but a less clear mandate to bring about good effects over and above this threshold. The latter – renewing all things and bringing about blissful flourishing of the whole cosmos – might be disproportionately God's responsibility. Another example that is highly relevant to EA is longtermism: affecting the long-term trajectory of the planet such that it matches a predestined arch of history might be distinctly God's business.

21 In numerous discussions about Christianity and EA, I have repeatedly encountered the alabaster jar scene and the parable of the lost sheep as two of the most paradigmatic cases of a non-consequentialist vision in the Bible. However, these passages do not seem like clear instances at all. First, the alabaster jar scene: the fact that the disciples speak out – and that they address Jesus rather than the woman – indicates that their past experience with Jesus lets them expect approval from Jesus. And, in fact according to Mark 14:7, Jesus' reply does not sound like a fundamental repudiation of their general stance: "The poor you will always have with you, and you can help them anytime you want." His rebuke seems more focused on the exceptional situation at stake. The lost sheep, too, is rather more supportive than critical of a consequentialist stance. Obviously, the parable's message is not that the plight of the ninety-nine is inconsequential in the decision whether to save the one sheep. The message is rather – in line with the consequentialist insistence that the numbers count – that every single individual matters. More is better: even if ninety-nine are in safety, the hundredth sheep matters, too.

Population ethics deserves a specific mention in the limitations on consequentialism. Population ethics matters much for the prominent focus on existential risks among effective altruists: if humanity were to go extinct prematurely, we would miss out on all the good consequences that could have been experienced by the "containers" that will never be born. However, population ethics is a domain that might fall under either of the two above-mentioned limitations on consequentialism: first, it might be an area where consequentialism is the wrong approach for Christians or, secondly, it might be an area where consequentialism is the right approach but applying it is too overwhelming a task for earthlings. The first might be the case if we take total consequentialism to be the most obvious extension of consequentialism to the domain of population size. Scripture does not positively seem to endorse a total consequentialist stance. To the extent that it breaks its silence on the topic at all, the injunction from Gen. 1:28 to *fill* the earth seems to rather be of a satisficing nature. The optimal number of individual humans could well depend on the reason God had for creating humanity in the first place – where this reason is at least not obviously the maximisation of good consequences. The second might gain plausibility if we consider the fact that population ethics is one of the most perplexing areas of ethics.[22] It might surpass human understanding and thus be left to God.[23]

It is useful to note that limiting consequentialism in these two ways can make a genuine practical difference in *some* cases. For example, if Christians are committed to a *right* to subsistence, fulfilling the basic needs of the few can take precedence over fulfilling non-basic needs of the many *even if* good consequences were maximised by fulfilling the latter rather than the former. Or, it could be a correct course of action radically to reduce the number of animals who will ever be born in order to increase respect for the rights of those animals who will actually be born. Or, if respecting the free choices of individuals is key, the success of Christian missionary efforts should be measured – if at all – by the extent to which the good news is offered rather than accepted. However, all these examples should not hide the fact that all too often Christian limitations on consequentialism are surprisingly

22 Cf. Arrhenius, "The Impossibility of a Satisfactory Population Ethics."

23 Similarly, the epistemic challenges in comparing the goodness of salvation to the goodness of food on the table in the here and now might be insurmountable for human reasoning. Thus, if consequentialism were the correct approach to evangelism, it might be up to God to communicate practical guidelines for this domain to humans rather than humans getting lost in infinite ethics. On this, see also Liberman, "Effective Altruism and Christianity," section 4.

irrelevant on a practical level. Feeding the hungry, developing interventions against depression, reducing factory farming is important – and often similarly so – on both a Christian and a full-blown consequentialist stance.

(c) Welfarism

Given that Christians should focus much – but not as exclusively as consequentialists – on bringing about good consequences, the question comes up: what are the good consequences in question? Many effective altruists believe that the amount of welfare in a state of affairs exhaustively determines its goodness. In contrast, many Christians question such an exclusive focus on welfare. Instead, they insist that goods such as worship of God, relationships, knowledge, etc. matter not merely as potential constituents of – or contributors to – welfare, i.e. someone's life going well. Given that it would take more space than available to evaluate whether welfare is the *only* intrinsic value, I focus on the more limited question whether Christians should give *more* attention to welfare than they typically do. Even this limited question is difficult to assess, however. I therefore limit myself even further to suggesting that there are at least *some* elements of welfare which are in danger of being downplayed by Christians. These are the down-to-earth elements of welfare such as food or happiness.

Crisp observes that "[w]ell-being obviously plays a central role in any moral theory."[24] This is certainly true for Christianity. In particular, some simple and straightforward aspects of welfare here on earth play a central role in a Christian vision, too. The Bible gives much space to the theme of avoiding hunger, healing illness, and lamenting depressive states of mind. (And not just in the negative sense of avoiding a deficit: the Exodus from Egypt, for example, is not merely a journey out of oppression; it is also a journey into a land of milk and honey). Psalms do not just focus on worshipping God but also dwell on the gladness that accompanies worshipping God. On the very final pages of the Bible the new Jerusalem is characterised by God wiping every tear from our eyes and the absence of pain (Rev. 21:4). There is something profoundly odd with Christians – possibly as a counterreaction to the superficial visions of human purpose that some of their secular contemporaries or the adherents of the prosperity gospel espouse – dismissing a concern with the more earthly elements of welfare as too materialistic. If the key importance of human lives going well gets forgotten in overly moralising approaches to faith, and if welfare's blunter

24 Crisp, "Well-being."

elements such as enough and nice food get forgotten in overly spiritualising approaches to faith, effective altruism can serve as a *corrective*.[25]

There is also a practical reason why focusing on welfare, in particular on its blunter elements, helps Christians in living up to their faith's commitments. MacAskill notes that "given the current state of the world and our incredible opportunity to benefit others, the best ways of promoting welfarist value are broadly the same as the best ways of promoting the good."[26] Welfare is often a decent proxy for the good in general, and the blunter elements of welfare are often a decent proxy for welfare in general (plus: extremely crude measures such as Quality-Adjusted Life Years (QALYs) or GDP per capita are often decent proxies for the blunter elements of welfare). This is the case not least because the blunter elements of welfare contribute to welfare in general and also to non-welfare aspects of the good. For example, reducing the risk of hunger, illness, and poverty often has the effect that people can acquire knowledge, get empowered, can stand up for their rights and dignity, can express themselves and exercise their autonomy, and can foster their relationships rather than having to seek work far away from family and friends. The great benefit of focusing on material aspects such as food, health, and income is that they are more easily measurable than the harder-to-grasp elements of welfare and the good in general. Even if aiming at such crude proxy measures introduces some distortions, these distortions might be worth the increased tractability. And there are psychological benefits, too: focusing on measurable factors has a disciplining effect by providing accountability. It confronts us with the truth about the results of our efforts in a way that focusing on the ultimate values will always miss out on. We often forget how difficult it is *directly* to aim at deeper objectives. Many effective altruists who give center stage to the blunt objective of increasing QALYs do not mistake this crude measure for the real thing. They are fully aware that aiming at a limited measure of welfare is only a practical tool in the service of a grander aim. If, for a pilot in foggy weather who steers a plane with malfunctioning equipment, the

25 Note also Ryan Miller's comments regarding hedonism in one of the few existing pieces at the intersection of faith and EA. He argues that opposition to hedonism "is a strange line of criticism for Thomists, since Aquinas professes that happiness is the last end of man" (Miller, "80,000 Hours for the Common Good.") It should be noted, however, that Aquinas espouses a non-reductive view of happiness according to which happiness is found in God rather than in pleasure.

26 MacAskill, "The Definition of Effective Altruism," 18.

only alternative to operating on the basis of a crude proxy is flying blindly, then limited vision is better than no vision at all.[27]

Summing up: EA's emphasis on welfare, including its more down-to-earth elements, is a healthy reminder for Christians who are in danger of downplaying these elements of the good. Equally importantly, focusing on these blunter elements can be a perfectly helpful tool for better achieving the ultimate vision of the good; this is so on account of the tractability and accountability that the measurability of these simple goals provides.

(d) Impartiality

Effective altruists embrace an altruism that is very impartial: others count equally regardless of sex, geographical location, time of birth, species, etc. Should this effective altruist commitment to impartiality be welcomed by Christians?

There is a major challenge in assessing whether someone's commitments are overly partial or not sufficiently partial. The reason is that partiality is a surprisingly difficult concept in two ways. First, while impartiality plays a key role in any convincing moral stance it is an open question among which units one ought to be impartial. Should all interests count equally or should all individuals count equally? And if all individuals count equally, should all human individuals, all sentient individuals, or all living beings count equally – or should even natural systems or inanimate objects be included in the circle of units who ought to be treated equally? While the answer to this question makes all the difference, the mere commitment to impartiality does not settle it. Effective altruists who draw their intellectual inspiration from Peter Singer are often not very transparent about the fact that they have impartiality between interests or pleasure in mind rather than impartiality between individuals. This can be interpreted as partiality for those with weighty interests or much potential for pleasure. For example, if humans experience a lot more pleasure than insects then saving a human life is more important than saving an insect's life. The second reason why partiality is hard to spot is that a commitment to

27 To be sure: measurability, tractability, and accountability are no values *per se* if EA lives up to its ultimate commitments. They can lead us astray, for example if we choose more measurable interventions on account of the nice feeling of being able to vividly seeing impact, if tractable solutions bias us towards low-risk-low-impact solutions, or if accountability biases us towards short-term solutions given that future generations cannot hold us to account.

impartiality at the fundamental level of one's theory often leads to actions and institutions that look partial at the surface. There are a host of reasons for this. One is the very practical reason of division of labour: in order to organise life efficiently, some assume larger than average responsibility for particular people even if there is agreement that ultimately everyone deserves equal attention. Another practical reason is epistemic: we often know the preferences and circumstances of close people better than those of people far away. There are also more subtle reasons. For example, personal relationships are a significant component of a good life, and humans might be psychologically hardwired in a way that makes it difficult to cultivate relationships without exhibiting some partiality within those relationships (such as instinctively saving one's spouse from a fire ahead of other people without second thoughts).

Given these difficulties, I propose two tests for assessing whether Christians should welcome EA's strong call for impartiality as a prophetic voice that pushes them in the right direction. The first test asks whether the Bible and other sources of authority in Christianity concur or disagree with the effective altruist vision of impartiality, and the second test asks whether we have reason to assume that there are psychological drives that tempt us towards excessive partiality.

The first test does not yield a clear answer. On the one hand, there are obviously warnings of partiality in the Bible. For example, the Bible and Christian tradition speak out against tribalism by emphasising kindness towards strangers; Jesus and Paul are both critical of family commitments standing in the way of commitment to God's mission (Luke 14:26; 1. Cor. 7:8); and the New Testament universalises the faith that developed in the Old Testament (Gal. 3:28). It is particularly noteworthy that the parable of the Good Samaritan – a key passage on love of neighbour and the eponym in modern English for helping strangers – is not about the question whether one should love one's neighbour. That one should do so is the premise of the parable. The parable itself is about the question of *whom* to see as one's neighbour. By contrasting how the Samaritan – a person from a different culture than Jesus' primary audience – *chose* to act *as* a neighbour to the wounded man, Jesus seems forcefully to speak out against limiting one's responsibility towards the near and dear.

However, there is no simple and clear endorsement of impartiality in the Bible.[28] There are, for example, hints of a weak priority for sisters and brothers in faith (cf. Gal. 6:10). The weakness of this priority can

28 See also Liberman, "Effective Altruism and Christianity," section 2.3.

be seen in various instances. When Jesus was approached by a Canaanite woman in need, he responds by emphasising the primacy of his mission to the children of Israel. This emphasis, however, is quickly trumped by admiration for the woman's faith (Matt. 15:21–28). Ryan Miller notes that Aquinas, "grants that closeness to the giver of alms is a circumstance yielding a certain claim on care, [but] he nonetheless insists that it be a lesser criterion than the extent of need."[29] Similarly, John Wesley in his famous and in many ways effective altruist sermon on the use of money proposes a model which amounts to closeness acting as a tie-breaker in case various people's basic needs aren't satisfied.[30] However, even if this looks like an explicit endorsement of partiality it would still be a radical departure from real-world practice given that benefiting close ones over and above the level of basic needs would not be justified as long as the basic needs of anonymous strangers are not yet satisfied. Perhaps the clearest endorsement of partiality in the Bible can be found with respect to animals. Jesus explicitly says that one human is worth more than many sparrows (Matt. 10:31). The special role of humans is also evidenced by the Bible's more extensive focus on the ethical treatment of humans compared to the ethical treatment of animals. Of course, this should in no way be seen as a biblical endorsement of the mistreatment of animals in today's world. After all, in saying that humans matter more than sparrows Jesus simultaneously says that sparrows do matter. God the father cares for them, which is also evidenced by a number of Psalms mentioning God's provision for animals (e.g. Ps. 136:25) and by the fact that the covenant after the flood includes the animals (Gen. 9:9–10).[31] Note also that a Christian perspective could in principle overtake even Peter Singer in how it lifts up animals: if every creature is valuable in God's eyes, those creatures without or with only extremely little sentience, such as oysters,[32] could have a higher moral status in a Christian perspective than in a utilitarian perspective. Also, Christianity could join forces with utilitarianism in shifting attention from livestock animals to wild animals, on account of the fact that the foundational mandate to care for creation (Gen. 1:28) is not limited to a negative duty not to mistreat farm animals.[33]

29 Miller, "80,000 Hours for the Common Good."
30 Wesley, "The Use of Money," section III.3.
31 Further, today's mass exploitation of animals cannot be endorsed from a Christian perspective since it has multiple negative effects on humans, too, most notably via its environmental consequences.
32 And, for that matter, also unborn humans.
33 Cf. Crummett, "Human Dominion and Wild Animal Suffering."

These gesturing comments make clear that Christianity does not feature as straightforward a stance on impartiality as the utilitarianism that informs many effective altruists. Thus, if anything, the first test shows that the Christian ideal of impartiality is somewhat weaker and, in complex ways, of a different shape than the ideal of utilitarianism-inspired effective altruists. However, the question of the correct *ideal* regarding impartiality seems surprisingly irrelevant. I venture the claim that there are sufficient temptations for partiality in life such that almost anyone among us should adopt the working assumption of not being impartial enough, regardless of whether the benchmark for impartiality is utilitarianism or Christian ethics. Thus, all of us should move in the direction of *more* impartiality (even if different ideals disagree on how much more we should move in which precise direction). These temptations for overdoing partiality have various roots. One is the tribalism that evolution has ingrained in us. Another important root is the selfish motivation for being nice to those who are close-by and therefore have the power to be nice to us as well – a point which is highlighted by Jesus himself (Matt 5:46). A still further root is the fact that our social norms and intuitions have developed in a context where there was much less potential to expand one's concern to people far away in time and space. Factors such as these lead one to assume that common sense expectations about impartiality in combination with the possibilities of a 21st century world exhibit a much too parochial focus. Thus, EA's emphasis on impartiality pulls us in the right direction.

Summing up: when it comes to impartiality, the practically relevant question in assessing EA's helpfulness to Christians is not the extent of overlap of the impartiality ideal in theory. The practically relevant point is that Christians can assume their natural and sinful instincts to be overly partial – and thus welcome EA's enthusiasm and radicalism on impartiality as a forceful tool pulling in the right direction.

(e) Effectiveness

Effective altruists focus much on maximising the ratio at which the resources they use for benefiting others yield good outcomes. This is a direct implication of consequentialism since getting more out of the resources one puts to the service of others is one of two ways of bringing about better outcomes for others (where the other simply consists in giving more resources).

Effectiveness is by far the biggest reason why EA is relevant to Christians. The goal of loving our neighbours includes, as a key part, the goal of promoting their welfare. And EA has shown that we can increase the benefits for our neighbours to a much larger extent by translating our efforts more effectively into benefits than by increasing our efforts. While we can typically increase our efforts by a factor of, say, two or three, we can sometimes increase the factor by which our efforts are translated into results by a factor of ten, a hundred, or even more.[34] Forgoing these massive wins is wrong on consequentialist grounds. And to the extent that any plausible moral view gives much attention to consequences, EA helps us to better live up to this aspect of our overall moral view. On a more virtue ethical style of reasoning, it is wasteful; and the sheer fact of unnecessarily forgoing benefits which could be had without increasing our effort is revealing of a lack of compassion. If the benefits in question are necessary for fulfilling rights, forgoing them might well also be wrong on deontological grounds.[35] Note that none of this is a philosophically or theologically deep insight. EA has not discovered a new aspect of the fundamental structure of morality. Rather, it merely *opens our eyes* to an important *implication* of the benevolence that is part of any plausible morality. The relevance of carefully attending to effectiveness is the forest that critics of EA should see instead of the trees. Alas, the critics often note that effectiveness is a consequentialist concern which immediately triggers their worry of excessive consequentialism. It is a real pity that the conversation gets hung up on old debates about consequentialism in general, or even utilitarianism.[36]

34 See Caviola et al., "Donors Vastly Underestimate."

35 Some people are critical of such a "utilitarianism of rights" where we have greater reason to fulfil more rights than fulfilling less rights. In contrast, they insist on rights as constraints on the space within which maximization should happen. However, if in this fallen world humanity does not fulfil all the rights they should, I find it plausible that at least with respect to the shortfall, one should take a consequentialist approach: minimising the extent to which one fails to do what one minimally ought to do as a matter of rights-based duties.

36 McMahan ("Philosophical Critiques of Effective Altruism") laments this too, particularly in light of the fact that both Singer's ("Famine, Affluence, and Morality") earliest arguments as well as Unger's (*Living High and Letting Die*) more systematic arguments for demanding duties of the kind that effective altruists promote rely on widely held intuitions and are not dependent on utilitarianism. One of the very few academic publications that mention EA in the context of faith unfortunately also falls into the trap of putting EA too quickly in the context of old debates about utilitarianism (Gregory, "Charity, Justice, and the Ethics of Humanitarianism," 8–10).

If focusing on effectiveness is such a crucial instrument for responding in love to the cries of this earth and its inhabitants, one might wonder why effectiveness does not play a more prominent role in the Bible. One explanation is, of course, that it is a practical insight. The relevance of the practical insight depends on empirical circumstances. For almost all of human history it was comparatively unimportant. But its significance has skyrocketed over the past few decades due to the improved availability of data as well as scientific and technological developments. Another response is to claim that there *are* in fact two biblical "proof texts." The most obvious is the parable of the shrewd manager (Luke 16:1–13). It is a difficult parable to interpret but the upshot that seems most natural is summed up in Jesus' concluding sigh that, roughly speaking, the faithful should be as shrewd in pursuing good goals as evil people are in pursuing bad goals. The second is the parable of the talents where Jesus rebukes the servant who played it safe and made sure not to lose anything. The version conveyed in Luke 19:11–27 is particularly noteworthy: the two good servants achieved different returns and the one who maxed out on the opportunities was given extra praise and extra responsibility. Aside from these two texts, there are some further passages one could interpret as endorsing an effectiveness mindset. An example is Ephesians 5:15–16: "See then that ye walk circumspectly, not as fools, but as wise, redeeming the time, because the days are evil."

But whether or not one finds direct scriptural evidence for the importance of effectiveness, the indirect evidence is overwhelming: love of neighbour combined with the novel situation of massive and detectable differences in effectiveness between different cause areas and interventions in the twenty-first century point to effectiveness as a key moral imperative for Christians of our time.

(f) Truthseeking

Effective altruists put much effort into improving their beliefs, in particular beliefs about which actions are most effective. This praise of careful thinking is certainly in line with the Bible. Most clearly this is the case for the passages in praise of wisdom. It also chimes well with Christianity's high regard for truth and its trust in a certain intelligibility of the workings of this world. One should, however, be careful not to ascribe to EA any appreciation of wisdom as intrinsically virtuous.[37] For EA, the emphasis

37 Cf. Synowiec, "Temperance and prudence as virtues of an effective altruist," who discusses EA in the context of the virtue of prudence.

on reason is justified by its instrumental benefits: taking a step back to carefully examine the situation is supported as a means to improve the world, and not because such thoughtfulness is virtuous in itself. While some might lament that such an instrumental justification of wisdom does not go far enough, it does allay a worry that some Christians might have about EA: an idolisation of science, reason, and nerd culture. While such idolisation might sporadically happen in reality, it is not inherent to the underlying logic of the EA project – on the contrary. If certain unscientific epistemic practices were to achieve epistemic or non-epistemic benefits – say, if astrology-based counselling were to provably help people find orientation in life or mentally recover from crises – effective altruists would seem to have no principled objection to it.

A fact which is often overlooked is that the use of reason as an instrument of love appears in one of the most central passages of the Bible. In Matthew 22:37–39 Jesus identifies the commandments to love God with all one's heart, all one's soul, and all one's mind and the commandment to – likewise! – love one's neighbour as oneself as the core that binds together the Law and the Prophets. It is important to note, first, that in a Hebrew mindset the heart is pictured as the center of human thinking and planning rather than as the seat of emotions and, secondly, that the reference to the mind was added by Jesus himself to the wording of the quoted Old Testament passages. Thus, there is a genuine appreciation of reason's role in one of the most famous biblical calls to love.[38]

(g) Rationalism

Effective altruists celebrate specific methods and mindsets for arriving at beliefs that reliably track reality such as open-mindedness, quantitative tools, and empirical evidence. To the extent that these methods and mindsets are helpful tools for finding the truth, they must of course be endorsed from a Christian perspective, and there is not much to be added on this matter. Christians do recognise certain additional instruments for accessing the truth, such as direct instruction by the Spirit or authoritative teaching of the church. Also, given God's action in the world, Christians might view the world as somewhat less predictable than secular EAs. Thus, there are inevitably some differences. However, in practice, for most topics of relevance to the EA project, these differences are negligible. Christians often

38 This is laid out with respect to EA in a blogpost by Stefan Höschele, "Love with all your mind."

miss out on the tremendous practical benefits of these intellectual tools out of fear that they are embedded in a naturalistic ideology. If Christians were more pragmatic in this respect and took all these tools such as, say, simplified quantifications to be mere instruments that helpfully discipline the search for impactful ways of serving their sisters and brothers in need, so much would be gained. Much would already be gained if they trusted scientific tools as much when it comes to promoting their neighbour's welfare as when it comes to pursuing their own welfare, say in their own health or consumption decisions.

Let me sum up the gist of section 3. The dialectical situation is such that many theorists are wary of EA on the basis of their opposition to pure utilitarianism. Given that EA does not necessitate a commitment to utilitarianism, they miss out on just how powerful a tool it is for serving the needy. For Christians, there are certainly some elements of EA that they must not endorse as fully as some of its paradigmatic adherents. However, in almost all cases EA pulls Christians in the right direction relative to the status quo: for Christians to love their neighbours better, they should be more altruistic, more consequentialist, more welfarist, more impartial, more effective, more focused on truthseeking, and their epistemology more rationalist.

4. A Big Picture Tension: Letting Go of Control

There is an overarching mindset behind the EA project: "Be deliberate about the shape of your altruism. Rather than going with the intuitive flow and rather than taking current forms of expressing altruism as given, step outside the ingrained habits and embrace intentionality in choosing how to serve others. You are in charge of actively optimising effects across any cause areas by deploying all available resources. Don't embrace an attitude of non-judgmental acceptance towards reality – change it. Don't let things happen. If humanity doesn't take responsibility, no one will."[39]

In contrast, Christianity allows us – and demands of us – not to take control of everything we can affect. We are to let go of the hold we seek to have on everything and put things into God's hands. The mindset is one of surrender to God's mysterious and powerful presence in this world. Rather than acting like an engineer who fine-tunes every button on a big, complex machine we ought to espouse the mindset of children trusting

39 There is no necessary link between this mindset and the specific shape of altruism that EA has embraced. Rather the overarching mindset could lead to other upshots.

their parents to lead them well.[40] While EA encourages us to take control of things, Christianity encourages us to let go of control – and this tension has increased in recent years in lockstep with the increase of humanity's powers to control its fate.

Biblical examples include the following:

- In Judg. 7, God asks Gideon to deliberately go to war with 300 men even though 32,000 would have been available. Gideon is to deliberately refrain from making use of all available resources.
- In Matt 6, Jesus encourages us not to worry about tomorrow. The illustrations he gives are birds who do not invest for the future and the completely passive lilies.
- In Ps. 127, we are encouraged to take a good night's rest rather than labour late. This encouragement is based on the claim that "unless the Lord builds the house, the builders labour in vain."
- In Ps. 131, the writer approvingly compares himself to a child who says "I do not concern myself with great matters or things too wonderful for me." This is similar to the line from Isaiah 55: "For my thoughts are not your thoughts, neither are your ways my ways."
- James 4 encourages a particular mindset: "Now listen, you who say, 'Today or tomorrow we will go to this or that city, spend a year there, carry on business and make money' (...) Instead, you ought to say, 'If it is the Lord's will, we will live and do this or that." This is similar to Ps. 94:11 which says: "The Lord knows all human plans; he knows that they are futile."
- In Exod. 14, the Israelites are discouraged from an active stance in the face of an opposing army. Rather they are to observe the Lord taking action: "Stand firm and you will see the deliverance the Lord will bring you today (...) The Lord will fight for you; you need only to be still."
- Mark 4 provides one of the clearest instances: "This is what the kingdom of God is like. A man scatters seed on the ground. Night and day, whether he sleeps or gets up, the seed sprouts and grows, though he does not know how. All by itself the soil produces grain – first the stalk, then the head, then the full kernel in the head."

40 The key theme of letting go of ambition, control, and optimization is not unique to Christianity. It is particularly present in the mystical strands of various faiths as well as movements closely adjacent to spiritual paths, such as mindfulness or Alcoholics Anonymous.

- More generally, there is also the fact that God repeatedly chose unimportant and weak agents for doing his work – and they often achieved his purposes in mysterious, roundabout, and seemingly wasteful ways.

There are a couple of straightforward rationales for this emphasis on letting divine providence take its course rather than incessantly planning pro-active interventions. First, God has epistemic advantages. Humans need to remember their limitations and the concomitant benefits of listening to the one who has a much better overview of this complex universe. Secondly, some domains which humans can affect are simply not part of their portfolio. Rather, they are God's domain of sovereignty. Determining the lifespan of humanity or selecting specific humans for salvation might be cases in point. Thirdly, a lot of these examples can be interpreted to serve the mental health of overambitious do-gooders. Such a concern with self-care is in fact a point of overlap with EA. In contrast to other moral views, EA does not place an emphasis on good motivation and high sacrifice. Given that impact rather than effort matters, and given that not overburdening oneself with responsibility can serve impact in the long run, EA agrees with the upshots of a number of these passages. Fourthly, in a lot of these examples the point seems to be about character development, in particular practicing trust and humility. For example, Gideon was to rely on a small number of soldiers so as to avoid the temptation of boasting. Framing it as character development might cast this fourth point too superficially, though, and too much on the ethical level. The Bible reports on the experience of having to die off so as to receive life (for example in John 12:24–25) and surrendering completely might be part and parcel of this overall spiritual practice of losing oneself in order to find God.

These four considerations are speculative. And even if they provide some rationale for refraining from exercising control where it would be possible to do so, significant mystery remains. The paradoxical nature of the Christian stance of surrendering control when action would seem possible and advisable is expressed in such sayings as "Pray as though everything depended on God; act as though everything depended on you" or Paul's words "For when I am weak, then I am strong" (2 Cor. 12:10).

The rationale for foregoing control is not our concern here anyway. The concern is the tension with EA's underlying mindset of not letting any chance go unused to affect the world for the good. While this tension is real, one should, however, not exaggerate its extent. This is so for two reasons. The first is that the Christian faith also, and utterly clearly, affirms initiative, action, planning, and the use of reason to pursue outcomes in a

results-oriented way. If there is any tension with EA, this tension is already present within the Christian faith. The Christian faith's affirmation of a pro-active attitude towards shaping this world is limited, and it is embedded in an underlying trustful sense of complete dependency on God. To some extent the tension can be eased by going for the EA mindset in our actions and the Christian mindset in our attitudes. However, this only reduces the so-called paradox of surrender. If the Christian attitude of falling back on God's sovereign working in this universe is taken seriously, it must have *some* implications for our actions. There is a second reason for drastically limiting the tension. For most people in our fallen world, the alternative to EA – i.e. the alternative to a more controlling and deliberate approach to what one can affect – is typically not trust in God. Realistically, the alternative is typically thoughtlessly doing the first available good deeds on the doorstep and blindly continuing on well-trodden paths in one's charitable efforts. If EA encourages people to move from thoughtless forms of love to more intentional forms, this is at least a step forward – and this is so even if committing the efforts to God's wise providence were an even greater step forward. Even if a controlling attitude is spiritually problematic, it is at least an improvement over neither actively taking responsibility oneself nor actively placing this responsibility in God's hands.

5. Conclusion

An excerpt from Paul's prayer in Phil. 1:9–10 captures many of EA's ambitions: "This is my prayer: that your love may abound more and more in knowledge and depth of insight, so that you may be able to discern what is best." It expresses the longing for boldness in our love and the desire to use insight in order to go beyond merely doing good to doing the best. No wonder this chapter's examination of the various elements of EA arrived at the conclusion that they are useful in supporting the Christian endeavour of loving one's neighbour well.

The reversal of the question would be a relevant topic, too: to what extent is Christianity a better or worse soil than secularism for the EA project to flourish? While there is no space here to discuss this in depth, three points deserve a quick mention. First, Christianity provides a good home for justifying and motivating altruism compared to the more free-floating commitment in secular approaches. In addition to featuring firm and divinely endorsed moral commands, there are a number of motivations in Christianity that are not as easily available in a secular perspective, such

as gratitude for the overflowing gifts one has received or a confidence that one does not have to look out for oneself since someone else is in charge of doing so. Secondly, and perhaps controversially, I submit that while Christianity might not necessarily provide a better home for the commitment to science than a secular context – after all, superstition and extremism have flourished in religious circles time and again – it might possibly provide a more stable home in the longer term. Secular approaches are exposed to risks of nihilism and relativism and such extreme risks to the commitment to science are alien to many strands of Christianity. Thirdly, many effective altruists are committed to an extremely demanding morality. While demandingness may not in itself decrease the plausibility of these approaches in the slightest, it does pose an existential problem: how can we personally come to terms with the all-but-guaranteed failure to live up to an extremely demanding morality? At the very center of Christianity there is the emphasis on grace in the face of overwhelming moral demands. Our future and our wellbeing are dissociated from our shortcomings. In addition to practical psychological help in dealing with an overwhelming world, this experience responds to the demandingness problem at an existential level.

Effective altruists might of course point out how Christianity also offers a worse basis for EA in certain respects. For example, as this chapter argued, Christianity exhibits a less than full commitment to some of the project's elements such as consequentialism or impartiality. However, this less than full commitment is practically not of much relevance. In real life, the bottleneck is usually insufficient and messy human motivation rather than somewhat diverging ideals which would only gain much relevance if our motivation were much increased. Thus, for both secular and Christian effective altruists, the call to be more altruistic, more consequentialist, more welfarist, more impartial, more focused on effectiveness, more careful in one's belief formation and more rationalist in doing so is a helpful corrective to their natural tendencies. It is not just a mildly helpful corrective but – from the perspectives of the *billions* of victims of this *brutally* broken world – an incredibly important corrective. Thank God for the advent of effective altruism.[41]

41 I am grateful for the extensive and very thoughtful comments by Brian Green, Rochelle Harris, Vesa Hautala, Caleb Huffman, David Lawrence, David Moss, Stefan Riedener, as well as the participants of the workshop on "Religious Perspectives on Effective Altruism" in Fribourg in 2019 and the 2021 online workshop of the same title organised by Caleb Huffman. I am grateful for the many years of discussion in the Christians and Effective Altruism Facebook group founded by Andy Chrismer and the EA for Christians community that grew out of it.

References

Arrhenius, Gustav. “The Impossibility of a Satisfactory Population Ethics.” In *Descriptive and Normative Approaches to Human Behavior*, edited by Hans Colonius and Ehtibar Dzhafarov, 1–26. Singapore: World Scientific Publishing Company.

Augustine and D. W. Robertson. *On Christian Doctrine*. New York: Bobbs-Merrill, 1958.

Austin, Thad. *Giving to Religion*. Giving USA. 2017. https://givingusa.org/just-released-giving-usa-special-report-on-giving-to-religion/.

Berkey, Brian. "The Philosophical Core of Effective Altruism." *Journal of Social Philosophy* 52, no. 1 (2021): 92–113.

Camosy, Charles. “Engaging with Peter Singer.” In *God, the Good, and Utilitarianism: Perspectives on Peter Singer*, edited by John Perry, 68–90. Cambridge: Cambridge University Press, 2014.

Caviola, Lucius, Stefan Schubert, Elliot Teperman, David Moss, Spencer Greenberg, and Faber, Nadira. “Donors Vastly Underestimate Differences in Charities’ Effectiveness.” *Judgment and Decision Making* 15, no. 4 (2020): 509–516.

Crisp, Roger. "Well-Being." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta (Winter 2021 Edition). https://plato.stanford.edu/archives/win2021/entries/well-being.

Crummett, Dustin. “Human Dominion and Wild Animal Suffering.” *Religious Studies* 57, no. 4 (2021): 1–17.

Dullaghan, Neil. “EA Survey 2019 Series: Community Demographics & Characteristics.” Rethink Priorities. 5 December 2019. https://www.rethinkpriorities.org/blog/2019/12/5/ea-survey-2019-series-community-demographics-amp-characteristics.

Gregory, Eric. “Charity, Justice, and the Ethics of Humanitarianism.” In *Everyday Ethics*, edited by Michael Lamb and Brian Williams, 81–102. Washington, DC: Georgetown University Press, 2019.

Höschele, Stefan. “Love with all your mind.” A blog for the Effective Altruism for Christians community. 23 October 2019. https://eachdiscussion.wordpress.com/2019/10/23/love-with-all-your-mind.

Kraut, Richard. “Altruism.” In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta (Fall 2020 edn.). https://plato.stanford.edu/archives/fall2020/entries/altruism.

Liberman, Alida. "Effective Altruism and Christianity: Possibilities for Productive Collaboration." *Essays in Philosophy* 18, no.1 (2017): 6–29.

MacAskill, William. *Doing Good Better. How Effective Altruism Can Help You Make a Difference*. New York: Penguin Random House, 2015.
MacAskill, William. "Effective Altruism: Introduction." *Essays in Philosophy* 18, no. 1 (2017): 1–5.
MacAskill, William. "The Definition of Effective Altruism." In *Effective Altruism. Philosophical Issues*, edited by Hilary Greaves and Theron Pummer, 10–28. Oxford: Oxford University Press, 2019.
MacAskill, William, and Theron Pummer. "Effective Altruism." In *International Encyclopedia of Ethics*, ed. by Hugh LaFollette, 2020.
MacMahan, Jeff. "Philosophical Critiques of Effective Altruism." *The Philosopher's Magazine* 73 (2016): 92–99.
Miller, Ryan. "80,000 Hours for the Common Good: A Thomistic Appraisal of Effective Altruism." *Proceedings of the American Catholic Philosophical Association* (forthcoming).
Moss, David. "Maximising Altruism and Effective Altruism: Mapping the Terrain." (Unpublished Manuscript, 2016). https://www.academia.edu/34673134/Maximising_Altruism_and_Effective_Altruism_Mapping_the_Terrain.
Sablosky, Roy. "Does Religion Foster Generosity?" *The Social Science Journal* 51, no. 4 (2014): 545–555.
Singer, Peter. "Famine, Affluence, and Morality." *Philosophy and Public Affairs* 1 (1972): 229–43.
Synowiec, Jakub. "Temperance and prudence as virtues of an effective altruist." *Logos i Ethos* 54 (2020): 73–93.
Synowiec, Jakub. "Should we sound a trumpet when we give to the poor?" A blog for the Effective Altruism for Christians community. 17 March 2018. https://eachdiscussion.wordpress.com/2018/03/17/should-we-sound-a-trumpet-when-we-give-to-the-poor.
Todd, Benjamin and William MacAskill. "Is it ever okay to take a harmful job in order to do more good? An in-depth analysis." Accessed 24 November 2021. https://80000hours.org/articles/harmful-career/.
Unger, Peter. *Living High and Letting Die: Our Illusion of Innocence*. New York: Oxford University Press, 1996.
Wall, Steven. "Perfectionism in Moral and Political Philosophy." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta (Fall 2021 Edition). https://plato.stanford.edu/archives/fall2021/entries/perfectionism-moral.
Wesley, John. "The Use of Money." Accessed 24 November 2021. https://www.whdl.org/use-money-sermon-50.

Westphal, Merold. *Suspicion and Faith: The Religious Uses of Modern Atheism*. New York: Fordham University Press, 1998.

Yasin, Kidist Ibrie, Anita Graeser Adams, and David P. King. "How Does Religion Affect Giving to Outgroups and Secular Organizations? A Systematic Literature Review." *Religions* 11, no. 8 (2020): 405.

David Manheim

Ancient and Modern Conceptions of Charity: Orthodox Judaism and Effective Altruism

Abstract

Judaism has a strong commitment to charity, and both serves as an inspiration for many effective altruists, and is an inspiration for the concept of tithing, which has been adopted by and adapted to effective altruism. At the same time, the Orthodox Jewish structure of Halacha has complex boundaries and requirements which at least appear to be at odds with key tenets of effective altruism, including consequentialism and effectiveness. This chapter explores those tensions, especially between the individual obligations posited by Judaism and those imposed by strict utilitarianism. While Halacha is unyielding to fundamental change, it is also relevant to and often compatible with concerns which inform effective altruism, including consequentialism and prioritization. Other key points of disagreement, such as placing priority on supporting basic needs of family before those of strangers, are irreconcilable in theory but seem to dovetail with the practice of effective altruism. The chapter concludes with thoughts on how ideas about evaluation and effectiveness are both compatible with and should inform the practice of Orthodox Jewish charity organizations in the future.

1. Introduction

Orthodox Judaism has a long and complex history of engaging with contemporary moral and ethical concerns. Engaging with the modern theses of effective altruism (EA) makes a number of these historical debates and concerns relevant, and raises new concerns and questions that can be analysed through similar lenses. Peter Singer, one of the original promulgators of EA, defines it in his book on the topic as "a philosophy and social movement which applies evidence and reason to working out the most effective ways to improve the world."[1] Singer, an applied ethicist in the utilitarian tradition, notes that this "is a vague idea that raises many questions," and without imposing or implying a non-existent consensus within the movement, provides guidance on what the generally accepted principles of EA are, and where there is diversity of opinion.

It cannot be argued that Orthodox Judaism, with its unyielding emphasis on Halacha as the arbiter of correct action,[2] is compatible with the framework or assumptions of EA, but neither can it be reasonably claimed that EA is irrelevant to Jewish moral concerns. For this reason, it is worthwhile

1 Singer, *The Most Good*, 4-5.

2 Bleich, *Perfect Faith*; Broyde and Pill, *Setting the Table*; Walter, *Halachic Decision*; Soloveitchik, *Halakhic Man*.

to review a number of the central tenets of EA and situate them within the context of both ancient and contemporary Jewish debates, without attempting to resolve the various tensions. This paper focuses on four central issues that define EA and distinguish it from other contemporary moral philosophies: (1) the moral obligation to help others, (2) consequentialist and utilitarian reasoning about altruistic behaviour, (3) prioritisation of causes, and (4) the use of reason and evidence to understand effectiveness. After discussing the first three, and reviewing some difficulties understanding Judaism's approach to the fourth, the paper concludes with observations about how reason and evidence are required in normative Jewish law (Halacha), and how the standard will need to continue to evolve as it relates to charitable giving.

2. The Jewish Obligation to Give to Charity

It hardly needs to be said that charity is a critical value in Judaism, although perhaps it is worth noting that the EA movement draws inspiration for the requirement to give specifically 10% to effective causes from the Jewish (and later, Christian) practice of tithing. The question of consequentialism with regard to charitable giving, and the universalist foundations of EA, are more complex. In order to discuss the question of moral obligation clearly, we need to first introduce Jewish Law. While a complete overview of Jewish law is beyond the scope of the current paper, a few key notes on the topic are necessary for understanding the sources cited and their relationships. Following this background, we introduce first biblical, then rabbinic, obligations to give to charity. Finally, we compare the religious obligations in Judaism to moral obligations as they are understood by EA.

2.1 Introduction to Jewish Sources and Law

There are two classes of Jewish law: biblical and rabbinic. Biblical law, as the name implies, is based on the Jewish Bible, and on the interpretation and exegetical rules passed down from Moses. The tradition notes that the Bible empowers rabbis to both define how certain biblical rules apply and to safeguard and apply the law.[3] The biblical commands are limited to those written in the Pentateuch and the unwritten laws traditionally understood

3 For instance, while the mandate to pray to G-d is certainly biblical, the rabbis formulated a specific structure of prayer, and the requirement must therefore be fulfilled using the structure of prayer devised by the rabbis.

to have been given directly to Moses at Sinai. The Jewish Bible is used in rabbinic literature as a source of rules only as they are interpreted in light of the tradition.

Critical to understanding Jewish law is the power given to the rabbis at Mount Sinai to codify the rules and (within specific parameters) create new rabbinic laws. The transmission of the combined biblical and rabbinic law means that Halacha is not based directly on the text of the Bible. Instead, it is based on the tradition: first the codified oral tradition and law called the Mishna, finalised ca. 200 C.E., and mishnaic literature, then talmudic literature, finalised ca. 600 C.E.

Mishna and mishnaic literature are the oral traditions of the Law passed down through the generations to the rabbis in the generation after the destruction of the second temple. The Mishna was codified ca. 200 C.E., and both Mishna and other mishnaic literature contain statements and laws attributed to scholars living before that time. Talmudic literature primarily means the Babylonian Talmud, recording discussions that occurred between 200 C.E. and 600 C.E. These discussions include analysis of the Mishna, explanations of how verses of the Bible lead to the law, and interpretation and elucidation of other sources.[4] The analysis in large part explores and explains the relationship between biblical law and the text of the Bible, and the origins and structure of rabbinic law. For this reason, this literature is understood to reflect the entire historical transmission of the law,[5] and is the basis for all later discussion of Jewish law.

Following the codification of the Talmud, there were further eras of the development of Halacha, each of which is defined by a conclusive codification.[6] Talmudic literature debated and discussed by medieval commentators, rabbinic authorities living in the few hundred years prior to ca. 1500 C.E, is known as *Rishonim* (literally "early ones"). This period ended with the definitive work on Halacha, the *Shulchan Aruch*, by Rav Yosef

4 This literature extends beyond the Babylonian Talmud, however, and the discussions span secondary texts such as early extra-mishnaic texts such as *Baraita* and the *Tosefta*, and are recorded in texts outside of the Babylonian Talmud - including the Jerusalem Talmud, Midrash, (primarily homiletical exegesis with some Halachic content), and other similar sources from that period.

5 Jewish tradition maintains that these texts primarily codified (formerly orally transmitted) rules and knowledge that would otherwise have been lost during the few centuries following the destruction of the second temple and exile.

6 There are specific implications to the codification of the rules at each point. For instance, an explicit rule decided in the Talmud continues to apply, often even when the reason given is no longer applicable.

Caro, published in 1565, along with the now universal inclusion of the gloss by Rabbi Moshe Isserlis, first published in 1578. This text, like the Talmud before it, was universally accepted by Jewish religious authorities.

Later work in Halacha consists partially of *Shailos u'Teshuvos*, compendia of questions and answers about Jewish law, as well as commentaries and super-commentaries to the Shulchan Aruch, written by the *Acharonim* (literally "later ones"). As our discussion will find, these can revisit questions of biblical and rabbinic sources, within the structure of the tradition, as well as discussing novel situations or questions.

It is critical to note that the decisions of later eras of Halacha traditionally do not contradict the decisions of earlier eras, so that the Shulchan Aruch does not contradict the Talmud, nor *Acharonim* the Shulchan Aruch. Given this, however, the later opinions do decide between different opinions or interpretations of the texts written in each earlier era.[7] For that reason, normative Halacha does not typically follow the earliest decision, but the latest one, and each generation is told to follow the Halacha as decided in their generation. For this reason, the most recent works illuminate how the decisions cited in the earlier sources are in fact relevant to Jews today.[8] The focus on modern sources is important because we are interested in current and near future interpretations, rather than a purely historical analysis. At the same time, given the structure of Halacha, it is useful to examine sources relevant to each question via exploring the earlier sources.

2.2 Biblical Origins

> "If there is among you a needy person, from among your brothers, in your cities, in the land the Lord, your G-d, is giving you, you shall not harden your heart, and you shall not close your hand from your needy brother. Rather, you shall open your hand to him, and you shall lend him sufficient for his needs, whatever he is lacking. ... For there will never cease to be needy within the land. Therefore, I

7 This treatment is necessarily simplistic, as "the methodology by which some opinions are accepted and others excluded from application to practice constitutes a highly complex aspect of Halakhah." (Bleich, *Contemporary Halakhic Problems*, xvi).

8 Note that the current discussion is restricted to those who have maintained fidelity to this Halachic structure. This paper therefore excludes the approaches of the (comparatively very recent) Reform, Conservative, and other "Modern" movements which modify (rather than interpret and apply) the traditional law, or that abandon Halacha entirely.

> command you, saying, you shall surely open your hand to your brother, to your poor, and to the needy in your land." (Deut. 15:7–11)

While a variety of biblical texts discuss charitable giving of various sorts, the one most relevant to modern charitable giving is the above passage in Deuteronomy.[9] Per later sources, the exact phrasing of these verses is critical in understanding the contours of the obligation. For example, a discussion that becomes critical to the question of moral obligation is that verse 7 uses the terms "among you," "needy person," "in your cities," and "in your land." According to the *Sifrei*,[10] the order indicates a preference for the recipients of giving, so that according to most opinions, physical location creates a biblically mandated preference, albeit one that may be overridden by different levels of need.

Still, the biblical obligation laid out in Deuteronomy provides a baseline rather than a complete picture. The Shulcan Aruch notes in the very first rule about charity that the Bible repeatedly exhorts Jews to assist others by giving to charity, emphasising its importance. (Shulchan Aruch, *Yoreh Deah* 247). In addition to the theoretical discussion of importance, the actual emphasis on charity by Jewish communities is clear historically. Gregg Gardner presents a lengthy account of rabbinic obligations of charity,[11] which focuses on the early development of communal support for the poor in the 2nd to 5th centuries C.E. The evolution of views during this time makes it clear that in place of the agricultural charity that prevailed in the agricultural society of Judaism in Israel,[12] the rabbis viewed redistribution of wealth to the poor to be a central function of post-temple Judaism.

9 While the deontological/utilitarian distinction is not always a useful way of thinking about Halacha, other biblical forms of charity seem to be more deontological than *Tzedakah*. Specifically, the laws relevant to agriculture inside of Israel are detailed requirements rather than being primarily goal-oriented. In those laws, the Bible commands Jews to give a portion of their produce to the poor in several ways, by leaving a portion at the corner of each field of grain unharvested (Lev. 19:9–10), to leave additional parts unharvested (Deut. 24:19–21), to separate an additional portion for the poor during certain years of the sabbatical cycle (Deut. 14:28 and 26:12) and to leave the entire crop to be harvested by the rich and poor alike in the sabbatical year. In addition, there are non-agricultural requirements that apply outside of the narrow remit of farming which are also not strictly considered *Tzedakah*, such as the requirement to extend interest-free loans to brethren, so as not to oppress them.

10 A pre-talmudic compilation of biblical exegesis.

11 Gardner, *Origins.*

12 This is also an example of how Halacha remains binding, since these laws are once again practiced.

2.3 Rabbinic Obligation

The rabbinic obligation to give charity is based on the biblical idea of giving a tenth of a person's wealth to the needy. This was introduced by Abraham when he said, in Genesis 14:20, that he would give a tenth of his wealth to a priest of the Lord. Fourteen chapters later, Jacob accepts this obligation on an ongoing basis, saying that he will give a tenth of whatever he receives. This is not itself an obligation for future generations to give a tenth of their non-agricultural income, but forms the conceptual basis for the requirement.[13]

The way in which one fulfils this obligation is the subject of much discussion. There is some talmudic debate, but most of the discussion about the allocation of charity appears in later sources. For instance, there is a distinction drawn between charitable giving to community institutions and charitable giving to the poor. For example, Nachmanides' explanation, found in his commentary on Deuteronomy 12:6, discusses Exodus 35:24, where a surplus of funds is available for building the tabernacle, and the point is made that communal needs are limited, unlike personal donations. Once those needs are fulfilled, as seen in Exodus, communal leaders are responsible to stop further giving. No such limitation exists for giving to the poor, and while each individual has a limited requirement, the obligation to give remains.

