

## 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.<sup>1</sup>

1 Westerhoff, *The Golden Age*.

Neither Indian Buddhist philosophy (henceforth, “Buddhism”)<sup>2</sup> 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.<sup>3</sup>

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.<sup>4</sup> 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.<sup>5</sup> (“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<sup>6</sup> and, more fundamentally, to *be* social movements at all.<sup>7</sup> 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;<sup>8</sup> 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<sup>9</sup> and promulgated their views within EA,<sup>10</sup> with the result that the Centre for Effective Altruism includes moral uncertainty as a key concept in its primer on EA topics,<sup>11</sup> “moral uncertainty and moderation” is a guiding value of 80,000 Hours,<sup>12</sup> and “worldview diversification” is central to Open Philanthropy Project’s grant-recommendation strategy.<sup>13</sup> 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;<sup>14</sup> 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).<sup>15</sup> 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 *samsāra*, the round of rebirth (lit. “wandering”). Beings are repeatedly reborn into the various realms of *samsāra*, 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 *samsāra*. *Samsāra* 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 *samsāra* 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 *samsāra*. Taken on the whole, *samsāric* existence is characterised by a preponderance of *duḥkha* over whatever conventional goods may be found within *samsāra*, such as transient pleasures or successes.

The reason *duḥkha* predominates in *samsāra* 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.<sup>16</sup>

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.<sup>17</sup> 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 *samsāra*. 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 *samsāra*, 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.<sup>18</sup> 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. Samsāric 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).<sup>19</sup> 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 *samsāra* – 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 *samsāra* 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 *samsāra*, are impermanent and ultimately subject to dissolution. We thus encounter what I will call the *samsāric* 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 *samsāra* is to help them out of *samsāra*. 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*,<sup>20</sup> a being who has attained (or nearly attained) awakening but, rather than passing out of *samsāra* upon the death of their physical body, elects to remain in *samsāra* and promulgate the *Dharma* – Buddhist teaching – until all sentient beings have themselves attained awakening.<sup>21</sup>

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 *samsāric* 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 *samsāra*. 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 *samsāra* 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 *samsāra* 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 *samsāra* by leading all sentient beings to awakening.<sup>22</sup> But helping sentient beings to achieve comparatively better short-run *samsāric* results is not going to bring about the end of *samsāra*. 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 *samsāra*). That worldly altruism plays some as yet unspecified role in the *bodhisattva* path does not, therefore, constitute a counterexample to the *samsāric* 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.<sup>23</sup>

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.<sup>24</sup> 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<sup>25</sup>). 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.<sup>26</sup> 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 *samsāra* 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.<sup>27</sup> 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<sup>28</sup> 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.<sup>29</sup> 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 leanness 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 (*vipassāna*) is generally as effective as combining insight meditation with calming concentration meditation (*śamatha*).

conventional welfare for those who are uninterested or unable.<sup>30</sup> 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).<sup>31</sup> 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,<sup>32</sup> 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 (*tṛṣṇā*)<sup>33</sup>, 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 Thānissaro Bhikkhu, "Dhammacakkavattana 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 *samsāra*. (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 *samsāra*.<sup>34</sup>

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.<sup>35</sup> This focus may seem to be more in line with the preoccupation with *dukkha* 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 *samsāra* 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.”<sup>36</sup> 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,<sup>37</sup> 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.<sup>38</sup> 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.<sup>39</sup> Buddhism would therefore agree with the majority<sup>40</sup> 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<sup>41</sup> 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).<sup>42</sup> 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).<sup>43</sup>

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.<sup>44</sup> Śāntideva appears to support this stance in maintaining that the *bodhisattva* should strive to eliminate the *duhkha* 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.*”<sup>45</sup> 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).<sup>46</sup> 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.<sup>47</sup>

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 (*buddhava-cana*) 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 *dukkha* 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.<sup>48</sup> 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)<sup>49</sup> 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.<sup>50</sup> 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,<sup>51</sup> killing a sentient being cannot, according to the Theravāda Abhidhamma, be entirely an act of compassion, and hence cannot be entirely wholesome or skilful.<sup>52</sup> 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<sup>53</sup> 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*<sup>54</sup> 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).<sup>55</sup> 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 *samsāric* 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.<sup>56</sup>

## 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.

