



Review

The Precipice: Existential Risk and the Future of Humanity (London: Bloomsbury, 2020), by Toby Ord

B.V.E. Hyde

Department of Philosophy, Durham University

Abstract

*This is a review of *The Precipice* (London, Bloomsbury, 2020) by Toby Ord.*

Keywords

Longtermism, Effective Altruism, Existential Risk, Extinction, Artificial Intelligence

Man's complacent assumption of the future is too confident. We think, because things have been easy for mankind as a whole for a generation or so, we are going on to perfect comfort and security in the future. We think that we shall always go to work at ten and leave off at four, and have dinner at seven for ever and ever. But these four suggestions, out of a host of others, must surely do a little against this complacency. Even now, for all we can tell, the coming terror may be crouching for its spring and the fall of humanity be at hand. In the case of every other predominant animal the world has ever seen, I repeat, the hour of its complete ascendancy has been the eve of its entire overthrow.

— H. G. Wells, "The Extinction of Man", 1894

Humanity is now on the precipice of extinction. According to Toby Ord, senior research fellow at the University of Oxford's Future of Humanity Institute, there is a 1 in 6 chance that civilization will come to an end in the next century. Throughout the twentieth century, that chance was 1 in 100.

The chance of 'existential catastrophe' in which intelligent life is completely annihilated is called 'existential risk' and, as mankind advances technologically, it has been increasing. We are now at a uniquely dangerous period in history characterized by unprecedented destructive capability with neither an understanding of this nor the global unity to do anything about it. Beginning with the development of the first atomic bomb in 1945, Ord calls this period the 'Precipice', but it will not last more than a few centuries: accordingly, we will either develop the necessary policy to reduce existential risk or humanity will end before we do.

What will cause the extinction of mankind? An asteroid, like the one that caused the mass extinction of all non-avian dinosaurs 66 million years ago? A supervolcanic eruption? Unlikely: all in all, natural risks together amount to only about a 1 in 10,000 chance of existential catastrophe per century in Ord's estimation. The existential risk associated with nuclear war, however, is 1 in 1,000, ten times higher than all natural risks put together. The risk of extinction by climate change is also 1 in 1,000. Much worse, though, is the threat posed by engineered pandemics, which have a 1 in 30 chance of ending the world and, the most dangerous of all, the existential risk of artificial intelligence unaligned with human values is estimated as 1 in 10 by Ord – a figure doubtless influenced by Nick Bostrom, who also judged artificial intelligence a serious threat to human existence in *Superintelligence: Paths, Dangers, Strategies* (Oxford: Oxford University Press, 2014). Bostrom's book, along with others like *Our Final Invention* (New York: Thomas Dunne, 2013) and *Human Compatible* (New York: Viking Press, 2019), brought

* Corresponding author.

E-mail address: b.v.e.hyde@outlook.com

concerns about existential risk from artificial general intelligence to public attention, and it is now commonplace to see public figures like Elon Musk and Bill Gates express concern about it. Although some have been skeptical of the risk it poses – like Michio Kaku in *Physics of the Future* (New York: Doubleday, 2011) – most recent books have not been, and several hypothetical takeover scenarios have been mapped, such as in *Life 3.0* (New York: Vintage Books, 2017) by Max Tegmark. The threat of artificial intelligence is more real than ever today, but it is perhaps no longer the biggest threat to human survival. Ord's figures are somewhat outdated already because, with the Russia-Ukraine War which has transpired after Ord wrote his book, the chance of nuclear warfare certainly ought to be estimated much higher. In 2023, the Doomsday Clock is set at ninety seconds to midnight, the closest to global catastrophe it has ever been, with the reasons cited being the war between Russian and Ukraine and the threat of nuclear warfare. Either way, what Ord's estimates mean is that the 1 in 6 chance that the world will end this next century is, pretty much entirely, manmade.

The bright side is that this means we can do something about it. In an ideal world, humanity needs to come together as a coherent agent to take responsibility for its future and make some strategic choices with the longterm future in mind. Less than 0.001% of gross world product is spent on targeted existential risk reduction interventions. For example, the Biological Weapons Convention, the global body founded to reduce risk of accidental or deliberate viral releases which, recall, have a 1 in 30 chance of extinguishing humanity, has an annual budget (\$1.4 million) smaller than that of the average MacDonald's. Motivation to fund existential risk mitigation is limited not only by ignorance of the dangers at hand, but also by insufficient global coordination.

Ord does not suggest that the survival of humanity is only a global policy issue. We can all play a role in safeguarding humanity's future, he thinks. Two of the most important ways individuals can change the world are through their careers and charitable donations. 80,000 Hours, a non-profit organization part of the Centre for Effective Altruism at the University of Oxford, conducts research on which careers have the largest positive social impact and provides career advice based on that research. Giving What We Can, another effective altruism organization based at the University of Oxford, set up by Ord and MacAskill in 2009, is a collective of individuals committed to donating a minimum of 10% of their income to the most effective charities, including those with a longtermist agenda. Ord also encourages public discourse about humanity's longterm future which, he would be right to think, is essential to an international, intergovernmental, unified response to existential risks.

One may express indifference or lack of concern towards the potential extinction of humanity. This perspective may arise particularly among individuals of advanced age or those who, for various reasons, believe that an existential catastrophe would transpire posthumously. Consequently, such individuals might call into question the relevance of the longterm future. This is where 'longtermism' comes in as an ethical position. The term was coined by Toby Ord and William MacAskill and refers to their view that positively influencing the longterm future is a key moral priority of our time. It was first popularized by Ord with *The Precipice* (London: Bloomsbury, 2020), but *What We Owe the Future* (London: Oneworld, 2022) by MacAskill has in the end been more influential. According to MacAskill, "distance in time is like distance in space". Your moral circle is big enough to donate to charities helping people across the world, so why not to people in the future? What makes MacAskill and other effective altruists care about future people so much is that there are so many of them. You might disagree with them here, but for effective altruists, numbers count. That is because effective altruism is a utilitarian movement, so what they are looking to do is the greatest good for the greatest number, and they are completely indiscriminate in this. Because the future is so large, and therefore so populous, "the early extinction of the human race would be a truly enormous tragedy", says MacAskill. This is also Ord's view: he thinks that existential catastrophe would betray the efforts of our ancestors, bring great harm upon those in whose lifetimes the end of the world comes about, and destroy the possibility of a vast future filled with human flourishing. "Longtermism", he says, "is animated by a moral re-orientation toward the vast future that existential risks threaten to foreclose".

The Precipice is yet another addition to the rapidly growing effective altruist – and, by extension, longtermist – movement. Whether or not you agree with his estimates, or what he proposes we do about them, it is difficult to take an ambivalent approach towards the existential risks Ord outlines in the book. If he has succeeded in one thing, it is drawing the attention of the human race to the risks it faces – risks that it has created and, crucially, can mitigate and eradicate.

References

- Barrat, James. 2013. *Our Final Invention*. New York: Thomas Dunne.
- Bostrom, Nick. 2014. *Superintelligence*. Oxford: Oxford University Press.
- Kaku, Michio. 2011. *Physics of the Future*. New York: Doubleday.
- MacAskill, William. 2022. *What We Owe the Future*. London: Oneworld.
- Ord, Toby. 2020. *The Precipice*. London: Bloomsbury.
- Russell, Stuart. 2019. *Human Compatible*. New York: Viking Press.
- Tegmark, Max. 2017. *Life 3.0*. New York: Vintage Books.