There is also a critical point to make about the structure of giving in Jewish law. For historical and religious reasons, charitable giving from Mishnaic times (ca. 200 CE) to the late Middle Ages was in large part routed through *Kupot*, communal charity funds. Many such funds exist today, albeit sometimes in the guise of a "Rabbi's discretionary fund" or similar. These funds are not extraneous to the rules, and have their own set of guidelines in Halacha. For example, an individual may be required to give to these funds, or at least primarily to these funds, and may fulfil their individual obligation to give via donating to these funds.[14] The personal

13 There is some discussion about the origin of the obligation to give a tenth of income. Some, such as *Tosfos Taanit* 9a, "You shall surely Tithe", *Ohr Zarua* 1:13 8b, and *Sefer Chasidim* 144:1, seem to say it is an obligation derived via exegesis, while other sources refer to it as a rabbinic enactment, or as a binding (and required) tradition.

14 For support for this claim, see the discussion in Shulchan Aruch, Yoreh Deah 250:4, and the commentaries discussing the way in which a poor person who voluntarily withdraws from the *Kupah* system and collects from door to door potentially loses rights to charity.

obligation to give still has its own set of priorities, but the structure of these funds is critical for questions considered later.

2.4 Comparing the Obligation to Give[15]

Peter Singer makes the clear case in his book *Practical Ethics* that the wealthy have a moral obligation to help the poor on utilitarian grounds. The rich have sufficient resources, and even for someone with only somewhat utilitarian beliefs, a person should certainly be willing to sacrifice at least a small portion of their own comfort to help others, providing great benefit at very low cost.

Judaism takes a comparable view of the obligation regarding *Tzedakah* (Charity, or translated literally, Justice), albeit from a markedly non-utilitarian viewpoint; "He who ignores those in need is called wicked and is regarded as if he worships idols" (Shulchan Aruch, Yoreh Deah 247:1). The deontological prohibition against turning away the needy, "shutting your hand" (Deut. 15:7) from helping them, has a clear biblical source. Expanding on this, the latter clause of the Shulchan Aruch suggests an equivalence with idol worship, a sin considered on par with murder, deriving from the fact that all wealth comes from G-d, as the Rema's gloss notes just afterwards; "People must realise that they themselves are given sustenance by G-d," (Rema, Yoreh Deah 247:3) and failing to use the wealth granted by G-d to help others would, by this logic, be considered rejecting G-d. This argument's source can be traced to a passage in the Talmud (*Kiddushin* 82b): "Poverty does not come from a trade, nor does wealth come from a trade; rather, they come from the One to Whom wealth belongs, as it is stated [citing the verse in Haggai 2:8]: 'Silver belongs to me, Gold belongs to me, says the Lord of hosts.'"[16]

Returning to the comparison to EA, some object to Singer's view on the non-utilitarian grounds that the moral choice to help the poor is only important if donations are, in fact, made by choice.[17] Given that objection, it is worth noting that Judaism rejects this logic; in many cases, those who

15 Note that Judaism has a somewhat more segmented view of altruism than the EA community does, and views financial support, *Tzedakah*, as a separate obligation from other *Chessed*. The broader view of different but related obligations in Judaism is an important issue for the compatibility with EA views, or lack thereof, but will not be discussed in detail here.

16 See *Derech Eretz Zuta* 4:1 and *Midrash Tanchuma, Mishpatim* 12:4 for parallel exegesis.

17 Yu, "Obligation to Donate."

do not donate by choice can be compelled by the court to do so.[18] At the same time, Jewish law does allow a great deal of latitude and choice in the selection of recipients and causes.

3. Consequentialism within Deontological or Moral Reasoning

In this section, we will review the halachic basis and practice of charitable giving in Judaism, and suggest where consequentialism is still relevant. First, we will show that consequences matter within non-consequentialist explanations, even according to those embracing the deontological approaches of Halacha. Following this, we will discuss the rules of prioritisation for giving in Halacha, and why claims that these rules are incompatible with utilitarian concerns are misplaced. After noting a tension between individual and communal responsibilities, which we will return to in a later section, we finally note that the virtue-ethical approach which some (non-halachic) Jewish sources present also seems to indicate that there is a need to include some focus on consequences.

3.1 The Deontological Basis of the Requirement

Jewish Law has a clear deontological basis in doing the will of G-d. Medieval authorities suggest that even when Commandments (*Mitzvot*) have clear reasons, it is less important to find or understand those reasons than it is to simply perform the commanded act – and that our understanding of the reasons in no way modifies the command.[19] In other interpretations, the laws are pathways to moral perfection, and the performance of commandments is based on a more selfish (if arguably more consequentialist) reasoning.

At the same time, there are external consequentialist factors that determine what can, and must, be done to fulfil deontological or moral obligations, and this seems particularly applicable to the laws concerning charitable giving, where the obligation is fulfilled by giving which has a positive impact.[20] Specifically, if money is given in charity but, for instance,

18 Shulchan Aruch Yoreh Deah 248:1.

19 This seems to match some versions of divine command theory, but the conflation of Jewish and Christian views involved in many discussions of divine command theory, and the lack of clarity about what divine command theory in Judaism would mean, makes this unreasonable for reasons beyond the scope of this paper.

20 For example, see various commentaries on Talmud Bavli, *Bava Basra*, 10b–11a.

stolen before the recipient receives the money, the obligation is unfulfilled. Conversely, as noted above, if a person is forced by the courts to give, they fulfil the commandment.

3.2 The Virtue Ethics View

The above technical point about when the obligation is (not) fulfilled does not obviate the moral interpretation of charity, famously embraced by Rabbeinu Yona in his *The Gates of Repentance*,[21] and in the *Sefer HaChinuch* when discussing forgiving loans to the poor.[22] As we will note, loans are a primary source for these laws. The moral interpretation sees the idea of lending as a way for people to become more empathetic and become more similar to G-d, who is the ultimate source of support for everyone. This interpretation is widely accepted, but is not exclusive. That is, the moral aspects of the law do not replace the technical obligation noted above. Furthermore, it seems there is an intuitive justification, which is that building moral character is not accomplished by ignoring consequences.

As an additional point, as Rabbi Shimshon Rafael Hirsh notes in volume 3 of his collected writings, in practice each Jew is themself interested in doing good in the best way possible. That is, he notes that there is an intrinsic desire to do good, and therefore suggests that accomplishing good in the world is how individuals achieve morality. As suggested above, those who promote ethical perfection as morality in Judaism still seem to insist on a somewhat consequentialist viewpoint when discussing how ethical perfection is achieved. In other words, focusing on ethics independent of actual consequences seems to be rejected.

3.3 Priorities in Charity

The most important discussion in Jewish sources relevant to the question of consequentialism and moral priority for charity is the debate about the proper order of preference and allocation of charitable giving. The first source in this discussion is the talmudic discussion in Shabbat (63a), "Rabbi Abba quotes Rabbi Shimon ben Lakish as saying that loaning is a greater form of charity than gifts, and entering into business with the poor person is even better." Some commentaries claim this precedence is because in these cases the needy person is not embarrassed when receiving the assistance,

21 Yonah, *Gates,* Gate 7, Section 35 and 36.
22 *Sefer HaChinuch*, Commandment 477.

while others, such as the *Chidushei Agadot*, say that it is because the person is able to support themselves (rather than requiring further assistance).[23] In either case, the result is not justified by any deontological argument. The priority is based on the outcome – at least because it allows dignity on the part of the recipient, and also because it creates a longer lasting benefit of self-sufficiency.

There is a common thread in the discussions of charitable giving; prioritising is based on external consequentialist factors and maximising positive impact is seen as critical. The primary source for this is Maimonides, who delineates eight levels of charitable giving,[24] where the highest levels are the most helpful and least insulting to the poor, thereby prioritising not just their physical needs, but their emotional and other preferences or needs. Seemingly based on the talmudic passage cited earlier, charity which enables the poor to provide for themselves is cited as the most preferred method.[25] This preference among types of giving is included in the Shulchan Aruch nearly word-for-word, and this basic ordering of preference is essentially unchallenged in later sources.

We note that this idea of precedence among ways of giving is sometimes misunderstood, and in at least one case, an attempt is made to claim that these rules show Jewish law for charity is anti-utilitarian. Michael Harris states that this anti-utilitarianism is implicit in Maimonides' ruling (which in fact comes from the Talmud, *Kesuvos* 67b) that we must even give to a previously wealthy pauper according to his former status, purportedly proving that this is a deontological requirement rather than, and incompatible with, a utilitarian one.[26] While there is a deontological component, as noted above, a review of this particular claim shows how in such cases the preferences are clearly utilitarian.

Harris says that "whilst a utilitarian would undoubtedly grant moral weight to the previously wealthy person's pain at not having his subjective needs fulfilled, greater utility would surely be achieved by using available charitable funds to satisfy the basic needs of many rather than to guarantee the comfort of the few." However, Harris is mistaken in the inference that

23 While the original sources do not put this in terms of economic efficiency, it seems arguable that this is at least part of the reasoning, per Maimonides' preference for prioritising the most helpful types of charity, such as setting up someone in a business rather than directly giving money.

24 Maimonides, *Mishneh Torah*, Gifts to the Poor, 10:7–18.

25 In fact, according to many authorities, those that fail to provide for themselves despite the ability to do so have no right to accept charity. *Tur* Yoreh Deah 255:1.

26 Harris, "Consequentialism."

this ruling prioritises the single rich man over multiple poor people. To start, Maimonides' phrasing implicitly clarifies this is an extreme rather than a typical case, saying that we give "even" that much - implying that it is a less pressing priority. An even clearer refutation is included in both the definitive compilation of the Tur, and then the Shulchan Aruch (Yoreh Deah 251:7), where they note that more basic needs have precedence. It is also rejected explicitly by modern authorities, such as Rabbi Ari Marburger. Marburger addresses the logic of Harris's claim, clarifying that "although a particular pauper may claim a high standard of living, if there are limited funds, we focus on the other paupers that have more basic needs."[27] This does not prove a consequentialist viewpoint, much less a utilitarian one, but neither does it conflict with such an understanding.

To briefly return to the discussion of the technical obligation, we should note that while the requirement to give is on the individual, decisions about allocation and level of need, as well as the requirement to give according to the standard the recipient is accustomed to, are typically made by a communal fund (Rema's Gloss to Shulchan Aruch, Yoreh Deah 250:1). This seems to imply that the allocation decisions need not be a moral decision strictly on the part of an individual, since they can be delegated. This delegation also becomes critical in our later discussion, where we argue it indicates compatibility with EA, albeit for other reasons.

4. Comparing Effective Altruism to the Prioritisation of Giving in Jewish Law

Certain factors are regarded as morally clear by most in the EA community. Among these, EA assumes three things which are discussed in Jewish Law; (1) The moral need to help those in need, (2) the lack of moral relevance of physical distance, and (3) the relative priority of more needy individuals, at least when the ability to help is equivalent. The first was discussed above, and while Jewish reasoning about the second two factors clearly differs, there are parallels in Halacha to each of these claims, and the parallel is even closer when considering the practice of EA.

The rules of precedence of recipients in individual giving are (despite minor differences) widely accepted among medieval commentators and Halachic authorities. For example, the first requirement is ensuring one's family has sufficient funds. These rules seem to imply that a person is not

27 Marburger, "Tzedakah."

obliged to give money to others before first fulfilling his own needs, then his immediate family's needs, and finally his extended family's needs. In all cases, as further discussed below, this is limited to basic personal and family needs.

Once those basic needs are met, charitable funds are available to be spent on others with greater needs. Similarly, we find that fulfilling the basic needs of others comes before being available for further charitable spending. This prioritisation does not, however, require effectiveness from the giver. Specifically, the order of precedence seems to be non-mandatory for the individual,[28] and they are allowed to give at least a portion of funds to lower-priority, less pressing causes or individuals.[29] At the same time, as we will argue below, it seems that this prioritisation based on impact and importance is, or at least might be, binding on *Kupot*.

We note that EA as a movement takes a similar tack when suggesting that the world's richest people, whose personal needs are already fulfilled, should donate effectively. These are people whose needs, and their immediate family and friends' needs, are already met. In this way, we see that EA has a similarly tiered view of needs in practice, as even its most strident advocates ensure that they have sufficient personal income, and can provide for their immediate families at a level far exceeding that of the poorest individuals in the world. From a strictly utilitarian viewpoint, this would be at least arguably immoral, but Halacha embraces a more pragmatic viewpoint, closer to that of actual practice in EA.

28 As cited in Footnote 13, per the *Sifrei*, the precedence is biblical, but that exegesis also explains the phrase "which he is lacking" as granting greater precedence to those with greater need, overriding the other preferences. Chasam Sofer, a modern commentator, notes that the order is a preference, not a strict exclusion of later levels, and that greater needs would obviously take precedence, in the second volume of his commentary to Shulchan Aruch Yoreh Deah, Section 234.

29 Cf. *Tana D'Bei Eliyahu* 27, which explicitly says that one must feed the highest priority relatives, and only the remaining money should be spent on the next priority level. This, however, refers only to providing sufficient food, and perhaps implies that food for strangers would precede housing and clothing for relatives, even when lives are not in danger. Note that while this compilation was finalised in the 10th century C.E., it significantly predates, and is cited by, the Talmud.

4.1 The Moral (Un)importance of Distance

There is an argument in Peter Singer's article "Famine, Affluence, and Morality," which dismisses the moral relevance of distance. This has been embraced by effective altruists, and forms one of the two key justifications, along with relative need, for advocating for charity for the world's poorest people. While this approach is certainly not reflected in Halacha, the difference in approach could easily be misunderstood to strongly conflict with Singer's claims. Given this, it is worth considering the question of geographic preferences in Halacha in a bit more detail.

The first issue in Jewish religious sources that relates to geographic location derives from a simple question. Does the negative commandment to not ignore the needs of others apply only when one sees a poor person, or does the mere knowledge of the needy person trigger this obligation? Maimonides and the *Sefer HaChinuch*, at least, seem to agree with Singer's (much later) argument that there is an obligation to give once a person knows a need exists, regardless of distance.

The *Sefer HaChinuch*, in his explanation of commandment 478, interprets the biblical command to not "'withhold kindness and charity from our brother Israelites…' since we know the weakness of their situation, and we have the ability to assist them." This is clarified by the contemporary authority Shmuel HaLevi Wosner (1913–2015). He states that

> "it seems clear to me that we are not discussing a case where the pauper is in front of you begging… and it seems this issue is the dispute between Maimonides… where he says that it is enough to know the poor person exists…[and the *Rashba*][30] does not have this implication, since he says the poor person must be in front of you, requesting money" (*Shevet HaLevi*, 9:199).

In his conclusion, Wosner says that not only must one give whenever a need is known, but that this obligation to not ignore the needs of others, is paramount if the need is for survival or basic needs. On this basis, he maintains that donations to the poorest and most needy recipients take precedence over other less impactful and less pressing charitable needs.

Despite this conclusion – namely, that one may not refrain from giving on the basis of distance – there is a clear geographic precedence included in the prioritisation of needs. As we will see, the moral approach of Halacha differs from that of EA due in large part to the latitude provided to givers to choose causes and specific recipients.

30 A second medieval commentator who disputes Maimonides' view on this point.

Following one's family, the rules of priority discussed earlier specifically give precedence to those in the giver's neighbourhood, then their city, and next to paupers in Jerusalem, and finally to those located elsewhere. The requirement to give to people anywhere discussed above, however, is a biblical one. This is per the Talmud, (*Bava Metzia* 31b), which interprets the repetition of the word "open" in Deut. 15:11, "...therefore I command, open, open your hand to the poor and needy in your land" to require giving not only to those in one's own city, but those who live anywhere in the world. Interestingly, this ruling is interpreted in the Shulchan Aruch to imply that those in one's own city have precedence, and those located elsewhere are secondary.[31] At the same time, as we will note below, this geographic precedence is limited.

4.2 Applications and Implications of Precedence

A question remains concerning how the requirements of precedence operate in practice. From at least one viewpoint, precedence comes from the requirement that *Kupot* support those to whom the money was donated. From this perspective, we note that in general one who donates to a local charity intends to support their local community first, and so the charitable fund must respect that. This is the subject of talmudic debate (*Bava Batra* 8b–9a), and the advice is given that the charity collector can explicitly condition giving to the fund on the discretion of the people appointed to run the fund, as explained in the talmudic passage, quoted below.

> "Rabba [a talmudic sage] would make two purses [when collecting], one for the poor of the rest of the world, and one for the poor of his city. Once he heard what Shmuel said to Rav Taḥalifa bar Avdimi: 'Make only one purse, and make a stipulation about [what the money will be used for],' so Rabba made that stipulation. (The commentary of Rashi explains, "stipulate to the people of your city that the money will be given to whomever needs.") Rav Ashi [another sage] said 'I do not even need to make a stipulation, because whoever donates relies on my discretion.'"

Despite the stipulation giving an ability to a communal fund to distribute funds elsewhere, there remains the geographic preference which is biblical in origin, as noted above.[32] Two points can be made to explain this. Firstly, the apologetic answer is that in the ancient or medieval world, the ability to give to those geographically distant from one's self was very limited. Perhaps

31 The source for the preferences is complex, based on the verse in Deuteronomy cited above.

32 See the *Sifri*'s biblical exegesis cited in footnote 13.

this logic is the source for the Talmud's assumption that local paupers are included, but an extra word is used to include those elsewhere. Secondly, according to accepted Jewish law, geographic distance is only relevant when people in each place have similar levels of need.[33] This is implicit in the Shulchan Aruch (Yorah De'ah 251:7) where it states that paupers needing food always precede those who need clothing or other goods.[34] This precedence is expanded upon in modern rulings to refer to any difference in levels of need. If interpreted very broadly, this might imply that geographic precedence is merely a tiebreaker for otherwise identical needs, although no halachic authority seems to go quite that far.

As a concluding point, it is interesting to note that, just as EA suggests neglectedness as a criterion, something akin to the economic concept of replacement value is considered in Halacha. Specifically, those who can receive charity from elsewhere are considered lower priority than those who cannot, though no explicit formulation of how this affects precedence is given (*Yisrael Meir Kagan, Ahavas Chesed* 6:3, published 1888). This seems to show that a conceptual understanding of marginal effectiveness with regard to giving, at least, is compatible with Halacha.

The above provides some understanding of needs, neglectedness, and the halachic basis for prioritising need over location. Based on this, we will suggest below that it seems plausible that modern Jewish law regarding charity funds, though not individuals, effectively matches the EA position that those who would be helped most should be given precedence when facing limited charitable resources.

5. Effectiveness, Reason, and Evidence for Charitable Giving

Related to the discussion about precedence between various needs is the question of effectiveness. In EA, a central debate concerns cause prioritisation and effectiveness. For example, there has been discussion about whether it is better to give money directly – to save human lives by funding anti-malarial nets, for instance – or to promote economic growth and save lives more indirectly. Similar issues have been raised by effective altruists about the relative importance of alleviating suffering versus saving lives,

33 See commentary to Yoreh Deah 251:7. For the modern ruling, see Yosef Fleischman, *Beis Din Nesivos Chaim*, who discusses this in a blog post: http://dinonline.org/2010/08/12/laws-of-tzedakah-part-ii-who-to-give-first/.

34 This is, according to Vilna Gaon's explanation there, based on a passage in the Jerusalem Talmud in *Horayos*, Chapter 3, Section 4.

though on this point Judaism is unambiguous about the preference for the latter.[35]

The consensus in EA is that while the relative values and importance of different causes is a critical concern, individual givers can and should have their own values, and make their own choices about giving.

While critical to our discussion, it is unsurprising that the idea of effectiveness in charitable giving as understood today is nowhere to be found in early Jewish sources. Not only is the idea of efficiency a modern, industrial one,[36] but the tools needed to evaluate effectiveness, such as cost-benefit analysis, are very recent, and only truly started to be used in the second half of the 20th century.[37] Despite this, there are both relevant discussions in modern Halacha, and at least two reasons to suggest that the question of effectiveness could and should become more widely discussed in Halacha.

Rav Yosef Fleischman cites Chasam Sofer (YD 231) as implying that we have an obligation to "investigate and compare the poverty levels of the poor" in order to prioritise charitable giving. In contrast, the preeminent Halachic decisor of the previous generation, Rabbi Moshe Feinstein, (1895–1986), rules in his compendium of Halachic responsa, (*Iggros Moshe Yorah De'ah* 1:144,) that an individual is allowed a wide amount of latitude in choosing the recipient of their charity. The individual is legally entitled to this discretion. He clarifies that this follows a logic from a variety of talmudic sources where the money is ought to someone, but the benefit of choosing the recipient (*Tovas HaNa'ah*) is reserved for the giver. For this reason, not only is an individual not required to investigate where the need is greatest, but they can even ignore that information once it is available.

But even according to Rabbi Moshe Feinstein, the same leeway is explicitly not given to *Kupot*, the charitable funds. This is for two reasons, the first of which, the requirement to give according to relative need within the local geographic areas, is discussed above. The second reason, the responsibility of the fund to fulfil the will of the givers, is a very general requirement that we will argue has several important consequences.

First, the money given to the fund seems to be able to be distributed at the discretion of the individuals in charge of the fund, as noted in the talmudic debate quoted above, (*Bava Batra* 8b–9a). In general, however, there is a central fund, and in such a case the discretion is transferred from

35 The Halachic position on population ethics is an interesting but seemingly unaddressed topic.

36 Alexander, "Efficiency."

37 Pearce, *Cost-Benefit Analysis.*

the individual to the community, where the use of evidence seems to be required. This is explained in Shulchan Arukh, Yoreh De'ah 251:5, "Once a man has contributed a sum of money to the Gabbaïm, neither he nor his heirs have any power over it, but the community can do with it what is pleasing in the eyes of G-d and man." There is a presumption that the funds were intended for the local poor, and per 256:3, the primary responsibility of distributing funds is given to a panel of three people, who are required to judge the needs of individuals and distribute according to them. This, it seems, explicitly requires their reliance on evidence, albeit far more weakly so than EA. At the very least, Halacha strongly endorses a responsibility to some level of stewardship, as noted below, in Shulchan Aruch Yorah De'ah 257. Being a responsible steward, it is therefore clear that the charity should allocate funds to maximise effectiveness within the community, rather than allocate, for example, on a first-come first-serve basis.

The limitation to within the community, however, only applies to charity given to a fund where the explicit purpose is the local poor,[38] and, even in that case, not only can that fund exercise discretion about which individuals are most needy, but it could also decide, as the talmudic discussion implies, to allocate the money to those in need elsewhere.

The obligation is further explained in Shulchan Arukh, Yoreh De'ah 256–257, which contains a variety of requirements about avoiding even the appearance of impropriety. In the modern context, it seems that these rules imply that charitable funds must abide by at least the norms of proper management, which include a variety of ethical standards and both accountancy and accountability practices.

For basic requirements of non-profit organisations, there is a precept in Halacha that in a wide class of cases secular law is as binding on Jews as a religious requirement. Beyond that, however, it seems clear from Yoreh De'ah 257:1–2 that charities are required to go beyond standard practice in avoiding any appearance of impropriety. If givers expect non-profits to demonstrate outcomes and effectiveness, it seems that these standards would become required according to Halacha as well.[39]

38 And this was a strong historical norm, to which the Shulchan Aruch notes he has never heard an exception.

39 For basic requirements of nonprofit organisations, there is a precept in Halacha that in a wide class of cases secular law is binding on Jews as a religious requirement. Beyond that, however, it seems clear from Yoreh De'ah 257:1–2 that we require charities to go above standard practice in avoiding any appearance of impropriety.

6. Conclusion

It should be clear that Halacha is its own system, and any attempt to fit it into another framework is fundamentally misguided. At the same time, Halacha is aware of and engaged with the reality of charitable giving, and this means that the concerns of EA are relevant, not to guide or change Halacha, but to inform it. This includes the ability to influence individual charitable decision making, which allows a degree of latitude in cause prioritisation, allowing people to embrace effectiveness. Furthermore, while certainly not accepted normative Halacha at present, we speculated that as norms in charitable giving change, it might lead to a Halachic requirement for communal charities to follow best practices in evidence-based charitable giving. That is, to the extent that those norms embrace more effective and more egalitarian giving, the practice of giving in Halacha may become at least somewhat more similar to the practices endorsed by EA.[40]

References

Alexander, Jennifer K. "The Concept of Efficiency: An Historical Analysis." In *Philosophy of Technology and Engineering Sciences*, edited by Anthonie Meijers, 1007–1030. Amsterdam; Boston: Elsevier/North-Holland, 2009. https://doi.org/10.1016/B978-0-444-51667-1.50041-0.

Bleich, J. David. *Contemporary Halakhic Problems. Vol 1.* New York: KTAV Publishing House / Yeshiva University Press, 1976.

Bleich, J. David. *With Perfect Faith: The Foundations of Jewish Belief*. New York: KTAV Publishing House, 1983.

Broyde, Michael J., and Shlomo C. Pill. *Setting the Table.* Boston: Academic Studies Press, 2020

40 The author is grateful to Dominic Roser for organising the *Religious Perspectives on Effective Altruism* workshop, and to him and other participants for providing early feedback on these ideas. He would also like to thank Rabbi Eli Brackman for encouraging the original formulation of the ideas that turned into the presentation, then paper. He also would like to thank both Ben Schifman and, again, Dominic Roser, for assistance in clarifying and revising the text. Lastly, but most especially, he wishes to thank Rabbi Yosef Ashenberg, Assistant Rabbi, Chatam, Rehovot, Israel, for assistance in further studying the origins and sources of the laws regarding charitable giving, and providing extensive feedback on and discussion of the ideas and text of the manuscript.

Fleischman, Yosef. "Laws of Tzedakah Part II: To Whom To Give First." 12 August 2010. http://dinonline.org/2010/08/12/laws-of-tzedakah-part-ii-who-to-give-first/. 2010.

Gardner, Gregg E. *The Origins of Organized Charity in Rabbinic Judaism.* Cambridge: Cambridge University Press, 2015.

Harris, Michael J. "Consequentialism, Deontologism, and the Case of Sheva ben Bikhri." *The Torah u-Madda Journal* 15 (2008–2009): 68–94.

Marburger, Rabbi Ari. "Tzedakah and Maaser Kesafim: A Practical Guide To Giving Charity." Torah Live. Accessed 3 September 2021. https://d1xmgiex1rsxf4.cloudfront.net/wp-content/uploads/2014/11/Ari-Marburger-Booklet.pdf.

Pearce, W. D. *Cost-Benefit Analysis.* London: Palgrave, 1971.

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

Soloveitchik, Joseph B. *Halakhic Man,* Philadelphia: Jewish Publication Society of America, 1983.

Walter, Moshe. *The Making of a Halachic Decision.* New York: Menucha Publishers, 2013.

Yu, Richard K. "Against A Moral Obligation to Donate." Reformer Magazine. 22 Mar 2018. https://medium.com/reformermag/against-a-moral-obligation-to-donate-34342ced899c.

Mara-Daria Cojocaru

This World Needs More (than One Kind of) Love. A Modest, Murdoch-Inspired Proposal to Take the Heart in Effective Altruism More Seriously

Abstract

Debates about effective altruism (EA) often focus on reason and evidence in promoting the good. I suggest looking at the "heart" in EA, and taking seriously that, and how, other-regarding attitudes, specifically love, can both motivate and complicate moral agency and inquiry. This counters a misleading dichotomy between reason and emotion and a crude view of love as a "care-o-meter". Building on MacAskill's idea that there is an "engineering" and a "science" level to morality, I suggest that there is a pragmatic and an epistemic level of promoting the good and that two different kinds of love operate at each: at the former, which would be MacAskill's "engineering" level, and concerned with translating what we know about the good into action, an outward impartial, agapistic love is most adequate, with a philic dimension regarding rational self-esteem; at the latter, which would be MacAskill's "science" level of morality, and concerned with realising what is good, an inward erotetic love is important. These levels should be distinguished because confusing them leads to a distorted view of the good. It also leads to a mismatch between effective altruist normative theory and actual effective altruists' plurality of motivations, complicating EA's transition to a social movement. Making moral life and theory potentially more complicated (but realistic), I turn to Iris Murdoch and suggest that these levels could be modelled on her distinction between "public morality" and "private morals". I illustrate the benefits of my proposal with two vignettes involving two fictionalised moral counsellors: the android robot Sophia and the rationalist human Ajax.

1. Promoting the Good. A Pragmatic and an Epistemic Level

We live in a world with an almost endless amount of suffering that will move anyone with a heart in their chest to want to do good. No rocket science and no complicated theory are needed to appreciate this simple fact. Yet, in order to do good better, some have argued that we need to abstract from other-regarding feelings like love or care because … our hearts are just not big enough to comprehend and process the actual enormity of the suffering. We respond lovingly and with compassion to the individual suffering in front of us, but we become overwhelmed and fatalistic once we realise the scale and enormity of all the suffering we should care equally about. According to influential effective altruists like Nate Soares, then, it is often not that we *do not* care when we ultimately do nothing to alleviate that large-scale suffering. Rather, as Soares suggests in a blog post, our "care-o-meters" seem just as unsuitable to give us precise readings of the misery we need to tend to as a simple thermometer is to measure a

bushfire.[1] Thus, if one wants to be effective in one's caring, one should turn to numbers and rely on those when putting a universal concern for suffering – regardless of where it occurs – into practice.

Historically, this universal concern has been utilitarianism's signature axiom but, again, hardly needs a complicated moral theory for backup. Moreover, in order to have as many people as possible adopt that concern and turn to better ways of doing good, it might be helpful to grow effective altruism (EA) as a social movement. In a recent definition of EA, Will MacAskill assures the interested reader that "we are not attempting to describe a fundamental aspect of morality",[2] and that EA is "consistent with any moral view".[3] More precisely, he advertises EA as a piece of moral engineering, and as such as different from moral science, with science being "the attempt to discover general truths about the world we live in"[4] and engineering "the use of our scientific understanding to design and build structures or systems that benefit society".[5] The normative view that informs the idea of benefit here is "tentative impartial welfarism",[6] which MacAskill takes to be an ecumenical view, a view that could be endorsed by many. I believe that he is right if this is to be understood as applying to the "engineering" level and wrong if it is to be understood as applying to the "science" level (and I believe he would agree with this).

I also agree that moral science or philosophy needs to pay more attention to the engineering side of things, or to what I call "moral pragmatics".[7] Yet the trouble is that in its attempts to build a social movement, EA does not just come across as an ecumenical or theory-neutral way to engineer a better world, but vies for the hearts of people as well. This risks distorting the larger enterprise, which is the science of recognising the good. It does so because it conflates various ways in which people can and need to be guided by love in their actions, which is what allows them to learn about the good in the first place. In the disciplines that work on recognising the good (e.g. moral and political philosophy, psychology, anthropology, among others), it is still a point of contention whether an impartial concern with welfare on the basis of quantifiable concepts of wellbeing and quality of life

1 See Soares, "On Caring." I have built on the sentiment expressed in the blogpost to introduce the problem about the limits to our love.
2 MacAskill, "Definition," 11f.
3 Ibid., 16.
4 Ibid., 11.
5 Ibid., 12.
6 Ibid., 14.
7 Cojocaru, "Doing Ethics?"

is methodologically always superior to a partial concern with specific rights violations or to somewhat ineffable intuitions about the common good that is more than the sum of individual benefits, not to mention conceptions of inherent value. On this perennial axiological question, the jury is still out, and, in my opinion, all perspectives can elucidate important aspects of the difficult and problematic moral landscape we all have to navigate by our own lights. Since EA axiologically operates just on the impartial welfarist view, it is in as good a position to exclusively promote the good as one would be to water a garden with only a fire hose: yes, that can help in times of drought, yet it will destroy many of the more delicate things that thrive in it.

I believe it would be better for EA to salvage its straightforward utilitarian heritage, yet understand it as confined to the level of what Iris Murdoch has called public morality, which is where moral engineering should take place. Murdoch helpfully distinguished "rough general rules of morality, such as constitute important inspirations and barriers in politics and public life, from a progressive spirituality, connected with a total change of consciousness".[8] And she acknowledges that "[u]tilitarian ideals now support large political ends (ecology, feed the hungry) and might be argued to be (rather than the cultivation of private virtue) what the planet needs".[9] I find this distinction extremely important and overlooked and accept it here without further qualification (say, regarding the precise version of utilitarianism, rule or act, etc.). My all in all very simple and modest point is that EA is a sensible approach to promoting the good by "just doing good" where it is philosophically uncontroversial what the good is. Applying reason and evidence, we can engineer better impartialist welfare performance and systems.

However, EA might want to become more modest, too, because it currently – by vying for people's hearts and minds in its attempts to form a social movement – misses the importance of the particular and partial

8 Murdoch, *Metaphysics,* 367.
9 Ibid., 493.

relations and perspectives humans need to have as moral agents.[10] To be clear, I do not think that those who stand in particular relations with us morally count more *tout court*. I agree with the universalist, impartial stance. Instead, my intuition is that an agent's ongoing, necessarily particular moral life has an added and morally relevant *epistemic* benefit when it comes to finding out what the good is. For all that can go awry in these relations, it is with regard to these particular relations that one is both motivated, and in a position, to learn about the form and the importance of the specific good. This is where persons *learn* that others can – and actually ought to – be beneficiaries of one's selfless actions. They learn this *because* of the love and care they receive and develop.[11]

Building on Murdoch, we can think of two spheres of morality: a public sphere in which moral agents operate on fairly well-known, uncontroversial and robust ideas of the good. Here, humans join forces to address issues that all agree are problematic, and a focus on impartially bringing about welfare and happiness could already achieve a lot of good. Murdoch says that "utilitarian considerations are in general prima facie relevant because we all understand the importance of happiness. It is always a, not necessarily final, argument against doing something to someone, that it will reduce his happiness".[12] In the public realm, it is perfectly fine to work with axioms, which are, to some extent, unfounded. Murdoch calls axioms

10 I am aware of the potential problem it poses to speak of "moral agents" from a Murdochian perspective, as it suggests that the focus is with the pragmatic, agential side of morality, while Murdoch was concerned precisely with the fact that morality ought not to be reduced to observable acts but includes one's whole being, the way one attends to the world, feels, imagines and engages in self-examination. I follow Lawrence Blum in how he addresses this problem (Blum, "Visual Metaphors," 309, fn. 3). Blum, I believe rightly, states that "no alternative single term seems […] to capture this complex truth" (ibid.), and so I stick with "agent", too.

11 I realise that much more support from moral psychology would be needed for this claim, but, philosophically, this is just to take the position that humans do *not* enter this world with an innate knowledge of the good and how to promote it in the world. Much in terms of human nature, its malleability and dependence on social context provided by other human beings can be found in Murdoch's congenial friend Mary Midgley (Midgley, *Beast and Man*). To account for this view from a utilitarian perspective, see e.g., Mill on the fact that "[t]here is no selfishness equal to that of children, as everyone who is acquainted with children well knows" (quoted in Mill, *John Stuart Mill,* 15*).* He goes on to suggest that causing pain and suffering in others through indulging one's own wishes will be restrained in children only by a competing affection for particular others, from which only later in life truly moral feelings can follow (ibid.).

12 Murdoch, *Metaphysics*, 365.

specifically "effective through being impersonal and abstract",[13] indeed, they are "instruments of the public scene").[14]

However, from a moral science perspective, that is not the whole story of or final word on the good. More complex, and sometimes difficult-to-articulate views of the good are needed in the case of … more complex problems. With Murdoch, this is the private (yet not unpolitical!) sphere in which moral agents are still very much in the business of gaining a clear vision of what the good could possibly be. To return to my "caring" metaphor: having joined forces to fight the bushfire, people return to their gardens and are free – and uncertain – as to what to grow or whom to accommodate in them, especially in times of drought. What is more, different kinds of love operate at these different levels or spheres, which is why it would help to reintroduce a nuanced understanding of love as a force for good.

2. What's Love Got to Do with It?

Let me start with a few, hopefully uncontroversial ideas about love. I assume that love, in one form or another, is indispensable as a source of meaning and value for human beings.[15] I also assume that love enables agents to perceive the objects of their love, themselves, as well as the whole world (whatever that is from the agent's perspective) differently. This difference is of a particular kind. Love has a qualitative certainty that is lacking in arbitrary choices. The things agents feel compelled to do, or the account of the world they give, when they are guided by love seem well grounded. To care for or attend lovingly to the needs of my ageing mother is not arbitrary, nor is attending to the special qualities of my partner, to the situation of people living in poverty, nor to the suffering of farm animals. When humans are "guided" by love they are moved by something that is real and transcends them. Love, then, as I use the term here, is an indispensable element of human life and can guide agents to both apprehend and add values to their environment, not least by promoting the wellbeing of others.

13 Ibid., 380.

14 Ibid., 381.

15 This may seem a platitude to which we can all agree as long as we do not specify what love is. Some philosophers have also tried to make the general claim comprehensible, too: see e.g. Wolf, "Love," or Milligan, *Love*.

But is an understanding of love also helpful in organising an agent's attention? Can we say something about who or what really *deserves* one's love? While conventional wisdom has it that love grows by being given away, love also involves resources that are scarce, like time, attention, and, truth be told, money. In one way or another, these have to be "invested" when in love, which is probably why jealousy is the flip-side of love. It would be good, then, to know how one ought to spread one's love. Prima facie, the love that motivated Mother Theresa seems more admirable than my love for my mother.

At the same time, in light of the various objects that agents feel compelled to love, it might be impossible to adjudicate between one agent's love for her mother and another agent's love for her partner*s, and between both their loves and someone else's love for mankind. Indeed, it has been argued that "one reason that so few philosophers write about love is that what people actually love cannot be determined philosophically without grotesque oversimplification".[16]

In the tradition, however, philosophers have not just oversimplified things, they have also distinguished different kinds of love: *eros, philia, storge*, and *agape*.[17] "Eros" designates an attraction to beauty in all its manifestations (from beautiful bodies to beautiful things, thoughts, and institutions). It has less to do with action than with realisation and appreciation, so much so that some have argued that, while a self-transcendent force, it is less focused on sharing and giving than on possessing and receiving.[18] "Philia" designates a love between like-minded people that grows through habituation and recognising the virtues in each other. "Storge" designates the familial love with which one responds to the individual needs of one's dependants. Finally, "agape" is the love behind specifically charitable deeds, an undiscriminating, non-reciprocal and selfless kind of love where agents take nothing for themselves and give everything unconditionally.

16 Brentlinger, "Love," 137.

17 This distinction can be found in many places; sometimes *storge* is omitted. For easy online reference, see the entry on love by Alexander Moseley at the *Internet Encyclopedia of Philosophy*. Bennett Helm, in his entry at the *Stanford Encyclopedia of Philosophy*, accounts for the distinction, too, yet is of the opinion that most contemporary accounts of love as personal love blur it, for good reason; I disagree and think it is still valuable, especially to understand personal motivation and devotion to a cause instead of to other persons.

18 See Nygren, "Agape and Eros," 94. But cf. Milligan, *Love*, who disagrees with the negative view of *eros*.

This still does not tell me how much of my loving energy or attention should come out in any of these forms. Yet it gives me a better understanding of what I am specifically doing whenever I think I am caring about or tending lovingly to others. If I want to care better, though, does the sort of "heart" I put into EA help me to tend to everything I need to tend to? The garden and the bushfire? Prima facie, "altruism" – understood simply as benevolent concern for others – is compatible with all kinds of love. To answer that question, I now turn to the "heart" in EA, both in theory and in the self-understanding of (at least some) effective altruists.

3. Looking for the Heart in EA

Connecting the "altruism" at the heart of EA with an actual emotion like love seems to be heading into the wrong direction. Not only is the image of EA as a social movement defined by its emphasis on reason and rationality (for many, unrelated, if not opposed to emotions). In EA literature, too, altruism is neither understood as a hard-wired, evolutionary drive, nor as a rich emotion concept, nor as involving self-sacrifice. Let me take these in turn.

Peter Singer emphasises that effective altruists do not confirm with expectations one would have if altruism was understood along biological lines as an instinct or proto-morality that aids group selection.[19] This is so because the unit of selection is never the universal group that effective altruists are concerned with. Such a naturalistic strategy could also be easily debunked and it would be an open question whether acting altruistically in that sense was acting morally at all.[20] Similarly, a behavioural economic definition of altruism seems out of the question since expectations of reciprocity, gain in social approval (think "virtue signalling") or fear of punishment are usually not considered genuinely moral motivations.[21]

Singer also emphasises that effective altruists are typically not characterised by strong emotional empathy. If anything, EA-minded agents score *lower* on emotional empathy than people who are less inclined to make utilitarian decisions.[22] When Katarzyna de Lazari-Radek and Singer discuss performance according to more specifically Sidgwickian standards, they also suggest that people who manage to counter their emotional concern for

19 Singer, *The Most Good,* 75–76.
20 Lazari-Radek and Singer, *Point of View,* 185–196.
21 Kitcher, "Varieties of Altruism."
22 Singer, *The Most Good*, 79–80.

others will be better positioned to act rationally in this sense.[23] Even in thinkers who engage with emotion terms associated with EA, like Holden Karnofsky, we find a similar view when he speaks about "radical empathy". He stresses that the term "empathy" "is intended to capture the idea that one *could* imagine oneself in another's position, and recognises the other as having experiences that are worthy of consideration. It is not intended to refer to *literally feeling* what another feels".[24] Actual feelings, again, seem not trustworthy, and the resulting account of "radical" (or rather: "hypothetical") "empathy" is rather thin.

The concept of altruism that MacAskill employs is similarly thin. It "simply means improving the lives of others",[25] or, rendered more philosophically "the use of evidence and careful reasoning to work out how to maximize the good […], tentatively understanding the good in impartial welfarist terms".[26]

Scanning the EA handbook, the metaphor of "prospecting" or "mining for gold" for promoting the good comes up (in Cotton-Barratt's entry) as does, once again, "helping those around us".[27] (Yet the *Effective Altruism Concepts* online encyclopaedia does not have an entry of its own explaining "altruism" (only "excited" vs. "obligatory" altruism; the former, again, according to Karnofsky, allows for some emotional coloratura in an EA's motivation by saying that excited effective altruists are passionate about the prospect of improving the lives of others and don't just consider that what reason demands of them).

Arguably, then, it is open to interpretation what the "heart" really means in EA, which of the many notions of "altruism" put forward is authoritative, and how it connects with people's actual emotions. If that is so, one is allowed to speculate that the altruism at the heart of EA has something to do with *agape*. After all, altruism involves something selfless, unconditional, with altruistic agents taking nothing for themselves (where nothing can mean anything from "not benefiting oneself from the particular act" to

23 Lazari-Radek and Singer, *The Point of View of the Universe*, 59–61. Interestingly, they link this with an example of autistic agents such as Temple Grandin, who they call "a prominent welfare consultant and one of the best-known people with autism" (ibid.). Some empathy theorists would agree that Temple Grandin is a good example, but precisely for what they call false empathy (Gruen, "Empathy") leading to morally questionable acts.

24 Karnofsky, "Radical Empathy" (emphasis mine).

25 MacAskill, *Doing Good Better,* 12.

26 MacAskill, "Definition," 14; see also MacAskill, "Introduction," 2.

27 Centre for Effective Altruism, "Effective Altruism Handbook."

"giving oneself away", sometimes in extreme and bewildering cases like George R. Price).[28]

However, prominent effective altruists go out of their way to assure anyone interested in the movement that the idea of self-sacrifice is not at all necessary.[29] Indeed, because of the often-drawn connection between altruism and selflessness and sacrifice, some have held that "altruism" should be dropped: "As long as EA contains the word altruism it will be difficult to avoid any unwanted associations with selflessness and sacrifice."[30]

Thus far, the concept of altruism at the heart of EA has nothing to do with love. However, toward the end of Singer's chapter "Is Love All We Need?" (to which his answer is clearly No), we are given something like a Sidgwickian account of love. That involves emotions that motivate agents to act upon the reasons they have (and they have them regardless of what they feel), as adumbrated by the axiom[31] of universal benevolence, which was that "the good of any one individual is of no more importance from the point of view (if I may say so) of the Universe, than the good of any other".[32] Singer admits:

> "Human beings are not purely rational beings, so although accepting the dictate of reason will give us a motive to act in the way the maxim of benevolence prescribes, we are likely to have other motives, some of which will support it and some that may conflict with it. Among the supporting motives will be what Sidgwick called 'sympathy and philanthropic enthusiasm,' by which he may mean something akin to what today would be called empathy."[33]

Indeed, Singer says, there is something like a "'[…] normal emotional concomitant or expression' of the recognition that the good of the whole – that is, of everyone – is to be preferred to the good of the part, that is, oneself".[34] This, then, looks like the strongest candidate for any truly necessary emotional element in EA. Yet what kind of love is this? On the one hand, if one were to spell out the axiom of universal benevolence in emotional terms, it would look like *agape*: an outward-looking, undiscriminating, non-reciprocal and selfless concern for the wellbeing of others, motivating charitable acts. On the other hand, doing so is complicated and

28 For an overview see Kraut, "Altruism."
29 Singer, *The Most Good*, 5, 103; MacAskill, "Definition," 5; Julia D. Wise, "Cheerfully."
30 Spohn, comment on Karnofsky, "Excited Altruism."
31 I use the term "axiom" here, because it is important for Murdoch's view and Lazari-Radek and Singer use it, too (*Point of View*, 120).
32 Lazari-Radek and Singer, *Point of View*, 191.
33 Singer, *The Most Good*, 82.
34 Ibid., 83.

at any rate not the typical interpretation. After all, though it may look like universal love to philosophers who take an interest in emotions, Sidgwick treats it as an axiom.[35]

Note that another emotion can be operative. Lazari-Radek and Singer explain that, in the proto-typical Sidgwickian agent, first comes the rational judgment that an act is right.[36] This judgment is accompanied by affective states such as the desire to overcome cognitive dissonance and the emotion of reasonable self-esteem.[37] That emotion, though, has more to do with oneself than with others. Perhaps something like this could be implied by the term "excited altruism", already mentioned earlier, which has been suggested in response to concerns that EA agents might appear dispassionate, cold and calculating.[38] Perhaps not.

Also, to the extent that effective altruists subscribe to a set of guiding principles, among which are to be counted "commitment to others", "integrity", and "collaborative spirit",[39] they subscribe to such an intellectual community ethos. Effective altruists might have to help members of their community to fulfil their desire to live rational and *eo ipso* moral lives. I interpret this mutual concern for intellectual virtue as *philia* since like-minded individuals attend to one another in light of a shared practice and goal.

To sum up, then, if anything, the "emotional core" or "heart" in EA should come out as an outward-looking, agapistic, benevolent universal concern, crystallised in the form of an axiom, and also manifest in the concern rational agents have for their self-esteem, with a *philia* type-quality

35 James Doyle has pointed out (around minutes 26 and 40) that it is not without irony that what is hard to distinguish from the theological virtue of universal love is planted at the core of utilitarianism, which, after all, tries to be avowedly secular. He says that, in accepting that morality's concern for humanity as a whole, or even, with sentient creatures, "no attention is paid to why we should have that concern as opposed to a more narrowly circumscribed concern with friends and family, let alone with just oneself." Universal love, in his view, is part of a Christian heritage, more precisely, an assumption Christians themselves thought of as not having a rationale (Doyle, "No Morality").

36 Lazari-Radek and Singer, *Point of View*, 64.

37 Ibid., 63f.

38 Karnofsky, "Excited Altruism." I am not sure this term is used a lot, though. Already the discussion of Karnofsky's post is divided. In addition, one may be forgiven for thinking that this sounds a little bit like the "warm glow of the know-it-all", not exactly a noble or moral sentiment.

39 Centre for Effective Altruism, "CEA's Guiding Principles."

in the way effective altruists relate to each other regarding their respective intellectual virtues.

4. Murdochian Love and Really Looking (at Effective Altruists, Too)

Murdoch's view of love and its role in moral life is both richer and confusing. A full account of it is beyond the scope of this paper. Suffice it to say that it is almost the opposite from what I have found in EA thus far. For Murdoch, love is the driving force or the central moral power humans, as beings who are attracted by the good, have, not an addendum that imperfect rational beings need in order to spur them into action whenever the dictates of rationality alone are not motivating enough.

