

# The Badness of Extinction

by Derek Parfit

(excerpted from *Reasons and Persons*, 1984, pgs. 453-454)

---

I believe that if we destroy mankind, as we now can, this outcome will be *much* worse than most people think. Compare three outcomes:

(1) Peace.

(2) A nuclear war that kills 99% of the world's existing population.

(3) A nuclear war that kills 100%.

(2) would be worse than (1), and (3) would be worse than (2). Which is the greater of these two differences? Most people believe that the greater difference is between (1) and (2). I believe that the difference between (2) and (3) is *very much* greater.

My view is held by two very different groups of people. Both groups would appeal to the same fact. The Earth will remain habitable for at least another billion years. Civilization began only a few thousand years ago. If we do not destroy mankind, these few thousand years may be only a tiny fraction of the whole of civilized human history. The difference between (2) and (3) may thus be the difference between this tiny fraction and all of the rest of this history. If we compare this possible history to a day, what has occurred so far is only a fraction of a second.

One of the groups who holds my view are Classical Utilitarians. They would claim ... that the destruction of mankind would be by far the greatest of all conceivable crimes. The badness of this crime would lie in the vast reduction of the possible sum of happiness.

Another group would agree, but for very different reasons. These people believe that there is little value in the mere sum of happiness. For these people, what matters are what Sidgwick called the 'ideal goods'—the Sciences, the Arts, and moral progress, or the continued advance towards a wholly just world-wide community. The destruction of mankind would prevent further achievements of these three kinds. This would be extremely bad because what matters most would be the higher achievements of these kinds and these highest achievements would come in future centuries.

## **Existential Risk, Artificial Intelligence, & the Future of Humanity**

As we discover and invent more and more powerful technologies, the question of what is called ‘existential risk’ arises (that is, something that threatens our future well-being; e.g., the permanent extinction of the human race).

Watch this clip from Nick Bostrom on technology and existential risk:

**Nick Bostrom, lecture [here](#)**

(relevant clip from 8:03 – 14:33)

Note that Bostrom’s list of threats includes ‘machine intelligence’. Could artificial intelligence pose one of the greatest existential risks to our species? Several of the biggest names in the field of science and technology have warned us about the dangers of AI. For some examples, watch these short clips from Stephen Hawking, Bill Gates, and Elon Musk on the dangers of artificial intelligence:

**Stephen Hawking, interview [here](#)**

(relevant clip from 4:06 – end of video)

**Bill Gates, interview [here](#)**

(relevant clip from 1:49 – 2:35)

**Elon Musk, interview [here](#)**

(relevant clip from 1:00:30 – 1:01:32)

Finally, watch this TED talk by Nick Bostrom on the coming “singularity”:

**Nick Bostrom**

**“What Happens When Our Computers Get Smarter Than We Are?” [here](#)**

(watch the entire video)