

The Role of Effective Altruism in Securing Humanity's Future

We inhabit a civilization on the verge of destruction. A hellish world smoldering in the wake of a nuclear holocaust, barren landscapes strewn with corpses from an engineered pandemic, or a dystopian planet ruled by mechanized overlords – any of these potential futures could soon become reality, Oxford philosopher Toby Ord warns in his book *The Precipice* – and we must do everything in our power to prevent them from occurring. If Ord is correct that humanity's existence hangs by a thread and that we bear an ethical responsibility to ensure the continuation of our species, then we must work tirelessly to avert the apocalypse. The Centre for Effective Altruism (CEA) and associated organizations are uniquely positioned to contribute to this effort. Yet doing so successfully would require them to adapt their message and actions. In this essay, I first explore how effective altruist organizations like the CEA ought to change in response to Ord's imperative. Then, I consider what incentives might inspire such an adjustment. Before addressing these points however, I will provide a brief overview of several key ideas presented in the book.

Ord explains how humanity has entered the most crucial juncture in our history, an era he calls *the precipice*, from which we can either move towards a flourishing future or plunge into an existential catastrophe that would irreversibly destroy our long-term potential (Ord). Such a catastrophe could manifest in a variety of ways, including anthropogenic disasters like nuclear war and emerging threats like unaligned artificial intelligence. Ord argues that we must prevent these existential catastrophes for the sake of ourselves and our fellow humans (Ord; Baum 2122).

The effective altruism movement originated with philosopher Peter Singer's seminal 1972 writings. Singer's main thesis stems from the intuitive observation that suffering and death are bad (Singer 751). He continues by claiming that individuals have a moral obligation to

prevent bad things from happening if doing so would not require them to sacrifice anything of comparable moral value (751). Importantly, physical separation from suffering does not eliminate the mandate to alleviate it (751). Singer points out that it is within the power of many people to prevent suffering by donating money to charities that mitigate famine, disease, and other hardships (751). Doing so, Singer claims, would not entail sacrificing anything of comparable moral worth if the money we donate would not be necessary for meeting our basic health requirements, and thus, charity represents a moral duty (751).

Modern effective altruists give vast sums to charitable organizations in order to save as many lives as possible. Effective altruist groups, including the CEA, encourage members to pursue careers like investment banking to maximize donations. They also create lists of the most effective charities and have spawned a series of affiliated clubs on college campuses. These organizations have made valuable achievements, but, if we take Ord's argument seriously, both their underlying message and their practices must change.

I propose two modifications to the guiding message of effective altruism, which I will address in turn. Singer's claim that physical separation from suffering does not remove our responsibility to help alleviate it represents an integral component of effective altruist ethics. Yet, in light of Ord's argument, this premise is far too weak. To adequately capture the weight of our moral responsibility to prevent an existential catastrophe, we must extend the claim so that it addresses not only *physical* separation but also *temporal* separation. Much of Ord's moral reasoning hinges on the argument that we have an ethical obligation to ensure our descendants can experience human flourishing. Ord even seems to imply that, since humanity has the potential to survive for countless millennia, most of the good our species will achieve lies ahead, and thus, the future deserves special moral consideration. Even if we do not live to see the

apocalypse ourselves, we owe our descendants an effort to ensure it never comes about.

Embracing this view requires effective altruist organizations to accept that temporal separation from suffering does not eliminate the obligation to prevent it.

Additionally, Singer appears to suggest that refugees, infected children in developing nations, and the like, are the primary ones suffering, whereas upper and middle class families in the western world live in comparable luxury and should donate not for themselves, but for the sake of their needy brethren. However, I propose a reframing of this argument so that it aligns with Ord's claims. Rather than focusing mainly on the dichotomy between the poor and the rich, effective altruists ought to consider humanity holistically and emphasize the need to invest in our collective future. If an existential catastrophe of the kind Ord describes ever transpired, all of us would suffer tremendously, regardless of our current station in life, and thus, we ought to be motivated by the need to protect everyone, poor or otherwise. Furthermore, the devastation such a tragedy would bring far exceeds the suffering of any one person or even of a moderately large group of people. To be clear, we must not lose sight of the disparities that exist in our society, but we cannot let these concerns overshadow the looming threats all of us may soon face.

Simply endorsing these claims is not enough, however. Effective altruist groups like the CEA must also change their actions. Firstly, they ought to include organizations dedicated to researching and preventing the existential risks Ord outlines on the lists of charities they send to their members. Moreover, effective altruist groups ought to invite experts like climate change scientists and engineers familiar with the dangers of AI systems to speak at their campus events. The CEA's website should include images and warnings that reflect the message Ord imparts.

Finally, it bears considering what may incentivize effective altruist organizations like the CEA to change their message and actions as described above. Fortunately, members already feel

motivated to promote the welfare of others, and the alterations I propose remain consistent with that goal. Hence, the main obstacle would be ensuring effective altruist groups understand the severity of the risks Ord describes and the toll they would take on humanity. Perhaps if the media focused more on environmental catastrophes and other existential risks (rather than local news stories), effective altruists would give them greater consideration. Moreover, since effective altruism remains deeply rooted in philosophy, and college students who embrace it likely have an interest in moral theory, perhaps the organizations these students join would change their message and actions if philosophy classes focused more on large scale ethics rather than mainly examining individual- or small group-level behavior. Ord himself seems to lament this deficit in the philosophical literature (Ord).

Ord argues that humanity stands poised at the edge of a precipice. The actions we take now will determine whether we reach our full potential or face an irreversible calamity. Effective altruist organizations like the CEA have the power to help avert such a disaster, but they must alter their message and actions to properly address Ord's concerns. If these groups, and others like them, make the necessary effort, we could secure our existence for generations to come. Indeed, our very future hangs in the balance.

Works Cited

Baum, Seth D. "The Precipice: Existential Risk and the Future of Humanity by Toby Ord. New York: Hachette Books, 2020." *Risk Analysis*, vol. 42, no. 9, 2022, pp. 2122–2124, <https://doi.org/10.1111/risa.13954>.

Ord, Toby. *The Precipice: Existential Risk and the Future of Humanity*. Hachette Books, 2021.

Singer, Peter. "Famine, Affluence, and Morality." *Ethics: Essential Readings in Moral Theory*, edited by George Sher, Routledge, 2012.