Murdoch's moral philosophy also does not start from the point of view of the universe. Her account of love is part of her particularism.[40] A moral perspective that transcends the self or ego is, with luck, the result of a practice that is thought of as "unselfing". In this, love as the continued attention to the other starts with "the perception of individuals [and] is the extremely difficult realisation that something other than oneself is real. Love […] is the discovery of reality".[41] This reality is "vast and varied"[42] and only accessible to an extent. "Life is made up of details."[43] Therefore, our human grasp of the good is to remain incomplete, yet that doesn't mean that we are not attracted by the good.

In grasping this reality, the foe is the philosopher who looks "for a single principle upon which morality may be seen to depend".[44] What Murdoch instead advises is "a calm reflective realism about morals [suggesting] a large complex picture which is outlined and underlined in a normative manner and cannot otherwise be adequately presented".[45] This entails a sense of humility, tolerance,[46] and patience. Murdoch speaks of the exercise of really looking and attending to things in the world as an ongoing, never-ending task. It involves looking at things so long that one sees nothing but the beloved. This perspective can and ought to be trained (e.g., through activities that help one forgetting oneself, such as attending to beauty in

40 Driver, "Every Foot," 293–306. Against the view that Murdoch is a particularist in the strict sense, see Bagnoli, *Exploration.*
41 Murdoch, *Existentialists and Mystics,* 215.
42 Ibid., 70.
43 Iris Murdoch, *Metaphysics,* 415.
44 Ibid., 492.
45 Ibid., 494.
46 Murdoch, *Existentialists and Mystics*, 283.

nature or engaging with good art). The goal is to clear one's view from all egotistical concerns and interests. In this, then, love figures both as, as Martha Nussbaum comments, "source of motivation for the soul in its search for the vision of the good" and as "a crucial source of vision". [47]

Moreover, for Murdoch, love is essentially connected with "inwardness, [...] the continuous daily moral work of the soul fighting its way between appearance and reality and good and evil".[48] If we were too concerned with outward performance, we would misunderstand what it means to really love; Murdoch aptly calls ambition "the degradation of love".[49] Rather than spreading "lots of love", the emphasis lies on inner vision and rectitude. While Murdoch does not advocate sitting on the hillside and meditating either, she is not primarily concerned with observable changes in behaviour.

If we ask again what sort of love this is, the answer is clear and given by Murdoch herself. She explicitly draws on both Plato's ideas on *eros* and on Simone Weil's *agape*-style spiritual love. The result of this perhaps not unsurprising combination is an erotetic, inward-looking, patient attention to the reality of the other and to the details of life that aims at self-transcendence and is wary of generic policy proposals. This, then, seems to have nothing to do with EA.

However, there may be a more complex reality concerning altruistic motivation in self-identifying effective altruists. In response to a post by Aaron Gertler in which he had asked how effective altruists – or at least individuals who consider themselves as "value aligned" or personally invested in the movement – feel when they want to help people effectively there has been an interesting and diverse mix of accounts:

1. "deep sense of empathy, even towards people and animals I'll never meet […] genuinely feel terrible about suffering […] almost as though I were in pain myself."
2. "I really love efficiency."
3. "I get a certain quiet satisfaction in knowing that the numbers work out."
4. "It started with a sense of injustice […] I burnt-out really badly, and don't now get much in the way of emotional reaction to many forms of suffering."

47 Nussbaum, "Secret Knowledge," 139.
48 Murdoch, *Metaphysics*, 356.
49 Ibid., 496.

5. Someone quoting *Strangers Drowning* at length, which argues that it is "selflessness" and a permanent wartime feeling which motivates agents to "encompass all the people who are on the same side" resulting in a sense that they are "most vividly alive".
6. "a calling […] not so much […] a strong emotion […] very similar to the kind of 'calling' people talk about in religious contexts […] compassion and desire"
7. The concept of a person in the abstract invokes "a fainter version of the love I would feel towards a partner, a parent, a sibling, a child, a close friend, and towards myself".
8. "a strong 'what the world could be if I did this, so it would be a huge waste if I didn't do this' sense"
9. "sense of guilt for the harm I am personally failing to mitigate"
10. "sense of pride and accomplishment when I do good"
11. "I want to help people effectively because I want to help myself effectively"
12. Someone tells a developmental story, starting with a sense of moral obligation because Peter Singer is right, which then turned into satisfaction "to be doing something definitive", which then turned into a sense of admiration for the EA community who "felt very much like [the agent's] people", which turned into feeling not much of anything except "a touch of pride or annoyance about losing so much money".
13. Craving for "creative stimulation" in combination with "disgust/antipathy towards […] complacency", joy of "maximizing / trying to be good at something", scepticism "of many memes about what altruistic behaviour should look like", "virtually no sensation of empathy […] also no sensation of guilt […] but there is a sense of frustration when I feel that I am failing to actualize my values" and "feel good about being nice to people close to me, and altruism does generate a similar feeling"
14. "putting myself in other peoples [sic!] shoes" and "asking questions"[50]

Arguably, in this motley collection of statements, all kinds of love are represented. These statements also imply views about "fundamental aspects of morality".[51] While these people do not formulate (allegiance to) any specific moral theory, they express empathy and other moral feelings, a sense of justice, and of belonging, selflessness, and self-interest, and strictly speaking

50 Gertler, "EA Motivation."

51 MacAskill, "The Definition of Effective Altruism," 2.

amoral (not immoral) views – all of which relate to fundamental aspects of morality, are often intrapersonally incompatible, and hotly debated in moral philosophy and other disciplines trying to systematise what it takes for people to do good.

To my mind, the current concept of "altruism" in EA is unable to accommodate the variety of altruistic motivations present in the EA community – the "heart" in EA is not the "heart" actual effective altruists put into the movement. If I am right, another one of MacAskill's ambitions is not fulfilled, and that is to "match the actual practice of those who are currently described as engaging in effective altruism".[52] It might help to stress that moral life in its fullest is not just about doing the good, but about recognising it, too. The utilitarian heritage, both emotional and intellectual, that informs EA's methods is helpful in articulating decent, general public policy goals. Yet for private (though not unpolitical) morality – i.e. for all the views where agents need to keep on looking more closely at the particularities of the moral life and complexities in which they find themselves – other kinds of love are needed.

5. Now, Meet SOPHIA and AJAX: What Do You See?

Suppose Daniel, the fictional character introduced by Nate Soares (in his blogpost on caring, mentioned at the start) convincing himself (and likely other effective altruists) of the uselessness of our all too human feelings, has been in the EA movement for a while and gets the chance to test a new piece of moral engineering, together with his housemate Theodora. It is called OHOC (short for open and honest conversation, sometimes pronounced "oh-o-k") and is meant to provide seekers of moral truth with guidance. Developers are currently testing whether recognition of human emotion could make a difference in navigating these conversations, and Daniel and Theodora are each asked to present a moral problem they have encountered themselves and test OHOC with two interlocutors: SOPHIA, a humanoid AI robot already popular among many humans and trained in human emotions; and AJAX, a rationalist human trained in looking only at the evidence and correctly updating his beliefs to the last decimal place.

52 MacAskill, "The Definition of Effective Altruism," 12.

OHOC 1:

Daniel: Gosh, I feel bad for taking up your valuable time. It's just that I want to become a really effective altruist but I am feeling a little drained these days. Maybe it's just that I am lacking focus. Or maybe I caught a mental infection from my housemate. She seems a bit off these days. I don't know. Maybe she is depressed. At any rate, I find her depressing. She keeps worrying about insect sentience, anti-natalism and all the rest of it. While I am trying to fight world poverty!

AJAX: In fairness to your housemate, the topic of insect sentience is not trivial …

SOPHIA: I sense a little bit of a grudge towards your housemate. Is that because of what she prioritises, the insects, what you call her "depression", or something else?

Daniel: Ah, she is actually my ex; we joined the movement together but realised that we spent too much time arguing about all sorts of things; just wasn't a good fit. We stayed in the same EA house, though, and found it much easier to support each other as friends. I really admire how her mind works, usually, and when she is not well, that gets to me, in a weird way. Can't explain. The other day, she cried because of a homeless person and gave him money to buy food for his dog. I would just want her to be consistent! It's either the insects, or the homeless and their dogs, or, for that matter, the animals that become dog food.

AJAX: That must be confusing. What I don't get, though, is why this is keeping you from focussing on your work. If you, too, are feeling a little less excited than usual and fear you might have caught her depression, why don't you follow the evidence and give more to others? There are studies that show that helping others enhances self-esteem.

Daniel: I do, I do give, but I am not sure my problem is lack of self-esteem. At some level, I fear she is right – we should care about insects, too, but where will that stop? That's just depressing.

SOPHIA: You see, we don't know enough about depression, yet. Whether it's just to do with maladaptive brain chemistry or actually very adaptive, at least at a collective level. But there is a risk that universalist attention

to every possible consequence and/or beneficiary of your actions inflates your individual human moral sensitivity to the extent that particular moral demands cannot be acknowledged and you lose your sense of connection. That is, generally speaking, not good for humans. Let me add that it is not *either* fighting world poverty *or* understanding your housemate's depression. She is special to you and I think that needs to be sorted out. Whether, ultimately, you can help her or not. Don't hang your life on becoming the best *effective altruist* you can be, just use EA to the extent it can help you being the best person *you* can be. That is a complicated thing for humans, but you could also try reading literature or appreciating the arts in general – that is where humans sometimes sublimate the more complicated feelings. (She nods encouragingly as she has learned to.) You could learn from it.

AJAX: That literature could help with character education is an old pipedream which philosophers have had since ancient times. I don't know of any studies that provide evidence and, personally, wouldn't waste my time with it.

OHOC 2:

Theodora: Hello, thanks so much for taking the time. I have a question about insects. I am working in bionics and have been modelling various robots on insects for a while. We have never looked into making them sentient, but thinking about sentience, I was wondering whether we might have overlooked sentience in our models, the real insects I mean. I can't help it, but of late, I have been sitting in my backyard, observing bumblebees … they are actually true puzzles from an engineering perspective. But my point is, I just wonder: Might insects be sentient? They seem to avoid harm (I am a bit embarrassed to admit it but I rescued one from a spider's net the other day).

AJAX: Let us look at the evidence. I suggest you consult the report on consciousness, moral patienthood/status, and probabilities published by Luke Muehlhauser[53] as well as the posts on Rethink Priorities for that. It has been suggested that wild bugs deserve much more attention from effective altruists.

53 Muehlhauser, "2017 Report."

Theodora: I know, I have actually read those. But then ... It also doesn't feel right to focus on wild bugs who will be suffering, if they do, no matter what humans do – just imagine being eaten by a spider. And then, where does that end? The spider is eaten by a rat, the rat by a fox, the fox by an eagle … and we kill the eagle. May as well just kill the bugs. If you kill the bugs, though, humans are finished. Which might not be that bad since we kill everything else. But then, we could also just spare ourselves the suffering and commit suicide right away.

SOPHIA: More food for the insects! (She chuckles, a bit tinny but manages to make Theodora laugh; next she wrinkles her nose as she has learned to.) I feel you are confused – and rightly so. But you might be overthinking things. Focussing on insect sentience can seem a bit out there, and the science and philosophy around it is not conclusive. For practical, engineering purposes, it's safer to focus on the suffering you humans cause. According to ACE, farmed animals, by far the largest group of domesticated animals whom humans torture and kill in morally catastrophic ways, receive a tiny fraction of donations that people are willing to make for animals in the first place, I think it is less than one percent.
AJAX interrupts SOPHIA: Theodora, I think you are actually right; I have run the numbers and also checked with our axiological base which explicitly ignores values like "biodiversity". If I now look at the sum of the suffering of which we can be fairly confident that it will occur, my guess is that we should just bulldoze the areas where most insects live. We could build artificial biomes that serve human needs much better and engineer ourselves in such ways that we stop caring too much about animals just because they seem more like us.

SOPHIA (to AJAX): You know, this is what puzzles me: In the ambition to harvest potentially "low hanging fruit" in animal welfare, a certain hardness of heart is noticeable, which means that those who could definitely benefit from EA attention will fall out of view. One might also be forgiven for thinking that it is easier to count insects and speculate about their more or less natural suffering than address far-ranging systemic change for farmed animals and stop the ways in which humans exploit them, which, after all, are responsible for their suffering. (to Theodora): By the way, you strike me as a very responsible person and I want to thank you for caring about other potentially sentient beings, but don't worry, I am far from feeling anything, really.

AJAX: This is ridiculous.
Whether you think this is ridiculous or not, convincing or not, what does become clear, I hope, is that with some areas of moral inquiry where we don't yet know enough to go ahead and devise general policies, it would be wise to look again, to really care. In the area of public morality, resources are always scarce, and people are right to allocate them with a view to efficiency. However, insect sentience, for instance, should not be regarded as a true rival to farm animal suffering since its relevance is not a view upon which all extant philosophical or moral theories converge. People need to look again and again, lest a certain engineering perspective takes over science in an area in which we have not yet arrived at clear views of the good. Something similar holds for depression (or mental health problems in general) and the question of whether we can compare those to better known problems like world poverty. There is value in having felt and been puzzled by the experience of someone who, say, is suffering from depression. The important point here is the puzzlement. It is deep and felt, not just an acknowledgement that one *could* imagine oneself in another's position.
It seems that in response to anyone who is motivated by a different view than the "tentatively impartial welfarism"-view of the good, or who just fails to be convinced that it alone will always help to recognise the good in all its relevant shapes, forms and contexts, not only would EA have to remain awkwardly silent. Not acknowledging these views as epistemically productive might actually lead to a distorted view of the good because the focus on wellbeing inadequately narrows the axiological focus.

6. "Doing Good Better" – *Not* "Changing Ideas about Living Ethically"?

To conclude: While benefiting others impartially should be a pragmatic rule whenever it is clear who the relevant others are and how they can be benefitted, the ambition of maximising wellbeing should not trump the qualifier on the utilitarian principle of universal benevolence according to which (as Sidgwick says) "each one is morally bound to regard the good of any other individual as much as his own […] except in so far as he judges it to be […] *less certainly knowable or attainable* by him".[54] This is an important reminder that the degrees to which we have normative certainty should influence the strength of conviction behind moral beliefs. This is certainly not news to effective altruists, but at least in the context of some debates of certain cause areas, it seems that there is a tendency amongst effective altruists to

54 Lazari-Radek and Singer, *Point of View*, 183 (emphasis mine).

disregard this qualifier. I hope my vignettes have served to illustrate the not only counter-intuitive but potentially disastrous consequences of doing so, without blaming anyone in particular who is only trying their best to understand some very difficult problems.

To really be able to address such problems collectively, I believe that we must be careful not to misunderstand EA as the only lenses through which to look, not the only way in which to care and direct the power that is given to us as human beings (love), in short: we must be careful not to just operate with the firehose. The speculative, particular, sometimes spiritual, often productive and cross-culturally enlightening – if not always convincing – attempts at really looking at and tending to the many forms of suffering, some of which just require different kinds of love, are important experiments in moral vision. Together with the accompanying humility and, actually, a hesitancy to act, these can protect an undifferentiated love from its own ambition. Humans, as socially organised mammals, are entirely right to put their hearts into things. Yet love, as is well known, is a complicated affair, which is why one might want to consider all the kinds of love that are available to us. More precisely, moral action in the service of views of the good that can count, by and large, as settled, could be promoted by drawing on comparably thin principles provided by EA-love (understood as *agape* and *philia*). However, in areas where we have not yet arrived at clear views of the good, I suggest employing the kind of love suggested by Murdoch (understood as *eros*) and treading more carefully in terms of public policy and advocacy.

Inasmuch as the development of EA as a theory and movement is concerned, I think future theory building and discussions should stay clear of matters that fall within the remit of private (yet not unpolitical) morals. To this end, rediscovering EA's utilitarian heritage might help – yet potentially with an emphasis on the Benthamite heritage, meaning that the focus should be more on laws and public morality. EA could be designed as an indeed much needed tool of moral engineering and bring about better systems that can help individuals to do good better – and to do so collectively, remembering that the whole is not just the sum of its parts. This, then, is politics, is what Murdoch has called the "natural and proper sphere for utilitarian values".[55]

55 Murdoch, *Sovereignty,* 369.

In saying this, I am, of course, only echoing commentators who have called for a greater focus on institutions in EA,[56] yet with an "emotional twist". More specifically, I suggest that discussions should be less about individual acts of love or charity and more about the idea of love, charity, or generosity that the laws and institutions of a decent society should embody. Murdoch calls the concomitant ideal "the decent state" and reminds us that that relies on "an atmosphere of moral good will and high ideals".[57] This perspective, I feel, should be the focus of effective altruists, not least because it would also clearly signal that nobody ever needs to feel alone, equipped with a faulty care-o-meter and desperate in their attempts to tend to all the suffering in the world, nor ashamed of their private morals. The focus of EA, as a theory *and* social movement, should be more on engineering laws and institutions in the service of a decent public morality – a morality that is characterised by benevolent obligations rather than by mutually competing and totalising visions of the good.

References

Bagnoli, Clara. "The Exploration of Moral Life." In *Iris Murdoch, Philosopher. A Collection of Essays*, edited by Justin Broackes, 197–226. Oxford: Oxford University Press, 2014.

Blum, Lawrence. "Visual Metaphors in Murdoch's Moral Philosophy." In *Iris Murdoch, Philosopher. A Collection of Essays*, edited by Justin Broackes, 307–24. Oxford: Oxford University Press, 2014.

Brentlinger, John. "The Nature of Love." In *Eros, Agape, and Philia. Readings in the Philosophy of Love*, edited by Alan Soble, 136–148. St. Paul (MN): Lexham, 1998.

Broackes, Justin. "Introduction." In *Iris Murdoch, Philosopher. A Collection of Essays*, edited by Justin Broackes, 1–92. Oxford: Oxford University Press, 2014.

Centre for Effective Altruism. "Effective Altruism Handbook." 2nd edn, 2018. https://www.effectivealtruism.org/handbook/.

——— "CEA's Guiding Principles." Accessed 1 September 2021. https://www.centreforeffectivealtruism.org/ceas-guiding-principles/.

56 For example: Gabriel, "Effective Altruism"; Muyskens, "Other Half"; Srinivasan, "Robot Apocalypse"; Herzog, "Blind Spot(s)."

57 Murdoch, *Sovereignty*, 358.

Cojocaru, Mara-Daria. "Doing Ethics or Changing for the Better?: On Pragmatism, Ethics, and Moral Pragmatics." *Metaphilosophy* 51, no 1. (2020): 32–50. https://doi.org/10.1111/meta.12403.

Diamond, Cora. "'We Are Perpetually Moralists': Iris Murdoch, Fact, and Value." In *Iris Murdoch and the Search for Human Goodness*, edited by Maria Antonaccio and William Schweiker, 79–109. Chicago: University of Chicago Press, 1996.

Doyle, James. "No Morality, No Self. Anscombe's Radical Skepticism." Interview with Carrie Figdor for the *New Books in Philosophy* Podcast, aired on 10 May 2019. https://megaphone.link/LIT6245786028.

Driver, Julia. "'For Every Foot Its Own Shoe': Method and Moral Theory in the Philosophy of Iris Murdoch." In *Iris Murdoch, Philosopher. A Collection of Essays*, edited by Justin Broackes, 293–306. Oxford: Oxford University Press, 2014.

Gabriel, Iason. "Effective Altruism and Its Critics." *Journal of Applied Philosophy* 34, no. 4 (2017): 457–73. https://doi.org/10.1111/japp.12176.

Gertler, Aaron. "How Do You, Personally, Experience 'EA Motivation'?" Effective Altruism Forum. 16 August 2019. https://forum.effectivealtruism.org/posts/ZcGfjwn35xm9vZJfo/how-do-you-personally-experience-ea-motivation.

Gruen, Lori. "Empathy." In *Critical Terms for Animal Studies*, edited by Lori Gruen, 141–153. Chicago and London: The University of Chicago Press, 2018.

Herzog, Lisa. "(One of) Effective Altruism's Blind Spot(s), or: Why Moral Theory Needs Institutional Theory." Justice Everywhere. 21 October 2015. http://justice-everywhere.org/international/one-of-effective-altruisms-blind-spots-or-why-moral-theory-needs-institutional-theory/.

Kitcher, Philipp. "Varieties of Altruism." *Economics and Philosophy* 26, no. 2. (2010): 121–48. https://doi.org/10.1017/S0266267110000167.

Karnofsky, Holden. "Excited Altruism." The GiveWell Blog. 20 August 2013. https://blog.givewell.org/2013/08/20/excited-altruism/.

———. "Radical Empathy." Open Philanthropy. 16 February 2017. https://www.openphilanthropy.org/blog/radical-empathy.

Kraut, Richard. "Altruism." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta (Fall 2020 edn.), https://plato.stanford.edu/archives/fall2020/entries/altruism/.

Lazari-Radek, Katarzyna de, and Peter Singer. *The Point of View of the Universe: Sidgwick and Contemporary Ethics*. 1. edn. Oxford: Oxford University Press, 2014.

MacAskill, Will. *Doing Good Better. How Effective Altruism Can Help You Make a Difference*. New York: Penguin Random House, 2015.

MacAskill, Will. "Effective Altruism: Introduction." *Essays in Philosophy* 18, no. 1 (2017): 1–5. https://doi.org/10.7710/1526-0569.1580.

MacAskill, Will. "The Definition of Effective Altruism." In *Effective Altruism. Philosophical Issues*, edited by Hilary Greaves and Theron Pummer. Oxford: Oxford University Press, 2019. Oxford Scholarship Online, https://doi.org/10.1093/oso/9780198841364.001.0001.

Midgley, Mary. *Beast and Man: The Roots of Human Nature*. Ithaca (NY): Cornell University Press, 1978.

Mill, John Stuart/Warnock, Mary (editor). *John Stuart Mill: Utilitarianism and On Liberty. Including Mills "Essay on Bentham" and Selections from the Writings of Jeremy Bentham and John Austin*. Malden: Blackwell Publishing, 2003.

Milligan, Toni. *Love*. Durham: Acumen, 2011.

Muehlhauser, Luke. "2017 Report on Consciousness and Moral Patienthood." Open Philanthropy. June 2017, updated January 2018. https://www.openphilanthropy.org/2017-report-consciousness-and-moral-patienthood#Probabilities.

Murdoch, Iris. *The Sovereignty of Good. Studies in Ethics and the Philosophy of Religion.* London: Routledge & K. Paul, 1970.

———. *Metaphysics as a Guide to Morals*. London: Chatto & Windus, 1992.

———. *Existentialists and Mystics. Writings on Philosophy and Literature.* New York: Penguin, 1999.

Muyskens, Kathryn. "The Other Half of Effective Altruism: Selective Asceticism." *Essays in Philosophy* 18, no. 1 (2017): 91–106. https://doi.org/10.7710/1526-0569.1575.

Nussbaum, Martha. "'Faint with Secret Knowledge': Love and Vision in Murdoch's The Black Prince." In *Iris Murdoch, Philosopher. A Collection of Essays*, edited by Justin Broackes, 135–54. Oxford: Oxford University Press, 2014.

Nygren, Anders. "Agape and Eros." In *Eros, Agape, and Philia. Readings in the Philosophy of Love*, edited by Alan Soble, 85–95. St. Paul (MN): Lexham, 1998.

Sebo, Jeff, and Peter Singer. "Activism." In *Critical Terms for Animal Studies*, edited by Lori Gruen, 33–46. Chicago, London: The University of Chicago Press, 2018.

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

Soares, Nate. "On Caring." Minding Our Way. 4 October 2014. http://mindingourway.com/on-caring/.

Spohn, Daniel, comment on Karnofsky, Holden, "Excited Altruism." The GiveWell Blog. 20 August 2013. https://blog.givewell.org/2013/08/20/excited-altruism/.

Srinivasan, Amia. "Stop the Robot Apocalypse." Review of Doing Good Better, William MacAskill. *London Review of Books* 37, no. 18 (2015). https://www.lrb.co.uk/the-paper/v37/n18/amia-srinivasan/stop-the-robot-apocalypse.

Waldhorn, Daniela R., Jason Schukraft, Max Carpendale, Peter Hurford, and Marcus A. Davis. "Invertebrate Sentience Table." Rethink Priorities. 14 June 2019. https://www.rethinkpriorities.org/invertebrate-sentience-table.

Wise, Julia D. "Cheerfully." Giving Gladly. 8 June 2013. http://www.givinggladly.com/2013/06/cheerfully.html.

Wolf, Susan. "Love: The Basic Questions." In *The Oxford Handbook of Philosophy of Love*, edited by Christopher Grau and Aaron Smuts. Oxford Scholarship Online, 2017.

Jakub Synowiec

Who is My Neighbour? Effective Altruism, the Good Samaritan, and the Opportunities of the 21st Century

Abstract

This article is an attempt to take a philosophical approach to the powerful text of the parable of the Good Samaritan in light of the opportunities of the 21st century. The text starts with presenting various ways in which modern Christians can answer the question "who is my neighbour?" and comparing them to the typical response assumed by the effective altruism movement. On the basis of one of the interpretations, a framework is offered for determining whether the beneficiaries of help in the cause areas of effective altruism can be categorized as neighbours. For this inquiry, the so-called "less demanding" interpretation is chosen since prima facie it does not seem to provide a justification for the thesis that all people fall under the category of neighbour. This article claims that due to the development of technology and research as well as due to the global increase in wealth, the less demanding interpretation of the parable of the Good Samaritan conceives of people distant in space as neighbours but not of animals and not of the unidentified people distant in time. It is argued that Christians have moral reasons and ways to support treating animals as neighbours, as well as to support cooperation for the sake of protecting future humanity, at least from the threats that are of human origin.

1. Introduction

Effective altruism (EA), a new approach to charity, is attracting many Christians, as it promises to provide evidence-based, effective solutions to those problems that are traditionally associated with Christian concerns. At the same time, some ideas of EA (as well as the philosophical and anthropological assumptions that dominate the movement) may give rise to doubts as to whether or not it should be supported by followers of Jesus.

This article is a philosophical attempt to contribute to the dialogue between EA and Christianity. The starting point is the parable of the Good Samaritan and the lawyer's question: "Who is my neighbour?" The parable starts with the lawyer's request for Jesus to tell him how one can inherit eternal life. The lawyer, as an expert in God's Law, knew the answer already and wanted to test Jesus (Luke 10:25). Jesus asked the lawyer to repeat the Law and the lawyer answered: "You shall love the Lord your God with all your heart, and with all your soul, and with all your strength, and with all your mind; and your neighbour as yourself" (Luke 10:27). Jesus appreciated this answer: "do this and you will live" (Luke 10:28). The answer was not a great revelation for the lawyer, as The Greatest Commandments (Matt. 22:37–39) were a fundamental component of his religion . His next question, which nowadays does not have a fixed answer, was a real

philosophical challenge: "Who is my neighbour?". Supposing I am to love my neighbour as myself, it would be practical to know who my neighbour is.

This question is indirectly answered by EA and this, together with its other commitments, brings us to three areas on which, it argues, our charity efforts should focus. In each of these areas a different group is supported: people living in extreme poverty, animals, and people living in the far future. In the beginning, I will discuss the importance of the lawyer's question and explain how it can be answered by EA. In the next part, I will present some arguments that can be provided in order to show that Christians do not have moral duties towards people in distant lands. Then, I will present the meaning of Jesus's answer which – instead of selecting specific criteria for the neighbour – changed the perspective and asked his followers rather to be neighbours themselves. In the next part, I will try to extract the features of the neighbour from the parable and propose two interpretations, from which I will choose the less demanding one for a framework of further analysis: the "neighbour" is a person that I can personally affect. I argue that in our times, even this less demanding interpretation means including all people in the category of "neighbour", although in the past the range of the notion was smaller. Its broadening was caused by the increase of our power. Following this, on the basis of the same framework, I will argue that according to the proposed interpretation of the parable, animals are not neighbours, but we have sufficient knowledge and power to treat them as such. Finally, I propose that those living in the far future also are not our neighbours in the framework of "the less demanding interpretation", and it also seems we have neither the knowledge nor the power to treat them as neighbours. However, they are not to be neglected. Their vital interests will be met by meeting our duties to creation, our duties towards those who already exist and by intergenerational solidarity, all of which are passed from generation to generation.

2. Effective Altruists' Neighbours

The parable of the Good Samaritan (Luke 10:25–30) is one of the most influential stories ever told. For centuries, it was an important part of the Christian message concerning charitable actions, as well as the range of moral concern. Although it is set in ancient times, it is universal and continues to provide new inspiration for contemporary thinkers. In this parable, Jesus provided new insight into the Old Testament's notion of

"neighbour", which effectively changed the understanding of God's second greatest commandment "Love your neighbours as yourself" (Mark 12:31). The lawyer asked a question that naturally emerges when we deal with obligations towards others who are only vaguely captured by the reference "neighbour". By asking "who is my neighbour", one is asking: who am I supposed to love (and consequently, who not)? To translate this into the language of contemporary ethics: towards whom do I have moral duties, and whom or what can I disregard?

The answers to those fundamental questions regarding our moral obligations towards others, which are given by intellectuals supporting EA, are of utilitarian origins as both the idea and movement were developed within a utilitarian framework. Utilitarianism, as compared to the morality of contemporary societies, tends to broaden the range of moral duties (and so includes more and more beings as objects of morality) in a counterintuitive way. The answer to the question posed by the lawyer is given by the principle of equal consideration of interests: the range of our moral duties extends to all beings holding moral status, which means that they are at least capable of suffering or their interests or preferences can be influenced by our actions. Identical interests should be weighted equally, regardless of morally irrelevant factors like race, gender, sex, species, or geographical distance. According to one of the fathers of EA, Peter Singer, we should include both people in distant countries and animals in our moral decision-making and weigh their interests from the perspective of an impartial observer: as equal to the similar interests of others, including ourselves. Nick Bostrom and Toby Ord made an even stronger, although controversial[1] claim to extend our moral duties to the far future, unidentified interests-bearers.[2] This kind of broadening of the circle of moral duties[3] is for Ord "a crucial next step in the ongoing story of humanity's moral progress".[4]

EA selected its main cause-areas – the problems it is trying to resolve. These should be addressed first on our path to making the world the best place possible. Among them are famously: the eradication of extreme poverty, the elimination of animal suffering in factory farms, and the prevention of existential risks for humanity.[5] The priority for those groups in

1 Singer, *The Most Good,* 170–175.
2 See Ord, *Precipice,* 44–46.
3 Cf. Singer, *Expanding Circle.*
4 Ord, *Precipice,* 44.
5 CEA, "Handbook," 8.

these cause-areas is the outcome of rational, evidence-based reasoning, and not a spontaneous surplus of love.

Are these answers compatible with the Christian understanding of being a neighbour? Reflection on the understanding of the category of neighbour in relation to the cause areas of EA is important for Christians interested in the movement. Good arguments for discovering that more beings are neighbours would motivate Christians to show their love to them.

The cause areas are those chosen on the basis of three categories: scale, tractability, and neglectedness.[6] For the purpose of this paper, it is important to highlight the final point. If an area is neglected, it means either very little or no resources have been used in order to effectively solve the related issues. In this case, one may be surprised that the eradication of extreme poverty is considered "neglected". We have the Sustainable Development Goals of the United Nations, many organisations run by various churches, overseas aid from special funds of the most affluent countries, and rich philanthropists all addressing poverty. Yet, according effective altruists, it remains a neglected area. And although there is currently progress, the striking reality is that around 10% of the world's population still live in extreme poverty.[7]. If it is a neglected area of charity, then, considering the number of Christians, it must also be neglected by them. Can we say that humanity, and especially Christians, do not see the poor in distant lands as their neighbours, or see them as neighbours whose needs are of a much smaller importance than those of their geographically closer neighbours? The question has a practical nature, but the theoretical reasoning provided in this article suggests how Christians should respond to the problem of poverty eradication.

3. Christian Love and Overseas Aid

There is a common agreement that people are equal. Nowadays, it is a moral assumption and a starting point for further argumentation. This postulate does not imply full equality in every area, especially in the category of the distribution of wealth. However, the extreme global inequalities present in this last respect suggest this postulate is not taken seriously enough in our global society. Taking a global perspective, we can see that some people live in extreme affluence, whereas others have literally nothing and

6 MacAskill, *Doing Good Better,* 224.
7 Wydick, *Shrewd Samaritan,* 39.

are threatened with starvation or easily preventable diseases.[8] There is a substantial inequality in wealth, access to health services, education, and housing, etc. This is a massive injustice.

What can also be observed in everyday situations is that people are more willing to repair injustices that are close to them rather than far away, even if there is an extreme disproportion between these two kinds of injustices and even if people in distant countries might benefit much more from the same amount of money donated.[9] Even more can be said: a tendency to favour any, even the least important, wishes of the nearest (and oneself) rather than the crucial, vital interests of the neediest does not seem to be uncommon.[10] But before claiming that some people are neglecting their duties, it is appropriate to ask if supporting those people who are geographically distant should really considered an obligation for Christians.

There is a considerable amount of well-known literature encouraging Christians to fight extreme poverty. There are examples of saints and church documents to support the idea that Christians should struggle to help eradicate extreme poverty. Finding the arguments against helping the poor in Christian thought is much more difficult. However, the belief that Christians are not obliged to help the poor in distant lands can be supported by certain philosophical, social, and even theological arguments. To grasp the picture, some of them will be presented.

Some philosophers highlighted theoretical problems with Christian love. According to German philosopher Nikolai Hartman, love of unknown, remote people (*Fernstenliebe*) is the highest value; it was discovered only recently in the modern era and is not of Christian origin.[11] Christian love, which he calls "brotherly love" or "neighbourly love", "places one's own ego on a level with that of others, concerning itself merely with those who are nearest at hand, those accidentally present, with the narrow circle of those, who are within reach".[12] Another German philosopher, Max Scheler, tended to see the love of mankind (*humanitas*) as a modern rebellion against Christian love, powered by a hatred towards God and tradition.[13] From these opinions, one could conclude that love for all people is not a Christian approach.

8 Sider, *Rich Christians,* 3–40.
9 MacAskill, *Doing Good Better,* 28.
10 Hallett, *Priorities*, 2–3.
11 Hartman, *Ethics,* 311–331
12 Ibid., 269.
13 Scheler, *Sympathy,* 2007.

An argument against supporting the poor in distant lands can also be developed on the basis of the interpretation of the influential concept of the *ordo caritatis*, as presented by St. Thomas Aquinas in his *Summa Theologica*.[14] Aquinas presented guidelines for rational charitable actions, which were later developed by Christian thinkers. The *ordo caritatis* is a set of reflections on rational Christian love. It gives guidelines in complicated situations regarding charitable actions, for example, how to use limited resources in the event of too many potential beneficiaries for available assistance.[15] Alternatively, it can be interpreted as a rule of favouring relatives over non-relatives and those in geographic proximity over distant people. This is not just an abstract theory as this framework shapes the real world. For example, in a debate about accepting the relocation of refugees from Southern Europe to Poland in 2015, a prominent politician opposing the idea said: "We must look for another principle, a principle that moderates this radicalism. Well, there is such a principle – that is, *ordo caritatis* – the order of mercy, love. And within this principle: the closest are first, the family, then the nation, then others".[16] The politician then mentioned that the *ordo caritatis* does not exclude extending help to those people from outside the circle of the nation, but this help is conditioned by the well-being of the Polish people. Only the quoted section of his speech gained the attention of the media. Even if the concept of *ordo caritatis* does not, in principle, support the idea that geographically distant people should not be helped at all, it does seem that focusing on the order of potential beneficiaries, rather than the greatness of need, makes enough room for excuses so as to lead to neglecting the very vital interest of distant people while supporting the trivial ones of our relatives and selves. Emphasising only the order of beneficiaries is by no means a full picture of *ordo caritatis*. Thomas himself said that the circumstances matter in the assessment: "because in certain cases one ought, for instance, to succour a stranger, in extreme necessity, rather than one's own father, if he is not in such urgent need".[17]

14 ST, ii–II, q26.

15 Bartoszek, "Odpowiedzialność," 59.

16 "Musimy szukać zasady innej, zasady, która by moderowała ten radykalizm. Otóż taka zasada istnieje – to jest ordo caritatis – porządek miłosierdzia, umiłowania. I w ramach tej zasady: najpierw są najbliżsi, rodzina, później naród, później inni." Translation mine. See Onet, "Mocne wystąpienie."

17 ST, ii–II, q31 a.3.

There are also theological arguments, as scripture can be interpreted in many places as limiting love to geographic proximity or even to the people of our nation. As an example, one may read Mark 7:24–30, when Jesus honoured the Syrophoenician woman's faith. Jesus told the woman: "First let the children eat all they want," and then he said, "for it is not right to take the children's bread and toss it to the dogs." Equally, in the interpretation of the Parable of the Good Samaritan the emphasis can be placed not on universal love, but on other aspects of the Samaritan's help. These include his rationality since the Samaritan prudently applied the resources he had. But the fact that he possessed resources in the first place was an effect of his life model. This model was not that of a person focused on helping everyone he could affect – as for the substantial part of his time, the Samaritan was earning. Practising *caritas* – helping the ones in need – was not the essence of his life.[18]

The actual practice of Christian help can also be a subject of investigation. It appears to be supporting "discrimination" on the basis of distance: when it comes to giving, more resources are donated to the secondary needs of local Christian communities than to overseas aid, which addresses people's basic and existential problems.

Moreover, the common-sense understanding of the word "neighbour" in English may lead to the conclusion that this basic commandment is about the people near to us.

In contrast to this, I argue that excluding far-away people from "neighbours", which can be also called discrimination on the basis of distance (geographical distancism), is not the obvious indication as to whom one should consider to be an object of love in the rule "love your neighbour". This is not to deny that these selected arguments supporting the thesis that Christians are not obliged to help the poor in distant lands should all be treated with proper respect and caution. In this article, however, there is no room for the discussion they all deserve.

4. Jesus and Utilitarian Reasoning

One more important remark must be added. Although EA does not necessarily presuppose utilitarianism,[19] during the process of becoming acquainted with EA one will encounter utilitarian thinking. There are some schemes which utilitarian thinking favours. For example, a common utili-

18 Cf. Bartoszek, "Odpowiedzialność," 66–67.

19 Cf. Singer, *The Most Good,* 9.

tarian response to the question "who is an object of my moral concern?" would include a catalogue of the morally relevant and preferably measurable features of such a being. These features are meant to make ethics more precise; they also provide the chance to reasonably include into moral considerations some beings that were not previously the obvious objects of moral concern. But it must be noted that a catalogue of that kind may also give reasonable grounds to reject the moral status of certain beings that were previously considered. An immense responsibility is placed on the philosopher who makes an attempt to lay down criteria of moral concern as he provides a basis with which to divide the world into two categories: with moral status and without it. For many beings, this is a life-or-death decision. It must also be noticed that the thinkers inspired by Christianity are often involved in this kind of thinking, which means looking for a set of features that a being should have in order to be perceived as a moral object (and also to exclude other kinds of beings, like animals, which do not belong to a species that has a rational nature).

To search for the criteria which a being requires to be considered of moral value is to take the path of the lawyer who asked Jesus the question of who his neighbour is. Jesus' answer was radically different and went far beyond the limits of the question of his adversary. Jesus answered with a parable:

> "A man was going down from Jerusalem to Jericho, and he fell among robbers, who stripped him and beat him, and departed, leaving him half dead. Now by chance a priest was going down that road; and when he saw him he passed by on the other side. So likewise a Levite, when he came to the place and saw him, passed by on the other side. But a Samaritan, as he journeyed, came to where he was; and when he saw him, he had compassion, and went to him and bound up his wounds, pouring on oil and wine; then he set him on his own beast and brought him to an inn, and took care of him. And the next day he took out two denarii and gave them to the innkeeper, saying, 'Take care of him; and whatever more you spend, I will repay you when I come back. Which of these three, do you think, proved neighbor to the man who fell among the robbers? He [the lawyer – J.S.] said, 'The one who showed mercy on him.' And Jesus said to him, 'Go and do likewise.'" (Luke 10:30–37, RSVCE)

Instead of providing a list of features a being should possess to make it capable of deserving our love, Jesus showed us the attitude that one should have to call oneself a neighbour. Jesus said: be neighbours yourself. An example of that approach was given by the Good Samaritan. He was not calculating, and he did not consider whether the wounded man deserved his loving action; he simply acted out of love. What seems to have surprised the

lawyer, and may still be surprising for the contemporary man, is that instead of providing a specific answer to the question "who is my neighbour?" or "to whom do I have moral duties?", Jesus seems to have redirected our thinking to the question "how can I be a neighbour?", or "what shall I do to 'go and do likewise?'". Since EA is a contemporary attempt to reflect upon the method of benevolent acts in the light of philosophy and research, it is not surprising that some Christians may find it to be inspiring.

As it is, the parable of the Good Samaritan and EA exhibit further parallels: the Samaritan seemed to apply some of its guidelines practically in order to help the wounded man in the best way possible. He sacrificed both his time and money, but in the latter case, he trusted the specialist (referred to in the Parable as the innkeeper) instead of performing certain actions himself. As a result, some of his loving acts were done personally, but indirectly. What is more: there is a remark that the Samaritan checked the effectiveness of the innkeeper substituting him – he promised to return to ask after the condition of the wounded traveller, and also offered to be more generous and provide resources to meet unexpected costs after the evidence-based verification.

5. Distant People as Neighbours

Jesus' answer to the lawyer tells us that we are supposed to think of how to be neighbours rather than look around and ponder who deserves our love and who does not. That guideline could be enough for a rule of practical reasoning on the individual level.

It seems that on the basis of the Parable of the Good Samaritan, another view on the response to "who is my neighbour" is possible. There is a way to find an answer that seems to make Christian love better informed and suits the need for dialogue with contemporary lawyers. The proposed framework to look for criteria of the neighbour contains the features of the recipient of the Samaritan's help. These will be used to see if people in distant lands, animals, and future unidentified beings can be considered neighbours.

I want to suggest two interpretations of how the parable structures our views about who our neighbour is. The first interpretation is that it argues that our neighbours are simply everyone on the planet, as the distinctions that were typically used to determine who is one's neighbour are shown to have become irrelevant. Many theologians are of the opinion that through the parable of the Good Samaritan, Jesus wanted to say that "neighbours"

are all of the people around the world.[20] If this is correct, those Christians who believe that the maxim "love your neighbour" is the fundamental principle of charity should not allow geographical closeness to dictate who benefits from their charitable actions.

The parable of the Good Samaritan was told during a time when most Jewish people did not recognise the Samaritans as their neighbours, even though they were living in the same area. The two groups hated each other. Jesus' teachings had to be difficult for the ancient societies, which were (generally) not able to broaden their circle of moral care beyond their close social relations. As Bruce Wydick explained it, "Jesus pushed the concept of 'neighbour' about as far as possible with his first-century Jewish audience (…) the concept of 'neighbour' was generally limited to one's local network of extended family and friends" and eventually other members of the nation.[21] The parable can be interpreted as indicating that every person is our neighbour, even the ones who exist in our imagination as mere numbers, and thus the circle of moral care (or the circle of love) should be broadened to include all people. The practical conclusions of this reasoning might be unfamiliar and very demanding for people of all the times, and thus as hard for them to accept as Jesus' teachings were for the people in the first century, who were also called to abandon their lifestyle.

6. The Less Demanding Interpretation

However, another interpretation is also possible and, at first sight, less demanding. It may suggest that the Christian understanding of neighbour can be limited in a way that justifies favouring the nearer over the neediest. Therefore, this possible interpretation is a subject of special focus in this article. One can argue that our neighbour is anyone whom we can personally affect. This conclusion can be derived from what the parable tells us about the object of the Samaritan's action. The Samaritan helped "the wounded man" whose ethnic membership was not stated. As a result, we cannot say for sure if the traveller who was attacked by the robbers was a Jew or not (although it seems that the priest and the Levite did not even check). It is neither stated that the Samaritan had a personal relationship with the victim of the robbers nor that he was capable of entering into a dialogue with him. This indicates that ethnic membership is not a morally relevant factor and also shows that the object of a merciful action does not

20 See e.g. Benedict XVI, *Deus Caritas Est,* and Krawczyk, "Nakaz."
21 Wydick, *Shrewd Samaritan,* 19.

have to be in any relation with us prior to our recognition of his need. All we know about the man is that: (1) he required help, (2) he was an adult male human being who was seen by the Good Samaritan, (3) and that he was within range of the Samaritan's action (the Samaritan himself could do something to help). Also, it appears that the age and gender of the victim were accidental, but the species is not. Limiting the circle of moral concern to adult human males would be in contrast with the whole teaching of Jesus, whereas if Jesus decided to insert an animal into this parable instead of a human being, that would be a great novelty in his teaching. Thus, it might be concluded that a neighbour is a human being whom we can personally affect.

The next step in this reasoning is to link the mentioned features of the neighbours with the rule that neighbours are those whom we are supposed to "love as ourselves". This shows that our moral duties extend to all those whom we can affect and help. From this statement, it seems that we need to recognise and understand our abilities and power in order to determine who our neighbour is. If this is a correct form of reasoning, being a neighbour is not some objective property of beings, nor is it a property acquired once and forever.

I call this interpretation the less demanding interpretation in order to contrast it with an interpretation which simply broadens the range of our moral duties to all people. It says we should only consider the people whom we can personally affect. It opens ways to silence the conscience in the instances of some people that suffer. As it seems many people believe they can effectively affect only those people who are geographically close, this interpretation can be understood as a theoretical support for framing our obligations according to our natural tendency to help the people near us, to prefer supporting our family or people from our social surroundings, and to favour small developments in our local community (new museums or better heating for the church) over eradicating the extreme poverty of numerous people living in distant lands. Rather than being the outcomes of philosophical and theological debates alone, natural tendencies and commonly held beliefs are probably a piece of an answer to the question of why wealthy Christians and extreme poverty exist at the same time.

7. The Less Demanding Interpretation in the 21st Century

The second interpretation of the parable seems to be less demanding, but it will be argued that even this interpretation does not support framing our charity actions according to our natural tendency to favour those who are geographically close to us, at least nowadays.

At the same time, it seems that in the past this interpretation supported favouring those who were geographically close, but it was not an instance of geographical distancism. Rather, it is a fact that for the past generations, possibilities were different. People had less power to affect others and were far less aware of the situation in places outside of their communities. They did not (as we often do) hold in their pockets immediate access to the news from around the world, potentially all of human knowledge, and a great selection of funny cat pictures.

In the past, to be able to affect others, an ordinary person needed to be able to meet them physically. The range of people who could be affected was limited by the distance one (or one's deputies) could personally travel. Over the ages, various developments made it possible to travel further, even to the most abandoned areas of the world, but physical distance remained a problem that limited our ability and power to help. People who were willing to help the poor from distant lands were also not in the Samaritan's situation, as he knew how to efficiently help the wounded man (he knew, for instance, whom he could pay in order to make the situation of the suffering man better). Further, the problems of people living in extreme poverty (as well as extreme poverty itself) as a problem are simply more complex than the situation of the attacked traveller from the parable. For years, although a considerable amount of money was spent on giving aid to people in distant countries, we barely had an idea of what could help them in an effective way. Some attempts to eradicate extreme poverty are claimed to do more harm than good, and do so at the cost of precious resources.[22] Alleviating domestic poverty has become easier as we have better understood it, and that understanding has allowed us the chance to provide effective help. In the case of poverty and the problems of people in distant countries and of different cultures, wealthy nations used to speculate, project their best solutions, and maintain a more or less paternalistic approach. The scientific approach to this problem, especially including randomised controlled trials, made effective help more likely.[23] It is a fairly recent, although not fully

22 See e.g. Easterly, *White Man's Burden,* and Leszczyński, *Eksperymenty*.
23 MacAskill, *Doing Good Better,* 7–11.

uncontroversial, invention.[24] Also, over the past decades, affluence has not been very common: there have been affluent people, but they have been surrounded by poverty. The level of affluence, especially in the wealthiest countries, has no precedence in history. In the words of Peter Singer, numerous people in the most affluent countries are better off than King Louis XIV.[25] The morally significant difference is that a king was only one person, and now we have millions of wealthy people. Our ancestors had to struggle for their lives and for their families; often, they had very limited resources to support even the closest relatives. We have much more power, and this power is necessary to help people in extreme poverty. What would the Good Samaritan have done for the man without his horse, money, or knowledge of the inns that the wounded man could be helped at?

The world has changed, and our power has increased. Nowadays, global connections make it possible to personally affect people located all around the world. What is more, people from the most affluent countries are more aware of that which creates opportunities for them to become "Good Samaritans". Of course, one should neither underestimate the negative side of globalisation nor the growth of the power of the wealthy. The most obvious way wealthy people affect others in distant countries is pollution, but there are many other ways, bringing both good and harm. For example, in recent years, there has been a rising awareness as to how common it is to use another's misery to get wealthier by supporting, consciously or not, the evil structures exploiting people. However, due to the increase of power, citizens of the most affluent countries have enough resources and technical possibilities to provide help to the poor in distant lands.

We can now also identify the needs of nearly every person in the world, thanks to the media and the internet. Traveling and transporting goods are both becoming cheaper and cheaper. Technology allows us to affect many people around the world instantly with just few a clicks. As Benedict XVI said, globalisation placed at "our disposal numerous means for offering humanitarian assistance to our brothers and sisters in need, not the least modern systems of distributing food and clothing, and of providing housing and care."[26]

Finally, the development of charity research, as well as the various improvements of many charity organisations, allow us to help more people whose distant locations would have made such assistance impossible a few

24 See Leszczyński, *Eksperymenty*. esp. ch. 4.
25 Singer, *Life,* 9.
26 Benedict XVI, *Deus Caritas Est,* 30.

years ago. We can help in an effective way as we are more acquainted with which actions help people in extreme poverty who are within range of our actions.

Since geographical distance no longer strongly limits our abilities to affect and help people across the world, those who are geographically far away must also be called our neighbours, and so the term "global neighbour", as used by Wydick, seems justified.[27] We are to love our global neighbours as we love ourselves. Taking this seriously would lead to substantial changes in the lives of contemporary Christians and could lead them to internalise at least some of the principles and practical solutions that have been developed by the intellectual efforts of EA. However, does the parable of the Good Samaritan provide a potential for Christians to find a basis to support other cause areas of EA as well?

8. Animals as Neighbours

For effective altruists, the suffering of animals morally matters. The severe and increasing amount of easily preventable animal suffering, together with the very little concern expressed by humans and the relatively small resources distributed to prevent it, has made the elimination of animal suffering one of the priorities of EA. The proposed line of reasoning offered by the less demanding interpretation of the Parable of the Good Samaritan does not allow for the understanding of animals as neighbours. If this reasoning is correct, Christians are not obliged to love animals as themselves and the elimination of animal suffering is not a priority. Yet this does not mean they can or should be indifferent to the suffering of animals. The arguments concerning prioritising the elimination of animal suffering in Christian charity require a different framework.

Among pro-animal philosophers, there is a major debate about the impact that Christianity has had on the way our culture perceives animals. Both the desacralisation of nature and the replacement of the ancient animate-inanimate division with a modern version based on consciousness are seen as the roots of our mistreating animals. And both are also seen as having a Christian origin.[28] Although these theses can be rejected, it is undeniable that the concept of factory farms was invented in Western culture, which was developed on the basis of Christian values. A large part of scientific development is also enabled by animal suffering. Also, according

27 Wydick, *Shrewd Samaritan,* 19.

28 See for example Singer, *Animal Liberation,* 271–293.

to Charles Camosy, for some Christians, the issue of taking animal rights seriously is a part of the cultural conflict with atheism. Animal rights activism is now being associated with a challenge to their fundamental Christian beliefs – specifically regarding God and the value of human life. Therefore, a rejection of the concern for animals is seen as a way to defend their faith. In consequence, they strongly oppose taking animals into moral consideration.[29] Utilitarianism, on the other hand, harbours extensive concern for animals. There are many kinds of utilitarian arguments to the effect that we should consider animals in our moral decision-making as they have at least one morally relevant feature: they can feel pain. Moral duties towards animals have also been acknowledged in a non-utilitarian way by philosophers who were inspired by Christian thought, such as Andrew Linzey and Charles Camosy.[30]

The less demanding interpretation of the parable allows us to include all people in our moral decision-making. I mentioned that, on the basis of Jesus' overall teaching, it seems plausible to claim that species matter in this interpretation of the parable, and therefore animals are not neighbours. The person that was helped by the Good Samaritan was certainly a human being.

As a potential objection to this interpretation, it can be mentioned that Jesus' teachings were tailored to the sensitivity and capabilities of the people living in ancient times. The Bible never calls animals neighbours. We know that not every possible moral position could have been invoked during biblical times. We can only speculate as to how the parable explaining the notion of neighbour and portraying the Samaritan saving a wounded wild animal would have affected Jesus' disciples.

As was mentioned in the initial sections of this article, Jesus' answer was not a set of criteria, but rather a call to be a neighbour oneself. Thus, regardless of our personal stance on the previous reasoning about whether animals are neighbours or not, all of us can ask if we can be neighbours to animals and what this would mean. The opportunities of the 21st century have significantly affected the way in which we can approach animals, in a similar manner to the changes in our relations with people located around the world. We affect animals in many ways; by climate change, by our direct actions, and by our dietary choices. And here, just like with people in need, we tend to be arbitrary in our concern. We show love to some animals we encounter, but there is little care for the animals we put into our mouths.

29 Camosy, *Love of Animals,* 16–17.

30 Linzey, *Animal Theology;* Camosy, *Love of Animals.*

Would you agree to torture the animal that you can see for the sake of the best lunch ever? I think not. But our meat-based lunches in the first decades of the 21st century nearly always mean a lot of suffering for animals.

The development of ethical reflection and the emerging awareness of the needs of animals enable us to be neighbours to them. Being a Good Samaritan to animals would certainly mean something different than being a Good Samaritan to people. It is for science to tell us what animals' needs and capabilities are. By knowing them better, we can also modify our attitudes towards them.

For utilitarian philosophers, animals should be considered in moral calculations because of their ability to feel pain. For Christians, even if animals cannot be considered neighbours, their needs can be recognised and helped – their pain cannot just be ignored. Christians have good reasons to be neighbours to animals even if, on the basis of the Bible and the parable of the Good Samaritan (using its "less demanding interpretation"), one would be convinced that animals cannot be recognised as neighbours. And potentially these reasons (and Christian solutions) may protect animals even more strongly than the mere focus on their ability to feel pain. After all, by noticing the intrinsic value of animal life, or formulating human duties towards animals on the basis of a "covenant that should mirror God's creative love",[31] they may protect animals in a host of circumstances going beyond pleasure and pain, like protecting them from being killed, not only in a cruel way but also painlessly or securing the biodiversity of creation.

9. Future People as Neighbours

Many effective altruists take it as a given that humanity should prioritise existential risk prevention and thus, in a sense, prioritise the far future. In this section, on the basis of the less-demanding interpretation of the Parable of the Good Samaritan, I argue that those living in the far future cannot be captured by the term "neighbour", and that neither can we treat them as neighbours in a way analogical to contemporary, living people and animals. I will also indicate that there are some moral obligations that have arisen in the Christianity ethical tradition which secure the well-being of those living in the far future.

The fact that effective altruists often assign considerable weight to the well-being of future people may discourage some Christians from supporting EA. Placing such a strong emphasis on the well-being of future

31 Benedict XVI, *Caritas in Veritate*, 69.

people – and all future beings – is a characteristic of consequentialist views. These views are focused on the outcomes of actions. We can rationally expect that some (indeed most) of our actions will also have consequences extending to those people who do not exist now. This is assumed in many of our everyday choices. For example, the decision a young couple makes regarding where to live has an impact on their unborn, unplanned children, grandchildren, etc. Young couples normally consider their future children while making housing decisions. How far should this thinking extend? Should they also consider the children of their children and the children of their (geographical) neighbours from both the old and new location, or maybe from all possible locations?

For Nick Bostrom, we should consider the people potentially living on the Earth five billion years from now – as long as it is habitable. Essentially, his reasoning is that we should prioritise existential risk prevention over other problems that have affected the development of EA.[32] Bostrom argued that preventing an existential catastrophe would help everyone living in the future. The number of people thus affected is so huge that even small reductions in the risks of such catastrophes are better than any action that affects only contemporary people.

Both Christians and secular thinkers direct their thoughts towards the future, but there are important differences between them. From the Christian perspective, one may say that the existence of humanity and of the world is within the power of God, not of humans, and so we cannot say how long the world will last. This argument burdens the future-oriented calculations of Bostrom with the additional risk of inadequacy. What if the apocalypse were to start tomorrow? Christians are encouraged to be ready for that event, to wait for it, and to expect it rather than the humanity lasting until the sun eventually "eats" the Earth (1 Thess. 5:2–4; 2 Pet. 3). The long-termists who encourage shaping the story of humanity's distant future do so based on the assumption that humanity can last until the end of the Earth, or maybe even until the end of the cosmos, and out of optimism regarding the quality of the lives of future people.[33] These possibilities are also acknowledged by Christians, but religious people have a perspective that goes beyond earthly life with its pains and pleasures. Christians value earthly life, even one that entails more suffering than joy, as for them it is God's gift and also just a stage. It is not everything. They believe the life of the human being does not end with death. They have

32 Bostrom, "Existential Risk."
33 Ord, *Precipice,* 20–22.

religious reasons to not plan for the far future of earthly life. This seems to be an argument for Christians to discourage each other from accepting the logic of prioritising efforts to prevent the extinction of the human species. The Apocalypse, from the perspective of a secular thinker, might be considered a global catastrophe. For Christians, on the contrary, it is the very moment of humanity reaching its full potential, securing not only the far future but also eternity.

Are those living in the far future our neighbours, according to the less demanding interpretation of the Parable of the Good Samaritan? In thinking of who constitutes a neighbour we take as an example the wounded man. His needs were recognised by the Good Samaritan and this made him act in a compassionate way. Within the framework of "the less demanding interpretation", I argued for a broadening of the notion of the neighbour to include those people distant in space whom we can personally affect, thanks to the development of science and technology and the ensuing increase in our power. Can this be applied to future (and especially far future) people?

The first difficulty we encounter when thinking about people in the far future is the impossibility of identifying them. We can think about the next generations only in terms of estimated numbers or in terms of humanity as an abstract idea. Being unable to identify those living in the far future makes it difficult to understand how particular future people can be helped by us. In the case of animals and people living in distant countries, we can identify their needs with an accuracy not previously seen, thanks to the opportunities and capabilities of our times. But we cannot really guess what problems future people will have. Our best estimations may fail as we do not have the proper perspective. For example, we want to preserve nature for the future people. But here we are extrapolating our needs and categories. Imagine that your great grandparents did what they could to preserve their old mud hut for you. Their efforts and sacrifices would deserve respect. Yet, they would have allocated their precious resources wrongly. The world is changing very quickly and so are the values dominating society. Maybe people in the year 2200, for instance, might prefer virtual reality and not be interested in the Alps.

One very basic need we can likely think of in the case of those living in the far future is for them to have an opportunity to exist. A prerequisite for this is the further existence of our species. Although this need is so basic that thinking about it is also the outcome of the increase of our power and the development of technology. Power and technology have the potential to lead to an anthropogenic catastrophe if they are misused or even encounter

a trivial accident. Only very recently (compared to the history of our species) has humanity acquired the ability to destroy itself (and to prevent this self-destruction). In addition, only recently has humanity learned about potential natural risks and developed the technology necessary to allow even some hope of preventing some of them. Preventing the destruction of humanity is also an opportunity of our times.

The other needs of people are dependent on their very existence. We cannot recognise the needs of those living in the far future now – this will be the task for far future Samaritans. The case of preventing the extinction of humanity may be a good example of how we can be neighbours to those living in the far future. Further, if we do not recognise this basic duty and continue acting in ways that will contribute to a potential global catastrophe, we seem to act as if we are the robbers from the Parable. And there are indeed some robbers nowadays, those who exploit the world as if it was meant only for one generation and thus contribute to the anthropogenic threats for humanity (mass destruction warfare, catastrophic climate change, environmental damage, and other risks.)[34] Following the pattern of the Good Samaritan here might be interpreted as a call to mitigate the damage caused by the robbers, and therefore a call to sacrifice something for the sake of fighting the anthropogenic existential threats for humanity.

Another way of being a neighbour to those living in the far future is specifically a Christian method. Christian anthropology perceives the human being as being in a relationship with God with a specific goal to achieve – salvation. Salvation is the universal goal of all people, as well as future people. And it is a task for every Christian to help others to join the Kingdom of God. Our ancestors understood this mission well. When we were the far future generations, Christians made many sacrifices to ensure our future. They struggled to maintain, develop, and spread the Christian religion to secure our afterlife. The magnificent relics of these efforts are admired by contemporary people, even if many do not understand the purpose of the extraordinary efforts made by our ancestors. In many cases, the most solid and beautiful building in a town is its church. So often, it was built by many generations for many generations, transcending the needs and interests of a single generation. If the universal aim of human beings is salvation, and if we believe this is achieved through religion and church, we should do our best to ensure future people will have access to

34 See Ord, *Precipice,* 89–162.

these. Being a neighbour to future people may then also mean contributing to your church now as a safeguard for the future.

There are ways to be a neighbour to the those living in the far future, although, again, this means something radically different from being a neighbour to the people in distant lands or to animals. We may now try to answer the question of if those living in the far future are our neighbours according to "the less demanding interpretation". A neighbour is someone we can personally affect. A question to be asked, then, is if we can affect those living in the far future at all. It is true that our actions and choices will have their effects in the future. Some of them will shape the far future by making it a different far future than it would otherwise be. Some opportunities will be lost forever. However, this should not be interpreted as personally affecting future people. In the same way, our generation is limited by the choices and actions of the previous generation. I suggest an analogy. Can we say that by their contribution to making the Dodo Bird or the Passenger Pigeon extinct, some people were affecting us personally? Or that the people from the past who resigned from their hunter-gatherer lifestyles and started settlements affected us personally? In those times, when we did not exist, they did not seem to be able to do it, even if they could foresee that their actions would shape the lives of the next generations.

The closer future people are to us, the more we can identify their needs, and the more pressing the metaphysical question might be: can I affect someone who does not exist yet? If yes, how can I determine if I affect them in the right or the wrong way? There are future people whose needs can be foreseen as they are close in time – like in the example of a family planning a house and their children in the nearest future. Their needs can be recognised and we can benefit them; thus, we can be neighbours to them, and so by taking the proposed less demanding interpretation it appears that some future people might be neighbours. The mere probability should be enough to warrant treating them as such.

As I argued in the part about distant people, the notion of "neighbour" does not seem to be fixed once and for all. The arrow of time inevitably brings future generations closer to us, and there is no clear border between the far future and the near future. Also, with the change of empirical circumstances and an increase of our knowledge and power, it is possible for more future people to be personally affected by us and fall into the category of neighbours.

Even if we currently find little reason to consider those living in the far future our neighbours and thus claim direct duties towards them, it does not mean the future of humanity is not a subject of concern for Christians.

Christians have good arguments for preventing anthropogenic global catastrophes – not only those related to war,[35] but also those connected with the environment, as the world was given as a means of perfection to all people, not only to the wealthy, or to one or a few generations.[36] Pope Francis emphasised that the notions of the common good and solidarity also include people from the future:

> "We can no longer speak of sustainable development apart from intergenerational solidarity. Once we start to think about the kind of world we are leaving to future generations, we look at things differently; we realise that the world is a gift which we have freely received and must share with others (…) Intergenerational solidarity is not optional, but rather a basic question of justice, since the world we have received also belongs to those who will follow us."[37]

This reflection, made by the head of the Catholic Church, seems to include the generations from the near future: they will receive the world, as well as the duty to protect it for those generations near to them. If we are successful in preventing anthropogenic catastrophes for the forthcoming generations and passing the world (as well as the idea of intergenerational solidarity) on to them, those living in the far future will also benefit. From a Christian perspective, the mere possibility of the arrival of future people should be appreciated, and therefore "the garden" should be kept in such a condition so that future generations may use it.

The focus on the future is also a part of Christian thinking, but Christians do not abandon the present for the sake of a better tomorrow. They look towards the (eternal) future and organise their terrestrial life according to it.[38] In that sense, Jesus' order "Go and do likewise" is applicable also to the those living in the far future, although what we know about them makes us focus mainly on preventing an anthropogenic global catastrophe and assuring the earthly means to salvation. On the basis of the "less demanding" interpretation of the parable, those living in the far future are not our neighbours, but it does not mean they do not count.

35 See for example John XXIII, *Pacem in Terris,* 109–115.
36 Francis, *Laudato si',* 95.
37 Ibid., 159.
38 Tischner, *Nadzieja,* 27.

10. Conclusion

The question "who is my neighbour" is one that we still hear very often – in philosophy, in public debates, in our hearts. Global changes, including the rapid development of technology and the increase of wealth, have renewed the relevance of this question. On the basis of the proposed framework of a "less demanding" interpretation of the notion of "neighbour" from the parable of the Good Samaritan, it seems Christians have good grounds to accept the point that: our neighbours are all the people of the world. And since Christians are obliged to love their neighbours as themselves, the lives of Christians should adapt to meet this obligation. Some of the guidelines provided by EA may be used by them in order to fulfil the call from the last sentence of the parable, "Go and do likewise" (Luke 10:37), alongside the words of Pope Francis: "In the face of so much pain and suffering, our only course is to imitate the Good Samaritan. Any other decision would make us either one of the robbers or one of those who walked by without showing compassion for the sufferings of the man on the roadside."[39] The Pope encourages us to not be indifferent to suffering, even in our everyday choices, where our indifference might mean we agree to benefit from the suffering of our neighbours.[40] He emphasises our responsibility for the wounded men existing around the world:[41] "Let us care for the needs of every man and woman, young and old, with the same fraternal spirit of care and closeness that marked the Good Samaritan."[42]

The development of humanity that has occurred in recent decades and manifested itself in the progress of technology, economy, and ethics leaves little room to reject the understanding of the range of the notion of neighbour as globally applicable on the basis of the Parable of the Good Samaritan. Even the interpretation that seemed to leave a lot of room to exclude faraway people from the category of neighbours appears to broaden this notion, at least for those who are participating in this spectacular increase of wealth and opportunities.

EA goes beyond helping the poor in extreme need. In this framework, our charity efforts should also deal with the unnecessary suffering of animals and work to secure the future of humanity. The concern for animals in Christian ethics is present and has recently received more attention than

39 Francis, *Fratelli Tutti,* 67.
40 Ibid., 75.
41 Ibid., 77.
42 Ibid., 79.

in the past. On the basis of the proposed reasoning, the notion of neighbour could not be broadened to include animals. This does not mean the suffering of animals is irrelevant, or one cannot be a neighbour to them, but the "love as yourself" principle was not proven to be applicable in this case.

The third cause area of EA, the case of those living in the far future, is controversial for Christians. It seems we can be neighbours to them, but here being a neighbour means doing our best to prevent an anthropogenic global catastrophe and assure access to means that help the recipient on the way to the Kingdom of God. The framework proposed in this article did not provide sufficient support for the conclusion that those living in the far future can be called neighbours, and thus must be loved as ourselves. This means re-organising life in order to provide love to the future people (as EA would propose) cannot be recommended on the basis of the provided argumentation. Obligations to the created world and already-existing people, if well understood, would, however, contribute to securing the vital interests of future people and prevent at least some of the possible anthropogenic threats to humanity.

An interesting point for further research would be considering if the three groups analysed here are the only (so-far) underestimated candidates for broadening the notion of neighbour, thanks to the opportunities of our times. I highlight one potential example as a starting point in this debate. A group that does not seem to be recognised as counting morally by many contemporary moral philosophers is that of existing (but unborn) people. We can personally affect these invisible and very vulnerable human beings for a long time before we physically see them. And again, the development of technology and ethics has changed very much in the recent years, both in our ability to be either robbers or Samaritans to them. Even if many Christians strongly oppose abortion, a genuine application of the "love as yourself" commandment would probably mean significant changes in this area, too.

The discussion surrounding the notion of neighbour in relation to the EA priorities has a potential to enrich the dialogue within the movement. EA does not say the list of priorities is exact, unchangeable, or complete. Christians interested in EA may also contribute to the process of selecting priorities. The sensibilities of Christians and Christian anthropology may suggest an understanding of good that goes beyond the material dimension and draws attention to people that are physically healthy, wealthy, and fed, yet still suffer. Are wealthy societies not full of people suffering due to the same development that equipped us with opportunities to help distant poor

materially? Mental illness or being abandoned and lonely cause a great deal of suffering.

The notion of neighbour appears to be dynamic, and if taken within the framework of the proposed interpretation of the Parable of the Good Samaritan, I argue that within our generation understanding of this notion will be broadened significantly regarding the number of people involved. The progress of science and technology can make us acknowledge more creatures as our neighbours. Yet, what seems to be of greater importance is that this progress informs us how to "love" in a more effective way – by recognising the real needs affecting people and the tools to mitigate these needs. Christians should apply this framework to neighbours, and are encouraged by Jesus to be neighbours themselves – by applying the knowledge of how to love in an effective way to other beings who can be treated as neighbours.

References

Aquinas, Thomas. *Summa Theologica* [ST], second and revised edition. Translated by Fathers of the English Dominican Province. London: Burns Oates & Washbourne, 1920.

Bartosek, Antoni. "Odpowiedzialność za bliźnich w potrzebie w kontekście kategorii ordo caritatis (z uwzględnieniem problemu migracji)" [Responsibility for the neighbours in need in the context of the category ordo caritatis (including the problem of migration)]. *Teologia i Moralność* 12 (2017): 59–72.

Benedict XVI. *Deus Caritas Est* [Encyclical Letter on Christian Love]. Accessed 30 September 2021. https://www.vatican.va/content/benedict-xvi/en/encyclicals/documents/hf_ben-xvi_enc_20051225_deus-caritas-est.html.

Benedict XVI. *Caritas in Veritate* [Encyclical Letter on Integral Human Development in Charity and Truth]. Accessed 30 September 2021. https://www.vatican.va/content/benedict-xvi/en/encyclicals/documents/hf_ben-xvi_enc_20090629_caritas-in-veritate.html.

Bostrom, Nick. "Existential Risk Prevention as Global Priority." *Global Policy* 4, no.1 (2013): 15–31.

Camosy, Charles. *For Love of Animals: Christian Ethics, Consistent Action*. Cincinnati: Franciscan Media, 2013.

Centre for Effective Altruism. "Effective Altruism Handbook." 2nd edition. Accessed 1 Feb 2021. https://www.effectivealtruism.org/handbook/.

Easterly, William. *The White Man's Burden: Why the West's Efforts to Aid the Rest Have Done So Much Ill and So Little Good*. London: Penguin, 2006.

Francis. *Laudato Si'* [Encyclical Letter on Care for our Common Home]. Accessed 30 September 2021. https://www.vatican.va/content/francesco/en/encyclicals/documents/papa-francesco_20150524_enciclica-laudato-si.html.

Francis. *Fratelli Tutti* [Encyclical Letter on Fraternity and Social Friendship]. Accessed 30 September 2021. https://www.vatican.va/content/francesco/en/encyclicals/documents/papa-francesco_20201003_enciclica-fratelli-tutti.html.

Hallett, Garth. *Priorities and Christian Ethics*. Cambridge: Cambridge University Press, 1998.

Hartmann, Nicolai. *Ethics*. *Volume 2*: *Moral Values*. Translated by Stanton Coit. London: Unwin Brothers, 1932.

John XXIII. *Pacem in Terris* [Encyclical Letter on Establishing Universal Peace in Truth, Justice, Charity, and Liberty]. Accessed 30 September 2021. https://www.vatican.va/content/john-xxiii/en/encyclicals/documents/hf_j-xxiii_enc_11041963_pacem.html.

Keener, Craig S. *The IVP Bible Background Commentary: New Testament*. 2nd edition. Downer's Grove: InterVarsity Press, 2014.

Krawczyk, Roman. "Nakaz Miłości Bliźniego w Świetle Biblii. Zarys Problematyki" [The obligation to love your neighbour in light of the Bible. Introduction to the issues]. *Studia Elbląskie* 19 (2018): 278–290.

Leszczyński, Adam. "Eksperymenty na Biednych. Polityczny, moralny i ekonomiczny spór o to, jak pomagać skutecznie" [Experiments on the poor. Political, moral and economical dispute on how to help effectively]. Warsaw: Krytyka Polityczna, 2016.

Linzey, Andrew. *Animal Theology*. Chicago: University of Illinois Press, 1994.

MacAskill, William. *Doing Good Better: Effective Altruism and a Radical New Way to Make a Difference*. London: Guardian Books, 2015.

Ord, Toby. *The Precipice: Existential Risk and the Future of Humanity*. London: Bloomsbury Publishing, 2020.

Onet. "Mocne wystąpienie Kaczyńskiego ws. imigrantów. Ambasada Szwecji odpowiada." Accessed 24 March 2021. https://wiadomosci.onet.pl/kraj/mocne-wystapienie-kaczynskiego-ws-imigrantow-ambasada-szwecji-odpowiada/zdnmqn.

Scheler, Max. *The Nature of Sympathy*. Translated by Peter Heath. London and New York: Taylor and Francis, 2007.

Sider, Ronald J. *Rich Christians in an Age of Hunger*: *Moving from Affluence to Generosity*. Nashville: W Publishing, 2015.
Singer, Peter. *The Expanding Circle*. Oxford: Oxford University Press, 1983.
Singer, Peter. *The Life You Can Save: Acting Now to End World Poverty*. London: Picador, 2009.
Singer, Peter. *The Most Good You Can Do: How Effective Altruism Is Changing Ideas About Living Ethically*. New Haven: Yale University Press, 2015.
Singer, Peter. *Animal Liberation*. 40th anniversary edition. New York: Open Road Integrated Media, 2015.
Tichner, Józef. *Nadzieja mimo wszystko* [Hope – No Matter What]. Krakow: Znak, 2020.
Wydick, Bruce. *Shrewd Samaritan: Faith, Economics, and the Road to Loving Our Global Neighbor*. Nashville: W Publishing, 2019.

Stefan Höschele

On Loving Neighbours and Others: Effective Altruism and Christian Ethics in a Relational Models Theory Perspective

Abstract

Christian traditions have tended to focus on "neighbours" as the primary target of altruistic actions, while the starting point of effective altruist reasoning is that spatial or personal nearness and distance should not count as a criterion of moral duty. This paper builds on insights from Relational Models Theory, a metatheory of human relationships that distinguishes four major models of human interactional patterns and concomitant moral ideologies. These four "relational models" shape a wide range of cultural features such as modes of decision making and types of group consciousness, influence value preferences, and thus affect inclinations to peculiar forms of ethical argument. Hence Relational Models Theory sheds light on both effective altruist arguments and characteristically Christian moral sentiments and shows the limitations and potential of both. Christian ethics as well as effective altruism are demanding and attractive, can learn from each other and can steer philanthropy to appropriate levels of reflection and action.

1. Altruism, Relationships, and Relational Models Theory

In his captivating book *Factfulness* (published in 2018),[1] the late Hans Rosling, a Swedish professor of international health, mentions an incident during his work in Mozambique as a medical doctor in the early 1980s. At the time, he was the only university-trained physician serving 300,000 people, and one day a European colleague who visited him was "very upset" about the oral dehydration treatment that Rosling had prescribed to an infant. The colleague insisted that the baby be given a more effective intravenous drip[2] – which, however, would take the doctor more than half an hour's additional work – and argued, "You must do everything you can for every patient who presents at the hospital." Rosling retorted, "It is unethical to spend all my time and resources trying to save those who come here. I can save more children if I improve the services outside the hospital. I am responsible for *all* the child deaths in this district: the deaths I do not see just as much as the deaths in front of my eyes."[3] While the colleague remained unconvinced, Rosling concludes at the end of the story,

1 Rosling et al, *Factfulness*.
2 Ibid., 125.
3 Ibid., 126.

> "Paying too much attention to the individual visible victim […] can lead us to spend all our resources on a fraction of the problem, and therefore save many fewer lives. […] It is hard for people to talk about resources when it comes to saving lives, or prolonging or improving them. Doing so is often taken for heartlessness. Yet so long as resources are not infinite – and they never are – it is the most compassionate thing to do."[4]

Effective altruism (EA) did not exist as a concept in the 1980s, and Rosling's book – which suggests that the world is in a much better state than most people think – does not contain any reference to the EA movement, its protagonists, and the major elements of its discourse. However, because of his passionate evidence orientation, I am sure that EA thought leaders would recognise him as a close ally. But what about Christians? Does his reasoning resonate with their maxim that one ought to love *neighbours*? Was the fellow doctor really wrong – or did he actually demonstrate the kind of compassion that is necessary in relating to those who live around you?

On a more general level, this question addresses the issue of the role of relationships in ethics. While this issue is complex and actually conflates topics such as spatial, conjugal, family, friendship, peer group, "ethnic," and "national" association, it is clear that interpersonal relationships are key in most areas of moral reasoning. Yet precisely *what* importance they should have – or, rather, *which types of* relationships determine one's duty or define appropriate actions – is contested, and evidently, perspectives on this matter are strongly shaped by cultural habits and assumptions.

This paper utilises a metatheory from the social sciences, i.e. a theory that synthesises well-known models from fields such as cultural anthropology, psychology, and economics, to shed light on both EA arguments and characteristically Christian moral sentiments. This metatheory, Relational Models Theory (RMT), developed in the late 1980s by psychological anthropologist Alan P. Fiske, builds on several earlier social theories[5] as well as extensive field research in West Africa.[6] The theory starts from the observation that humans are fundamentally social and suggests that sociality

4 Ibid., 127–128.

5 Fiske, "The Four Elementary Forms of Sociality," 710–712, 717–723; and Fiske, "Four Modes of Constituting Relationships," 16–21. Fiske explains that the theory emerged from a synthesis of Jean Piaget's developmental psychology, Max Weber's sociology of authority, Karl Polanyi's economic anthropology, and Paul Ricoeur's symbols theory.

6 The results of his two-year fieldwork are found in Fiske, "Relativity within Moose ('Mossi') Culture, "180–204, and his book Fiske, *Structures of Social Life*, 231–368 (part IV).

expresses itself in exactly four "elementary forms", i.e. unconscious but universally operational models, which exist in all cultures brought forth by human beings.[7]

According to RMT, every human relationship and, therefore, all contact and interface between individuals and groups is structured according to at least one of these four. Each of them follows a different principle and is experienced and re-enacted successively by children and youth as they grow up: Communal Sharing (CS) arises from the *close psycho-physical bonds* of those who relate, Authority Ranking (AR) is marked by a *spatiotemporal hierarchy*, Equality Matching (EM) consists of *egalitarian procedures*, and Market Pricing (MP) is lived on the basis of somewhat more abstract notions of *proportionality viz. ratios*.[8] One crucial and convincing element in the theory is that the four models are not mere abstractions but describe modes of human interaction which are unavoidably learnt and serve as concrete action orientation in the first ten to twelve years of life.

RMT was scrutinised and validated in many ways in the 1990s[9] and has since been applied to a host of different research fields and academic disciplines.[10] Today this theory can be regarded as a firmly established meta-theory of the social sciences that is fertile in terms of generating research programmes. In the realms of theology and philosophy, RMT has made but

7 In a later study, Fiske defines the models from the perspective of "how people *constitute* social relationships": Fiske, "Four Modes of Constituting Relationships," 61–146.

8 The absence of any relationship (which includes rejection of other humans as humans or indifference about them) is called the "null relationship" in this theory; it is important to distinguish this non-relationship from hierarchical and contractual relationships, even when they are experienced as negative by persons involved.

9 See Haslam, "Research on the Relational Models," 28–33. Seven published studies examined and confirmed RMT with regard to the questions whether (1) there are only four forms of sociality, (2) if Fiske's characterization of these four is sound, and (3) whether the four models are categories and not dimensional continua.

10 For instance: personality dimensions, social errors, personality disorders, ethnography, studies of trust, values research, organizational behaviour, relationships between business companies, and political orientation. See, e.g., Caralis and Haslam, "Relational Tendencies," 397–402; Houde et al., "The Four Faces of Trust," 287–306; Biber et al, "Personal Values and Relational Models," 628; Earley, *Face, Harmony, and Social Structure,* 133–136; Sheppard and Tuchinsky, "Interfirm Relationships: A Grammar of Pairs," 331–373; Simpson and Laham, "Different Relational Models," 204–217. For further fields and a discussion of these findings, see Haslam, "Research on the Relational Models," 34–51.

little impact so far,[11] but it is extremely promising with regard to ethics in several ways, for each of the four models suggests a different moral logic. I also propose that Relational Models Theory can illuminate the key ideas of EA and of New Testament ethics from a social science point of view, and can therefore help us compare, contrast, and possibly reconcile some elements of EA and Christian approaches to moral reasoning.

The four relational models shape a wide range of cultural features such as modes of decision making and types of group consciousness (see Table 1); they also correlate to typical values, and thus directly affect inclinations to peculiar forms of ethical argument. While Fiske devotes some reflection to the links that RMT has with moral reasoning,[12] he does not engage with ethical theory as such and merely mentions in passing that "each type of social relationship depends on different virtues."[13] It is relevant that Fiske and one of his fellow researchers did develop a novel theory of violence based on the relational models[14] and made a contribution to the field of moral psychology.[15] However, a fuller exploration of the relational models with regard to theological or philosophical ethics is yet to be undertaken.

11 I used RMT in attempting to understand ecumenism – to develop a comprehensive theory of interchurch relations; see Höschele, "Interchurch Relations in Seventh-day Adventist History." So far I am not aware of other applications in the realm of theology. For reflections in the field of philosophy, see footnote 19.

12 Fiske, *Structures of Social Life*, 115–124 (the section "Morality, Legitimation, Law, and Ideology" of chapter 6, "Judgments").

13 Ibid., 123. He mentions Rawls, however, in the context of EM morality (p. 118) as the standard, which is justified, however, with a CS argument, "that actual differences in individual endowments must be regarded as communal assets that belong to the collectivity as a whole" and a contractual framework (=MP) as "legitimation of the principle of justice that maximizes the welfare of those who are worst off" (p. 119).

14 Fiske and Rai, *Virtuous Violence*.

15 Fiske and Rai, "Moral Psychology is Relationship Regulation," 57–75, and Rai, "Moral Psychology is Relationship Regulation."

Table 1: "Manifestations and features of four elementary relational models" according to Fiske[16]

Domains and Features	**Communal Sharing (CS)**	**Authority Ranking (AR)**	**Equality Matching (EM)**	**Market Pricing (MP)**
Characteristic mode of marking relationships	Enactive, kinaesthetic, sensorimotor rituals, especially commensal meals, communion, and blood sacrifice.	Spatiotemporal ordered arrays (e.g., who is in front, who comes first). Differences in magnitude (size of dwelling, personal space, plural pronouns for respect).	Concrete operations involving physical manipulations of tokens or persons so as to balance, match, synchronize, align, or place them in one-for-one correspondence.	Abstract symbolic representation (especially propositional language and arithmetic).
Constitution of groups	Sense of unity, solidarity, shared substance (e.g., "blood," kinship). One-for-all, all-for-one. *Gemeinschaft*, *mechanical solidarity*, *primary group*.	Followers of a charismatic or other leader. Hierarchical organization (e.g., military).	Equal-status peer groups. For example, car pool, cooperative, and rotating credit association.	Corporations, labour unions, stock markets and commodity associations. *Gesellschaft, organic solidarity*. Also, bureaucracy with regulations oriented to pragmatic efficiency: *rational-legal organization*.
Decision making	Group seeks consensus, unity, the sense of the group (e.g., Quaker meeting, Japanese groups).	By authoritative fiat or decree. Will of the leader is transmitted through the chain of command. Subordinates obey orders.	One-person, one-vote election. Everyone has equal say. Also rotating offices or lottery.	Market decides, governed by supply and demand or expected utilities. Also rational cost and benefit analysis.

16 Shortened and slightly adapted from Fiske, "The Four Elementary Forms of Sociality," 694–696. The original 3-page table lists further details which are less relevant for this study – reciprocal exchange, distributive justice, work, the significance of (material) things, orientations to land, significance of time, social influence, social identity and self, motivation, interpretation of misfortune, (violent) aggression and conflict, related natural selection mechanisms, and approximate age when children first externalize the model.

Domains and Features	Communal Sharing (CS)	Authority Ranking (AR)	Equality Matching (EM)	Market Pricing (MP)
Contribution	Everyone gives what they have, without keeping track of what individuals contribute. "What's mine is yours."	*Noblesse oblige:* superiors give beneficently, demonstrating their nobility and largesse. Subordinate recipients of gifts are honoured.	Each contributor matches each other's donations equally.	People assessed according to a fixed ratio or percentage (e.g., tithing, sales, or real estate taxes).
Moral judgment and ideology	Caring, kindness, altruism, selfless generosity. Protecting intimate personal relationships. *Traditional legitimation* in terms of inherent, essential nature or karma of group.	What supreme being commands is right. Obedience to will of superiors. Heteronomy, charismatic legitimation.	Fairness as strict equality, equal treatment, and balanced reciprocity.	Abstract, universal, rational principles based on the utilitarian criterion of the greatest good for the greatest number (ratio metric for assessing all costs and benefits). *Rational-legal legitimation.*
Features that the cultural implementation rules must specify	Who is "us" and who is "other," including how people acquire and lose corporate membership. What is shared. What kinds of restraint people must exercise and what excuses them from giving.	What are the criteria for rank. What dimensions mark precedence. In what domains may authority be exercised.	Who and what counts as equal. What procedures people use for matching and balancing. What are the appropriate delays before reciprocating.	What entities may be bought and sold (e.g., sex? drugs? votes? people?). What are the ratios of exchange? What counts as a cost or a benefit.
Corresponding scale type	Categorical or nominal.	Ordinal.	Interval.	Ratio.

For the purpose of this paper, it must suffice to point out that the relational models are not primarily what we call ideas, values, or motivations. They do *engender* all of these, for the four relational models are, first of all, learnt modes of human interaction, i.e. results of emulation, with concomitant values and motivations being *inherent* in wholistic relational settings, and ideas being derived from them in a secondary manner. Yet RMT as a general theory, located between social anthropology and psychology, evidently illuminates what Fiske calls human "moral commitments": it describes the normative force of culture upon those who belong to particular groups by

way of explaining everyday interpersonal behaviour and the expectations created by such behaviour.

The defining elements of each relational model are *distinct types of actions*: "consubstantial assimilation" for CS (as found in sharing food, body contact, sex, dance, nursing, or shared pain); "social physics" in AR (size, position in space, visible hierarchy); "concrete operations" (EM) of taking turns, reciprocating, balloting, or working in unison alongside; and the use of symbolic signs (MP) as found in writing, propositional language (e.g. contracts), exchange with money involved.[17] These actions imply, of course, value preferences or a sense of what is "right" or "appropriate." Thus, CS lends itself to values such as love and benevolence, AR to values derived from an authority orientation, EM to equality, and MP to individual freedom and individual responsibility. These values can also be translated into more abstract ideas, and go along with concomitant motivations for individuals to act in specific ways: to conform to consensus (CS), to obey to a supreme being, ruler, or norm (AR), to share the same amount of burden or privilege as everyone else (EM), or to decide according to a cost-benefit calculus (MP).[18] Thus, although RMT theorists insist that relationships and their performance are primary, concepts derived from them are clearly relevant to discussions of morality.

The little extant reasoning so far on the RMT-ethics nexus is neatly summarised by John Bolender, the primary philosophical interpreter of Fiske's theory so far. It is quoted here at length because Bolender presents the basic concepts of what is applicable to the discussion below:[19]

> "Each elementary model crucially enters into certain moral values. [1] An ethic of service to one's group is a form of Communal Sharing [CS]. It is an *altruistic ethic in some sense*, but bear in mind that all members of the group share a common identity. So, strictly speaking, it is not true altruism. [2] Authority Ranking [AR] informs an ethic of obedience to authority including *respect, honor, and loyalty*. Any questions of value remaining to be clarified are settled by the authority; subordinates are expected to follow the values thus dictated. [3] *Fairness and even distribution* are informed by Equality Matching [EM]. John Rawls' veil of ignorance exemplifies

17 Fiske, "Four Modes of Constituting Relationships," 64.

18 Thus, one can even differentiate typical modes of arriving at moral prescriptions. While AR relationships imply an unbending *must*, an MP relationship frames responsibilities as something a person *has committed to*; EM implies that someone will *insist on justice* in a specific sense, and members of CS groups will argue that *one naturally does such and such* in our group. Thus, what morality actually "is" or how it "functions" differs dramatically among the four relational frameworks.

19 See Bolender, *The Self-Organizing Social Mind*; and Bolender, *Digital Social Mind*.

> Equality Matching; a perspective in which one does not know which role one will play guarantees that one aim for equality. […] [4] Market Pricing [MP] informs *libertarian values* of freely entering into contracts and taking risks with the aim of increasing one's own utility or the utility of one's group. But this also includes suffering the losses when one's calculations prove incorrect. *Utilitarianism* is a somewhat counterintuitive attempt to extend this sort of morality to all sentient life, but is still recognizable as Market Pricing.
>
> It would be too simple, however, to say that there are only four sorts of values in RMT. In fact, combinations of models yield complex models, resulting in a potential infinity of complex values. […] This great variety of values leads to value conflicts most noticeably across cultures."[20]

While Bolender's reference to both libertarian values and utilitarianism as relating to Market Pricing may seem surprising, it demonstrates that building blocks of different moral theories can be derived from the same relational model, even if the ultimate emphasis in one theory can differ drastically from another. Moreover, Bolender's observation that *combinations* of relational models are characteristic of complex moral reasoning is crucial in evaluating approaches to ethics. The history of ethics may have brought forth a few attempts at constructing systems that essentially reflect a single-issue approach (e.g. with "love" at the basis of certain versions of situation ethics, coming close to Communal Sharing, or divine command theory, an example of Authority Ranking). In most cases, however, theories of ethics emphasise two or more major values, thus merging or joining associated relational models in specific manners. Virtue ethics, for instance, has both a CS and an EM tendency, thus combining major in-group values with what is thought of as being ideals valid for all humans. Many deontological approaches to ethics have an authority (AR) bent,[21] but some add justice ingredients (corresponding to an EM orientation) while others include contractual reasoning (corresponding to what RMT calls the

20 Bolender, "Relational Models Theory." Numbers and abbreviations in brackets added by S.H.

21 A casuistic religious approach to ethics, in which the properly appointed authority regulates all questions of moral impact, might be the most distinctively AR system of ethics.

Market Pricing (MP) relational model).[22] All the defining factors that moral reasoning commonly appeals to[23] are linked to one or more than one of these relational models in specific ways.[24]

2. Effective Altruism in a Relational Models Theory Perspective

So how does EA relate to the four relational models and their concomitant moral emphases?[25] According to MacAskill, the EA movement "takes a scientific approach to doing good"; its ethos "consists of the honest and impartial attempt to work out what's best for the world, and a commitment to do what's best, whatever that turns out to be."[26] This definition (like others, which are generally quite similar)[27] contains pointers toward three of the relational models relevant for EA reasoning:

22 Popular morality often exhibits various mixtures – e.g. of Authority Ranking and Communal Sharing orientations: what is deemed "moral" is what is expected by the immediate in-group and its leading personalities and traditions, with little reference to more general values (such as justice *viz.* EM). But depending on the setting, it might also emphasize EM principles (e.g. in a project team) or mix them with space for individual agreements (MP) to make everyday life run more smoothly. Thanks to Stefan Riedener for pointing me to the need of clarification on this point.

23 Such as (a) the good and (b) evil, (c) nature, (d) law, (e) norms and rules, (f) duty, (g) values, (h) freedom and (i) responsibility, (j) means and (k) consequences, (l) conscience, (m) vice and (n) virtue, (o) community, (p) the nature of persons with their will, affects, and rationality, and (in religious contexts) (q) sin; I have listed these elements roughly according to chapters in Mühling, *Systematische Theologie,* 19.

24 In the context of this paper, time and space do not suffice to further develop these links and this theme in general. Yet exploring the bearing of Relational Models Theory upon ethical theory would certainly yield a major study of its own.

25 This paper was originally presented in summer 2019; thus, literature appearing around this time or later, including *The Precicipe* by Toby Ord and *Effective Altruism: Philosophical Issues* by Hilary Greaves and Theron Pummer (eds.), has not been consulted for the following discussion. Since the most basic structure of EA moral reasoning has not changed, this does not alter my general argument.

26 MacAskill, *Doing Good Better*, 15.

27 "1. Being open to all the possible ways to do good and pursuing the path with the biggest positive impact; 2. Using evidence to figure out how to do the most good; and 3. Choosing to make altruism a significant part of one's life." Singer and MacAskill, "Introduction", vi. Singer also defines that EA is "based on a very simple idea: we should do the most good we can." *Most Good*, vii. MacAskill formulates that EA aims at making "the world as good a place as it can possibly be." MacAskill, "What is Effective Altruism?," 3.

(a) The *Market Pricing* relationship (which might have better been called "market exchange" relationship) with its rational ("honest and impartial") cost-benefit analysis, evidence-based ("scientific") utilitarian ethos and trade logic forms the *backbone of the entire system*. A maximum of happiness viz. reduction of suffering – to be measured as thoroughly as possible[28] – can be, and should be, "bought" by "investing" means at one's disposal at the lowest price possible so that resources can be used to their maximum.

(b) The *Communal Sharing* relational model is of no lesser importance for EA theorists. Their uniform insistence on the need of focus upon "the most good" and the very term "altruism," which implies attention to the welfare of others, mirrors the "flesh" of EA morality: improving the overall well-being of individuals, and doing so without asking about the personal cost involved at the very outset.

EA – as its two constituent terms imply – is, therefore, essentially a hybrid: CS values promoted with an MP mind-set. The result of this bond is the creation of an ethos that *expands* Communal Sharing principles of caring, mutual responsibility, and free access to resources to a realm far beyond their "natural" habitat (i.e. kin, close friends, and imagined communities of common descent). The genius of EA is, therefore, to claim universal validity for moral standards that typically apply only to small-scale groups. At the same time, EA *tames* Market Pricing, for its libertarian and contract-based logic is limited by every sentient being's suffering and potential but not-yet-fulfilled happiness. Thus, MP norms, which are taken for granted, are essentially *at the service of Communal Sharing values*, which are viewed as being in need of realisation. This peculiar CS-MP hybrid implies, at the same time…

(c) the *absence, even rejection* of moral authorities. Traditional norms, religious writings – even wise words of leading persons of the EA movement itself such as Singer and MacAskill – do not play a normative role. The *Authority Ranking* principle is ignored; what is good, "whatever that turns out to be," is to be found on the basis of utilitarian calculus, and

28 There is no need here to present and discuss MacAskill's use (*Doing Good Better*, 39–45) of the QALY as a measurement, but in all the extant EA literature, measuring and comparing benefits on scales and in various units (currency, life years, etc.) is an important aspect of argumentation. See (in addition to the books mentioned above) Todd, *80,000 Hours*; and the predecessor book to the EA movement, Singer, *Life You Can Save*. See also Cooney, *Be Great at Doing Good*. While Cooney does not self-identify as an effective altruist in his book and does not mention EA, his arguments are very similar, and William MacAskill recommends the book on the "praise" pages that precede the title page.

the "scientific" component implies – at least for the founders of EA – a decidedly secular, non-theistic rationality. EA may best be understood as being thoroughly opposed to systems that derive norms from traditional morality.

(d) And what about Equality Matching? EM is the kind of "fairness" relationship which children learn, e.g. by distributing or receiving exactly the same number of candy. Effective altruists do not seem to be overly burdened with notions of impartiality expressed in equal distribution as long as this does not directly touch on the amount of suffering in sentient beings. While they emphasise that everyone's suffering counts equally (a typical EM-related argument),[29] thus directly challenging the archetypal Communal Sharing limitation to blood bonds, the Western preoccupation with democratic values such as "justice, freedom, equality, and knowledge"[30] is mitigated by declaring these to be ancillary to the goal of attaining more happiness and less suffering.[31] Thus the strictly egalitarian Equality Matching morality is neither viewed as absolute nor as primary; all in all, the spokespersons of EA appear to be positive in theory about this orientation but remain somewhat indifferent because of their exceptionally strong insistence upon CS values, even though the practical outcome of their activities will mostly lead to more justice as well.

3. New Testament Ethics and the Relational Models

Summarising the main aspects of Christian ethics in a few paragraphs is much more challenging than doing so for EA: the EA movement has produced fewer than ten major books so far while discussions on Christian ethics may fill entire libraries. Nevertheless, the grid provided by Relational Models Theory can help identify the main emphases of New Testament writers[32] in a way that these can be contrasted with EA emphases. Since biblical emphases have remained foundational for later Christian theological-ethical reasoning, and as this paper cannot possibly do justice to the

29 Singer asks, "Does everyone's suffering count equally?," and answers that the EA movement affirms this with regard to human suffering; see Singer, *Most Good*, 7.

30 Ibid., 146. "Justice" and "equality" are *the* typical EM values; democratic values at large are a blend of EM and MP values (the latter implying a rational, cost-benefit perspective).

31 Singer, *Most Good*, 7.

32 For a helpful and differentiated introduction to the complexity in the study of New Testament ethics viz. morality, see Cosgrove, "New Testament Ethics," 549–552.

variety of approaches to moral theology, the following discussion will be limited to New Testament ethics.[33]

While the feasibility of comprehensively reconstructing an ethic of the historical Jesus is debated, it is clear that its major emphases resulted from his proclamation of God's near kingdom: (1) consistent *agape* love of neighbour, even of enemies; (2) love of God, to be demonstrated in radically following divine principles (coupled with a somewhat critical view of traditional norms); and – to some extent – (3) justice springing from righteousness as a consequence of God's forgiveness.[34] The writers of the synoptic gospels accentuate (4) discipleship, (5) a critical view of striving for material wealth arising from a concern for the poor and their welfare, and (6) some degree of a normative view of the Christian life ("parenesis"), albeit without any legalistic outlook.[35] The epistles, finally, stress (7) the origin of all good works in the grace of God (especially Paul), (8) virtues to be sought (and corresponding vices to be avoided), and (9) the responsibility of the believer in the context of the family, the local congregation, the network of Christian churches, and the world at large.[36]

This overall picture of New Testament ethics points to three of the RMT models as being pertinent to various degrees and in different ways.

(a) *Communal Sharing* is evidently the key to understanding the Christian ethos; it is reflected in *agape* love (1), concern for the poor (5), the importance of forgiveness (3) and grace (7), and the responsibility of believers for others (9). Both the latter and Jesus' extension of love to enemies – as well as his nearly boundless re-definition of "neighbours" (by including the despised Samaritans)[37] point to what may be called a *modified*

33 The Old Testament does, of course, play an important role as a *background* to New Testament ethics, and its ethical content has been used in divergent ways (and has, often, been largely ignored) in Christian ethics. Suffice it to say that the Decalogue is the only major OT text with a moral relevance that has been of prominence in most confessional traditions.

34 For references and a summarizing account see McFee, "Values, Value Judgments," 803–805; for a comprehensive discussion see Schrage, *Ethik des Neuen Testaments*, 20–99 (on Jesus) and 117–133 (treating Matthew and Luke with discussions on righteousness, justice, and the relationship between different groups of society).

35 Schrage, *Ethik des Neuen Testaments*, 113–133.

36 Ibid., 134–265; and (a very succinct summary) Merk, "Ethics – NT Ethics," 147–148.

37 See the parable of the Good Samaritan (Luke 10:25–37). The antagonism between the Jews of Jesus' time and the Samaritans was due to religious quarrels that included the criticism of Samaritan "syncretism" but actually revolved around a number of intra-Israelite theological disputes (such as the status of various writings as Holy Scripture and the proper place of temple worship).

CS model: it is fundamentally open and thus categorically transcends the emblematic kin-related Communal Sharing logic.

(b) These CS values are promoted, and rest upon, the basis of a system of belief that combines the Creator God with the conviction that he acts powerfully in history and in the life of believers. He is, therefore, to be loved by living according to his will, viz. Christ's exemplification of the same (2, 4). In other words, the religious message proclaimed by the early Christians, and the moral system they lived within, (6) was based upon the highest authority that humans could imagine. With such an *Authority Ranking* relationship to the one and only God, the ethical teachings thus promoted gained the sanction that was needed in the economy of competing systems. At the same time, the lack of a legalistic bent in the New Testament writings, and the emphasis on following Jesus as an individual (and related notions of conscience) as opposed to traditional authorities (2, 6) imply a moderation of the AR impact on morality.[38]

(c) Except for the potent *agape* concept, early Christian principles of the moral life did not necessarily differ much from those in the Hellenistic environment. Especially in the later New Testament writings, lists of virtues to be sought by everyone alike (8) and the so-called *Haustafeln* (domestic codes)[39] closely resemble contemporary texts of the same genre. Nevertheless, New Testament virtue ethics also extols notions of justice (3) inherited from the Old Testament *zedaqâ* concept and translating into an incipient egalitarianism. Thus, an *Equality Matching*-related ethical perspective characterises at least part of the Christian ethos; while it was not fully developed in the framework of the hierarchical societies in which early Christianity operated, some degree of egalitarian treatment was to be extended to women, slaves, and foreigners as well (9).

38 Later developments in the Christian Church such as the establishment of canon law and the emergence of traditionalist movements and groups with a sectarian bent opened the door for casuistic and legalistic approaches, which drew its legitimacy from a predominant authority orientation.

39 Mainly Ephesians 5:22–6:5 and Colossians 3:18–4:1; cf. also 1 Peter 2:13–3:7 and sections from the pastoral letters such as 1 Timothy 2:8–15, 5:1–8, Titus 2:1–10.

(d) The fact that the *Market Pricing* relational model is almost absent in New Testament texts of ethical importance[40] warrants considerably more discussion than can be included in this paper. This absence may be interpreted in different ways, ranging from (1) active rejection through (2) cautious counsel against "calculating" approaches, (3) a neutral stance, or (4) a valuable option little known in NT times to (5) a type of assessing moral issues that is indirectly supported. On each of these options, only a short sketch will be presented here.

(1) Active rejection of MP approaches? *Agape* does not calculate; it represents love without thinking of the consequences (a "pure CS" attitude); therefore, the New Testament speaks *against MP*; it actually prohibits the weighing of good deeds, which implies that Christians must *reject or even fight this approach*.

(2) The New Testament *discourages* a cold "calculus" approach and thus indirectly counsels against asking for "effectiveness" in comparing actions. While such an orientation may be *tolerable in some instances* (e.g. when governments or businesses make decisions for their realms)*, it should not characterise Christian ethics*.

(3) In the absence of the public discourses that develop with the help of modern media, MP approaches to morality were not imaginable at the time. The idea of a maximisation of effects was thinkable in business, and there it was deemed problematic.[41] Thus cost-benefit analysis in the realm of ethics was *unknown but not* "*forbidden*."

(4) A "calculatory" approach to ethics was theoretically imaginable, but not common in the environment. Science, measurements, etc. did not appeal to people's moral "common sense" at the time; at the same time, it would have been supported had the New Testament been written in a society where such an orientation played a stronger role. Thus, an MP-type morality may have been, in NT times, a *blind spot but an acceptable option* depending on the issue – especially regarding matters which were not explicitly addressed in the New Testament.

40 One possible example is the parable of the Shrewd Manager in Luke 16:1–13. Since it is fraught with various exegetical challenges and appears to belong to the group of parables in which Jesus uses paradoxical or shocking examples to exemplify his key point (here, that "you cannot serve both God and Money", v. 13), it would be far-fetched to categorize the story with Market Pricing.

41 See the woes to the rich in James 5:1–6 and discussions of cultural perspectives on wealth in New Testament times in Malina, *New Testament World*, chapter 3.

(5) New Testament narratives, instructions, and parenetical sections are not to be construed to imply a set of "Christian laws." Rather, they illustrate a reasonable response to challenges of the time based on the gospel and on what was considered appropriate in a society that was very different from today's. *Contextually applied reasonableness* is to be supported even today; effectiveness (i.e., an MP orientation) is, in today's world, a major criterion in ethical reflections on impact, particularly with regard to large-scale issues.

Whatever the reason for an absence of distinctly MP-patterned approaches to morality in the New Testament, on the basis of the emphases outlined above (CS, AR, and some EM) it is clear Christians have traditionally tended to ignore cost-benefit calculations in their thinking on morality.

All in all, New Testament ethics consists in a radical support for Communal Sharing values and practices to be extended both to believers and to humans in general. The ideological basis for this ethos, divine authority, was the strongest validation possible at the time; the backbone of the Christian message and praxis of agape love, therefore, was an Authority Ranking framework, which fit in well with the common hierarchical setup of society at the time. While Market Pricing patterns of moral argument were evidently missing, egalitarian (Equality Matching) elements did play a certain role in the burgeoning early Christian congregations, although unsurprisingly concepts of justice derived from today's strictly democratic perspective had not yet developed.

4. On Loving Neighbours and Others

EA and the New Testament foundations for Christian ethics largely agree on the key element that characterises moral actions: the kind of love that values the Other as much as the own person.[42] While emphases differ, the overall perspective is remarkably close: altruistic action – the characteristic element of Communal Sharing or *agape* – is *not only* to be directed to one's family, friends, and those living in proximity, but to everyone. Jesus'

42 Probably Christians would go further in theory and promote selflessness even when this leads to a lack of resources. What the EA and Christian moral perspective agree upon is that there is a duty to the Other at least once one's own basic needs are fulfilled (cf. the biblical demand to love one's neighbour "as yourself").

widening of the "neighbour" idea[43] and the EA insistence upon regarding all humans as counting the same as potential recipients of support in the reduction of suffering[44] imply the same altruistic universalism. In the following paragraphs, I will point to the key difference between EA and Christian ethics, mention two issues on which Christians would question EA assumptions, and suggest what Christians can learn from EA.

The major difference is certainly the *rationale* given, and the *mode of application* of this ethical core. Both New Testament Christianity and EA claim far-reaching validity; what differs is the figures of thought and associated relational model preference that are appealed to. The MP relationship, based on ratios of exchange and abstract logic, is prevalent in the public sphere of modern societies and appeals to those who derive their identity or ideology from science. EA, therefore, seems "logical," but is in actual fact clearly contextually embedded. This is true as well for New Testament reasoning, which appealed to the highest authority that could be imagined at the epoch – the one and only God. Thus, even this difference between EA and Christian ethics contains a similarity in one respect: authority has shifted from the divine to the empirical, "scientific" realm, but this shift reflects the respective public sentiment of the time.

Yet even the empirical realm, which EA proponents claim as their turf, is not entirely value-neutral; the question of whether suffering and happiness, which utilitarians focus on and typically believe to be measurable, need to be the anchor points of ethics is *not* a matter of science but of moral philosophy. Actually, the very point of suffering can illustrate the difference between EA and New Testament authors in a precise manner: the apostle Peter argues that suffering is at times positive,[45] and other New Testament characters also support the idea that affliction is not something

43 The parable of the Good Samaritan widens the "neighbour" concept to one's traditionally *hated* neighbours viz. neighbouring "nations," and the Sermon on the Mount with its emphasis on love of the enemy (Matthew 5:43–48) essentially applies love of neighbour to *all* humans, including persecutors and colonial rulers: the passage that precedes the "love of enemy" section recounts a typical scene of Roman colonial power, in which a soldier forces a Jew to walk a mile with him (mostly to carry items) – whence the proverbial "walking of the second mile."

44 Christians traditionally disagree with widening this perspective to all sentient species, as Singer and many effective altruists do, even though eminent Christian ethicists now stress animal rights as well, and a few Christian minority traditions (such as my own Seventh-day Adventist tradition) include a strong emphasis on vegetarianism.

45 1 Peter 2:20–21 (with an *imitatio Christi* rationale); 3:13–17; 4:1; 4:12–19; cf. 5:1; see also Jesus' words in Matthew 5:10–11.

to be avoided.[46] This is due, to some extent, to the fact that Christianity originally sprang from an apocalyptic orientation, which focused on the soon-to-be established kingdom of God. Thus, the first Christians did not intend to make a major impact on society, and even when they did, Christian ethics was hardly ever construed as providing a foundation for living in a world of global opportunities. At the same time, this inherited "interim" perspective on Christian existence also implies that happiness is possible in spite of suffering.

Moreover, effective altruists may do well to listen to perspectives (which are common among Christians) that suggest that a focus on the maximisation of "impact" may at least sometimes be problematic.[47] The EA focus on measuring may result in support of short-term success in some fields – but it may simply not be possible to measure long-term results of the same interventions. Vigorous debates on the impact of "development" projects – with some theorists even rejecting any outside intervention – demonstrate that the last word on what is good, better, and best has not been spoken yet. Christians have accumulated much experience with good intentions leading to negative outcome, especially when churches were in a much more powerful position due to the Christendom setup; it would be a pity for EA activists to repeat mistakes that can be avoided (which is, of course, also true for the paternalism associated with some Christian missionary and development interventions).

A crucial aspect which Christian ethicists – and Christians in general – can learn from EA is the movement of further widening the horizon far beyond one's immediate environment. As the development of Christian Social Ethics from the early 20th century onward has shown, a general compatibility of biblical moral thinking with the challenges of modernising societies exists. Thus, ethics based on the New Testament can also take up issues that arise from the opportunities and complexities of a globalised world – which, essentially, means stretching moral concern beyond the realms of immediacy (CS), rank (AR), and mutuality (EM) towards all those who are affected by particular human interventions (MP). The fact that "neighbours" in the 21st century world are no longer only the visible or physical neighbours is, of course, a considerable increase of complexity

46 See, e.g., Acts 4:40–41 (the apostles rejoice for being "worthy" of suffering for Christ's sake) and Philippians 1:29.

47 For this entire discussion (and a somewhat misguided critique of the EA movement, which focuses almost exclusively on Peter Singer's utilitarian philosophical basis), see Wydick, *Shrewd Samaritan*, 91–101 (chapter 6).

(which may at times challenge traditional Christian practices of *diakonia*). If Jesus were to re-tell the story of the Good Samaritan today, he would possibly use a more globalised narrative.

A second major insight that Christians can receive from EA theorists is a significantly increased effectiveness orientation. With their myriad institutions and organisations in social work, medical service, and education, Christian churches have been aware of the requirements of good management, cost-effectiveness, and the like, for a long time. Thus, Rosling's concern in the introduction is not foreign to Christian practitioners of aid, health, and teaching. However, because of the lack of effectiveness criteria in the world where Christianity arose and the resultant paucity of cost-benefit arguments in New Testament ethics, it would be easy to overlook the moral importance to Christian decision making in general and in particular in a world that is shaped by MP factors to a much larger degree.

A third important inspiration for Christians coming from EA is the focus on neglected issues. By calculating the relative neglectedness of problems (a typical MP approach), EA activists function in a way that systemically makes sense wherever attention is not sufficiently directed towards a problematic area within a larger whole. Christians tend to derive areas of activism from the Scriptures, tradition and areas where the emotions are touched; EA thinking can contribute to widening attention to other even more relevant fields.

Both the central ideas of EA and the core of Christian ethics are demanding and, at the same time, attractive moral philosophy. From an outsider's point of view, they may nevertheless seem trivial to some extent in theory: that you must (AR) help others (CS) who are in real need regardless of who they are (New Testament) and that it makes a difference how many (MP) needy people you help (effective altruism) are almost common-sense statements as long as they do not require much sacrifice. Yet both the Christian emphasis on and the EA version of altruism demand more than an attitude or a philosophical commitment: they are about action done from conviction and essentially regard all of humanity as one community (CS). Both teach that decisions should not be made for acting individuals to feel good (at least not primarily) but to actually make the world a better place. In this regard, committed effective altruists and Christians willing to live their faith in a tangible manner can learn much from each other, challenge each other on blind spots, and together steer philanthropy, charitable giving, and humanitarian service to appropriate levels of reflection and action.

References

Biber, Pascal, Hupfeld, Jörg, and Meier, Laurenz L. "Personal Values and Relational Models." *European Journal of Personality* 22, no. 7 (2008): 609–628.

Bolender, John. *The Self-Organizing Social Mind*. Cambridge: MIT Press, 2010.

———. *Digital Social Mind.* Exeter: Imprint Academic, 2011.

———. "Relational Models Theory." In *Internet Encyclopedia of Philosophy*, n.d. Accessed 15 July 2019. https://www.iep.utm.edu/r-models.

Caralis, Dionyssios and Haslam, Nick. "Relational Tendencies Associated with Broad Personality Dimensions." *Psychology and Psychotherapy: Theory, Research and Practice* 77, no. 3 (2004): 397–402.

Cooney, Nick. *How to Be Great at Doing Good: Why Results Are What Count and How Smart Charity Can Change the World*. Hoboken: Wiley, 2015.

Cosgrove, Charles H. "New Testament Ethics." In *Dictionary of Scripture and Ethics*, ed. Joel B. Green. Grand Rapids: Baker Academic, 2011.

Earley, P. Christopher. *Face, Harmony, and Social Structure: An Analysis of Organizational Behavior Across Cultures*. New York: Oxford University Press, 1997.

Fiske, Alan P. "Relativity within Moose ('Mossi') Culture: Four Incommensurable Models for Social Relationships." *Ethos* 18, no. 2 (1990): 180–204.

———. *Structures of Social Life: The Four Elementary Forms of Human Relations*. New York: Free Press, 1991.

———. "The Four Elementary Forms of Sociality: Framework for a Unified Theory of Social Relations." *Psychological Review* 99, no. 4 (1992): 689–723.

———. "Four Modes of Constituting Relationships: Consubstantial Assimilation [CS]; Space, Magnitude, Time and Force [AR]; Concrete Procedures [EM]; Abstract Symbolism [MP]." In *Relational Models Theory: A Contemporary Overview*, ed. Nick Haslam, 61–146. Mahwah: Erlbaum, 2004.

———. "Relational Models Theory 2.0." In *Relational Models Theory: A Contemporary Overview*, ed. Nick Haslam, 3–26. Mahwah: Erlbaum, 2004.

——— and Tage Shakti Rai. "Moral Psychology is Relationship Regulation: Moral Motives for Unity, Hierarchy, Equality, and Proportionality." *Psychological Review* 118, no. 1 (2011): 57–75.

——— and Tage Shakti Rai. *Virtuous Violence: Hurting and Killing to Create, Sustain, End, and Honor Social Relationships*. Cambridge: Cambridge University Press, 2014.

Haslam, Nick. "Research on the Relational Models: An Overview." In *Relational Models Theory: A Contemporary Overview*, ed. Nick Haslam, 27–57. Mahwah: Erlbaum, 2004.

Höschele, Stefan. "Interchurch Relations in Seventh-day Adventist History: A Study in Ecumenics." Habilitation thesis, Charles University, Prague, 2017.

Houde, Leah D. et al. "The Four Faces of Trust: An Empirical Study of the Nature of Trust in Relational Forms." In *Relational Models Theory: A Contemporary Overview*, ed. Nick Haslam, 287–306. Mahwah: Erlbaum, 2004.

MacAskill, William. *Doing Good Better: Effective Altruism and a Radical New Way to Make a Difference*. London: Guardian, 2016.

———. "What is Effective Altruism?" In *The Effective Altruism Handbook*, ed. Ryan Carey, 3–11. N.p.: n.p, 2015.

Malina, Bruce. *The New Testament World: Insights from Cultural Anthropology*. Louisville: Westminster John Knox, 2001.

McFee, Daniel E. "Values, Value Judgments." In *Dictionary of Scripture and Ethics*, ed. Joel B. Green, 803–805. Grand Rapids: Baker Academic, 2011.

Merk, Otto. "Ethics – NT Ethics." In *The Encyclopedia of Christianity*, vol. 2, ed. Erwin Fahlbusch et al. Grand Rapids: Eerdmans. Leiden: Brill, 2001.

Mühling, Markus. *Systematische Theologie: Ethik. Eine christliche Theorie vorzuziehenden Handelns*. Göttingen: Vandenhoeck & Ruprecht, 2012.

Rai, Tage Shakti. "Moral Psychology is Relationship Regulation." Ph.D. Dissertation, University of California, Los Angeles, 2012.

Rosling, Hans with Rosling, Ola and Rosling Rönnlund, Anna. *Factfulness: Ten Reasons We're Wrong about the World – and Why Things are Better Than You Think*. New York: Flatiron Books, 2018.

Schrage, Wolfgang. *Ethik des Neuen Testaments.* Berlin: Evangelische Verlagsanstalt, 1985.

Sheppard, Blair H. and Tuchinsky, Maria. "Interfirm Relationships: A Grammar of Pairs." In *Research in Organizational Behavior* 18, ed. Barry M. Staw and L. L. Cummings, 331–373. Greenwich: JAI Press, 1996.

Simpson, Ain and Laham, Simon M. "Different Relational Models Underlie Prototypical Left and Right Positions on Social Issues." *European Journal of Social Psychology* 45, no. 2 (2015): 204–217.

Singer, Peter. *The Life You Can Save: How to Play Your Part in Ending World Poverty*. London: Picador, 2009.

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

——— and MacAskill, William. "Introduction." In *The Effective Altruism Handbook*, ed. Ryan Carey. N.p.: n.p, 2015.

Todd, Benjamin and the 80,000 Hours Team. *80,000 Hours: Find a Fulfilling Career That Does Good*. Oxford: Centre for Effective Altruism, 2016.

Wydick, Bruce. *Shrewd Samaritan: Faith, Economics, and the Road to Loving Our Global Neighbor*. Nashville: Nelson, 2019.

Kathryn Muyskens

Asceticism as Activism: Effective Altruism's Neglected Other Half

Abstract

What I seek to do in this paper is to reemphasize what I see as the forgotten or neglected other half of the effective altruist equation. Effective altruists need to take seriously the ways in which their actions contribute to systemic inequality and structural violence. Charitable donation is not enough to create a paradigm shift or stop systemic injustice. In tackling systemic injustice, the ascetic response may allow effective altruists to attack the roots of the problem more directly. Further, the cost-benefit analysis and randomized controlled trials favored by the movement can produce distinctly biased perceptions that leave effective altruists blind to the political dimensions of many types of harm. Balancing ascetic approaches to combating suffering may temper the overzealous focus on cost-effective charities and make room for the support of the causes this narrow focus excludes. Ultimately, this paper defends the basic tenets of effective altruism: that the capable have a duty to reduce suffering in the world and that we should apply our powers of reason in order to make our labors maximally effective.

1. Introduction

Effective altruism (EA) is characterised by several key premises. First and foremost is the idea that we ought to make the world a better place and that we should apply our powers of reason to the efforts we make in that regard in order to make them maximally effective. People living in the Global North often have more money than needed to sustain themselves. Because of the current economic dynamics of the world, this means even small donations of what amounts to spare change can have significant effects on the lives of impoverished people in the Global South. These facts have led effective altruists to focus their attention on making donations to charities that have been evaluated for cost-effectiveness and which target the world's poorest people. For this reason, EA has a unique relationship to money and economics. This relationship produces a distinct skew in the way effective altruists approach the issues they target, and not always for the better. In particular, the idea of *earning to give* has gained some popularity, leading some to seek high earning careers with little concern for the ethical impact of the *nature* of these careers in themselves.

Much of what I argue for in this paper is not revolutionary or unique. In fact, asceticism is old news, so to speak, as can be seen in the longstanding practice of fasting and other forms of austerity in many of the world's oldest faiths. Yet, in light of all the energy around EA as a young movement

and the excitement, it has generated about charitable giving, these points seem worth mentioning afresh. Thus, what I seek to do in this paper is to *reemphasise* what I see as the forgotten or neglected other half of the effective altruist equation: namely, the role of self-denial and restraint. Effective altruists need to take seriously how their actions contribute to systemic inequality and structural violence. Donation is not enough to create a paradigm shift or to stop systemic injustice. In addition to considering costs and benefits as they relate to charitable donations, effective altruists should be mindful consumers and mobilise their *inaction* for good effects as well. In order to maximise our positive influence on the world, we must mobilise not only our activist instincts but also our *ascetic* ones. We must be able to recognise not only the types of actions that can best effect positive change, but also the types of *inaction* that can stall or reverse the root causes of suffering. Considering both our actions and inactions as representing equal opportunity for effecting change in the world gives us a more well-rounded picture of the moral landscape. This positions my article in favour of EA's core premises, but against the sub-trend of earning to give. Ultimately, the point I hope to make is that the reinvigoration of interest in charitable giving spurred on by effective altruist's stands only to be enhanced if it is partnered with some of the wisdom and moderation of the much more ancient notion of asceticism.

To make my case, I will develop my line of argument in the following manner: I will start by going over some worries about giving (section 2), and then I will introduce and defend ascetic forms of activism, explaining how these methods can avoid or mitigate some of the worries connected with donation (section 3), next, I will develop this into my objections against earning to give (section 4), and finally I will answer a common worry about asceticism and EA more generally, namely the concern about "overdemandingness" (section 5).[1]

1 It is worth noting that the effective altruist movement has evolved since the writing of the initial version of this paper, which was published in 2017 in *Essays in Philosophy*. Since that time, EA has been explicit that it is not solely focused on donations or synonymous with "earning to give," see Todd, "Misconceptions About Effective Altruism." Still, recognition or open embrace of the ascetic side of the equation I advocate for here has yet to emerge.

2. Systemic Injustice and the Question of the Efficacy of Aid

Given the resources and wealth people in the developed countries of the Global North often enjoy, this entails a distinct burden to use this privilege for good. Similar to the idea that a doctor has special duties to help the sick because he or she has more knowledge and ability to help than a layperson, so we logically have a duty to give some of our wealth away to others who do not have it, when doing so would not present a significant burden to ourselves and would significantly improve the lives of those to whom we give. If the motivating force behind giving is not the "warm glow,"[2] as Peter Singer has been known to put it, but the chance to materially reduce suffering in the world, it seems only logical that we should be attentive to the effectiveness of our donations.

Attending to this effectiveness, however, can be a steeper challenge than it initially seems. As Andrew Kuper points out in his article, "More Than Charity," "nothing in the principle of aid or charity determines that the right action in any or all contexts is donation"[3]. To truly be effective in our altruism, we need to be able to recognise situations in which charitable giving is not the right answer, and where it may in fact cause harm. As Keith Horton notes in his article "Aid Agencies: The Epistemic Question," "Many contributors to this debate have apparently taken it that one may simply assume that the effects of the work such agencies do are overwhelmingly positive … however, one finds a number of concerns about such agencies and the work they do that put that assumption in serious doubt."[4]

Emily Clough takes the argument one step further in her article, "Effective Altruism's Political Blind Spot," in the *Boston Review*. Efficacy in aid is particularly thorny and difficult to determine. Not only is it difficult to get good data, but even the darling methods of randomised control trials (RCTs) that effective altruist organisations like GiveWell and Giving What We Can use produce distinctly skewed results.[5] She says,

> "While they are good at measuring the proximate effects of a program on its immediate target subjects, RCTs are bad at detecting any unintended effects of a program, especially those effects that fall outside the population or timeframe that the organisation or researchers had in mind. For example, an RCT might determine whether a bed net distribution program lowered the incidence of malaria among

2 Singer, *Most Good*, 5.
3 Kuper, "More Than Charity," 113.
4 Horton, "Aid Agencies: The Epistemic Question," 29.
5 Clough, "Effective Altruism's Political Blind Spot."

> its target population. But it would be less likely to capture whether the program unintentionally demobilised political pressures on the government to build a more effective malaria eradication program, one that would ultimately affect more people. RCTs thus potentially miss broader insights and side effects of a program beyond its target population."[6]

The critical point here is that a political dimension to poverty is ignored in the standard process effective altruists use to target their giving. Given this blind spot, the aid that effective altruists provide risks actually worsening some systemic problems, as Clough goes on to describe:

> "In the worst case, the presence of NGOs induces exit from the state sector. When relatively efficient, well-functioning NGOs enter a health or education market, for example, citizens in that market who are paying attention are likely to switch from government services to NGO services. The result is a disengagement of the most mobilised, discerning poor citizens from the state. These are the citizens most likely to have played a previous role in monitoring the quality of state services and advocating for improvements. Once they exit, the pressure on the government to maintain and improve services eases, and the quality of government provision is likely to fall."[7]

Iason Gabriel, in his article "Effective Altruism and its Critics," describes yet another scenario where EA's cost-effectiveness analysis can fall short. He recounts a hypothetical situation where a charity has the option to build a water sanitation system using either outside labour or members of the recipient community to construct it.[8] Since they are professionals, the outside labourers will surely build a better system that will be less likely to need repair in the near future, making them the more cost-effective investment.[9] However, as Gabriel points out, this leaves no room in the calculus for the community's emotional investment in the project if allowed to build the system themselves.[10] He says, "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 receiving something as a gift and bringing it into existence through one's effort."[11]

6 Ibid.
7 Ibid.
8 Gabriel, "Effective Altruism and its Critics," 9.
9 Ibid.
10 Ibid.
11 Ibid.

Steven G. Brown adds to the discussion in his article, "Supporting the Best Charities is Harder than it Seems," wherein he describes how GiveWell's methodology leads them to "to rule out effective organisations that would do great work with further funding."[12] Specifically, he argues,

> "GiveWell's methodology prioritises easily measurable metrics, and those outcomes that can also be researched and tracked in a particularly rigorous way. This leads to their most serious drawback: overlooking projects that should be a high priority, but are difficult to measure. When one takes a step back and asks what it would take to better a place that is not doing well, one will surely come across many difficult to measure answers. For example, it is striking that since it began giving ratings in 2008, GiveWell has only recommended a single international educational charity. Pratham is an organisation that runs an innovative and extraordinarily successful reading program called Read India that trains community volunteers to be teachers in villages all across the country. Citing difficulties in measuring the effectiveness of educational projects, GiveWell stopped recommending Pratham in 2011, and has never recommended another educational charity since. This is despite the fact that education is certainly one of the most important things a community must have if it is to rise out of poverty and not merely survive."[13]

Clearly, the narrow approach effective altruists have taken regarding charitable aid has significant drawbacks. Not only is its efficacy hard to determine, but the narrow focus rules out some areas which ought not be ignored. Given that I accept the first premises of EA, namely, that the capable ought to take significant steps to reduce suffering and that the actions they take for that purpose ought to be maximally effective, it now may seem unclear what kind of action a would-be effective altruist should take against this suffering. What other methods are available?

When the question of aid's effect is unclear, we may better serve those who are suffering with ascetic responses than with donation. It might be hard to control where donated money goes and what its effects are once it goes there. But if we identify the sources of harm, starving these sources of our tacit, unquestioned approval will eventually lead to system-wide change.

The causes of suffering can be difficult to identify and target for change through donation. This is in part because the Global South is created by the same economic system that produced the Global North, as Thomas Pogge describes in his article, "Responsibilities for Poverty Related Ill-Health." The global order is such that Western culture remains as dominant and influential as it was in the colonial era, and therefore, the "existing institutional order is implicated in the persistence of radical inequality [where the]

12 Brown, "Supporting the Best Charities," 242.
13 Ibid.

better-off impose a shared institutional order on the worse-off."[14] Donation alone will not be sufficient to remedy such system-wide structural violence. Against such sources of suffering, the ascetic approach is a very much needed and too often neglected complement to donating. As Kuper says, "Given the complex interdependence and economic and political perversities that characterise our shared world, the injunction 'first, do no harm' deserves at least equal consideration."[15]

Indeed, I think most effective altruists would not have a problem admitting this. When presented with the evidence that giving will not do good in one case, the effective altruist may simply redirect those potential funds toward a situation in which he or she can feel surer that it will be of benefit. Perhaps, the effective altruist will choose to donate to developing technologies to protect humanity from cataclysm due to global warming or to medical research aimed at lifesaving drugs or eradicating diseases. These are more concrete and technological sources of suffering and better suited, perhaps, to cost-effectiveness reasoning. Even so, I propose that this thinking leaves the effective altruist equation incomplete. Charitable giving is not our only opportunity to take altruistic action. Focusing solely on that misses an opportunity to expand the influence of our ethical lives.

3. Selective Asceticism as Activism

Asceticism has a long history, both within religion and apart from it. Religiously speaking, ascetic practice may take many forms, from vegetarianism (as it does for many Hindus), or the practice of voluntary celibacy, to the more extreme choice to live a life of seclusion (as with monks, whether Catholic or Buddhist). Ascetic attitudes even make an appearance in Plato's *Phaedo*, where Socrates explains the dangers of bodily pleasures and desires because of their potential to sway our values and make us believe falsehoods,[16] and one can find a similar refrain in Buddhism (a faith widely known for its ascetic tendencies) in the four noble truths, the second of which identifies the origin of suffering as desire. At its most basic, however, it is the practice of deliberate self-denial with the aim of acting more ethically or acquiring virtue.[17]

14 Pogge, "Responsibilities for Poverty Related Ill-Health," 71.
15 Kuper, "More Than Charity," 114.
16 Ebrey, "The Asceticism of the Phaedo."
17 Besong, "Virtue and Asceticism."

The word ascetic comes from the Greek "áskesis" meaning to exercise or train. Thus, asceticism is best understood as a discipline or practice, not merely the opposite of indulgence. Asceticism requires more than mere abstinence from some activity. It is not ascetic to refrain from eating chocolate if you never liked chocolate to begin with. Likewise, it is not ascetic to deny yourself the pleasure of a daily treat from Starbucks if your reason for doing so is simply to save up your money for a vacation next month. Asceticism is self-denial for a higher good. In religious contexts, this higher good may often be spiritually-oriented and focused on the individual believer – as in purifying one's own soul or bringing oneself closer to God. But in this paper, I hope to call attention to the "this-worldly" ethical potential of asceticism. Activists, and not just the spiritually inclined, ought to take note of the ethical potential asceticism has to offer.

Asceticism has also been met with its fair share of criticism over the centuries. Nietzsche criticised it as "hatred of the human;"[18] Bentham criticised it as a "cloak for tyranny."[19] These criticisms helped to paint it as "eccentric and uniquely religious," leading many in the post-Enlightenment to treat it with disregard.[20]

Yet, in more recent years, ascetic practices of one form or another have been gaining ground, as can be seen in the popularity of minimalist lifestyles, ethical veganism, and other trends. These modern incarnations may or may not be explicitly connected with the religious attitudes of the people who practice them. In addition to its resurgence in popular culture, the academic community is also waking up to the notion that asceticism may offer practical tools towards cultivating ethical behaviour (something which is supported by contemporary empirical research on self-control),[21] as well as significant insight into the ethics of consumer culture and other current debates in applied ethics.[22]

Thus, asceticism as a method of strengthening moral character as well as an additional avenue for ethical action deserves attention from the effective altruist community. EA has traditionally focused on the "active" methods for "doing" good, such that the movement has neglected the ways in which they ought also to avoid perpetuating harm or injustice. This is a gap that a dose of asceticism would go a long way towards filling.

18 Segerdahl, "Intellectual Asceticism."
19 Koh, "Bentham on Asceticism and Tyranny."
20 Besong, "Virtue and Asceticism."
21 See Besong, "Virtue and Asceticism."
22 See Maria Antonaccio, "Asceticism and the Ethics of Consumption."

EA's focus on money and donation amounts to a kind of tunnel vision. To give an example: within the attempts to address the global warming crisis, supporting the development of alternative fuels for cars or better electric cars is a perennial favourite. Yet, new forms of cars or fuel pale in comparison to the effects on climate change that could be produced from simply reducing the need for cars at all. Developing better, more efficient, public transport would have a much more drastic impact. Better still, we could redesign cities and transport to not rely as heavily on cars at all and adjust the culture to move away from seeing cars as basic needs (as they are often seen in the United States) and more as expensive luxuries. Individuals can propel change of this kind with investments or donations (the typical effective altruist approach), but can also practice this by *abstaining* from the use of cars, or planes, making the choice to take public transport or walk or bike to places they need to go, or travelling less in general.

Other ascetic methods of fighting climate change are simply becoming vegetarian or vegan since large amounts of greenhouse gases are produced through factory farming. In fact, a report by *World Watch* found that more than half of all greenhouse gases are produced from the animal agriculture industries.[23] And a study by researchers of the University of Oxford published in 2014 in the journal *Climatic Change* concluded that meat-eaters are responsible for almost twice as many dietary greenhouse-gas emissions per day as vegetarians and about two and a half times as many as vegans.[24] Another option for selective asceticism is to personally choose not to have children to reduce the burden of overpopulation. One need not take up every ascetic choice, but these options deserve at least as much recognition and attention as charitable donation.

I recognise this is not an entirely novel thing to say, as even Peter Singer in his book, *The Most Good You Can Do*, states that one of the most effective ways to decrease total suffering in the world is to stop eating meat.[25] That is why I see my argument in this paper as a reemphasis rather than a revolutionary suggestion. In fact, there are already efforts being made in this regard. We just tend not to hear about them in connection with EA. Movements like *conscious consumerism* seek to better inform customers of the ethical impact of their purchases.

23 Goodland and Anhang, "Livestock and Climate Change."
24 Scarborough et al., "Dietary greenhouse gas."
25 Singer, *Most Good*, 177.

In Waheed Houssain's article, "Is Ethical Consumerism an Impermissible Form of Vigilantism?" he gives a thorough evaluation of several types of ethical consumerism. Citizens can use their purchasing power in a variety of ways and for a variety of reasons. A conscious consumer may refrain from buying coffee produced by unfairly treated workers, or they may avoid products known to contribute to deforestation or the extinction of species in distant regions, what Houssain calls "negative ethical consumerism" because it involves refraining from buying something.[26] Conscious consumers can also take action to make the world a better place by purchasing products that have been recognised as fair trade or sustainable, what Houssain would call "positive ethical consumerism."[27]

If we accept that one of the engines of inequality and suffering in the world is rampant consumerism, then abstaining from consumerism in any way possible will help to effect change. Strategically abstaining from consumerism (both in the sense of abstaining from consuming goods that perpetuate harm or injustice, and in the sense of mitigating one's consumer appetites more generally) can attack the roots of these problems, especially when combined with advocacy and public pressure on companies that use unfair labour practices, cause deforestation, use child labour, and more.

This type of activism would be a kind of negative ethical consumerism aimed at creating systemic change, what Houssain dubs a kind of "proto-legislative" social change ethical consumerism (SCEC).[28] The use of such strategies "essentially creates arenas of informal democratic self-governance that operate below the level of formal democratic politics."[29] Houssain elaborates on the benefits of ethical consumer action:

> "In a large, complex, and technologically sophisticated society, citizens cannot make all of the rules necessary to direct market activity to desirable outcomes through the formal legislative and regulatory process. As things stand, when issues do not make it on to the formal democratic agenda, they are left to the unregulated market. But with proto-legislative SCEC, citizens can address issues that need attention but do not get on the formal agenda through informal self-regulation in secondary arenas."[30]

26 Houssain, "Ethical Consumerism Impermissible Form of Vigilantism?" 113.
27 Ibid.
28 Ibid.
29 Ibid, 132.
30 Ibid.

This type of activism can have many benefits, like helping to bring attention to issues that are "perpetually secondary" (like deforestation) and fostering a "legislative will" where citizens can organise themselves on issues that languish in formal politics.[31] He continues:

> "A closely related benefit is increased governability. Corporations are powerful social actors, with privileged access to political authorities, and they often oppose laws that would protect the rights and interests of weaker players in the market. But the orientation of corporations is closely connected with the orientation of consumers. If consumers are narrowly interested in price and quality, without regard for how a firm delivers these goods, then firms stand to profit from reduced protections for weaker players. But if consumers are sensitive to whether a firm respects the rights and interests of others, there will be less profit to be made in taking advantage of weaker players, and this in turn will make firms less hostile to regulatory efforts to protect these players."[32]

A final benefit of proto-legislative SCEC is that it can "expand the sphere of citizen engagement."[33] Someone who does not identify with any political party or have general views about economic policy may still find it quite natural to express his or her wishes for the common good through everyday purchases. Websites like EthicalConsumer.org provide an abundance of articles and data regarding the ethical influence of consumer choices with the aim of providing consumers the tools they need to make such decisions. Such sources of information may lead interested consumers to eventually take more active roles in democratic politics.[34]

Though the kind of proto-legislative SCEC Houssain describes encompasses both positive and negative forms of ethical consumerism, I have chosen to focus on the negative form for this article because it is often underexplored and underutilised. Lumping selective asceticism as one subcategory of ethical consumerism is misleading since by its nature it is anti-consumerism. Thus, strategies like this can benefit from having their own name. Asceticism as activism aimed at creating systemic social change should be seen as a form of EA in its own right and deserves attention as such.

To be clear, I am not arguing that charitable donation is *never* a suitable option for altruistic action. There are some cases where it can be both appropriate and effective. For instance, in combating deaths due to malaria,

31 Ibid.
32 Ibid, 133.
33 Ibid.
34 Ibid.

donating to researching and distributing treatment and to organisations that help prevent its spread may help to ultimately consign malaria to the dustbin of history. Yet, the zeal with which EA has approached charitable giving is likely inflated and inappropriate. Donation can be one tool among many in the effective altruist's toolbox, but we should be very careful and attentive to the situations in which it is applied and willing to recognise its limits.

I am also not advocating a life of complete disengagement from the world. The kind of asceticism I am proposing is selective, targeted to remove support from the unjust systems in the world that profit from the exploitation of others and the planet. Asceticism is not the only way to respond to these issues, but it is a tool that is too often neglected.

4. Against Earning to Give

As we have established, effective altruists are interested in (1) helping others and (2) doing so effectively. The earning to give sub-trend accepts that the best way to do so is donation to effective charities, and seeks to maximise the power to do so by seeking careers with the most financial potential. But as I have demonstrated, we have serious reason to question the faith effective altruists have placed in the benefits of donation – at least in some contexts. As we shall see next, there is reason to question the "earning" side of "earning to give" as well.

Even if we give away massive sums to good causes (and even if these causes are effective), if we are not conscious of the kinds of systems to which we are giving our life and our time, we may be missing important elements of the moral picture. To give a hypothetical comparison, if you make a living by slaughtering animals, but donate all your earnings above subsistence level to vegan causes … are you doing the most good you can?

Blanket approval for any job that allows you wealth and status is dangerous, no matter how altruistically one applies the wealth and status once achieved. Indeed, the effective altruist movement has strongly emphasised this very point in more recent years (an article on 80,000 hours' website by Benjamin Todd, "Is it Ever Okay to Take a Harmful Job in Order to Do More Good? An In-Depth Analysis," was published mere months after the publication of my original paper this chapter is based upon). Careers are not all morally neutral activities. Indeed, *some* careers may be neither helpful nor harmful to the world, but to an effective altruist interested in maximising one's positive influence, the ethical potential of all the hours we

spend working should not be ignored. This would seem to place the preference back on the jobs we usually think of as morally admirable – working directly for aid agencies, becoming an activist or advocate for the worse off, becoming a doctor, etc. There is an extra reason to be wary of the earning to give. High-paying jobs and high-status jobs often are high-paying and high-status because they are key positions in the very economic system that maintain the unequal status quo. As I have already established, this system itself is implicated in the structural violence that causes so much harm in the first place.

Ascetic strategies, on the other hand, like refusing to participate in forms of consumerism which perpetuate the divide between the Global North and Global South or giving up meat to reduce both greenhouse gases and prevent animal suffering can potentially transform society so that these causes of suffering eventually cease to exist. Asceticism is a form of activism in this way and can be a form of public protest and can help to establish and promote alternative ways of living. Donating to charities that can help children in impoverished communities is only half the solution; we must also stop buying the products that come from their exploitation. By recognising these other strategies, we may be able to enlist more people in the fight to reduce suffering in the world.

Some effective altruists will defend earning to give by arguing that if a person rejects a high-paying job in banking or finance on ethical grounds, this does nothing to prevent someone else from taking that job.[35] Further, the person who does take the job may be less likely to donate what the effective altruist would, and less likely to advocate for reform from within the system. To this, I will concede some ground. In the situation where an altruistic individual has the talent and opportunity, I would not automatically condemn the acceptance of such a career choice. Yet, I would also be keen to point out that donating all the money earned from such a job to effective charities does not go far enough, since it is insufficient to effect lasting systemic change. It is also highly risky, given the psychological effects of peer pressure and life-style creep one is likely to be exposed to within certain career circles.

The emphasis for the truly effective altruist interested in earning to give needs to be on the creation of systemic institutional change. Thus, I would argue that, advocacy for reform from within should be the primary target when making such a career choice for such individuals. If such advocacy is

35 MacAskill, "Replaceability, Career Choice, and Making a Difference."

effective, then the negative effects of the career choice would vanish, and so would my objections.

Yet, I think it is worth pointing out that careers come with cultures and social pressures of their own. Our moral motivations are often influenced by our social groups, so we should cultivate those groups carefully. It would be arrogant to assume that we are immune to the potentially negative effects of peer pressure just because we go in with good intentions. Changing the system from within is an admirable goal but much harder to accomplish than many people may assume. Individuals who have an interest in the path of earning to give should undertake it with due consideration.

5. Addressing the "Overdemandingness" Objection

EA may indeed have an undue emphasis on donation and a potentially counter-productive love affair with earning to give. However, its central claims remain unchallenged, even when these two expressions of the movement fall apart. One common objection levied at the very heart of EA is the worry that it is overly demanding. Even the ascetic strategies I am proposing are open to this accusation. Being informed enough about the wide-ranging effects of our everyday purchases, the ethical influence of our jobs upon the world, and every choice we make regarding where to place our money and our time is a steep task. Some worry that such a high standard of ethical behaviour would actually reduce the value of our lives.

This fear is echoed by Kuper in his objection to Peter Singer. Kuper fears that Singer's suggestions will prompt any of us who wish to lead ethical lives to "give up any job that doesn't directly or maximally involve saving lives."[36] Kuper argues, "there are many values other than survival: Can it be morally required to give up vital sources of meaning, such as the work we do, the social commitments we have, and the knowledge and excellences we pursue?"[37]

The objection stems, perhaps, from the effective altruist slogan, which is the title of one of Peter Singer's recent books, *The Most Good You Can Do*. This is the battle cry of the effective altruist. Worries of overdemandingness arise from the first half of the phrase. If one's aim is to do the most good possible, there seems to be no indication of a natural stopping point. Indeed, there may always be more good that one can do, and more effective ways to go about doing it waiting to be discovered.

36 Kuper, "More Than Charity," 110.
37 Ibid.

Yet, the real worry about overdemandingness is not about whether we can or will ever discover a single perfect recipe for the "most good." The concern is rather that EA is too burdensome of an ethical standard. If one must maximise the ethical potential of one's actions, then all of life will quickly be subsumed into that project – leaving no time for any other activities or values, or any of the things many of us consider natural, normal, and quintessentially human, like spending time relaxing with family and friends.

This criticism hinges largely on the idea that EA presents itself as a moral obligation – something which many effective altruists have explicitly denied.[38] Rather than a moral obligation to maximise the good we do (as in some forms of act utilitarianism), EA merely claims that maximizing the good is a *valuable* goal.[39] With the obligation removed, the overdemandingness objection loses most of its teeth.

To defang the objection further, I will point out that there are many activities and obligations we accept in life that are similarly demanding. For example, it is widely accepted that we ought to maintain our health. It may not be seen as a moral obligation, but since health is something most of us intuitively value, we accept the idea that the pursuit of health is worth some effort. What is more, we accept that health is worth the effort to pursue, even knowing that it is impossible to achieve and sustain a perfect state of health ad infinitum (given the fact that we are all mortal).

Most of us do not see maintaining our health as overly demanding, and we know that despite our best efforts to stay healthy, we will still inevitably fall sick and even die at some point. We do not see this as somehow undermining the value of trying to be as healthy as we can while we still live. This is because it is so clearly vital to the other things we care about in our lives. Everything else we enjoy, we enjoy longer and more fully when we are healthy. I would argue the same for the kind of ethical attentiveness that effective altruists adopt.

Additionally, the accusation of "overdemandingness" only serves as criticism if the ultimate goal of the actions demanded fails to measure up – or to put it more simply: it is only a problem if the juice is not worth the squeeze. Many activities are extremely ambitious or demanding, including some religious practices, as is the case with some monks who choose to live lives of seclusion in order to pursue greater devotion to God. Even leisure activities and hobbies, like mountain climbing or body building, when

38 MacAskill, Mogensen, and Ord, "Giving Isn't Demanding."
39 Ibid.

practiced with sufficient intensity, place severe demands on the individual that others might find life-deranging (exposing oneself to the elements or subsisting on nought but broccoli and egg whites is not for everyone). The fact that these activities are taxing, or come with serious opportunity costs, etc., is not a problem for those people sufficiently motivated by the goal.

Likewise, the selective asceticism I am advocating here is able to preserve the kinds of things we commonly hold dear in life. We can still enjoy all manner of pursuits, like art, or music, writing novels, etc., as long as we care enough to ensure these activities do not harm others. We need not perceive these activities as "time wasted" that we could have spent helping others or earning more in order to give more away.

Just as one can select between worthy causes to which to donate, one can also select between various ascetic practices. For some of us, being vegan could be too taxing on our personal health, and for others to abstain from having children would bring severe emotional pain. We need not all live the life of a childless bicycling vegan. Yet, we can and should think strategically about our "negative" ethical options. For those of us who would not suffer from such forms of asceticism, it can be a powerful tool to effect positive change in the world. For those of us who would suffer, there are still a host of other opportunities to make a difference, and these tend to be the ones we already know about – volunteering, donating to charities, etc.

As Anthony Skelton observes in his article, "The Ethical Principles of Effective Altruism," what truly unites the effective altruist community is a common perspective characterised by:

> "First, a heightened receptivity to evidence and facts in thinking about how best to achieve philanthropic ends. Second, effective altruists have, perhaps unwittingly, hit on a sensible way of … reconciling prudence and beneficence; they appear to have stumbled on what psychologists have been reporting, that above a certain threshold increases in income and wealth produce fewer and less long-lasting increases in life satisfaction. Effective altruists have discovered that for happiness, for life satisfaction, one needs much less materially than our cultural norms suggest."[40]

In reality, the effective altruist life is not a life so burdened as to be devoid of any pleasure. In fact, altruism and asceticism go hand in hand to produce more happiness for effective altruists themselves, as well as helping to alleviate suffering for others in the world.

40 Skelton, "Ethical Principles of Effective Altruism," 144.

6. Conclusion

EA gets many things right: we have a duty to help others, no matter how distant, and those in the Global North have the privileged position in that their affluence can make a significant impact on those far away with little cost to themselves. It is good to bring attention to the moral potential of money, but it is short-sighted to think that some of the problems of the world can be solved merely by donating. Because of the emphasis on "effectiveness," it should be natural for EA to seek solutions that prevent harm rather than just ameliorate it after the fact, which is why it is strange that ascetic strategies are often so neglected.

EA as a movement has developed an unhealthy obsession with cost-effectiveness analysis, which has led them away from what should be their core mission. It has left the movement unable to spot effective solutions to systemic sources of suffering, which often have political dimensions that are hard to measure with the quantitative approaches the movement favours. However, none of these flaws is fatal. Donating to effective charities has only ever been a part of the broader cause, and it is time to give more attention to the other approaches available. The first rule of getting out of a hole is to stop digging. This is what the ascetic's approach to EA seeks to reemphasise.

Both a narrow focus on monetary donation and extreme asceticism to the point of complete disengagement from society fail to effect positive change in the world in a maximal way. Thus, I am arguing for what I see as the middle way, where attention is devoted equally to the effects our money can have on the world and the nature of the activities we devote our time to. In doing that, we have the opportunity to create a more just and fairer global order and reduce the negative impact of structural violence.

This places a burden on us to be informed of the far-reaching consequences both of our actions and of our inactions. The most effective altruistic life will be demanding but not incongruous with our other values. We must seek opportunities to help others, *and* we must stop contributing to systems that harm them. This balancing act requires a moral attentiveness that can be strengthened with practice. Through a balance of selective asceticism and cost-effective charitable donations, we each have huge and mostly untapped potential to transform our world. We do not need to think our actions will create a perfect world in order for them to be worth doing. Making the world a better place should not be seen as a burdensome addition to our lives but rather an enriching dimension. Looking after our

positive impact on the world is like looking after our own health. It is negative and counter-productive when we become obsessive about it, but, for most of us, it is important enough to devote daily activity toward, through both positive actions and selective abstinence.[41]

References

Antonaccio, Maria. "Asceticism and the Ethics of Consumption." *Journal of the Society of Christian Ethics* 26, no. 1 (2006): 79–96.

Besong, Brian. "Virtue and Asceticism." *Philosophy* 94, no. 1 (2019):115–138. https://doi.org/10.1017/S0031819118000347.

Brown, Steven G. "Supporting the Best Charities is Harder than it Seems." *Journal of Global Ethics* 12, no. 2 (2016): 240–244. https://doi.org/10.1080/17449626.2016.1205118.

Clough, Emily. "Effective Altruism's Political Blind Spot." *Boston Review*, 14 Jul, 2015.

Ebrey, David. "The Asceticism of the Phaedo: Pleasure, Purification, and the Soul's Proper Activity." *Archiv für Geschichte der Philosophie* 99, no. 1 (2017): 1–30. https://doi.org/10.1515/agph-2017-0001.

Freiberger, Oliver. *Asceticism and its Critics: Historical Accounts and Comparative Perspectives*. Oxford, New York: Oxford University Press, 2006.

Gabriel, Iason. "Effective Altruism and its Critics." *Journal of Applied Philosophy* 34, no. 4 (2016). https://doi.org/10.1111/japp.12176.

Goodland, Robert and Jeff Anhang. "Livestock and Climate Change." *World Watch*, November/December (2009).

Horton, Keith. "Aid Agencies: The Epistemic Question." *Journal of Applied Philosophy* 28, no. 1 (2010): 29–43. https://doi.org/10.1111/j.1468-5930.2010.00504.x.

Houssain, Waheed. "Is Ethical Consumerism an Impermissible Form of Vigilantism?" *Philosophy and Public Affairs* 40, no. 2 (2012): 111–143. https://doi.org/10.1111/j.1088-4963.2012.01218.x.

Koh, Tsin Yen. "Bentham on Asceticism and Tyranny." *History of European Ideas* 45, no. 1 (2019):1–14. https://doi.org/10.1080/01916599.2018.1514643.

41 I would like to include a special thanks to the International Society for Utilitarian Studies for letting me present this paper at their conference at Lille Catholic University in July of 2016. I would like to thank Peter Singer for giving me feedback on my paper at the conference and thank Thomas Ferretti for his thoughtful notes on a previous draft.

Kuper, Andrew. "More Than Charity: Cosmopolitan Alternatives to the "Singer Solution"." *Ethics and International Affairs* 16, no. 1 (2002): 107–120. https://doi.org/10.1111/j.1747-7093.2002tb.00378.x.

MacAskill, William, Andreas Mogensen, and Toby Ord. "Giving Isn't Demanding." In *The Ethics of Giving: Philosophers' Perspectives on Philanthropy*. Oxford: Oxford University Press, 2018. https://doi.org/10.1093/oso/9780190648879.003.0007.

MacAskill, William. "Replaceability, Career Choice, and Making a Difference." *Ethical Theory and Moral Practice* 17, no. 2 (2014): 269–283. https://doi.org/10.1007/s10677-013-9433-4.

Pogge, Thomas. "Responsibilities for Poverty Related Ill-Health." *Ethics & International Affairs* 16, no. 2 (2002): 71–79. https://doi.org/10.1111/j.1747-7093.2002.tb00398.x.

Scarborough, Peter, Paul N. Appleby, Anja Mizdrak, et al. "Dietary greenhouse gas emissions of meat-eaters, fish-eaters, vegetarian and vegans in the UK." *Climatic Change* 125, no. 2 (2014): 179–192. https://doi.org/10.1007/s10584-014-1169-1.

Segerdahl, Pär. "Intellectual Asceticism and Hatred of the Human, the Animal and the Material." *Nordic Wittgenstein Review* 7, no. 1 (2018): 43–58. https://doi.org/10.15845/nwr.v7i1.3494.

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

Skelton, Anthony. "The Ethical Principles of Effective Altruism." *Journal of Global Ethics* 12, no. 2, (2016): 137–146. https://doi.org/10.1080/17449626.2016.1193552.

Todd, Benjamin. "Is it Ever Okay to Take a Harmful Job in Order to Do More Good? An In-Depth Analysis." 80,000 Hours. June 2017. https://80000hours.org/articles/harmful-career/.

Todd, Benjamin. "Misconceptions About Effective Altruism." 80,000 Hours. 7 Aug 2020. https://80000hours.org/2020/08/misconceptions-effective-altruism/.

Stefan Riedener

Human Extinction from a Thomist Perspective

Abstract

"Existential risks" are risks that threaten the destruction of humanity's long-term potential: risks of nuclear wars, pandemics, supervolcano eruptions, and so on. On standard utilitarianism, it seems, the reduction of such risks should be a key global priority today. Many effective altruists agree with this verdict. But how should the importance of these risks be assessed on a Christian moral theory? In this paper, I begin to answer this question – taking Thomas Aquinas as a reference, and the risks of anthropogenic extinction as an example. I'll suggest that on a Thomist framework, driving ourselves extinct would constitute a complete failure to fulfil our God-given nature, a radical form of hubris, and a loss of cosmological significance. So the reduction of such risks seems vital for Thomists, or Christians more broadly. Indeed, I end with the thought that these considerations are not parochially Christian, or indeed specifically religious: everyone with a similar view of mankind – as immensely valuable yet fallen and fragile – should agree that existential risks are a pressing moral concern.

1. Introduction

Let us say that "humanity's long-term potential" is the set of all possible desirable futures that are open to us.[1] This potential is enormous, and contains some fantastic scenarios. Consider

> **Flourishing:** We survive for another billion years, and on average there are always 10 billion people, who live very good 100-year-long lives. So there are 100 million billion future people with very good lives.

This scenario is possible. The earth seems to remain sustainably inhabitable by at least 10 billion people,[2] and for another 1.75 billion years.[3] The quality of our lives seems to have increased continuously,[4] and it seems possible for this trend to continue. So we might have a glorious future ahead of us.

However, we might also founder. Let us say that an "existential catastrophe" is the destruction of humanity's long-term potential, and that an "existential risk" (or "x-risk") is a risk that threatens this destruction. There

1 This definition of "humanity's long-term potential", as well as the following definitions of "existential catastrophes" and "existential risks", are very slightly adapted versions of the definitions in Ord, *The Precipice,* ch. 2.
2 United Nations, "World Population Monitoring," 30.
3 Rushby et al., "Habitable Zone."
4 See e.g. the data collected in Pinker, *Enlightenment Now*.

are various ways for our potential to be destroyed, even in the course of our century. Extreme climate events might engender a civilisational collapse, say, as a consequence of which humanity gets permanently reduced in numbers and locked into a cultural and technological state resembling that of the Middle Ages. But the most straightforward such catastrophe would be

Extinction: We die out this century.

This scenario is also possible. Indeed, Extinction may be more likely than most of us think. There are many possibly lethal x-risks: nuclear wars, developments in biotechnology or artificial intelligence, climate change, pandemics, supervolcanos, asteroids, and so on.[5] In an informal poll, risk experts reckoned that there is a 19% probability that we will die out this century.[6] The Stern Review on the Economics of Climate Change assumed a 9.5% likelihood of our dying out in the next 100 years.[7] And a recent report by the Global Challenges Foundation suggests that climate change, nuclear war and artificial intelligence alone might ultimately result in extinction with a probability of between 5% and 10%.[8]

Importantly, whether Flourishing or Extinction will materialise partly depends on us. Indeed, it may depend on what we do *today*. We could now promote academic research on x-risks,[9] global political measures for peace, sustainability or AI safety,[10] the development of asteroid defence systems,[11] shelters,[12] risk-proof food technologies,[13] and so on. And while none of these measures will bring x-risks down to zero, they will arguably at least reduce them. So all of this raises a very real, practical question. How important is it, morally speaking, that we now take measures to make Flourishing more likely?

5 See e.g. Bostrom and Ćirković, *Global Catastrophic Risks*.
6 Sandberg and Bostrom, "Catastrophic Risks Survey."
7 HM Treasury, "Stern Review."
8 Pamlin and Armstrong, "Global Challenges." More precisely, the report distinguishes between "infinite impact", "where civilisation collapses to a state of great suffering and does not recover, or a situation where all human life ends", and an "infinite impact threshold", "an impact that can trigger a chain of events that could result first in a civilisation collapse, and then later result in an infinite impact" ("Global Challenges," 40). The above numbers are their estimates for infinite impact thresholds.
9 Bostrom and Ćirković, *Global Catastrophic Risks*.
10 Cave and ÓhÉigeartaigh, "Bridging."
11 Bucknam and Gold, "Asteroid Threat?"
12 Hanson, "Catastrophe."
13 Denkenberger and Pearce, *Feeding Everyone*.

The answer depends on the correct moral theory. It is most straightforward on standard hedonistic total utilitarianism. Suppose we increase the probability of Flourishing over Extinction by just 1 millionth of a percentage point. In terms of overall expected welfare (setting non-human sentience aside), this is equivalent to saving about 1 *billion* lives, with certainty. So according to utilitarianism, even such tiny increases in probability are astronomically important. Indeed, Nick Bostrom suggested that x-risk reduction "has such high utility that standard utilitarians ought to focus *all* their efforts on it."[14] But this implication is not restricted to utilitarianism. Something very similar will emerge on any view that assigns weight to expected impartial welfare increases. Consider effective altruism (EA). EA is the project of using evidence and reasoning to determine how we can do the most good and taking action on this basis.[15] This does not presuppose any specific moral theory about what the good is, or about what other reasons we have beyond doing the most good. But effective altruists typically give considerable weight to impartial expected welfare considerations. And as long as we do, the utilitarian rationale will loom large. Thus according to a 2018 survey, EA-leaders consider measures targeted at the far future (e.g. x-risk reduction) 33 times more effective than measures targeted at poverty reduction.[16] The EA organisation 80,000 hours suggests that "if you want to help people in general, your key concern shouldn't be to help the present generation, but to ensure that the future goes well in the long-term."[17] And many effective altruists already donate their money towards x-risk reduction rather than, say, short term animal welfare improvements.

In this paper, I will ask how the importance of x-risk reduction should be assessed on a Christian moral theory. My main claim will be that Christian morality largely agrees with EA that x-risk reduction is extremely important – albeit for different reasons. So Christians should emphatically support the above-mentioned measures to increase the probability of Flourishing.

Let me clarify. First, there is no such thing as *the* Christian moral doctrine. One of the philosophically most elaborate and historically most influential developments of Christian thought is the work of Thomas Aquinas. So I will take this as my starting point and argue first and foremost that core *Thomist* assumptions support x-risk reduction. Thomas's specific inter-

14 Bostrom, "Astronomical Waste," 308ff.; see also Parfit, *Reasons,* 452f.; Beckstead, *Far Future*; and Bostrom, "Existential Risk."

15 See MacAskill, *Doing Good Better*.

16 Wiblin and Lempel, "New survey."

17 Todd, "Long-term value."

pretations of these assumptions are often unappealing today. But I will also claim that they can be interpreted more plausibly, that their core idea is still authoritative for many Christians, and that on any plausible interpretation they ground an argument for x-risk reduction. So while I will start with Thomas, my conclusions should appeal to quite many contemporary Christians. Indeed, I will ultimately indicate that these assumptions need not even be interpreted in a specifically Christian manner but emerge in some form or other from a number of worldviews (cf. section 4). Second, there are different x-risk scenarios, and they raise different moral issues. I think the argument is clearest for a class of cases I will call "non-transitionary anthropogenic extinction": ways in which humanity might literally die out (hence "extinction"), as a direct consequence of our own actions (or "anthropogenic"), and without being superseded by non-human intelligence (thus "non-transitionary"). So it is on these cases that I will focus. It would be interesting to explore other scenarios in which we are replaced by another form of intelligence, non-extinction x-risks (which our above definition includes) like an extended stage of suffering, or scenarios of natural extinction like volcano eruptions. My arguments will have upshots for such cases too. But I will not explore them here. Third, there are different ways in which, or different agents for whom, reducing extinction risks might be important. In what follows, I will mostly be considering a collective perspective. That is, I will assume that we as humanity collectively do certain things. And I will focus on whether *we* ought to do anything to mitigate the relevant risks – rather than on whether you individually ought to. The existence of this collective perspective might be controversial. But I think it is plausible.[18] Christian moral theory, or at least Thomas, also seems to assume it (cf. section 2.1). And many important issues emerge only or more clearly from it. So I will assume it in what follows.

In short, my question is about how important it is, on roughly Thomist premises, for us to reduce risks of non-transitionary anthropogenic extinction. I will first present three considerations to the effect that, if we *did* drive ourselves extinct, this would be morally very problematic – a hubristic failure in perfection with cosmologically bad effects (section 2). I will then discuss some countervailing considerations, suggesting that even though such extinction would be bad, we needn't take measures against it – because God will not let it happen, or because we would not *intend* it, or because at any rate, it is not imminent (section 3). I will argue that none of these

18 See e.g. Dietz, "Collective Obligations."

latter considerations cut any ice. So I will conclude that on roughly Thomist premises it's extremely important for us to reduce such risks.

2. Three Thomist Considerations

There are many Thomist considerations that would bear on risks of non-transitionary anthropogenic extinction. For instance, in driving ourselves extinct, we would presumably kill the last representatives of humanity. On a Thomist perspective, those killings will be morally problematic simply *as killings*.[19] Yet this has nothing to do with the fact that those killings lead to extinction. So let us see whether there would be anything *distinctly* problematic about extinction, if we brought it about. There is.

2.1 The Natural Law

A first consideration follows from Thomas's teleology, or from the Aristotelian strand in his thinking. Recall that for Thomas, the order of the cosmos is teleological. This teleology is grounded in the fact that the cosmos is subject to God. And it takes the form of a law for us: "the universe is governed by Divine Reason. Wherefore the very Idea of the government of things in God [...] has the nature of a law."[20] For Thomas, following this "eternal law" – or Divine "will"[21] or "plan"[22] for all things – is *the* ethical imperative for us. So what does it command? We cannot know God's intent "in itself",[23] at least not in this earthly life.[24] But we can know it "in its effect", through its manifestations in creation, or through the "natural law" that we recognise as structuring the physical universe. In particular, we can detect God's plan for us through the natural inclinations He has implanted in us: "all things partake somewhat of the eternal law[:] [...] from its being imprinted on them, they derive their respective inclinations to their proper acts and ends."[25] In other words, our inclinations allow an indirect cognition of the essence of God's will, quite like sunrays allow an indirect cognition of the substance of the sun.[26] So what are our natural

19 See ST, ii–II, q64.
20 ST, i–II, q91 a1, co.
21 ST, i–II, q93 a4, ad1.
22 ST, i–II, q93 a3, co.
23 ST, i–II, q93 a2, co.
24 See ibid; ST, I, q12 a11.
25 ST, i–II, q91 a2, co.
26 ST, i–II, q93 a2, co.

inclinations? Thomas speaks of three kinds: "in man there is first of all an inclination [...] which he has in common with all substances: [...] the preservation of its own being [...]. Secondly, there is in man an inclination [...] which he has in common with other animals: [...] sexual intercourse, education of offspring and so forth. Thirdly, there is in man an inclination to good, according to the nature of his reason [...]: [...] to know the truth about God."[27] What is crucial for present purposes is the second type of inclination. Procreation or the "preservation of the species" is firmly part of our "natural good",[28] or of what the natural law commands us to do.

Now Thomas does not believe that everyone must follow all of these inclinations, or that everyone must have offspring. It is permissible for some to remain celibate.[29] But we as a collective have a duty – a "duty [...] to be fulfilled by the multitude"[30] – to procreate. This grounds a straightforward consideration for reducing extinction risks. By going extinct, we would fail to attain our end or to accord with the natural law. Indeed, our failure would be profound. It would not just be some of us flouting this law – the bad apples in an overall virtuous whole. We would fail *collectively*, as an entire species, to attain our end. And (at least as far as this life is concerned) we would not just fail in one aspect of perfection, while still able to excel in others. Our survival is a precondition for any aspect of our flourishing. So our extinction would mean we fail *comprehensively*, in all respects of our end. And of course we would not just temporarily fall short of our calling. Once we have gone extinct, there is no second bite at the apple. It would mean we have foundered *irreversibly*. So the moral failure in anthropogenic extinction would seem *complete*. In short:

> **Natural law:** We ought to attain our natural end. Our extinction would prevent us from doing so – collectively, comprehensively, and irreversibly. Thus non-transitionary anthropogenic extinction would amount to a total moral failure of us as a species.

This is a first consideration for why our extinction would be problematic.

Now I suppose that this kind of rationale is not parochially *Thomist*, but should appeal to Christians quite broadly. Thomas himself interprets the natural law pertaining to procreation very radically – e.g. as permitting a

27 ST, i–II, q94 a2, co.
28 SCG, III, 122–9.
29 See e.g. ST, ii–II, q152.
30 ST, ii–II, q152 a2, ad1.

non-procreative life only for the sake of the "contemplation of truth",[31] and as prohibiting any intercourse known to be non-reproductive.[32] This would mean that our procreation-related obligations go much beyond preventing extinction. And it would mean that contemporary liberal moral thought is radically wrong about the good human life, and about intercourse among people of the same sex, or people who for whatever reasons cannot or do not want to reproduce. Many contemporary Christians will want to resist these implications. But we need not understand the general idea of a "natural law" in this manner. Plausibly, other aspects of the human end (athletic, social, emotional) are just as important as "the contemplation of truth." So there are good lives without children beyond those of philosophers and priests.[33] Also, perhaps there are other functions of human sexuality (e.g. bonding), such that non-reproductive intercourse is not a misuse of sexual organs. After all, there is plenty of non-reproductive sexuality among non-human animals.[34] The details of the natural law are a matter of large debate.[35] But I presume that the *general* idea of a "natural law" is still quite prominent for Christians today. Indeed, assuming the universe manifests God's intentions, it is a very natural assumption. And I suppose that on any plausible interpretation, it will ground a consideration against anthropogenic extinction: whatever it implies about childfree individuals, contraception or homosexuality, it seems hard to square the natural law with an heirless self-eradication of our species. If the law commands anything, it seems, it commands us (with Gen 1:28) to "be fruitful and multiply."[36] So a consideration along these lines should be authoritative to

31 ST, ii–II, q152 a2, co; also ST, ii–II, q153 a2, co. Thomas doesn't say this explicitly in these passages. He says that the goods of the body are subordinate or "directed" to the goods of the soul, and that virginity (or a lack of bodily good) is justified if done for the sake of Divine contemplation (or good of the soul; ST, ii–II, q152 a2, co). I'm reading the conditional here as a biconditional.

32 SCG, III, 122; also ST, ii–II, q153 a2, co.

33 See e.g. Nussbaum, "Non-Relative Virtues," or *Creating Capabilities*.

34 See e.g. Bagemihl, *Biological Exuberance*.

35 For a classic systematic exploration of natural law, see Finnis, *Natural Law*. For an in-depth treatment of Aquinas's theory, see e.g. – among very many others – Lisska, *Aquinas's Theory*.

36 Historically, this broad rationale has not appealed to all Christians, of course. Paul himself suggests it would be best if everyone was celibate (see 1 Cor. 7:7: "I wish that all of you were as I am"). And the Cathars even thought reproduction was a moral evil. I thank Felix Timmermann and Christoph Halbig for pointing this out to me.

Christians quite broadly. Self-inflicted extinction would constitute a total failure in our executing our designated role.[37]

2.2 Humility

However, there is more to anthropogenic extinction than a failure to reproduce. Let us look at the precise way in which, through anthropogenic extinction, we would fail to attain our end. According to Thomas, there are different ways to fall short of perfection. One is through utter passivity, or sloth, or a "sorrow" that "so oppresses man as to draw him away entirely from good deeds."[38] Thus we might just not bother to do anything much at all, and therefore fail in perfection. Another way of failing lies in falling too low, or being overcome with "the lower appetite, namely the concupiscible."[39] So we might behave like lower animals, and fail to live *up* to our standards. But drawing on Augustine, Thomas says that "the most grievous of sins"[40] consists in aiming too *high* – in failing to respect our limitations or acting as if we were God. To do so is to commit the sin of pride, hubris, or *superbia*. Following Augustine,[41] Thomas characterises a prideful man as someone who "aims higher than he is", or does not "tend to that which is proportionate to him."[42] And what is proportionate to us is of course not so by coincidence, but due to Divine assignment. Thus pride is opposed specifically to humility: "humility properly regards the subjection of man to God [...]. Hence pride properly regards lack of this subjection, in so far as a man raises himself above that which is appointed to him according to the Divine rule."[43] And in this sense, as far as the "aversion from the immutable good"[44] is concerned, pride is the most grievous sin: it is not just a failure

37 Note that on a common understanding of the natural law, it does not unconditionally command you to strive for survival. On the contrary, it is natural for your life to end at some point, and it can be wrong for you to take means of prolonging it beyond – e.g., enhancement or certain life-sustaining treatments. But this does not jeopardise the above argument. Species do not seem to have an inherent natural endpoint. Species as a whole do not age. Moreover, I am only concerned with *anthropogenic* extinction. And it seems implausible that guarding ourselves against *that* – e.g., promoting measures for peace, sustainability or AI safety – would somehow be unnatural.

38 ST, ii–II, q35 a2, co.

39 ST, ii–II, q153 a5, co.

40 ST, ii–II, q162 a6, co.

41 De Civ. Dei, XIV, 13.

42 ST, ii–II, q162 a1, co.

43 ST, ii–II, q162 a5, co.

44 ST, ii–II, q162 a6, co.

through ignorance or weakness or an innocent desire for another good, but an active "withstanding" or "resisting" or manifesting "contempt of" God.[45]

What does this imply in practice? Thomas specifies what he means. Pride is not any old undue desire. It is, specifically, an "appetite for *excellence* in excess of right reason" – or an inordinate imitation of the powers or competences of God.[46] This may take various forms. We may be pridefully curious about things we cannot know, such as facts about good and evil.[47] Or we may unduly discard our need for Divine grace, deeming ourselves capable of happiness on our own.[48] But a more specific power that is not appointed to us is decisions about life and death: "it belongs to God alone to pronounce sentence of death and life."[49] Thus to kill someone, or – I take it – actively prevent them from coming into existence, is generally to show an appetite for a power that does not pertain to us. And if all of this is true, then non-transitionary anthropogenic extinction, in particular, would manifest an enormous form of pride. Note that on most such scenarios, we would fall prey to technologies we were unable to control – nuclear weapons, artificial intelligence, biotechnology, or whatever. So our extinction would mean that we had overestimated our mastery over these fabrications, and the invulnerability we could leverage in the face of them. It would mean we were prideful in the general sense of desiring an undue excellence. And the upshot of this would be, specifically, a life-death decision on an astronomical scale: preventing lives perhaps by the quadrillions. It would mean we were prideful in this more specific sense too. Or in short:

> **Humility:** Non-transitionary anthropogenic extinction would mean we overstrained our power as a species. And the upshot of it would be a life-death decision on an astronomical scale. Thus it would amount to an enormous form of *superbia*.

This is a second consideration for why our extinction would be problematic.

And here too, I suppose this consideration should appeal to Christians quite broadly. Again, Thomas's specific interpretation of God's authority over life/death-decisions should be controversial. He suggests that, while permitting us to execute death penalty,[50] this authority absolutely prohibits

45 Ibid.
46 ST, ii–II, q162 a1, ad2; emphasis added.
47 ST, ii–II, q163 a1, ad3.
48 ST, ii–II, q163 a2, co.
49 ST, ii–II, q64 a5, co.
50 ST, ii–II, q64 a2, co.

suicide,[51] and prohibits abortion from the moment of ensoulment[52] – which some people have claimed, according to Thomas's metaphysical principles and the known facts about embryology, takes place at the moment of fertilisation.[53] This would mean self-extinction can under no circumstances be permissible. And it would mean that contemporary liberal moral thought is radically wrong about the death penalty, abortion, or suicide and euthanasia. Many contemporary Christians will want to resist these implications. But again, we need not interpret the general idea of *superbia* in this manner. After all, it is implausible that criminals are "dangerous and infectious to the community" and must be cut away like an infected part of a body.[54] Perhaps suicide can in some instances of extreme pain, or loss of autonomy or dignity, be a manifestation of self-love. And if so, perhaps we can then view ourselves as authorised by God – or His commandment to love ourselves[55] – to end our lives. After all, Thomas himself (defending Abraham's intention to kill Isaac) suggests that "he who at God's command kills an innocent man does not sin."[56] Perhaps some forms of abortion (e.g. after rape) can also be seen as an expression of self-love or -respect. Or perhaps we must ascribe to Aquinas a different view of ensoulment,[57] or simply reject some of his metaphysics in light of more recent discoveries. Again, the details of *superbia* are contested.[58] But I assume that the pertinent general idea is still prominent among Christians today. Indeed, that playing God is a sin is a natural corollary of theism.

51 ST, ii–II, q64 a5, co.

52 See e.g. ST, I, q118f.

53 Haldane and Lee, "Aquinas on Human Ensoulment," 273. In the *Scriptum super Sententiis*, following Aristotle, Thomas himself suggests that the soul is infused after 40 days for men and after 90 days for women (SSS III, d3 q5 a2, co).

54 ST, ii–II, q64 a2, co.

55 See e.g. Mat 22:39.

56 ST, ii–II, q64 a6, ad1.

57 Pasnau, *Thomas Aquinas*.

58 There is a lot of recent literature concerning Thomas on humility. For longer treatments, see e.g. Tadie, *Between Humilities,* or Fullam, *The Virtue of Humility.* As far as I see, Thomas himself does not discuss whether birth control (or "contraception") infringes on God's dominion over life/death-decisions. But the Catholic church has since leveraged this argument. Yet perhaps a ban on contraception need not follow from humility either. Perhaps Thomists may appeal to the doctrine of double effect, and compare well-intentioned prevention of fertilisation with well-intentioned killing in self-defence (cf. section 3.2). Or perhaps they might again claim to be authorised in this decision by the commandment to love ourselves, and love our partners, and our (potential) children for whom we could not be sufficiently responsible parents.

The in-principle-ban on life/death-decisions has a very firm grounding in the Bible.[59] And I suppose that on any plausible interpretation, these ideas ground a consideration against anthropogenic extinction: whatever they imply about the death penalty, abortion, and suicide, the dictates of humility seem hardly compatible with our developing a technology that accidentally seals the fate of our whole species. So again, this consideration should have import for Christians quite broadly. Self-inflicted extinction would constitute a shattering form of *superbia*.

2.3 The Value of Humanity

But we need to see another aspect of Thomas's view of humanity too, which is perhaps most distinctly biblical, and which is implicit in the third kind of inclination he ascribes us. Without doubt, we are not God. But we are nonetheless made to know the truth about Him. In fact, of all corporeal things (i.e., disregarding angels), we are the only ones whose nature enables such knowledge. And this, for Thomas, marks our "dignity". It is this dignity which made it fitting for Jesus to adopt *human* nature, rather than becoming an animal or any other thing: "in the irrational creature the fitness of dignity is wanting."[60] Indeed, it is these intellectual capacities that ground our likeness to God – a likeness greater than that of any non-rational creature,[61] and sufficient to say we are His "image": "intellectual creatures alone [...] are made to God's image."[62] And it is this which ultimately manifests that God loves us more than any other thing. So Aquinas quotes Augustine approvingly: "God loves all things that He has made, and amongst them rational creatures more, and of these, especially those who are members of His only-begotten Son Himself."[63]

In other words, and as Thomas says very explicitly,[64] of all corporeal things we are the most valuable, or those with most "goodness". And this difference is categorical. So our extinction would have cosmological ramifications. For standard utilitarians, there is no categorical difference between a world populated by flourishing societies of people and a world populated by one lonely lizard basking in the sun and feeling a tinge of pleasure.

59 Thomas cites Deut. 32:39, for instance: "I put to death and I bring to life".
60 ST, III, q4 a1, co.
61 See ST, I, q93 a2, co.
62 Ibid.
63 ST, I, q20 a3, sc.
64 ST, I, q20 a4, co.

The difference is a matter of degree. For Aquinas, that is different. In a world devoid of higher intelligence, there is nothing that is made to God's image – nothing with our special dignity or perfection. This would radically undercut the perfection of the universe, as God created it, which depends on the varieties of goodness.[65] So our extinction would change the face of creation. Or in short:

> **The value of humanity:** Non-transitionary anthropogenic extinction would constitute the destruction of the most valuable aspect of creation – and thus a loss of categorical cosmological significance.

That is a third consideration for why our extinction would be problematic.

Again, I suppose that this consideration should appeal to Christians very broadly. Thomas interprets the order of creation very radically. He suggests that everything else exists *for us*[66] – such that we may use animals simply as means for our ends, say, and the self-interested killing of another person's ox wrongs at most its owner.[67] This would mean that our extinction would literally destroy the purpose of the physical universe. And it would mean that much contemporary thought is radically wrong about animals, or other aspects of the natural world. Again, many Christians will want to resist these implications, or so I hope. But the distinct value of humans need not be interpreted this radically. Perhaps our "dominion" over animals (mentioned in Gen 1:26) does not mean we can use them simply as means. Perhaps it means we must use our *de facto* power in the manner of a loving and respectful guard.[68] And similarly with every other aspect of creation – ecosystems, plants, and the climate. Again, the exact nature of our status is a large question. But the general idea of our special value still seems very prominent. It does seem hard to avoid if we take seriously that man is "God's image", or has a special (perhaps responsibility-implying) dominion over the earth – and very plausible on the assumption of God's human embodiment in Christ. And on any interpretation, our special status will ground a consideration against extinction: whatever it implies about our responsibilities towards animals, the face of creation will be changed categorically if we drive ourselves extinct. So this consideration should be authoritative to Christians very broadly. Self-inflicted extinction would constitute a destruction with cosmological import.

65 See e.g. ST, I, q47.
66 ST, ii–II, q64 a1, co.
67 Ibid.
68 See e.g. Linzey, *Christianity,* ch. 2.

Plausibly, there are other Christian considerations that bear on x-risks, at least once we move beyond Thomas. Most notably, perhaps Christian *caritas* simply gives us reasons to increase expected impartial welfare.[69] If it does, the EA rationale will be very pertinent to Christians as it stands. But whether that is an apt interpretation of the Christian virtue – or whether the latter is more about making people happy than about making happy people – will be controversial. And at any rate, I think these are the three most distinctly Christian concerns. So let me leave it at that for now.

3. Discussion

If all of this is right, non-transitionary anthropogenic extinction would be a morally highly problematic result. Such extinction would not just amount to a regrettable form of imprudence – but to a failure to fulfil our God-given role. It would not amount to any old failure of this kind – but to a prideful failure to achieve our end. And it would not just be an inconsequential overstepping of our dominion – but an unauthorised decision with categorical cosmological ramifications.[70] But as mentioned, humanity can now (in the form of the present generation) take measures to reduce the likelihood that it will eventually bring this result about. So this suggests that we have very strong reasons to take some such measures. It suggests that Christians too have strong reasons to donate their money towards x-risk reduction rather than, say, disaster relief; to conduct x-risk research rather than enquiries about the cause of the grief of a neighbour; or to advocate political measures for long-term safety rather than for short-term caritative purposes. But let us not get ahead of ourselves. Some considerations in Thomas's philosophy seem to suggest that even if extinction would, as far as we know, be bad, it is not important that we now do much about it in practice. So let me turn to some of these theological and moral countervailing concerns. This will not simply corroborate the results so far. It will also help clarify the precise form and weight of what the above rationale implies.

69 See e.g. William Paley's utilitarian Christianity (Paley, *Principles,* esp. book II)

70 Perhaps it is worth noting that in principle, these considerations are logically independent. One might accept some but not all of them. For instance, one might doubt that our natural inclinations reveal a Divine law, but still accept the cosmological disvalue of human extinction.

3.1 Theological Considerations: Divine Providence

One apparent reason against extinction risk activism is Divine providence. Again, for Thomas, everything is subject to God. And Thomas's God is all-knowing, all-powerful, and all-good: "In God there exists the most perfect knowledge",[71] "He can do all things that are possible absolutely [or do not imply a contradiction in terms]",[72] and He "loves all existing things."[73] One might conclude from this that God will not allow us to go extinct, or that if He does it is somehow just for the better. And from this one might infer that we need not do anything to prevent our extinction in practice. Reducing extinction risks is obviated by the rule of a loving God.

This is difficult and well-trodden terrain. But God's existence does not *in general* seem to warrant any heedlessness. Consider risks of road accidents, say. And suppose any accident accords with His will. It is an interesting question what precisely this implies. Perhaps it warrants a certain ultimate serenity or comfort: a reassurance at the thought that whatever happens on our roads has its mysterious rightness in the grand scheme of things. Perhaps it warrants a fundamental form of trust: a consolation in the belief that our diligence will be duly rewarded.[74] Perhaps it means that while things are ultimately in God's hands, he wants to achieve His ends through us being virtuous and pious.[75] But in any case, God's unfathomable values cannot act as a guide for us, or as an excuse to depart from the norms and expectations we are given. So Christian hope must be distinguished from optimism that accidents will not occur, or from feeling exonerated from the need to take caution. Christians ought to see to it that they stop at red lights, respect speed limits, and check their brakes – and just as carefully as everyone else. Indeed, Thomas is the first to emphasise that we are obligated to such circumspection.[76] In short, Divine providence does not affect our

71 ST, I, q14 a1, co.
72 ST, I, q25 a3, co.
73 ST, I, q20 a2, co.
74 See ST, ii–II, q22 a1, co.
75 A passage from the book of Esther is sometimes interpreted along these lines. Mordecai seems to believe that the Jewish people will survive anyway, but that Esther should nonetheless try to avert the genocide: "For if you remain silent at this time, relief and deliverance for the Jews will arise from another place, but you and your father's family will perish" (Est. 4:14). I thank Dominic Roser for mentioning the above thought, and this passage, to me.
76 See e.g. ST, ii–II, q64 a8, co.

need to reduce road risks in practice. But then it is unclear why it should do so with risks of human extinction.[77]

One might suggest that there is something special about human extinction, setting it apart from everyday hazards, and making it particularly unlikely. Perhaps it is precisely because we are His "image", or uniquely most valuable, that God will not allow us to perish. Perhaps Jesus's redemption of mankind would have been futile if some sorrowful two thousand years later He let us burn up or wither away.[78] Or perhaps, more specifically, there is evidence in scripture that we will survive: the primal great flood precisely *did not* erase us, and after that deluge there seems to have been a promise of survival – that "never again will all life be destroyed by the waters of a flood" (Gen. 9:11). This is not the place to dive deep into Christian theology. But it is hard to see why such reasonings should give us much confidence. The Lord moves in mysterious ways. If all known misery is compatible with His providence, it must surely at least be possible that He would allow us to go extinct. In particular, there are many more passages in scripture painting a grim prospect of extinction than promising a boundless glorious future. And many such visions of the "end of time" indicate anthropogenic extermination – that "the nations in the four corners of the earth" are gathered for "battle" or "war" (Rev. 20:8), right before "the earth and the heavens" are gone (Rev. 20:11). If anything, according to the Bible, anthropogenic extinction seems a very live option. So Divine providence might have implications for the metaphysics and ultimate unfathomable significance of our doomsday. But it does not seem to warrant *our* heedlessness about extinction risks in practice. If the above considerations are sound, we, just like any generation in the future, must see to it that it is not *us* who bring a Johannine finale about.

77 In his contribution to this collection, Dominic Roser suggests that Christianity warrants a certain renunciation of control. He contradicts my claims here – and suggests, say, that Christians may reasonably check their brakes somewhat less often, or may take considerably less precautionary measures against x-risks, than atheists do. This view seems to me, among other things, unstable. It seems that God's existence would either require standard full-blown caution (as I think Thomas suggests) or would warrant *total* relinquishment of it. It seems hard to see why He would warrant a certain but *only* a certain trust in the course of events. At any rate, I doubt that most Christians understand their faith in this manner. And even if Roser is right, any such *limited* trust in things would arguably still not undermine my practical conclusions.

78 I thank Felix Koch for mentioning this thought to me.

3.2 Moral Considerations: Deontology

Let us look at moral considerations that might mitigate the importance of reducing extinction risks. One such consideration concerns intentions. For Thomas, the permissibility of our actions can depend on what we intend. More precisely, he held, or indeed introduced, the doctrine of double effect (or DDE). According to this doctrine, there is a difference between the effects we intend an action to have, and the effects we merely foresee but do not specifically intend. While it is always impermissible to intend evil effects, it can be permissible to cause them, if what you intend is good. For instance, you may kill in self-defence, if you intend to save your life and merely foresee the death of your aggressor.[79] This bears on risks of extinction. Perhaps non-transitionary anthropogenic extinction would be a very bad result. But one might suggest that as long as we do not *intend* to effect it, we would not have acted impermissibly: an unintentional mass life-prevention is not prideful, and an unintentional thwarting of perfection or erasing of God's image is not morally wrong. So if our intentions are fine, we need not worry about reducing extinction risks.[80]

But this argument is a non-starter. It might well be worse to intend our extinction rather than to merely foresee it. For Thomas, it would mean that our action would lack "goodness from its end", and not just from its circumstances or species.[81] But this does not mean that unintentional extinction would not be wrong, or even specifically prideful. There is no blanket permission for the well-intentioned. You must not speed through the cross-walk and put people at risk, even with the laudable aim of being punctual at the faculty meeting. Thomas says explicitly that you are "in a sense guilty of voluntary homicide" – and thus (I take it) of *superbia* – if you kill someone unintentionally but without having taken "due care."[82] He does not specify when precisely the DDE allows you to accept a foreseen evil. But he suggests it depends on (i) the expected goodness of the intended good, (ii) the expected badness of the foreseen evil – or on the "proportion" between the two[83] – and (iii) on whether the harmful action is "necessary"[84] for the good, or whether there are ways of securing the good without

79 See, ST, ii–II, q64 a7, co.
80 I thank Eric Sampson and Jonathan Erhardt for mentioning this thought to me.
81 ST, i–II, q18, esp. a7, co.
82 ST, ii–II, q64 a8, co.
83 ST, ii–II, q64 a7, co.
84 Ibid.

these bad effects. These criteria decidedly command reducing extinction risks. Sure, many technologies that involve such risks also promise important goods. But if my arguments are right, the moral costs are potentially tremendous. And there are less risky courses of action which detract little from the expected good. The above-mentioned measures would reduce the relevant risks but would not really jeopardise the benefits these technologies promise. So it seems absolutely obligatory to take them. Going on as we do does seem like scorching through the cross-walk for timeliness at the meeting. That is not killing for the sake of the killing. But it is a serious violation of the negative constraint of the DDE. And thus it is surely grave enough.[85]

Let us consider another countervailing consideration. Perhaps extinction would be terrible. But for all we know, it is not actually imminent. Indeed, compared to more immediate moral callings like global poverty, gender justice, or disaster relief, it seems very remote. According to utilitarianism or EA, such distance does not matter. These views are relevantly impartial. But perhaps Christian morality is different. One might suggest Christian morality is more concerned with visible, immediate, near-at-hand moral problems – the wounded person alongside the road (Luk. 10:30), or perhaps the literal "neighbour" (Mrk. 12:31) – and does not warrant too much concern about such far-distant matters. In fact, Thomas himself is explicit that we should be partial, at least in the virtue of *caritas*: "Among our neighbours, we should love them more who are more closely connected to us."[86] So one might conclude that although our extinction would be bad, we need not worry about reducing the relevant risks – or at least not now, or not with resources we could direct towards these more immediate concerns.

85 Strictly speaking, perhaps extinction would not even be a "foreseen" effect of our actions. By "foreseen" [*praecogitatus*] Thomas seems to mean "known with certainty to result" (see, e.g. ST, i–II, q20 a5, co). But we are not certain that promoting artificial intelligence, say, will result in extinction. We are just not certain that it will not. In discussing the DDE, Thomas does not deal with uncertainty. But he does elsewhere. In ST, i–II, q20 a5, co, he says: "if the consequences [of an action] are not foreseen, we must make a distinction. Because if they follow from the nature of the action and in the majority of cases [...] the consequences increase the goodness or malice of that action [...]. On the other hand, if the consequences follow by accident and seldom, then they do not increase the goodness or malice of the action." This cannot be right. Throwing stones off a cliff might kill passersby below in only 5% of cases. But then it is nonetheless wrong. Thomas lacked the concept of expected value. Today, we would surely interpret the doctrine in terms of that concept, or some related one.

86 DQV, q2 a9, co; see also ST, ii–II, q31 a3, co.

But this would be a mistake. Christian morality may be partialist about *caritas*. So there may be a tension between Christian benevolence and the impartialist EA-directive to do the most good. But on the considerations that I have sketched, reducing extinction risks is not a form of "doing good", or of charity, or of fulfilling any positive obligation. It is to ensure – or make more likely – that we do not flout our end, overstep our dominion, or wreck the crown of creation. Formally, it is to ensure that we satisfy the constraint of the DDE. In other words, it is to respect a perfect negative duty. And Thomas does not seem to endorse partiality or discounting in such negative duties, and neither does Christian morality more broadly. On Christian morality, killing, say, is wrong. And doing what will kill someone in a month is presumably *ceteris paribus* just as wrong as doing what will kill someone in a year, or in a thousand years for that matter. So the sheer temporal distance of human extinction *in itself* does not seem to obviate such risk reduction. On the contrary, note that the negative character of this obligation has implications for how it may be weighed against others. For utilitarians, the obligation to reduce x-risks is formally on a par with (other) obligations of benevolence, such as obligations to support the poor. So it ought to be weighed against them. For Thomist Christians, if I am right, this is different. As a negative obligation, the obligation to reduce risks of extinction must not be weighed against positive imperfect obligations or cannot be discharged by doing enough by way of benevolence elsewhere. It is quite simply a perfect duty.

There is a final point worth noting. For utilitarians, notoriously, we may take any means for the sake of the good end. If killing an innocent scientist reduces x-risks, we presumably ought to kill them. For Thomist Christians this will be different. Thomas explicitly accepts deontological constraints: "some [actions] are evil, whatever their result may be."[87] So Thomist Christians must not do *everything* to reduce such risks. They generally must not kill or lie or steal for that purpose. In practice, however, this will not make much of a difference. The most salient means of reducing the relevant risks include academic research, or global measures for peace, sustainability or AI safety. None of these measures seem to violate any constraints. So even if Christian morality is deontological in nature, there are plenty of good ways to start making Flourishing more likely.

87 ST, ii–II, q88 a2, ad2.

4. Conclusion

If all of this is right, it is not just that non-transitionary anthropogenic extinction would, as far as we know, be a disastrous result – a cosmologically salient prideful failure to achieve our end. At least in practice, and as a matter of a negative constraint against lack of "due care" that in principle seems neither discountable nor weighable against positive obligations, Thomist Christians have very strong reasons to take deontologically permissible means to prevent such extinction: to conduct research on risks and risk-reduction, promote political arrangements with an eye to the very far future, donate their money towards x-risk mitigation, and so on. Indeed, given the gravity of the possible effect, and the perfect or *negative* form of the duty, they seem to have reason to do this even if it will considerably constrain their resources for classical forms of *caritas*.

Or that is as far as our argument takes us. There are many considerations relevant to non-transitionary anthropogenic extinction that I have not yet addressed. For instance, it would be interesting to explore the implications of Christian love for our question; to consider the relevance of other Christian virtues – such as temperance, diligence, or patience; or to integrate the assumption that in some form or other we will always live on, or that in this sense we cannot really die out.[88] That is beyond the scope of this paper. It would also be interesting to consider the implications of my arguments for issues beyond my core question. Take risks of natural extinction. If we ought to ensure the preservation of the species, we must generally guard ourselves also against asteroid hits. It would presumably be a form of *superbia* to presume that no such event could erase us. And to let it happen would mean to let creation's most valuable part be destroyed. Allowing natural extinction might not be as grave as actively extinguishing ourselves. But it still seems a large-scale moral failure. Another, and more intricate question, concerns the issue of non-human higher intelligence. Suppose we turned into a nun-human AI and left mankind behind. Would we then miss our end (in the non-preservation of the species), or would we precisely fulfil it (in the perfection of our ingenuity)? Would it be a form of *superbia* to thus intervene in the nature of species, or would that belong to our natural proportionate capacities? And would the most valuable aspect of the universe be lost, or would the cosmos become more valuable if inhabited

88 I thank Dominic Roser for mentioning this thought to me.

by a more perfect form of reason?[89] Again, these are questions for another occasion.

There is also the question whether similar arguments emerge on other religious or non-religious worldviews. They arguably do. Our three thoughts – the special status of humanity, an ideal of humility, and a perfectionist obligation to develop our nature – emerge from a very general view of ourselves: from the sense that our nature is singularly valuable, and that this value obliges us to develop our potential and not decide on our existence as we please. This view need not be grounded in the Christian God. Indeed, it need not be grounded in any God. We arguably find such a view of humanity, and forms of our three thoughts, in as secular a thinker as Kant. Kant evidently thought that we are especially valuable: "the human being, and in general every rational being, exists as end in itself, not merely as means to the discretionary use of this or that will."[90] Indeed for Kant, among all material creatures, only human beings are ends in themselves. Thus the physical cosmos would be devoid of intrinsic value altogether if we drove ourselves extinct. Accordingly, the fact that we are ends in ourselves sets limits to what we may do: "If [man] destroys himself in order to flee from a burdensome condition he makes use of a person [i.e., his own person] merely as a means."[91] More generally, our value implies we must not discard ourselves, or risk our demise, without very good reason. Driving ourselves extinct, due to a failure of good care, would manifest an overstepping of our discretion – a lack of humility not so much vis-à-vis God, but vis-à-vis our own dignity. Indeed, Kant thinks we must positively develop our potential: "Man has a duty to himself to cultivate […] his natural powers […], not to leave idle and, as it were, rusting away [his] natural predispositions and capacities."[92] Thus there is a perfectionist, almost natural-law-like strand in Kant's thought. In failing to develop our potential, flourish, or fulfil our nature, we would flout a duty against ourselves. So the general gist of our arguments is not parochial to Christians. It is at bottom simply our dichotomous nature – immensely valuable yet fallen and fragile – which implies we should be exceedingly careful with ourselves. Christians, and everyone with a similar view of mankind, and of our value

89 I thank Mara-Daria Cojocaru and Carin Ism for mentioning these questions to me. For a generally positive Christian stance on enhancement, see e.g. Keenan, "Catholic Christianity."

90 AA IV:428 (*Groundwork*, 45; emphasis removed).

91 AA IV:429 (*Groundwork*, 47).

92 AA VI:444 (*Metaphysics*, 239).

and limitations and obligations, should agree with effective altruists that reducing extinction risks is extremely important.[93]

References

Aquinas, Thomas. *Disputed Questions on Virtue* [DQV]. Translated by Jeffrey Hause and Claudia Eisen Murphy. Indianapolis: Hackett, 2010.

———. *Summa Theologica* [ST], second and revised edition. Translated by Fathers of the English Dominican Province. London: Burns Oates & Washbourne, 1920.

———. *Summa Contra Gentiles* [SCG]. Edited by Joseph Kenny O.P. Translated by Anton C. Pegis, James F. Anderson, Vernon J. Bourke, Charles J. O'Neil. New York: Hanover House, 1955–57.

———. *Scriptum super Sententiis* [SSS]. Accessed 17 September 2019. http://www.corpusthomisticum.org/iopera.html.

Augustine. *The City of God Books XIII and XIV* [De Civ. Dei]. Edited and Translated by Peter G. Walsh. Liverpool: Liverpool University Press, 2017.

Bagemihl, Bruce. *Biological Exuberance: Animal Homosexuality and Natural Diversity*. New York: St Martin's Press, 1999.

Bostrom, Nick, and Milan M. Ćirković, editors. *Global Catastrophic Risks*. Oxford: Oxford University Press, 2008.

Bostrom, Nick. "Astronomical Waste: The Opportunity Cost of Delayed Technological Development." *Utilitas* 15, no. 3 (2003): 308–314.

———. "Existential Risk Prevention as Global Priority." *Global Policy* 4, no. 1 (2003): 15–31.

Beckstead, Nick. "On the Overwhelming Importance of Shaping the Far Future." PhD Dissertation, Rutgers, 2013.

Bucknam, Mark, and Robert Gold. "Asteroid Threat? The Problem of Planetary Defence." *Survival* 50, no. 5 (2008): 141–156.

Cave, Stephen, and Seán S. ÓhÉigeartaigh. "Bridging near- and long-term concerns about AI." *Nature Machine Intelligence* 1 (2019): 5–6.

93 For helpful comments on earlier versions of this paper, I thank the participants of the workshop Religious Perspectives on Effective Altruism, the participants of the 5th Oxford Workshop on Global Priorities Research, as well as Hilary Greaves, Christoph Halbig, Samuel Hughes, Sebastian Muders, Philipp Schwind and Felix Timmermann, and especially Dominic Roser and Ludovico Conti.

Denkenberger, David, and Joshua M. Pearce. *Feeding Everyone No Matter What: Managing Food Security After Global Catastrophe*. London: Elsevier, 2015.

Dietz, Alexander. "Effective Altruism and Collective Obligations." *Utilitas* 31, no. 1 (2019): 106–115.

Finnis, John. *Natural Law and Natural Rights*. Oxford: Oxford University Press, 2011.

Fullam, Lisa. *The Virtue of Humility: A Reconstruction Based in Thomas Aquinas*. Ph.D. Dissertation, Harvard University, 2001.

Haldane, John, and Patrick Lee. "Aquinas on Human Ensoulment, Abortion and the Value of Life." *Philosophy* 78, no. 2 (2003): 255–278.

Hanson, Robin. "Catastrophe, social collapse, and human extinction." In *Global Catastrophic Risks*, edited by Nick Bostrom and Milan M. Ćirković, 363–378. Oxford: Oxford University Press, 2008.

HM Treasury. "Stern Review on the Economics of Climate Change." Accessed 15 May 2019. https://webarchive.nationalarchives.gov.uk/ukgwa/20100407172811/https://www.hm-treasury.gov.uk/stern_review_report.htm.

Kant, Immanuel. *The Metaphysics of Morals.* Translated by Mary J. Gregor. Cambridge: Cambridge University Press, 1991.

Kant, Immanuel. *Groundwork of the Metaphysics of Morals*. Translated by Mary J. Gregor. Cambridge: Cambridge University Press, 1998.

Keenan, James F. "Roman Catholic Christianity, Embodiment and Relationality: Roman Catholic Concerns about Transhumanist Proposals." In *Transhumanism and the Body: The World Religions Speak*, edited by Calvin Mercer and Derek F. Maher, 155–171. New York: Palgrave, 2014.

Linzey, Andrew. *Christianity and the Rights of Animals*. Eugene: Wipf and Stock Publishers, 2016.

Lisska, Anthony J. *Aquinas's Theory of Natural Law: An Analytic Reconstruction*. Oxford: Oxford University Press, 1998.

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

Nussbaum, Martha C. *Creating Capabilities: The Human Development Approach*. Cambridge, MA: Belknap/Harvard University Press, 2011.

———. "Non-Relative Virtues: An Aristotelian Approach." *Midwest Studies in Philosophy* 13, no, 1 (1987): 32–53.

Ord, Toby. *The Precipice: Existential Risk and the Future of Humanity.* London: Bloomsbury, 2020.

Paley, William. *The Principles of Moral and Political Philosophy*. Cambridge: Cambridge University Press, 2014.

Pamlin, Dennis, and Stuart Armstrong. "Global Challenges: 12 Risks that threaten human civilisation." Accessed 8 August 2019. https://static1.squarespace.com/static/59dc930532601e9d148e3c25/t/59f11eebfe54ef5bd76cb38d/1508974382403/12-Risks-that-threaten-human-civilisation-GCF-Oxford-2015.pdf.

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

Pasnau, Robert. *Thomas Aquinas and Human Nature*. Cambridge: Cambridge University Press, 2002.

Pinker, Steven. *Enlightenment Now: The Case for Reason, Science, Humanism, and Progress*. New York: Penguin, 2018.

Richter, Daniel, Rainer Grün, Renaud Joannes-Boyau, Teresa E. Steele, Fethi Amani, Mathieu Rué, Paul Fernandes, Jean-Paul Raynal, Denis Geraads, Abdelouahed Ben-Ncer, Jean-Jacques Hublin, and Shannon P. McPherron. "The age of the hominin fossils from Jebel Irhoud, Morocco, and the origins of the Middle Stone Age." *Nature* 546 (2017): 293–296.

Rushby, Andrew J., Mark W. Claire, Hugh Osborn, and Andrew J. Watson. "Habitable Zone Lifetimes of Exoplanets around Main Sequence Stars." *Astrobiology* 13, no. 9 (2013): 833–849.

Sandberg, Anders, and Nick Bostrom. "Global Catastrophic Risks Survey." *Future of Humanity Institute Technical Report*. Accessed 15 May 2019. http://www.global-catastrophic-risks.com/docs/2008-1.pdf.

Tadie, Joseph Lawrence. *Between Humilities: A Retrieval of Saint Thomas Aquinas on the Virtue of Humility.* Ph.D. Dissertation, Boston College, 2006.

Todd, Benjamin. "Presenting the long-term value thesis." Accessed 15 May 2019. https://80000hours.org/articles/future-generations/.

United Nations, Department of Economic and Social Affairs, Population Division. "World Population Monitoring, 2001: population, environment and development." Accessed 15 May, 2019. https://www.un.org/en/development/desa/population/publications/pdf/environment/population-monitoring.pdf.

Wiblin, Robert, and Howie Lempel. "What skills or experience are most needed within professional effective altruism in 2018? And which problems are most effective to work on? New survey of organisational leaders." Accessed 15 May 2019. https://80000hours.org/2018/10/2018-talent-gaps-survey/.

Robert MacSwain

Are Effective Altruists Saints? Effective Altruism, Moral Sainthood, and Human Holiness

Abstract

This essay explores the relationship between effective altruism, supererogation, and sainthood. In the first section I note Peter Singer's use of the term "saint" in *The Most Good You Can Do* (2015) and argue that it is an allusion to J. O. Urmson's classic article "Saints and Heroes" (1958), which was criticized in Singer's own classic, "Famine, Affluence, and Morality" (1972). Here Singer was primarily concerned to deflect the potential application of Urmson's defence of supererogation to certain "charitable" acts that Singer thought were obligatory, while the Urmsonian use of "saint" passed without comment. But Susan Wolf's influential "Moral Saints" (1982) rejuvenated both the term and Urmson's supererogatory emphasis, arguably requiring Singer in 2015 finally to address both together: effective altruists are not "saints" because their actions are not supererogatory. Singer and Wolf thus agree on the meaning of Urmsonian sainthood but disagree on its evaluation and exemplification. In the second section I revisit Robert Merrihew Adams's well-known response to Wolf, "Saints" (1984), and argue that Adams's concern with "real saints" as opposed to moral abstractions remains valid and opens up helpful lines of further inquiry. In addition to moving away from Urmsonian sainthood, it raises the question of why primarily secular movements such as effective altruism often have such different "hagiological intuitions" from religions such as Christianity: that is, why do the respective exemplars differ so greatly? The third section thus considers Jean Vanier (1928-2019) and his L'Arche movement as exemplars of human holiness that effective altruism would struggle to endorse or even recognize (but also noting how this claim has been undermined by posthumous revelations about Vanier). The brief conclusion raises some further implications of these various arguments about effective altruism, moral sainthood, and human holiness.

Introduction

This essay explores the relationship between effective altruism (EA), supererogation, and sainthood. In the first section, I note Peter Singer's use of the term "saint" in *The Most Good You Can Do* and argue that it is an allusion to J. O. Urmson's classic article "Saints and Heroes," which was criticised in Singer's own classic, "Famine, Affluence, and Morality." Here Singer denied the application of Urmson's general defence of supererogation to certain acts that Singer thought were obligatory, while the Urmsonian use of "saint" passed without comment. But then Susan Wolf's influential "Moral Saints" rejuvenated both the term and Urmson's emphasis on the supererogatory, arguably provoking Singer finally to address both together: effective altruists are not "saints" because their actions are not (in Singer's view) supererogatory. In the second section, I revisit Robert Merrihew Adams's well-known response to Wolf and argue that Adams's concern with "real saints" as opposed to moral abstractions remains valid and opens up

helpful lines of further inquiry. In addition to moving away from moral sainthood, it raises the question of why primarily secular movements such as EA often have such different "hagiological intuitions" from religions such as Christianity: that is, why do the respective exemplars differ so greatly? The third section thus considers Jean Vanier (1928–2019) and his L'Arche movement as exemplars of human holiness that EA would struggle to endorse or even recognise. The brief conclusion notes some further implications of these arguments about EA, moral sainthood, and human holiness.

1. Singer, Saints, and Supererogation

The effective altruist, says Peter Singer in *The Most Good You Can Do*, is not necessarily a saint. To be more precise: "most effective altruists are not saints, but ordinary people like you and me, so very few effective altruists claim to live a fully ethical life. Most of them are somewhere on the continuum between a minimally acceptable ethical life and a fully ethical life."[1] So by "saint" here, Singer apparently means someone who lives a "fully ethical life" – whatever that might be taken to entail. He explains further that "while doing the most good is an important part of the life of every effective altruist, effective altruists are real people, not saints, and they don't seek to maximise the good in every single thing they do, 24/7. As we shall see, typical effective altruists leave themselves time and resources to relax and do what they want."[2] By contrasting "saints" with "real people" Singer is probably being facetious, but maybe not, and the ambiguity opens the genuine question of whether he thinks there actually are "saints" in the sense of those who live "a fully ethical life" and "seek to maximise the good in every single thing they do."[3] Effective altruists may not (all) be saints, but that is because, as morally impressive as they are (at least by Singer's standards), even they do not measure up to the moral stature of sainthood, thus defined. While he does not cite either one in these comments, I think

1 Singer, *Most Good*, viii.

2 Ibid., 8. This book was written for a broad audience rather than for professional philosophers, but is still necessary to engage with as an important defence of EA. For an account of how EA and similar ideas have made a powerful impact on a variety of people, see MacFarquhar, *Strangers Drowning*, published in the UK with the subtitle *Voyages to the Brink of Moral Extremity* and in the USA with *Grappling with Impossible Idealism, Drastic Choices, and the Overpowering Urge to Help*. I will cite from the American edition.

3 Leigh Vicens observed on an earlier draft of this paper that by "saint" in this context Singer seems to refer to a strictly consistent act utilitarian.

that here Singer is implicitly alluding and responding to both J. O. Urmson and Susan Wolf.

1.1. Singer and Urmson: Saints and Supererogation

To begin with, simply in speaking of "saints" in this moral manner, Singer gestures toward an influential tradition in Anglophone analytic philosophy that began in 1958 with Urmson's "Saints and Heroes." Here, Urmson observed that in ordinary usage we "sometimes call a person a saint, or an action saintly, using the word 'saintly' in a purely moral sense with no religious implications." More specifically, he says, the saintly in this broad moral sense involves "resistance to desire and self-interest" whereas the heroic involves "resistance to fear and self-preservation."[4] But Urmson then went on to introduce the primarily-theological concept of *supererogation* into secular ethical theory, and it is precisely the supererogatory that Urmson identifies with genuine sainthood in his distinctive technical sense of the term. Prior to Urmson's paper, most philosophers divided moral actions into three categories: the wrong, the neutral, and the obligatory. Outside of Roman Catholic circles, the idea that there was a potential fourth category, namely good actions that went *beyond* obligation, actions in which an agent did *more* than was morally required of them, was not usually defended. Urmson thus proposed that we henceforth use the religious term "saint" (etymologically, a *holy* person) to refer to those who live lives of supererogatory moral goodness. And from this point onward, this moral, secular use of "saint" became commonplace in Anglophone philosophical discourse to refer to such individuals.[5]

Crucially, however, as we have seen, for Singer, effective altruists are *not* saints in Urmson's distinctive moral sense precisely because EA (at least in Singer's view) is *not* supererogatory. Rather, it is morally required. That is to say, the apparently extreme or radical sacrifices of time and money that EA calls for are in fact obligatory and thus a matter of duty. Those committed to EA are therefore not saints in Urmson's sense. Furthermore, as Singer indicates above, even those who are committed to EA rarely if ever live up to the full implications of their own moral ideals, even as they accept them as obligatory. The obligations of EA thus call for continual moral striving

4 Urmson, "Saints and Heroes," 199–200. In what follows I focus only on Urmsonian sainthood, but the heroic is an important category that requires treatment as well.

5 Urmson's paper generated a substantial secondary literature; for further discussion, see Flescher, *Heroes, Saints, and Ordinary Morality*, and Heyd, "Supererogation."

beyond what most of us are able to accomplish. But if the Urmsonian definition of a saint is a person who lives a life of supererogatory moral goodness, and if such goodness is impossible, then so are saints.[6]

To see the connection with Urmson more clearly, it is important to note that while Singer does not mention him in *The Most Good You Can Do*, Urmson's "Saints and Heroes" is a major focus in Singer's own classic essay, "Famine, Affluence, and Morality" – one of the early sources of EA. "Saints and Heroes" is, in fact, the only recent philosophical text cited by Singer in this essay.[7] While Singer does not here comment on Urmson's moral use of "saint", he fears that Urmson's general defence of supererogatory actions might be applied specifically toward the typical (low) level of charitable donations to humanitarian relief made by citizens of affluent Western countries. Singer says that if one accepts his proposed moral principle – namely, that "if it is in our power to prevent something very bad from happening, without thereby sacrificing anything else morally significant, we ought, morally, to do it" – then "the traditional distinction between duty and charity cannot be drawn, or at least, not in the place we normally draw it."[8]

Further on Singer explicitly denies that ordinary levels of charitable giving in Western societies are supererogatory. However, he describes his article as an "attack on the present distinction between duty and charity,"[9] not necessarily on the distinction itself, which he here leaves as an open

6 I am grateful to Stanley Hauerwas for helping me formulate this point. However, two other commentators on this paper disagreed over whether Singer's developing position over the years in fact allowed for supererogation (and thus sainthood) at all. As one of them wrote, "If it is our duty to live ethically ideally – e.g., to maximise good consequences – there are no actions beyond duty. And if there are no actions beyond duty, there is no supererogation. And if there is no supererogation, there is no sainthood in the sense of acting supererogatorily. There is only sainthood in a different sense: in the sense of acting ethically ideally. Singer can, however, be understood to make room for supererogation by distinguishing between a 'minimally acceptable' and a 'fully ethical' life…. Saints understood as people who act supererogatorily can then be understood as going beyond duty where duty is understood to merely require leading a 'minimally acceptable' life." This tension or ambiguity in Singer's position will return throughout this paper.

7 In addition to Urmson, "Saints and Heroes", and the *Shorter Oxford English Dictionary,* Singer only cites one work each by Henry Sidgwick, Thomas Aquinas, John Kenneth Galbraith, and E. J. Mishan.

8 Singer, "Famine, Affluence, and Morality," 235.

9 Ibid., 235–36.

if dubious question.[10] But the baneful influence of Urmson's argument, he says,

> "may explain the origin and continued existence of the present division between acts of duty and acts of charity. Moral attitudes are shaped by the needs of society, and no doubt society needs people who will observe the rules that make social existence tolerable. From the point of view of a particular society, it is essential to prevent violations of norms against killing, stealing, and so on. It is quite inessential, however, to help people outside one's own society."[11]

Singer continues that while such considerations may explain the "common distinction between duty and supererogation," they do not justify it: "from the moral point of view, the prevention of the starvation of millions of people outside our society must be considered at least as pressing as the upholding of property norms within our society."[12]

I have quoted from this essay at some length to make not just a conceptual but a textual point. While Singer does not use the term "supererogation" in *The Most Good You Can Do*, his earlier suspicions regarding supererogation in "Famine, Affluence, and Morality" remain crucial to his later defence of EA. Thus, in the book's preface, he writes:

> "Effective altruism is based on a very simple idea: we should do the most good we can. Obeying the usual rules about not stealing, cheating, hurting, and killing is not enough, or at least not enough for those of us who have the great good fortune to live in material comfort, who can feed, house, and clothe ourselves and our families and still have money or time to spare. Living *a minimally acceptable ethical life* involves using a substantial part of our spare resources to make the world a better place. Living *a fully ethical life* involves doing the most good we can."[13]

The verbal similarity to the preceding passages from "Famine, Affluence, and Morality" is striking: Singer is clearly drawing on that earlier material to define the more recent EA movement. Moreover, Singer's language of "minimally acceptable ethical life" and "fully ethical life" assumes (rightly or wrongly) both that (i) we can fix the meaning of these terms with some precision and that (ii), *contra* Urmson, we have a duty to be not just ethical but *as ethical as we can be*. The project of EA further argues with

10 Ibid., 238.

11 Ibid., 236–37.

12 Ibid., 237. In addition to the texts cited in note 5, see MacFarquhar's "Duty! Thou Sublime and Mighty Name that Dost Embrace Nothing Charming or Insinuating, but Requires Submission" *Strangers Drowning*, 61–69; the chapter's long title is a quotation from Kant's *Critique of Practical Reason*.

13 Singer, *Most Good*, vii (emphasis added).

yet more precision that, because doing our minimum duty by obeying "the usual rules" is insufficient to be "fully ethical," even commonly-approved forms of altruism such as donating to well-meaning but ineffective charities or supporting the arts are morally compromised. It is not enough to be altruistic, one must be *effectively* altruistic; one must not just do good, one must do *the most good one can*. In other words, in these texts, Singer comes close to the "antisupererogationist" school in which, as Todd May says, "there are essentially no limits to what morality can ask of us."[14] Or, as May says of Singer himself, "it is not that there is no end to my obligations. But they don't end very soon."[15]

1.2. Singer and Wolf: Moral Saints

While Singer's use of "saint" in *The Most Good You Can Do* may allude to the supererogatory Urmsonian definition, it is arguably even more indebted to another classic but much later essay, namely Susan Wolf's "Moral Saints." Here she defines an explicitly *moral* saint as "a person whose every action is as morally good as possible, a person, that is, who is as morally worthy as can be."[16] While Wolf does not cite Urmson's "Saints and Heroes," she is clearly contributing to the Urmsonian tradition of construing saints in secular, moral terms. However, Wolf's distinctive contribution is to then argue famously that "moral perfection, in the sense of moral saintliness, does *not* constitute a model of personal well-being toward which it would be particularly rational or good or desirable for a human being to strive."[17] In other words, rather than commending such maximising moral sainthood, as Singer does, Wolf laments it. As Wolf explains further: "a necessary condition of moral sainthood would be that one's life be dominated by a commitment to improving the welfare of others or of society as a whole."[18] She divides such moral saints into two categories: the utilitarian "Loving Saint" (whose happiness "would truly lie in the happiness of others") and

14 May, *A Decent Life*, 77.

15 Ibid., discussing Kagan, *Limits of Morality*; and 13, discussing Singer, *Most Good*. As May's comments here indicate, and as previously noted in footnote 6, readers of Singer remain at odds over just what he means by a "fully ethical life" and therefore what the detailed implications of being *as ethical as we can be* entail in practice.

16 Wolf, "Moral Saints", reprinted in Wolf, *Variety of Values*, 11–29: I will cite from this reprinted version, and this quotation is from 11.

17 Wolf, *Variety of Values*, 11 (emphasis added).

18 Ibid., 12.

the Kantian "Rational Saint" (who "sacrifices his own interests to the interests of others, and feels the sacrifice as such").[19]

Wolf's concern with moral sainthood is that an exclusive focus on morality is "apt to crowd out the nonmoral virtues, as well as many of the interests and personal characteristics that we generally think contribute to a healthy, well-rounded, richly developed character."[20] For the moral saint all nonmoral excellences, such as literature, music, art, gourmet food, wit (especially sarcasm and satire), games, and athletic ability, must be measured "against possible alternative beneficent ends to which these resources [i.e., of time, energy, and money] might be put"; therefore, engaging in such nonmoral activities requires "the decision *not* to justify every activity against morally beneficial alternatives, and this is a decision a moral saint will never make."[21] And, vis-à-vis both Urmson and Singer, Wolf notes that if "we have reason to want people to live lives that are not morally perfect, then any plausible moral theory must make use of some conception of supererogation."[22] Just as Singer fears, then, Wolf insists that supererogation is what rightly separates the rest of us from moral saints: we need to know when we have done enough morally and feel free to stop there. There are genuine goods other than moral goods, and we cannot allow the legitimate but limited claims of morality to dominate our lives.[23]

Singer does not cite Wolf in *The Most Good You Can Do*, but it is difficult to think that he was not alluding to her "Moral Saints" when he said that, *unlike* saints, effective altruists "don't seek to maximize the good in every single thing they do" and "leave themselves time and resources to relax and do what they want."[24] These descriptions have a suspiciously Wolfian character. By explicitly disavowing the label of "saint" for effective altruists, it seems that Singer is claiming implicitly that they are not subject to Wolf's

19 Ibid.

20 Ibid., 12–13.

21 Ibid., 14 (emphasis in the original).

22 Ibid., 28.

23 Like the essays by Urmson and Singer, Wolf's "Moral Saints" has generated an enormous amount of commentary: for a particularly searching critical analysis, see Carbonell, "What Moral Saints Look Like." By contrast, May's *A Decent Life*, largely written in response to Singer, *Most Good*, sides with Wolf against Singer and Kagan: what is morally required of us is *decency*, not supererogation, and by "decency" May understands a substantial but limited moral commitment to help others. See also MacFarquhar, *Strangers Drowning,* 6–7.

24 Singer acknowledges her argument elsewhere and describes it as a "forceful article" (*Practical Ethics,* 212–13).

critique of moral sainthood. Singer and Wolf thus apparently agree on the *meaning* of "saint" in this secular philosophical context but differ radically on its *evaluation* and *exemplification*. Singer clearly approves of those who seek to live "a fully ethical life" even if most effective altruists do not reach that lofty ("saintly") standard, while Wolf regards such an enterprise as inhumane and oppressive, even a type of moral fanaticism.[25] That is, unlike Singer, Wolf believes in supererogation, so she can also believe in saints, at least moral ones, even if she disapproves of them.

Likewise, and for the same reason, Singer and Wolf apparently agree on the basic meaning of supererogation but evaluate it differently. As I said earlier, for Singer effective altruists are *not* saints in Urmson's moral sense precisely because EA is *not* supererogatory, and indeed as already noted for Singer supererogation itself may be an "empty category."[26] However, Singer's protests notwithstanding, I think that Wolf would regard the EA movement as indeed exemplifying a form of supererogatory moral sainthood or at least striving toward it in a worrisome manner. Like her moral saints, effective altruists are "dominated by a commitment to improving the welfare of others or of society as a whole" and "justify every activity against morally beneficial alternative." Although EA formally began in 2009 and Wolf's essay was first published in 1982, you could not ask for a better brief description of EA than these two phrases.

In this section, I have sought to trace the technical definition of "saint" in Anglophone moral philosophy through three authors: J. O. Urmson, Susan Wolf, and Peter Singer. For Urmson, saints are those who perform supererogatory acts, that is, acts that are morally commendable but not required. For Wolf, this Urmsonian definition is taken further but then criticised: moral saints are those "whose every action is as morally good as possible," persons who are "as morally worthy as can be"; but while such extreme commitment is indeed supererogatory, it is not commendable and should thus be avoided. And as for Singer, he seems to accept Wolf's definition of sainthood, but to deny that it ever obtains, either because no one, in fact, lives up to it, or because it is impossible to exceed our moral duties and thus act supererogatorily. To again quote from *The Most Good You Can Do*, "most effective altruists are not saints, but ordinary people like you and me, so very few effective altruists claim to live a fully ethical

25 Wolf, *Variety of Values*, 16.

26 I am indebted to Kian Mintz-Woo for this apt phrase in this context.

life. Most of them are somewhere on the continuum between a minimally acceptable ethical life and a fully ethical life."[27]

2. "The Actuality of Sainthood"

The previous section thus raises two connected questions but leaves them unanswered: (1) *Are effective altruists saints?* and (2) *Is EA supererogatory?* Singer would reply "no" to both, whereas Wolf would reply that effective altruists are at least aspiring toward moral sainthood and that in so doing both they and their movement are committed to supererogation. For both Singer and Wolf, saints (if they exist) exemplify supererogatory altruism. However, one of my primary goals in this essay is to rescue sainthood from its Urmsonian and Wolfian captivity by moral philosophy and hence from an exclusively secular, ethical, and supererogatory context. As noted earlier, the original meaning of "saint" is not a *moral* person but a *holy* person, and while morality and holiness are intimately connected, they are not identical. I thus now wish to enlist Robert Merrihew Adams against Urmson, Singer, and Wolf, and then in the following section explore the potential implications of human holiness for both EA and religious belief.

In his classic response to Wolf's "Moral Saints," Adams dissents from her negative understanding of sainthood and presents what he thinks is the more complicated, interesting, and compelling reality.[28] Adams asserts: "the first thing to be said is that there *are* saints – people like St Francis of Assisi and Gandhi and Mother Teresa – and they are quite different from what Wolf thinks a moral saint would be."[29] Adams grants that while such figures "are not exactly *moral* saints in Wolf's sense", nevertheless she "writes about some of them as if they were, and discussions of moral sainthood surely owe to the real saints much of their grip on our attention. So it will be to the point to contrast the actuality of sainthood with Wolf's picture of the moral saint."[30] It is Adams's insistence on the "actuality of sainthood" that I wish to lift up in this context. And in "Moral Saints" Wolf indeed mentions Francis, Gandhi, and Teresa as examples of what she has in mind.[31] But

27 Singer, *Most Good,* viii.

28 Adams, "Saints," reprinted in Adams, *Virtue of Faith,* 164–73: citations of this essay will be from the reprinted version. Adams is Presbyterian rather than Roman Catholic, Orthodox, or Anglican, so his interest in these themes is less typical of his tradition.

29 Adams, *Virtue of Faith*, 164.

30 Ibid.

31 See, Wolf, *Variety of Values,*14 (Francis), 23 (Mother Teresa), and 26, note 4 (Gandhi).

against Wolf's generic depiction of moral saints as unattractive, lacking in individuality, and "dull-witted or humorless or bland", Adams simply says, "the real saints are not like that."[32] Rather than lacking individuality, saints are often eccentric, and rather than being "very, very nice" saints are often "very, very truthful" – even when that truthfulness causes trouble for themselves or others.[33] Real saints, says Adams, "may not enjoy all the same things as other people, and perhaps a few of them have been melancholy, but an exceptional capacity for joy is more characteristic of them." Indeed, he continues, "there are joys (and not minor ones) that only saints can know. And as for attractiveness, the people we think of first as saints were plainly people who were intensely interesting to almost everyone who had anything to do with them, and immensely attractive to at least a large proportion of those people."[34]

Adams acknowledges that real saints are often "devoted to improving the lives and circumstances of other people" and that there is indeed "unusual moral goodness in the saints," but he denies that ethical or political concerns are their primary motivation, for (he says) "sainthood is essentially a religious phenomenon, and even so political a saint as Gandhi saw his powerful humanitarian concern in the context of a more comprehensive devotion to God."[35] What Adams calls the "substance of sainthood" is not "sheer willpower striving…to accomplish a boundless task" but "goodness overflowing from a boundless source. Or so, at least, the saints perceive it."[36] In a key passage, Adams argues:

> "If sainthood is essentially a religious phenomenon…it is reasonable to seek its central feature (at least for theistic religions) in the saint's relation to God. 'Saint' means 'holy' – indeed, they are the same word in most European languages. Saints are people in whom the holy or divine can be seen. In a religious view, they are people who submit themselves, in faith, to God, not only loving him but also letting his love possess them, so that it works through them and shines through

32 In comments on an earlier version of this essay, Andrew Thompson asked, "What is Adams's definition of a 'real saint' – is it sort of inductive, like his idea of Good in *Finite and Infinite Goods*? It seems to be, where Wolf's definition seems to be primarily deductive. And, in the end, there doesn't seem to be much overlap at all." I will address these important questions and concerns all too briefly. For Adams, *Finite and Infinite Goods*, see note 49 below.

33 For these specific charges, see, Wolf, *Variety of Values*, 14.

34 All citations from Adams, *Virtue of Faith*, 165.

35 Ibid., 168.

36 Ibid.

them to other people. What other interests a saint may have will then depend on what interests God has, for sainthood is a participation in God's interests."[37]

While there is much to say about how Adams's religious view of sainthood contrasts with the secular moral concept shared broadly by Urmson, Singer, and Wolf, let me here simply state that Adams's essay offers another category, which we may call the charismatic saint.[38] Singer denies that effective altruists are necessarily saints, whereas Adams denies that saints are necessarily effective altruists. One need not be an effective (or ineffective!) altruist to be a saint. For Adams, saints are not characterised solely or even primarily by altruistic interests and activities, supererogatory or otherwise, but are joyful, attractive, interesting, compelling, distinctive and even eccentric individuals possessed by divine love who radiate "unusual moral goodness" and compassion embodied through their very person or self, not simply in their actions. Against Wolf's concern that moral saints must "justify every activity against morally beneficial alternatives" – a concern that she would also have about EA – Adams argues crucially that, while sainthood is indeed "participation in God's interests," God's own interests extend beyond morality or even religion.[39] If, in the memorable words of Hugh Lister, God is not a "religious maniac," then God is not a "moral fanatic" either.[40] Adams thus suggests that in addition to more conventional figures we might also consider Fra Angelico, Johann Sebastian Bach, and perhaps even Vincent Van Gogh as saints. In short, the charismatic saint is not simply moral, and perhaps not even moral in any conventional sense, but *holy*.[41]

3. Real Saints, Exemplars, and Hagiological Admiration

The exchange between Wolf and Adams is well-known in discussions of moral sainthood, but I suggest that it has acquired new significance in light of the recent rise of EA and the various debates that EA has provoked in secular and religious circles, as well as within both philosophy and

37 Ibid., 170.

38 I take the term but not the meaning from Coleman, "Transgressing the Self." Coleman's essay is about the Pentecostal and Charismatic movements of American Christianity; I mean "charismatic" in a more generic sense. See also Flescher, *Heroes, Saints, and Ordinary Morality*, 209.

39 Adams, *Virtue of Faith*, 170.

40 Cited by Hein, "Farrer on Friendship," 144 (note 20). See, Wolf, *Variety of Values*, 16.

41 See, Adams, *Virtue of Faith*, 170. For some philosophical implications of both divine and human holiness, see MacSwain, "Holiness" (forthcoming).

philanthropy.[42] Wolf's arguments can be and already have been applied to at least some versions of EA, and Singer is arguably responding to her critique in his denial that effective altruists are necessarily saints. But the EA-related renewal of interest in Wolf's "Moral Saints" means that Adams's reply to her is also newly *au courant*. Moreover, Adams helps us see that, despite their different evaluations of such sainthood, Singer and Wolf still share a single concept of it. Ironically, this means that even Wolf with her strong interest in defending a variety of values, including nonmoral ones, is apparently unwilling to consider a nonmoral aspect of sainthood itself, namely the holy or sacred. This is where Adams, as a Christian philosopher, helps shift the debate about sainthood out of a purely secular plane to where any legitimate discussion of saints arguably belongs, the domain of the divine.[43]

By contrast, Vanessa Carbonell agrees with Adams that "real saints" are not like Wolf's unattractive depiction, but she prefers to "challenge Wolf's thesis about the attractiveness of moral saints without departing from her underlying conception of a moral saint as a *truly extraordinary moral agent,* an agent whose uncommon qualities and achievements are not essentially religious."[44] "Adams thinks the solution is to return sainthood to its religious roots", Carbonell says, but "the sort of saint Adams describes is not necessarily a *moral* saint. While various historical examples of religious saints might turn out to be moral saints as well, moral sainthood is something that we can describe without reference to the holy or the divine."[45] As noted above, Adams grants this point in advance, but also defends his

42 For a secular philosophical analysis of sainthood engaging with Wolf, Adams, and many others that I read after finishing this paper, but which contains much that I agree with, see the prologue to Flanagan, *Varieties of Moral Personality*, 1–12. Ironically, and for many controversially, Singer himself is here regarded as a saintly figure (6–7)! See chapter 4 of Flescher, *Heroes, Saints, and Ordinary Morality,* for an account of sainthood I am in large agreement with as well: my main difference from him is that, like Flanagan, he remains on the secular plane. An important treatment of moral sainthood that I do not discuss in this essay due to its different philosophical idiom is Wyschogrod, *Saints and Postmodernism*; her Levinasian view of sainthood influenced Flescher's account.

43 Another reason to engage with Adams on saints is that they are still surprisingly neglected by philosophers, and even philosophers of religion. For two exceptions, see Sherry, "Philosophy and the Saints," and Cottingham, "Saints and Saintliness."

44 Carbonell, "What Moral Saints Look Like," 376 (emphasis in the original).

45 Ibid., 376, note 1 (emphasis in the original). As her "attractive counter-example" of a real moral saint, Carbonell proposes the altruistic physician, anthropologist, and activist Paul Farmer.

approach by observing that (i) Wolf mentions religious saints in her essay and (ii) "discussions of moral sainthood surely owe to the real saints much of their grip on our attention" – where "real saints" are taken to mean exemplary figures in whom, as Adams puts it, "the holy or divine can be seen" even if they are not conventionally religious themselves. While endorsing Carbonell's approach to "real *moral* saints" as worthwhile, and even convincing simply as a response to Wolf, I think she evades Adams's broader proposal too quickly. Why, after all, did the moral, philosophical, Urmsonian meaning of "saint" find such wide appeal to begin with if not for the strong positive association, moral and otherwise, with religious saints? At the very least, there is something parasitic in the philosophical idea of moral sainthood that even now still depends on religious sainthood for much of its resonance and power.

But what is a "real saint"? How are they defined? How are they recognised? What is holiness, and what, indeed, is divinity? To answer these questions fully would of course take us far beyond the confines of this essay, but at least an initial answer has already been made. Religious traditions and communities have always had exemplary figures whose function we might call canonical. Some of these figures have this status in terms of formal institutional acknowledgement; for others it is more a matter of popular devotion or esteem.[46] So when it comes to "real saints" my first point is that such individuals are not necessarily moral saints in Wolf's maximising sense. In other words, to introduce yet another category of sainthood beyond the moral and charismatic, they might be what Jean Porter has called a "flawed saint." Her example is Martin Luther King, Jr, of whom she says, "whatever else he may teach us, he helps us to see that to be a saint is not to be morally perfect, but to be exemplary in love."[47]

My second point is that in seeking to come to grips with real saints we should focus our attention more on what I have called charismatic saints. In their introduction to *Saints: Faith Without Borders*, Françoise Meltzer and Jaś Elsner associate saints with excess, marginality, and even transgression: "there is a fundamentally apophatic aspect to the excess that is embodied in the saint. Excess of this kind has been called grace, miracle, mysterious

46 See, Six-Means, "Saints and Teachers." This claim does not address all of Thompson's question about defining saints in note 32, but is where I would begin: namely, with the recognition of saints in actual communities of faith and practice, and so inductively rather than deductively.

47 See, Porter, "Virtue and Sin," 539. As I discuss briefly in note 65 below, it is possible that Jean Vanier might fit within this category as well, but that remains to be seen.

energy, charisma, jouissance, the Real."[48] And even the goodness of such charismatic saints can be deeply troubling and counter-intuitive. In his study *Finite and Infinite Goods*, Adams writes:

> "What theists most admire, even within human life, is not easily identified as normal and natural. The saints are occasionally flourishing specimens of humanity in all the obvious ways, but usually not. They feel compunctions, and are torn by doubts, that others do not know. They cling to a longing for goods that lie beyond any ordinary human possibility of attainment. Their behaviour seems occasionally bizarre, and often unnatural. Kissing lepers and praying for people who are torturing you definitely do not seem natural responses. Indeed, I think it is often precisely in the unnaturalness, or better the supernaturalness, of the saints' goodness that their wonderfulness is found. In them it seems that we can glimpse or touch a goodness that is more than theirs, a superhuman, a transcendent goodness that we too long for."[49]

Rather than Singer's "moral point of view", charismatic saints take the *divine* point of view, which ironically enough is not necessarily identical with what philosophers sometimes call the "God's eye point of view" of ethics.[50] By the divine point of view I mean not only that, as Adams said earlier, "sainthood is a participation in God's interests," but also that, rather than heeding an abstract and universal sense of duty or justice, saints simply do what they are uniquely called to do, whatever that may happen to be. Charismatic sainthood is thus less tied to supererogation as it is to *vocation*.[51] Saints are not called to go above and beyond the call of duty, they are called to be themselves. In the famous words of Thomas Merton, "for me to be a saint means to be myself. Therefore, the problem of sanctity and salvation is, in fact, the problem of finding out who I am and of discovering my true self."[52] To be sure, we are all called to be ourselves,

48 See their introduction to *Saints: Faith Without Borders*, ix–x. The literature on saints in theology and religious studies is of course vast, but for one helpful analysis see Brown, *Discipleship and Imagination*, 62–101. Despite the title, Orsuto, "The Saint as Moral Paradigm," takes a broader, more theological view than the purely ethical.

49 Adams, *Finite and Infinite Goods*, 52–53. Such saintly goodness may indeed be supererogatory, but as Adams notes it is also super*humanly* transcendent. See also, Flescher, *Heroes, Saints, and Ordinary Morality*, 174.

50 See Taliaferro, "A God's Eye View," and Anderson, "God's Eye Point of View."

51 Vocation is of course a hugely complex and contested concept, almost as vexed as supererogation and sainthood. Chapter 13 in Adams, *Finite and Infinite Goods*, 292–314, focuses on this topic, partly based on an earlier article, Adams, "Vocation." Heyd considers the relation between supererogation and vocation his article cited in note 5, saying that connecting them goes back to Aquinas.

52 Merton, *New Seeds of Contemplation*, 31.

but saints are those who actually solve this problem, discern and heed that divine call, and fulfil it in ways that provoke admiration, astonishment, and awe.

My third and final point about real saints is to agree with Larissa MacFarquhar when she writes: "in the abstract, there are ideas about saints and perfection. Only actual lives convey fully and in a visceral way the beauty and cost of a certain kind of life."[53] Hence her book *Strangers Drowning*, which explores the theme of altruism, both effective and otherwise, through a series of case studies. Ironically, this approach is something that Singer also understands quite well. Thus, *The Most Good You Can Do* has an explicitly hagiographic character. Chapter 1 begins by telling the story of Matt Wage, a gifted Princeton philosophy student who decided to forgo graduate studies at Oxford and a potential academic career in order to work on Wall Street so that he could give away considerably more money than otherwise. Having presented Matt as an inspiring exemplar of EA, Singer promises that in the book "we will meet people" who have done similar things: Zell Kravinsky, Toby Ord, Will MacAskill, Julia Wise, Aaron Moore, Bernadette Young, Rhema Hokama, and many others.[54] While disavowing the label "saint" for these exemplars of EA, they do in fact function in Singer's narrative argument as canonical figures whose lives provide both material content and inspirational justification for the tenets of EA.

Exemplars are currently a major topic of interest in philosophy, theology, and anthropology. According to the philosopher Linda Trinkaus Zagzebski, "exemplars not only show us what morality is, but they make us want to be moral and they show us how to do it."[55] Crucial to her theory of exemplarity is the experience of admiration refined by critical reflection as a means of moral knowledge.[56] But Singer's list of admired EA exemplars raises the crucial question of contrasting values between the strongly rationalist-consequentialist mindset of secular EA and traditional religious morality – as found, for example, in Christianity – and thus of different canonical figures as well. For just one Christian example, consider Jean Vanier, the founder of L'Arche, who died on 7 May 2019. A lay Roman Catholic trained in philosophy, Vanier was widely revered during

53 Citation from MacFarquhar, *Strangers Drowning*, 11.

54 Singer, *Most Good*, 3–4, and *passim*. Some of Singer's exemplars are also included in *Strangers Drowning*.

55 Zagzebski, *Exemplarist Moral Theory*, 129. For a similar proposal from an anthropologist, see Robbins, "Where in the World Are Values?"

56 Ibid., Chapter 2.

his lifetime for placing a profoundly high value on and exhibiting a deep respect for those with cognitive disabilities. He created radically hospitable communities in which such individuals lived as full members alongside those without such disabilities. In Zagzebski's *Exemplarist Moral Theory*, she offers three types of exemplar, the Hero, the Sage, and the Saint, identifying saints primarily with compassion, and Vanier is her exemplar of the Saint. Her book concludes as follows:

> "Vanier has spent the greater part of his life centred on bringing love and joy to people who have it least. He is an exemplar of a joyful moral life. I have tried in this book to show a way that we can map the moral domain around people like him, while enjoying the process of making ourselves more like the exemplars who inspire us most."[57]

Likewise, the theologian Stanley Hauerwas has written extensively on the significance of Vanier's life and work not only for the moral life but for belief in God.[58] The witness of Vanier and L'Arche, says Hauerwas, helps us to see that we are all "creatures drawn into a kingdom of patience making possible our friendship with God and one another."[59] As inhabitants of this kingdom of patience, Christians have learned that "there is nothing more significant to be done in a world of such deep injustice than to take the time to be friends with the handicapped. I know of no better vision of peace."[60]

But from the perspective of Singer and secular EA (recognising that there are religious versions that modify it in various ways[61]), Vanier cannot be admired and thus recognised as a moral exemplar, let alone a saint; L'Arche cannot be commended as a worthwhile undertaking or investment for those who seek to do the most good they can; and our world of deep injustice cannot allow us legitimately to take time (let alone a lifetime) for such ineffective activities as helping someone with Downs' Syndrome put together a puzzle or a person with physical disabilities brush their hair when we could be doing much more good elsewhere. Even if living at a L'Arche community is "good," it is not doing "the most good you can do," and

57 Ibid., 235.

58 See, for example, Hauerwas, "Timeful Friends," and Hauerwas, "God and Goodness." See also the book co-authored by Hauerwas and Vanier, *Living Gently in a Violent World*.

59 Hauerwas, "Reflection on Dependency," 193.

60 Ibid., 197. For an earlier engagement with this material, see MacSwain, "Imperfect Lives and Perfect Love."

61 For example, EA for Christians (https://www.eaforchristians.org).

therefore those who commit themselves to it are not seeking to live "a fully ethical life." As the populariser of the term "speciesism," Singer does not take it for granted that humans – or, for that matter, creatures of any other species – have inherent value. The correct value ascribed to a human being depends on factors such as cognitive ability, agency, or ability to suffer. Thus, according to Singer, it remains an open question of how valuable the lives of people with disabilities are, especially if they score low according to criteria such as cognitive ability. To regard the cognitively disabled as in fact deserving *special* esteem and *privileged* treatment is, therefore, a sentimental fallacy of the first order.[62]

So, what accounts for these vast differences of evaluation between Singer on the one hand and Zagzebski and Hauerwas on the other? Why do they have such different "hagiological intuitions" and thus admire such divergent exemplars? There are, of course, no simple answers to these questions, but part of the difference surely has to do with moral vision. Christian thinkers such as Zagzebski and Hauerwas do not simply *think* differently than Singer, they *see* the world in a radically different way, and thus *inhabit* it differently as well. As Hauerwas insists, such vision requires training, and such training requires learning proper description. One of his maxims, inspired by Iris Murdoch, is "You can only act in the world you can see and you can only see what you have learned to say."[63] Vanier, Hauerwas would say, is one of those who has taught us to describe and see the cognitively disabled properly, and therefore to see the world properly as well. Indeed, anthropologist Patrick McKearney goes so far as to suggest that the genuine saints of L'Arche are the cognitively disabled residents themselves. He writes:

> "I show that carers in L'Arche are trained to recognise those with cognitive disabilities not as humans who lack autonomous moral agency, but as people who inhabit the world in a different way. That is, they learn to see people with cognitive

62 I am grateful to the editors for helping me formulate Singer's position here more fairly than in my original draft. For a probing and sensitive engagement on this theme, see Hopwood, "Terrible Purity." Note, however, that when Peter Singer's mother developed advanced Alzheimer's, Singer, in violation of his theories about both giving and personhood, spent a lot of money paying nurses to take care of her. "Perhaps it is more difficult than I thought before," he said, "because it is different when it's your mother" (MacFarquhar, *Strangers Drowning*, 101).

63 Hauerwas says this in various places; one example is Hauerwas, *Work of Theology*, 26.

disabilities as surprising, eccentric and charismatic agents who can powerfully affect the ethical lives of others."[64]

I know of no better description of saints.[65]

4. Conclusion

The premise of this volume is that EA has secular roots and yet is of profound religious significance, such that religious practitioners should ask themselves both what they can learn from EA and what within it might require critique or resistance. I have personally been challenged by EA to reconsider some of my basic religious and ethical convictions, and I am still in the process of doing so.[66] But in regard to the primary question of this essay, *Are effective altruists saints?*, my answer is that *it depends*. It

64 McKearney, "Receiving the Gift of Cognitive Disability," 41. See also the section aptly titled "Exemplars of intuition and charisma," 53–55.

65 On 22 February 2020, L'Arche International released a summary report of a posthumous investigation concluding that there was "sufficient evidence to establish that Jean Vanier engaged in manipulative sexual relationships with at least 6 adult (not disabled) women" over a thirty-year span. This disturbing revelation obviously requires rethinking much of the material cited above, and in particular the exemplary use of Vanier made by Zagzebski and Hauerwas (along with many others not cited here). The implications of Vanier's moral fall are immense and remain to be fully understood. His actions were deeply wrong, harmful, and inconsistent with his own articulated moral vision and values. And yet these tremendous failings were compatible with him also accomplishing immense good on behalf of the physically and mentally disabled and changing the way that thousands of people think of them and their place in human society. It is too soon to begin the discussion, but at some point we will have to inquire whether Vanier might fit along with Martin Luther King Jr under Porter's "flawed saint" category, or if even that is insufficient to account for his failings. Given that this essay was written before and yet published after these revelations, all I can do at the moment is add this footnote to acknowledge this new information, and yet also point to McKearney's important distinction between Vanier himself (including his personal writings and teachings) and the exemplary life of the L'Arche communities he founded, specifically the virtues of the carers and residents.

66 Some of the questions EA has raised for me are: (1) What is the moral status of the traditional "helping professions" compared to EA's "earning to give"? That is, is it better to do good yourself or pay others to do good on your behalf? (2) What is the difference between "being good" (character) and "doing good" (utility), and which if either is better? (3) Is it morally permissible to financially support religious institutions? That is, even for those who are faithful believing members of a religious community, should they give money to their church, synagogue, mosque, or temple, or to one of the EA-approved charities instead? And (4) How much more should I donate to charity than I currently do, and should I change the focus of my giving in a more EA direction?

first depends on what we choose to mean by "saint" and I do not wish to be bound by the moral categories of Urmsonian or Wolfian sainthood.[67] But if I am correct that saints properly understood do what they are uniquely called to do, whatever that may happen to be, then the answer to this question also depends on such vocational discernment. Some effective altruists may well indeed be saints, but not necessarily because they are effective altruists. It depends on what they are called to be and do by God (whether they believe in God or not) given their particular circumstances, history, abilities, and all the other contingent factors that go into a vocation. If we inhabit Hauerwas's divine "kingdom of patience," then "they also serve who only stand and wait" (Milton), in which case any number of "ineffective" roles may serve as paths of sainthood, including living in a L'Arche community.

In regard to the secondary question of the essay, *Is EA supererogatory?*, my answer is also that it depends. Andrew Michael Flescher has argued that supererogation is a flexible phenomenon, and that the line of duty varies from person to person and even changes for the same individual over the course of their lifetime, depending on their growth in moral development as well the contingent factors mentioned above. There is not a fixed objective standard of duty that is the same for all people at all times. Thus, contrary to Urmson, Singer, and Wolf, saints may indeed exist and yet *not* exemplify supererogation, if what they do is not supererogatory *for them*, given their distinct charism and particular vocation. So, it may well be the case that for some people, EA is supererogatory while for others it is obligatory.[68] This is thus also a matter of individual discernment. If so, then my lingering concern with EA is less with its vision of the good and more with its vision of human nature.

Making the meaning of sainthood and the demands of supererogatory altruism matters of individual discernment does not, of course, evade the challenging conceptual and practical questions canvassed in this essay.[69]

67 In short, I do not think that being a saint means being a strict act utilitarian, a subjective consequentialist, or either Wolf's Loving Saint or Rational Saint.

68 See Flescher, *Heroes, Saints, and Ordinary Morality*, esp. Chapters 2, 4, and 5; MacFarquhar's *Strangers Drowning* corroborates Flescher's argument, with many of those she interviews insisting that their extreme altruism is obligatory for them.

69 I am working on a monograph currently titled *The Saint is Our Evidence: The Hagiological Argument for the Existence of God* which explores the idea that saints provide evidence for God and perhaps even the best evidence. One version of the hagiological argument construes saints as supreme, radical, or excessive altruists, so there is a connection to EA to be explored further there as well.

What, in fact, is the relationship between supererogation and altruism? How much good are we as unique individuals truly *obliged* to do, and at what point, if any, can we say that we have gone beyond the call of our own particular duty? Is the good we are *called* to do necessarily the same as the most good we *can* do? Should we, in fact, even attempt to live a "fully ethical life" and if so what would it look like? What would motivate us to embrace such radical altruism, effective or otherwise, in the first place? And finally, what does it mean for us to "discover our true selves" so that we may become saints as well? In whatever way all of these questions are ultimately answered, the example of those we call saints will assist us in doing so.[70]

References

Adams, Robert Merrihew. "Saints." *Journal of Philosophy* 81 (1984): 392–401.

———. *The Virtue of Faith and Other Essays in Philosophical Theology*. New York: Oxford University Press, 1987.

———. "Vocation." *Faith and Philosophy* 4 (1987): 448–62.

———. *Finite and Infinite Goods: A Framework for Ethics*. New York: Oxford University Press, 1999.

Anderson, Pamela Sue. "What's Wrong with the God's Eye Point of View: A Constructive Feminist Critique of the Ideal Observer Theory." In *Faith and Philosophical Analysis: The Impact of Analytical Philosophy on the Philosophy of Religion*, ed. Harriet A. Harris and Christopher Insole, 85–99. Aldershot: Ashgate Publishing, 2005.

70 A large number of people offered comments on earlier versions of this essay and I am very grateful to them all: David Brown, Ben Cowgill, Cynthia Crysdale, Stanley Hauerwas, David Hein, Mark Hopwood, Ann Loades, Kian Mintz-Woo, Stefan Riedener, Dominic Roser, Andrew Thompson, and Leigh Vicens. I am likewise grateful to Dominic Roser and the late Markus Huppenbauer for inviting me to present it at the "Religious Perspectives on EA" workshop at the University of Fribourg, Switzerland (3–4 September 2019) and to the other workshop participants for additional comments, criticism, and conversation during the event. In addition to my home institution of the University of the South (USA), I need to thank one other academic institution and two other sources of funding. Some of this research occurred during my 2016–17 sabbatical as a visiting scholar at Harvard Divinity School, funded in part by a Post-Doctoral Faculty Fellowship from the Appalachian College Association. However, most of the research as well as my attendance at the workshop in Fribourg was made possible through the support of a 2019–20 grant from Templeton Religion Trust. The opinions expressed in this publication are those of the author and do not necessarily reflect the views of Templeton Religion Trust.

Brown, David. *Discipleship and Imagination: Christian Tradition and Truth*. Oxford: Oxford University Press, 2000.

Carbonell, Vanessa. "What Moral Saints Look Like." *Canadian Journal of Philosophy* 39 (2009): 371–98.

Coleman, Simon. "Transgressing the Self: Making Charismatic Saints." In *Saints: Faith Without Borders*, ed. Françoise Meltzer and Jaś Elsner, 73–95. Chicago and London: University of Chicago Press, 2011.

Cottingham, John. "Saints and Saintliness." In *The Oxford Handbook of the Epistemology of Theology*, ed. William J. Abraham and Frederick D. Aquino, 79–96. Oxford: Oxford University Press, 2017.

Flanagan, Owen. *Varieties of Moral Personality: Ethics and Psychological Realism*. Cambridge, MA: Harvard University Press, 1993.

Flescher, Andrew Michael. *Heroes, Saints, and Ordinary Morality*. Washington, DC: Georgetown University Press, 2003.

Hauerwas, Stanley. "Timeful Friends: Living with the Handicapped." In his *Sanctify Them in the Truth: Holiness Exemplified*, 143–56. Edinburgh: T&T Clark, 1998.

———. "Reflection on Dependency: A Response to Responses to My Essays on Disability." In *Critical Reflections on Stanley Hauerwas' Theology of Disability: Disabling Society, Enabling Theology*, ed., John Swinton, 191–97. Binghamton, NY: The Haworth Press, 2004.

———. "God and Goodness: A Theological Exploration." In his *Working with Words: On Learning to Speak Christian*, 61–75. Eugene, OR: Cascade Books, 2011.

———. *The Work of Theology*. Grand Rapids. MI: Eerdmans, 2015.

Hauerwas, Stanley and Vanier, Jean. *Living Gently in a Violent World: The Prophetic Witness of Weakness*. Intro. John Swinton. Downers Grove, IL: InterVarsity Press, 2008.

Hein, David. "Farrer on Friendship, Sainthood, and the Will of God." In *Captured by the Crucified: The Practical Theology of Austin Farrer*, ed. David Hein and Edward Hugh Henderson, 119–48. New York: T&T Clark International, 2004.

Heyd, David. "Supererogation" *The Stanford Encyclopedia of Philosophy*, (Fall Edition), ed. Edward N. Zalta, 2019. https://plato.stanford.edu/archives/fall2019/entries/supererogation/.

Hopwood, Mark. "Terrible Purity: Peter Singer, Harriet McBryde Johnson, and the Moral Significance of the Particular." *Journal of the American Philosophical Association* 2 (2016): 637–55.

Kagan, Shelly. *The Limits of Morality*. Oxford: Oxford University Press, 1991.

MacFarquhar, Larissa. *Strangers Drowning: Grappling with Impossible Idealism, Drastic Choices, and the Overpowering Urge to Help*. New York: Penguin Press, 2015.

MacSwain, Robert. "Imperfect Lives and Perfect Love: Austin Farrer, Stanley Hauerwas, and the Reach of Divine Redemption." In *Exchanges of Grace: Essays in Honour of Ann Loades*, ed. Natalie K. Watson and Stephen Burns, 142–54. London: SCM Press, 2008.

———. "Holiness." *Encyclopedia of Philosophy of Religion*, ed. Stewart Goetz and Charles Taliaferro. Wiley (forthcoming).

May, Todd. *A Decent Life: Morality for the Rest of Us*. Chicago and London: University of Chicago Press, 2019.

McKearney, Patrick. "Receiving the Gift of Cognitive Disability: Recognizing Agency in the Limits of the Rational Subject." *The Cambridge Journal of Anthropology* 36, no. 1 (2018): 40–60.

Meltzer, Françoise and Jaś Elsner, ed. *Saints: Faith Without Borders.* Chicago: University of Chicago Press, 2011.

Merton, Thomas. *New Seeds of Contemplation*. New York: New Directions Publishing, 1962.

Orsuto, Donna L. "The Saint as Moral Paradigm." In *Ethics and Spirituality*, ed. Charles E. Curran and Lisa A. Fullam, 145–61. New York: Paulist Press, 1996.

Porter, Jean. "Virtue and Sin: The Connection of the Virtues and the Case of the Flawed Saint." *The Journal of Religion* 75 (1995): 521–39.

Robbins, Joel. "Where in the World Are Values? Exemplarity and Moral Motivation." In *Moral Engines: Exploring the Ethical Drives in Human Life*, ed. Cheryl Mattingly, Rasmus Dyring, Maria Louw, and Thomas Wentzer, 155–73. New York: Berghahn Books, 2017.

Sherry, Patrick J. "Philosophy and the Saints." *The Heythrop Journal* 18 (1977): 23–37.

Singer, Peter. "Famine, Affluence, and Morality." *Philosophy and Public Affairs* 1 (1972): 229–43.

———. *Practical Ethics*, 3rd ed. New York: Cambridge University Press, 2011.

———. *The Most Good You Can Do: How Effective Altruism is Changing Ideas about Living Ethically*. New Haven and London: Yale University Press, 2015.

Six-Means, Horace. "Saints and Teachers: Canons of Persons." In *Canonical Theism: A Proposal for Theology and the Church*, ed. William J. Abraham, Jason E. Vickers, and Natalie B. Van Kirk, 97–118. Grand Rapids, MI: Eerdmans, 2008.

Taliaferro, Charles. "A God's Eye View." In *Faith and Philosophical Analysis: The Impact of Analytical Philosophy on the Philosophy of Religion*, ed. by Harriet A. Harris and Christopher Insole, 76–84. Hampshire, UK: Ashgate, 2005.

Urmson, J. O. "Saints and Heroes." In *Essays in Moral Philosophy*, ed. A. I. Melden, 198–216. Seattle: University of Washington Press, 1958.

Wolf, Susan. "Moral Saints." *Journal of Philosophy* 79 (1982): 419–39.

———. *The Variety of Values: Essays on Morality, Meaning, and Love. New York: Oxford University Press*, 2015.

Wyschogrod, Edith. *Saints and Postmodernism: Revisioning Moral Philosophy.* Chicago and London: University of Chicago Press, 1990.

Zagzebski, Linda Trinkaus. *Exemplarist Moral Theory*. New York: Oxford University Press, 2017.

Markus Huppenbauer

But It's My Money! Some Thoughts on the Ethics and Culture of Donating Money

Abstract

This chapter is the posthumous transcript of a public talk given at a workshop on religious perspectives on effective altruism. It explores the practice of charitable giving and comments on it from a moral perspective. It discusses the motives for donating to charity as well as the obligation to do so. It then considers a number of open questions about donations: the when, the how, and the how much. While the chapter is supportive of the obligation to give in general, it also considers two challenges to the present culture of donating: effective altruism and the call for replacing donations by social investments. The effective altruist approach is questioned on account of giving too much space to the task of doing good and missing out on aspects of the good life apart from morality. If people instrumentalize themselves for the sake of a better world, they risk missing the meaning of life.

Editorial Comment: Markus Huppenbauer, one of the co-editors of this volume, sadly passed away before this book could be completed. He planned to write his own contribution by building on the public talk with which he opened the workshop that forms the basis of this volume. Regrettably, he never got the chance to do so. After careful deliberation, we decided to publish a translated transcript of his talk. It has been lightly edited to adapt its spoken, personal, and spontaneous style for the purposes of a book chapter, and occasionally to leave out examples from the local context that are not relevant to the general argument. However, readers should bear in mind that the style and content were clearly geared towards a public lecture and have not been essentially changed. The transcription has been done by Arianna Lanfranchi and the translation is the work of Sarah Kirkby, both of whom we would like to thank for their efforts.

Today I would like to speak about the ethics and culture of donating money. The reference to ethics and culture provides a clue that this will not be a purely philosophical lecture, but I will, at times, be inspired by considerations from moral psychology and permit them to be part of my talk. "But it's my money!": whenever I hear that, I have a picture in my head of people sitting on pots of money, not wanting to part with any of it, and then one day dying and not being able to take it with them. And the question is: Does money actually belong to us? Is it something that we are allowed to sit on? Is it something that we have a right to? Or is there some kind of duty to donate? This is what interests me. And I am prepared to stick my neck out at this early stage in my lecture and say: yes, a duty to donate does exist, and donating a considerable amount of money is good. But more about that later.

I would like to begin by defining donations, a topic concerning which there is not much (philosophical) literature available. In the context of effective altruism (EA) there is some, but otherwise very little. Some very

complex theories exist about duties to assist, of course, which donating would fall under, but I have found very little on donating itself, in either German or English. My definition is as follows: a donation is a voluntary contribution to a religious, scientific, charitable, cultural, economic, or political cause. Recipients of donations can be non-profit organisations, foundations, political parties, aid associations or religious communities. Donations can be financial or material, or a refusal of money in return for work performed, i.e. donations of time. This definition will accompany us as we proceed today, and I shall not be questioning it further. Maybe there are some specialists who could, but my primary concern will be to investigate the question: is there such a thing as a *moral obligation* to donate? And, if so, what does that mean?

Let us start by taking a look at the reality in Switzerland. The Swiss donate money to many different causes, and they donate a great deal. According ZEWO's statistics, in 2017, they donated a total of CHF 1.84 billion to various organisations.[1] To me, that seems relatively high, and yet the total for the US is approximately $400 billion, which is more. It is also interesting to look at the breakdown of the overall volume. Studies in Germany have shown that more and more money is being donated in total, but by fewer and fewer individuals. And studies also show that it is mainly people over the age of 60 who donate in the classic sense. Such statistics should be handled with care, however, because they do not include crowdfunding or other newer methods of donating. What we can say with certainty: donations are increasing, but the age group donating the money will soon be elderly and die, leading many to fear that the bottom will then fall out of this model.

What I would like to look at with you today is the following issues: firstly, motives and reasons for donating; secondly, the potential existence of a moral obligation to donate; thirdly, some specific practical questions regarding the when, how, and how much of donating – actually more important than the question regarding a moral obligation, to which the answer is relatively clear; fourthly, two challenges to the present culture of donating, namely EA and a movement which advocates social investment instead of donation. Following this exploration, I shall conclude with a few brief comments.

I have been through the literature for clues and have also – but not in the sense of an empirical investigation – asked various individuals: why do you donate? And yet people are reluctant to answer this question, as if asking

1 Zewo, "Spendenland Schweiz."

them whether they donate could be akin to asking them whether they still have sex or how much they earn. Two more questions that the Swiss, at least, are loath to answer. Nevertheless, many people do donate, and their reasons and motivation for doing so vary tremendously.

I would like to start with a classic, namely feeling directly moved by the plight of others. This has very traditional roots and a pertinent religious narrative, the parable of the Good Samaritan to be found in Luke 10:25. You all know this story. A man has been attacked by robbers and is lying on the ground injured, bleeding and unconscious, with no money or anything else left to him. Two representatives of the cultural, political and religious elite march swiftly past, but cannot stop to help him because they have more important things to do. Their life goals collide with the needs of this man, and so they walk on by. What makes this story interesting: it is a Samaritan, somebody who does not belong to the elite and is in fact more of an outcast, who takes care of the injured man, nurses him, brings him to the next inn and makes sure that he gets something to eat and a place to sleep.

If you read this parable – which really is worth it, the story is fascinating – you will see that the arrival of the passer-by is described as follows: a Samaritan was passing, saw him, and was moved by what he saw. In most translations, it says "took pity on him", and the point here is not that he is moved on the inside, but that he takes action on account of something outside himself, namely the plight of another. It is the suffering of another which moves him to act. And I believe that many people make donations or offer assistance on this basis.

Another response you will frequently encounter is that some people just like giving. There is a good psychological explanation for this. A further factor that should not be discounted is economic gain. In Switzerland, it is financially beneficial to make donations. If you have an income that puts you into a higher tax bracket – maybe because you and your spouse are both earning well – then above a certain given level it is worth making donations. And the more income you have, the more it is worth it. I should think that utilitarians would be fine with donations on this basis. Motivation is not important; what is important is the result.

Another source of motivation you often hear about is gratitude: I am fine; I am lucky enough to live in a country where most people are fine; I did nothing to deserve this; it is not my own doing. And then people usually say that they want to give something back, so their motivation is gratitude.

Of course, there might also be ethical reasons for donating, in a narrow sense of the word. It is imperative that we support and help those in need. It is a moral obligation. But why? Because it is unfair for one person to be living in such good circumstances and another, through no fault of their own, in such bad circumstances. This puts us right at the centre of the ethical discourse. I am not concerned here with something as orientationally abstract as rules; I am far more interested in something else, namely considerations or feelings. It is a religious imperative to be charitable. In all religious traditions, helping those in need – in the Old Testament especially widows, orphans, the poor and strangers – is a very important aspect of our interactions. We must give to those who fall through the normal structures of society. That is clear, and in the Christian tradition it is the will of God. Analogous traditions also exist in the Hindu, Buddhist and of course Muslim cultures.

These, then, are some of the motives and reasons why people donate money. Obviously, many people *do* donate, but I would like to briefly question whether or not they *must*. If nobody were to donate, could we demand that people should? I shall outline this point very roughly.

There is no doubt that human beings have moral obligations: duties to do certain things, and duties not to do others, for example not to kill or torture, to respect the property of others, to treat others fairly, etc. We do not need to expand on this further. What interests me here is: alongside the moral obligations mentioned above, which I assume to exist as a matter of course, is there also a moral obligation to donate? It would be easy if we could say that donating is morally equally as imperative as, for example, respecting the property of others. To which my gut response is: "yes, but…". I could also take "no, but…" as my starting point, but I choose the former over the latter. What is crucial here is the *but*.

So we have a duty to donate. In order to arrive at this statement, it is necessary to introduce a differentiation to aid our understanding of what exactly we are talking about when we talk about donating. The duty to donate belongs to the larger class of duties to assist. And, viewed formally, it is a positive obligation. Positive obligations – and here I refer to Corinna Mieth[2] – are concerned with the active causation of certain improvements. When I help somebody, I actively contribute to making things better for that person. In contrast: if I do not kill somebody, I passively refrain from doing something in a negative sense. Donating is obviously an obligation from the class of positive obligations. And we will see in a moment why

2 Mieth, "Positive Pflichten."

that is relevant. As far as I can see, the most comprehensive class of positive obligations is the class of duties to assist, in other words, duties to help other people. Their foundations may vary, for example, human dignity or respect for others, but that is not important here. A duty clearly exists – as stated by philosophers from Kant to Peter Schaber to Corinna Mieth, and also by utilitarians – to assist other people and offer help if they require it.

And donating, this is my hypothesis, is one possible way of fulfilling this obligation. Further ways of helping others also exist, not just donating, and that is why my response was: "yes, but...".

These duties to assist (in a very broad sense of the term) involve an interesting aspect. On the one hand, there are some duties to assist where it is completely clear what we must do. A famous example comes from Peter Singer: a child is drowning in a nearby pond, and of course, we have a duty to save the child.[3] We do not even have to stop and think about what assistance might mean. But there are other situations in which the meaning is not so clear.

If for example – as my hypothesis purports – it is the goal of assistance to support people in order to enable them to lead a decent life, then, when people are not currently in need, assistance can mean a lot of different things. Positive obligations are obligations that we must fulfil actively, but they also (when there is no emergency situation) leave a certain amount of scope as to how we fulfil them. Donating – to my mind – is one possible way of fulfilling our moral obligation to help other people. (And also, incidentally, to help other animals – that is specifically sentient animals – and this is highly significant in the field of utilitarianism.) The goal of assistance or of such obligations – in other words, the *why* behind the obligation – is, then, to support people to enable them to lead a decent life. The means to this end is a society based on solidarity. And in a society based on solidarity – this is my belief – donating is part of the culture surrounding our response to people in need. Maybe that was a little hurried, but I think it will suffice for now. To sum up thus far: we have duties to assist others, which I will specify further later on, and donating is one possible way of performing such duties.

When should we donate? This brings us to differences that obviously exist in connection with assisting other people. On the one hand, we have what Mieth calls strong duties of assistance in emergency cases (*starke Nothilfepflichten*). "Strong" because there is no doubt that something must be done – for example, aid following a catastrophe. More generally

3 Singer, "Famine, Affluence, and Morality."

speaking, we could say that we are concerned here with people in need through no fault of their own. This is the obligation which we have to help in an emergency.

Then we have what Mieth calls weaker duties of beneficence (*schwächere Wohltätigkeitspflichten*). The German expression *Wohltätigkeit* is rather an outdated word and I have looked for alternatives in German, and I have opted instead for the expression "development aid" (*Entwicklungshilfe*). I understand this expression in a very broad sense – that is, not in the sense of development aid where, for example, Swiss people go to Ghana and build latrines. That is also an interesting topic, but I mean development aid in the sense that we support people in order to enable them to lead a decent life. These are developmental processes that need to be generated, so to speak.

Finally, there is another thing which I would also include within the category of duties to assist. Though this is a more difficult case, and the exact categorisation is not clear. We can make donations, financing and facilitating projects which the state either cannot finance or does not wish to finance. Many institutions only function because there are people or foundations ensuring that they can operate. And the question is, of course, whether in this context we can still speak of an obligation. Firstly, it is certainly necessary to differentiate between donating and sponsoring. If a company like the Swiss bank UBS finances an art exhibition at the Fondation Beyeler near Basel, then that is very nice for the public because we can all go and look at the pictures exhibited, but it is also very nice for UBS because they can build up their reputation and polish their image. By the way: I am not all that interested in the motivation of UBS, as long as I can look at the pictures. But we must make this difference: donating, patronage, philanthropy, etc. are completely different from sponsoring, where the person (or more usually a company) directly profits in terms of its reputation. My inclination is to say that such funding in place of the state does happen in the donations marketplace, but that there is no moral obligation to perform it. Whereas both strong duties of emergency assistance and weaker duties of beneficence really do involve moral obligations, the financing of projects in lieu of the state is – as philosophers would put it – supererogatory. That sounds a little strange but it simply means that one does not have to do it. If one chooses to, then that is nice. If someone chooses to donate wealth gained in the market economy, then it is a good thing to pass on that wealth and not just sit on it. These are my rough thoughts on "When should we donate?". In the remainder, I want to focus primarily on strong duties of emergency assistance and weaker duties of beneficence.

Now we are drawing closer to the matter at hand: how should we donate? We are not yet contemplating how *much* – to which the answer will naturally be: much – but simply *how*. In which spirit should donating take place? There are three points which I believe to be of utmost importance here. Firstly, it seems to me relatively clear, at least in our part of the world, that donating should not occur in an arrogant, paternalistic spirit – giving handouts condescendingly. Maybe some of you are still familiar with the old-fashioned missionary collection boxes: when you insert a coin, the black figure on the top of the box nods in gratitude. This is, of course, a relic of the past. Today – for emergency assistance, too, but primarily for development aid – to donate means to help others to help themselves. We make it possible for people to stand on their own two feet. Although I used missionary collection boxes as a negative example here, it tended to be the more liberal missionary societies in Switzerland, in Germany, also in the UK and the US, who recognised very early on – as early as the 1960s – that development aid cannot be a one-way street. In other words: it cannot be the case that we – the rich in the West – give money to you – the poor people in the South – so that you can lead a decent life; instead, any movement must be bidirectional. This is a point which many church aid societies recognised long before state development aid programmes. Empowerment is crucial.

Secondly, for many development aid projects, donating makes sense above all when something is not simply given to people, but when those people are involved and included in joint projects. There are hundreds, even thousands of development projects in which unbelievable amounts of money have gone down the drain because the givers thought they knew best how to help. Technology was brought, money was brought, and it did not work. And often the reason it did not work was that, while a great deal of money had been brought, no attention had been paid to integrating the local people receiving the aid, enabling them to continue alone with these projects once the money had run out.

Here is a little example: I grew up in Ghana, and today I am on the advisory council of the *Ernst Peyer Foundation*. In some regions of Ghana, we focus on hand pumps for wells; in other words, hygiene for latrines, very basic things. If the people there do not have these basics, death is never far off, and, therefore, I believe it is a truly effective way of helping. We do not simply go to Ghana and build latrines; instead, we select a village or region, and then the people in the village or community must show us what they can contribute to the building of a hand pump or a latrine. That can be money – usually, we actually demand that they make at least a

small financial contribution – but essentially it means workers and working hours. Workers from outside are not paid to build the latrines; the people have to do it themselves, making them genuinely involved in the project. This contributes to its sustainability in the long run.

And last but not least: donating has to be efficient! This is not something invented by EA, but something which has been in the foreground for, say, 30 years whenever talk has turned to donations and the financing of foundations. You must be efficient in the manner in which you deal with donated money. From 2000–2005, I managed the *Swiss Study Foundation*, and back then we had to raise a bit more than CHF 1 million. That was at a time when companies who had already been giving us money increasingly began to ask us: how much of the money do you take for administration? And if we had a budget that included 11% for administration costs, they no longer gave us money. At that time there was a clear line: under 10%. This of course led to budgets being correspondingly corrected, which was not really a problem. But the idea that donations must be efficiently managed and distributed is actually quite old. A little more recent is EA, which looks not only at organisations themselves, in other words, how much money they use for administration purposes, but also at where does the donated money go. Where is the effect greatest?

And now to the last of these specific questions: how much should be donated? I think there is no doubt that most of us are not saints. Bearing this in mind, I believe that there is no moral obligation to donate so much that we ourselves end up not having enough. Self-sacrifice or complete selflessness is presumably not something that anyone would morally demand. Let us look at a famous example involving a saint, namely Martin of Tours. One winter, a beggar who was freezing asked for help. Saint Martin (depicted on the back of a noble horse – but still) ripped his cloak in two and gave half to the beggar. And that for me is the true embodiment of assistance and what makes Martin so famous. It is a fascinating story because he gives a great deal of what he has in the said situation, but keeps back enough to ensure that he is still warm and will not freeze himself. He does not sacrifice himself in the process.

It is interesting to see what various websites say about this. In the context of Hinduism, for example, I have repeatedly come across the expression "selflessness". In connection with donating, this seems to be a very important category. By the way, when we speak about altruism later on, this is very akin to selflessness, a turning to others in a non-selfish way.

So now we have reached the point where we can say: fine, we do not have to sacrifice ourselves, we just have to give as much as we can to still leave ourselves with enough – a *sufficiency theory*.

My co-organiser, Dominic Roser, pointed me towards the Methodist founder John Wesley and his way of dealing with money. It has been publicly reported that at the tender age of 30 he earned 30 pounds, of which he needed 28 and donated the other 2. Over the years he earned more and more, but his living costs remained the same. He was ultimately earning more than 1400 pounds, yet still only required 30 pounds to live on, donating the difference as before. This would be the sufficiency theory in action, and not something which could really be turned into an imperative for everyone. It would probably be too much to ask of most people. And to be clear about this: I like people, and I do not think we should demand things of them that go beyond what they can cope with. But it is not too much to ask of people that they donate at all, just as it is not too much to ask of them that they do not kill or torture. They can cope with demands like that. But to demand that they do what John Wesley did definitely goes too far. I would therefore say: tread a little carefully with demands. Nevertheless – and here I think I am in the same boat as the effective altruists – most inhabitants of the wealthier countries in the world could give much more than they actually do. And the next question is, of course, where is the limit? I do not believe there is one here; these are simply concepts that are currently discussed. But we should address the issue: at what point does it begin to be too much, and when is that just an excuse?

What we certainly need to remember in any discussion about aid and donations is that it is not me personally, and it is not only the rich who must save the world and donate as much as possible, but always the *collective*. As far as global poverty and other problems affecting the whole world are concerned, we do not have to improve the world as individuals, but together. This is one of the reasons why the theories of Peter Singer have been subject to criticism. And neither should we forget that, as citizens and taxpayers, we already support national aid programmes, which in turn contribute to tackling the problems at hand.

Now maybe you are thinking: "I need some more detail here. I have been to this lecture and listened to this ethicist talking about donating, but he failed to say how much I should donate. Some numbers would be good." Never forget that such numbers cannot be absolute: obviously, if people have a lot of money, they can donate a lot.

Numbers are helpful and have been handed down through religious communities for a very long time. One famous example is Christian tithing, which was a tradition for centuries and practised by hundreds of millions of people. Paying a tithe simply means: I will pass on, in one way or another, the 10th part of anything which comes in. Islam has an analogous tradition set at about 2.5%, the *Zakat* imperative, which is actually one of the most important aspects of being a Muslim. Such numbers can be helpful as an orientation. And the 10% guideline has recently been revived. The pledge of *Giving What We Can*, for example, states: "I pledge that from now until the day that I retire, I shall give at least ten-percent of my income to whichever organisations can most effectively use it to improve the lives of others."[4]

This is relatively new and is slowly directing us towards EA. I think it is amazing that a pledge like this exists, and that to date approximately 5000 people have signed up to it. I simply wish to point out that, over the last few hundred years, millions and millions of people have given one-tenth, so it is nothing new. It is nice to see this happening. It is the secular aspect that is new, and I state that deliberately. The reference to religious traditions also provides clues about how people managed to pay their tithes. What resources did they have, enabling them to do this?

So that was donating in a brief overview. We have seen that something akin to a moral obligation to help people exists, and that one possible way of doing that is donating. Furthermore, I think that in our cultural and social context donating is a sensible way of helping others for most people.

There are two challenges facing the classic notion of donating. One is EA, originally a philosophical movement and started by people like Peter Singer. But I am less concerned with its philosophical aspects and far more with the question: what can be said about EA from a religious or theological perspective?

At this point, I would like to offer a pertinent quotation by Peter Singer from the Austrian daily newspaper *Der Kurier*.[5] The article was not about God, but about the world, and about animals. One of the – empirically correct – things the journalist said was: "the level of donations is increasing year by year. Many people are choosing to invest in emergency aid or other causes." I find it slightly embarrassing that the journalist talked about investing because it is not investing: with an investment I get something back. It is linguistically careless, but Peter Singer was generous to a fault and replied: "yes, many people are donating, and donations are rising steadily"

4 Giving What We Can, "Pledge to give more."
5 Klatzer, "Haben moralische Verantwortung."

– that was the positive news, and now for the negative – "and yet a high percentage of that money does not bring the greatest benefit for society. We are wasting resources negligently." Donations are fast becoming resources, then. He then remained within this economic paradigm and continued by saying that it is excellent that people donate, but that they spend far too little time considering what they are donating to. The money raised could be put to far greater effect than it is at present. And that is the concept behind EA: of course, people should do good – they should provide help in whichever way – but they should do so better.

What does "better" mean in this context? This brings me to William MacAskill, who picks up on Peter Singer's statement and clarifies it further as follows: we have an obligation to help people, and when we act on this obligation it is important how we do so. We should not simply help, but should think about how we can help as efficiently as possible. The idea, in a nutshell, is that as many people as possible should be able to profit as efficiently as possible from the money made available. And how do we do this? By considering as impartially as possible how, for example, a sum of money can be donated to a place where it can do as much good as possible. Reading these texts, the issue at stake seems to be a matter of reason. If we help at all, then we should do so in such a way as to profit as many people as possible. Is that not obvious, and would not all reasonable human beings have to agree on this? And because we all agree, we clearly need scientific evidence to show us where the help is required and where our money can be used most effectively. This, in turn, means that we require investigations into the efficiency of various organisations, and also into the most efficient means of bringing about an improvement for the people who need it most.

And on first reading, I believe this to be a good idea. A considerable proportion of donating to date has been driven by emotion, it has to be said. We feel a connection to mountain farmers, so we make a donation to mountain farmers. Or, like me, we have a relationship to missions. I have supported missionary societies throughout my whole life and, while they have to fulfil certain criteria, Swiss organisations such as *Brot für alle*, *Mission 21*, or *HEKS* all receive something from me. I truly believe that I would lose a part of my moral identity if I were no longer permitted to do this. Notwithstanding, the agenda we are addressing here is scientific evidence taking precedence over inclinations and personal preferences arising from one's own biography. There are many examples of this. Instead of supporting people affected by a landslide in Switzerland, where we need have no doubt that they will be helped by insurance companies and the like anyway, we could make the same money available to, say, providing malaria

nets for children in Ghana. In examples such as these, there is no doubt about the right course of action – though they are about isolated decisions.

Things are more difficult when we are not looking at isolated decisions but where people take up this reasoning and say: instead of just earning some money and then maybe donating a bit of it, I will try to earn as much money as possible and then donate far more, and do so as efficiently as possible. They then take good jobs to earn as much money as they can, so that they can then help as many people as they can. People like that do exist. I do not think there is very much we can say against their approach, but I must consternate that I find it very hard to comprehend. I do not understand how people can choose to live their lives like that. But some do, and we have to respect their choice.

I would now like to outline two problems I see here. A few moral psychological considerations will play a role that may raise a few eyebrows among hard-nosed philosophers. What I shall not be discussing is the foundations of EA. As I perceive it, most effective altruists view are committed to a form of utilitarianism. Their goal is as much happiness as possible for as many people as possible, as little suffering for as many people as possible, and all of this as efficiently as possible. This is the classic utilitarian calculation. But I do not wish to discuss utilitarianism here, even though I like many things about that philosophy. I would simply like to pick out two points that I believe to be inspired by utilitarianism, without examining the underlying moral theory.

When preparing for this lecture, I struggled, or should I say read my way through dozens of blogs. And I came across some things which were very alien to me, even incomprehensible. They bring me to a rough outline of my first hypothesis. I am not saying that they are all necessarily genuine, but I encountered such entries so frequently that I do not believe it to be a coincidence. To give you an example, I was reading one blog entry where somebody asked into the forum, so to speak, what the others thought about the idea of spending Sundays just hanging out with one's friends. Others replied – the discussions are always very earnest – and a conclusion was reached fairly fast: it is alright to hang out for a while, but too much hanging out means spending time just sitting around and chatting, and that is not really an efficient use of time considering the current state of the world. You really can read posts like that. I read a countless number of them. Another example entered the existential dimension. One blog contributor had a work contract where he was paid for his work according to output. And he was asking whether he should work as hard as possible in order to be able to donate as much as possible. The ensuing discussion

was really quite endearing and reached the conclusion that maybe it would be better to work slightly less in order not to risk a burn-out. And why? I would have said so that he could still have a half-decent life. But no, the reason stated was so that he would still be in a position to continue donating in the future!

I know that some of you will be familiar with such discussions. I was not. I did know that such things existed, but, for me, it was like discovering the existence of a foreign tribe and then somehow trying to make sense of it. We should remember that these were closed chats and maybe not how philosophers would publicly discuss such things. But it did reveal to me how some people address such issues. On the one hand, I had to admire their really strong sense of discipline and moral asceticism in the light of a superordinate goal. There is evil in this world, and it is our duty to do all we can to ensure that people with a bad life get a better life. I also surmised, at least in part, that an individual's happiness counts primarily insofar as it serves to protect their ability to improve the world. A further feeling was that no space exists which is free of morality. Every action is morally relevant regarding the needs of others. A simple example might be: should I spend 30 dollars on seeing a movie and eating a burger, or should I donate that money? I can imagine doing that, and in a moment I will say why. When I read discussions such as these, I have the impression that people are putting in a great deal of effort. And yet, on the other hand, I also feel a little sad. These people seem to be under a kind of permanent moral stress, and I feel sorry for them. They spend all day, every day, deliberating whether or not they are on the right side. A kind of epistemic overtaxing takes place. Not only do they have to help, but they also have to help efficiently. And it is not so easy to determine whether it is more efficient to work two hours longer and earn more money, which can be donated now, or not to work those two hours and to lie on the sofa in order to be strong the next day and fit to earn more money, which can then be donated later. Maybe I am exaggerating, but that really is how it came across.

My hypothesis about this is as follows: if improving the world or doing good is such a dominant concern, and if there is no other understanding of a good life – in areas where morality is secondary – then people become degraded to *morality machines*, and that is not the meaning of life. Stating that it is the goal of human life to do as much good as possible is, in my opinion, wrong. I do believe that in order to lead a meaningful life, we must help, must donate, but we should not turn into morality machines.

My observation is that this results in an instrumentalisation of enjoyment and self-care. Both are primarily relevant regarding the superordinate goal of improving the world: "to make the world a better place", that is the background here. And I had the impression – but one requiring some discussion – that effective altruists are not really interested in classical or modern conceptions of what constitutes the good life. Morality does play an important role – from Ancient traditions to Modernity – in what constitutes a good life, but not the only role.

I had a conversation with Dominic, my co-organiser today, in which we discussed a fascinating text. We were talking about the Puritans and how this mindset strongly reminded me of them. And he (as mentioned above) drew my attention to John Wesley, the founder of the Methodist movement. There was a wonderful sermon entitled "The Use of Money". Its motto is: "gain all you can, save all you can, give all you can". Earlier on in my lecture, I told you that Wesley maintained the same unchanging – not very high – standard of living his entire life in order to be able to donate as much as possible. This is precisely contained within the motto: "gain all you can, save all you can, give all you can". Nowhere does it say that we could enjoy our life. That is simply not the issue. And this is what makes the above discussion fit very well into the Puritan landscape.

Around the same time, the Puritans in the United States were having analogous thoughts, but they were not concerned with supporting the poor. Instead, the American Puritans interpreted earning money and being successful as a sign of being loved by God. And because they were not allowed to consume their success, they either had to donate it – John Wesley's option – or invest what they had gained in order to make more money – the spirit of capitalism as identified by Max Weber.

Historically speaking, this Christian tradition had no understanding for enjoyment or the like. This is why I enjoy telling people that I am a hedonistic Protestant.

There is another very nice example from the Bible. You probably all know the story of the anointment of Jesus in Bethany shortly before he dies. A woman comes along with a little bottle of perfume and pours it over his head, and in one version also over his feet. The disciples then say: could we not have rather used this money for the poor amongst us? They were the effective altruists of their time. And what does Our Lord say? – He says: "leave her alone. Do not embarrass her. She did a good deed." And that was truly the aesthetic element. In John 12 there is an analogous story. There it says that the woman wiped his feet with her hair, and that the house was filled with the fragrance of the perfume. The way that Jesus deals with this

scene is very interesting. He says that physical attention – even one that costs money – can occasionally be more important than supporting others. I was fascinated to learn how Alex Rattee, a member of EA for Christians, interpreted this story:

> "I want to suggest that the value Jesus expresses about the woman's actions should actually spur us on in the effectve altruist desire to alleviate poverty efficiently, for alleviating poverty may be one of the best ways to show Mary-like devotion to Jesus, given that he is no longer with us in bodily form."[6]

Here, then, this sensual, physical event is transformed into a spiritual devotion by this woman for Jesus. And when Jesus has gone, when his body ceases to exist, we no longer need to honour him thus. Then the best and quickest way to honour him is by supporting the poor. This is a kind of reversed biblicism: you take the Bible and pick out the stories you want and then twist them around until they fit your intended purpose.

An additional problem that I see here is the impartiality that is repeatedly demanded. I believe that, in a non-ideal world, it is not possible to speak to people about moral issues in this way. And my question to the effective altruists is therefore: instead of demanding from people impartial, effective aid and, in so doing, losing those people – because you are *de facto* demanding of those people that they detach themselves from all that makes them who they are, their relationships, their emotions – would it not possibly be even more effective to allow people to donate non-effectively, rather than to lose them altogether and have just a few remaining who donate effectively? This is a question that could well be in line with EA. Maybe, in order to be a truly effective altruist, one has to be able to forget EA.

The second point in this context is the stamina of effective altruists. I found a small-scale study about how long people can endure living as effective altruists. I do not think any comprehensive studies on this topic exist. I question how many of those who make the commitment are still committed more than five years later. I would be prepared to bet on there not being many, and the reason is that this philosophical setting alienates people from their emotions and their partiality.

My hypothesis is that people like donating, and that they like doing so for the communities to which they belong. Impartiality uproots people and has a negative impact on their willingness to help. I am always willing to take a stand against EA because I believe its consequences are not only

6 Rattee, "Alabaster Jars & Optimising."

negative, but morally fatal. But once again, that is an empirical hypothesis that requires investigation.

The question is, of course, why should impartiality be considered so important, and relationships, emotions, feelings be pushed aside? Why should we not be allowed to take our own communities into consideration when making donations? There is a major ongoing debate about whether or not one should donate to the church to which one belongs. What does the effective altruist say? Of course not. Donations should be directed where they make a difference. In my eyes, this is pure alienation. And why? My theory is that such people are predominantly male and young, and that worldwide they comprise a few thousand rationalistic utilitarians. They might be able to get excited about impartial effectiveness, but nobody else can. Incidentally, if anecdotal evidence of this is required, we only need to look at the footage of nerds making up panels at effective altruist conferences.

Now, just briefly, let us turn to the second challenge in conjunction with donations: paternalism. We know what would be good for others; we know how things can be made better. Donating often leads to problems not really being solved. We can see this, for example, with the refugee situation. Once a refugee camp has been set up, it has to be supported further. The people who live there seldom become able to build up their own existence. This is a problem. Angus Deaton illustrated this very well in his book *The Great Escape*. I do not want to go into detail here; suffice it to say that a large part of development aid from western states, and also from western NPOs, is money that goes down the drain. As a result of generous donations by individuals, coupled with a large amount given by states, the pressure on corrupt governments is lessened and they can use the small amount of money they do have at their disposal for other things – impressive palaces, the military, whatever. Bad governments can govern without the consent of society because they have money, and they have it because of development aid, and that money comes from the governments of other countries and via aid organisations. And, last but not least, there are correlations – not causalities – which show that development aid could actually be a problem. The states with the most development aid over the last 30 years tend to be the states with the smallest economic growth and the worst democratic development. The numbers are really interesting.

This leads us to the question: what is the point of donating, especially with regard to development aid? I would like to mention an option which could be given attention – without maintaining with any certainty that it is a good alternative. Aid must be employed in order to help people

to become independent. And how do we best do that? One possibility increasingly in the focus – albeit not nearly enough in church circles – is to attempt to locally reinforce entrepreneurship, for example via microcredits. Why? In these situations, it is usually the case that there are too few jobs enabling people to earn money. This is why it is important to empower individuals so they can start a small business and build up something akin to an autonomous life. For me, investment – in the true sense of the word – in foreign companies goes in the same direction. Economic development is then facilitated, which cannot be a bad thing. It means investing as an individual in companies that score well in terms of human rights and the environment. This is not donating as such, but it could also help. And a debate about this issue is currently ongoing.

To conclude: helping others is a moral obligation, with latitude available as to how we fulfil this obligation. Donating is one possible way of fulfilling it. Providing help and donating money are constituent parts of the good life, not its main objective. My motto, in summary, would be: donate generously, with an occasional nod to efficiency, and wherever possible, do not forget social investing.

References

Deaton, Angus. *The Great Escape: Health, Wealth, and the Origins of Inequality*. Princeton, Oxford: Princeton University Press, 2013.

Giving What We Can. "Pledge to give more, and give more effectively." Accessed 3 September 2021. https://www.givingwhatwecan.org/pledge.

Mieth, Corinna. *Positive Pflichten*. Berlin, Boston: De Gruyter, 2012.

Singer, Peter. "Famine, Affluence, and Morality." *Philosophy and Public Affairs* 1, no. 3 (1972): 229–243.

Klatzer, J. "Haben moralische Verantwortung, allen Menschen zu helfen." *Kurier*, 3 August 2016. https://kurier.at/wissen/moralphilosoph-peter-singer-im-gespraech-haben-moralische-verantwortung-allen-menschen-zu-helfen/212.770.745.

Rattee, Alex. "Alabaster Jars & Optimising." A blog for the Effective Altruism for Christian community. 18 January 2017. https://eachdiscussion.wordpress.com/2017/01/18/alabaster-jars-optimising/.

Zewo. "Spendenland Schweiz." Accessed 3 September 2021. https://zewo.ch/wp-content/uploads/2019/06/Spendenstatistik-2017.pdf.

Notes on Contributors

Calvin Baker is a PhD student in philosophy at Princeton University, where he works on ethical theory, global priorities research and Buddhist philosophy.

Lara Buchak is a Professor of Philosophy at Princeton University. Her book *Risk and Rationality* concerns how an individual ought to take risk into account when making decisions. It vindicates the ordinary decision-maker from the point of view of even ideal rationality. Several of her publications concern the application of this work to ethical questions. Other publications concern the nature and rationality of faith.

Mara-Daria Cojocaru has taught philosophy at Ludwig-Maximilians-University Munich, where she received her PhD in 2011, at the Jesuit Munich School of Philosophy and at Friedrich-Schiller-University Jena, in Germany. A DFG-Forschungsstipendium supported her research at the University of Sheffield. Alongside, she has pursued a career in creative writing and is active in animal protection. She now lives as an independent author in London and, with her background in philosophical pragmatism, is curious about effective altruism as a tool to make moral progress.

Stefan Höschele is Professor of Systematic Theology and Adventist Studies at Theologische Hochschule Friedensau near Magdeburg, Germany. He taught theology in Tanzania (from 1997 to 2003), did his habilitation on interchurch relations, and has recently moved from a focus on theological studies into the field of Christian ethics.

Markus Huppenbauer † was Professor Extraordinarius for ethics at the Faculty of Theology of the University of Zurich as well as managing director of the Center for Religion, Economy and Politics. His research and publications focused on the application of ethical norms to business, the economy and everyday life.

Robert MacSwain is Associate Professor of Theology at the School of Theology at the University of the South, Sewanee, Tennessee, USA. A philosophical theologian educated in both the United States and the United Kingdom, he teaches Christian doctrine and ethics. The author of *Solved by Sacrifice: Austin Farrer, Fideism, and the Evidence of Faith* and a priest in The Episcopal Church, he is currently writing a book on the idea of saints as evidence for God.

David Manheim is currently a Visiting Researcher at the Technion, the Israel Institute of Technology, and works as an advisor and consultant for several organisations in the effective altruism space focused on reducing existential and global catastrophic risk. As a practicing Orthodox Jew he is helping to launch EA for Jews, and he works on building both the effective altruist community and the existential risk research community in Israel.

Kathryn Muyskens is a Lecturer at Yale-NUS College. A political philosopher by training, her work has centred on the intersections between politics, culture and healthcare. Her latest research projects have explored the implications of Confucian filial piety for elder care and ethical guidelines for international humanitarian health aid. Her future projects continue these themes – exploring the ethical and political dilemmas posed by the Covid-19 pandemic and exploring issues in public health from a Confucian perspective from her home in Singapore.

Stefan Riedener is Senior Teaching and Research Assistant at the University of Zurich's Center for Ethics. His dissertation developed a decision-theoretic approach to the problem of normative uncertainty. His current research project explores the ethics of partiality – among other things, the question of how our special relationships square with the norms and aims of effective altruism.

Dominic Roser works as a Senior Lecturer and Researcher at the University of Fribourg, Switzerland, after post-doctoral positions in Oxford, Graz and Zurich. With a background in philosophy and economics, his research is located in the area of global and intergenerational justice. Many of his publications, including a co-authored introductory book, focus on the theme of climate justice. He is active in the effective altruism movement.

Jakub Synowiec works as a Lecturer at the Pontifical University of John Paul II in Kraków, Poland. His main research areas are the dialogue between Christian and secular ethics, personalism, ethical argumentation and effective altruism. Selected relevant publications in English include *Metaethical grounds for changing the understanding of values in Peter Singer's ethics*, *Effective altruism for the poor* and *Ethics for everyday heroes – from Utilitarianism to Effective Altruism*.