Opinion Page

Opinion Page: Pessimistic Humanism and Existential Threats, *Robert Nola*

Robert Nola, Philosophy Department, University of Auckland

In this article I wish to explore some connections between a humanism which is pessimistic and claims that we humans are facing impending disasters of our own making (like aspects of the current pandemic or nuclear war).

Recently I was pleasantly surprised to find an account of *pessimistic humanism* by my PhD supervisor, John Passmore, written in *Free Inquiry* 1997 (seven years before his death in 2004). He had been invited to comment briefly on 'Why I am a Secular Humanist' in the illustrious company of other commentators such as Herman Bondi, Richard Dawkins, E. O. Wilson, Yelana Bonner, Arthur C. Clarke. But instead he gives reasons for being a pessimistic humanist. I wish to comment on these points.

Then I wish to develop the pessimistic side of his position by drawing upon Martin Rees' book *Our Final Century: Will Civilization Survive the Twenty-First Century?* (2004). Lord Rees, who works at Cambridge University, was at one time the Astronomer Royal, President of the Royal Society and Master of Trinity College. He is a distinguished scientist who has thought deeply about what future humanity might have in the technological world it has created - if it has a future at all! "I think the odds are no better than fifty-fifty that our present civilisation on Earth will survive to the end of the present century without a serious setback." Also: "…misdirected technology could jeopardise life's potential, foreclosing its human and posthuman future" (p8).

Rees does not say how he calculates the fifty-fifty odds. But he clearly endorses the view that the technology we have created can well lead us to disaster. He tells us writing in his 2004 book: 'I staked one thousand dollars on a bet: "That by the year 2020 an instance of bio-error or bioterror will have killed a million people"' (p74). If COVID-19 is originally a bio-error made by the Chinese, then he is well on the way to winning his bet. For Rees, humankind has a doubtful future! But how doubtful remains unclear without further evidence.

Passmore on Pessimistic Humanism

Like many humanists Passmore is an atheist. But the aspect of religion he chooses to reject is its morality rather than its theism: 'I rebelled as a young boy against the view that the whole of humanity suffers because a single person was disobedient. This I saw as tyranny of the first order'. This is a good point. If one person allegedly per-

forms a morally culpable act, it does not follow that others who did not perform that act are also culpable and deserve to be punished. It is a central Christian doctrine concerning Eve's alleged disobedience in the (alleged) Garden of Eden that all other members of humanity are made to suffer for her act (this is part of the doctrine of original sin). For the young Passmore, and for me, the alleged culpability of the rest of humanity is morally unacceptable and repugnant. And this holds even if one thinks that this story is a fable (Catholics often treat it non-literally). What is mistaken is the fable's moral presuppositions.

The second point Passmore makes has two aspects, the first moral and the second epistemic: "I am willing to admit that there is no deed so dreadful that we can safely say 'no human being could do that' and no belief so absurd that we can safely say 'no human being could believe that.'" This is the pessimistic side of his humanism. We are invited to face the disturbing facts that some of us humans can be grossly immoral and evil, and some of us can be so stupid as to believe any rubbish. Passmore insists that there is nothing about us humans which would rule out these two possibilities; to suppose that we could rule them out would require a quite misleading account of the nature of humans. Rather as the philosopher Kant claimed pessimistically: "Out of the crooked timber of humanity nothing straight can be made". To establish this position, Passmore wrote a book entitled *The Perfectibility of Man* (1970) in which his account belied the title. We are not perfectible, hence his pessimism.

His third point strikes an optimistic note: "But on the other side I point to the marvellous achievements of human beings in science and art and acts of courage, love, and self-sacrifice." As a matter of fact, some of us humans are wretchedly evil; but also as a matter of fact some humans can, in some respects, be admirable. Since we humans are a mixed bag, it would be wrong not to recognise those aspects of ourselves which make us irredeemably bad or those which make us worthy. Importantly we need to recognise that for the pessimistic humanist there is no rosy picture to be painted of us humans and our natures; to do so would be to ignore some of our worst traits, many of which have been exhibited in our recent history.

In an earlier article 'Atheism Without Humanism' (*The Open Society*, March 2019) I described how many versions of humanism can be set out in the style of a manifesto involving a number of principles. We can readily approve of most of these principles. However often an attempt is made to found the features of our morality upon the natures we humans possess. But there is much unclarity and some confusion about how this foundational link is to be made (the case for this will not be pursued again here). But if Passmore and Kant, amongst others, are right, then for the pessimistic humanist, whatever account we might give of morality, it cannot be founded on our flawed human natures or characteristics. The degree to which humans can be perverse undermines the optimism which many humanists suppose; so, another approach to morality must be found.

Rees on our Existential Risks

Like other humanists, Martin Rees is an atheist; but he is more sympathetic to religion than most and wishes to find an accommodation with it. His humanism can also be described as "pessimistic". His book *Our Final Century* makes a disturbing case for this. But his position is not entirely pessimistic since he is also a co-founder of the Centre for the Study of Existential Risk (CSER) at the Uni-

versity of Cambridge. This is a research group devoted to the study of extinction-level threats posed by our technology (see http://cser.org/). Here Rees investigates ways in which we can fend off the worst of the possible existential threats he envisages.

Though Rees does not make the point, it is important to distinguish the many non-fatal threats which confront us from those threats which are existential, i.e., threaten our very existence (however this might be further defined). Though there is some similarity to the pessimism of Passmore, Rees explores different kinds of threats, both existential and non-existential, which have faced humanity due to the technology humans have created. Passmore focuses on our flawed morality and belief systems which need not necessarily be fatal to us humans.

Rees tells us: "I'm a technological optimist in that I do believe that technology will provide solutions that will allow the world in 2050 to support 9 billion people at an acceptable standard of living. But I'm a political pessimist in that I am concerned about whether the science will be appropriately applied." ☒ Note that his technological optimism extends only to technology's ability to provide for a large population. It remains part of his view that technologies, even though some of them have marvellously enhanced our human existence, also contain aspects which can lead us to be pessimistic about our future. And this is separate from his political pessimism about the application of science.

The following is a quick survey of the global catastrophes that Rees claims have a fair chance of undermining our existence. Most are connected in some way to us humans and our evolving technologies. Let us set aside some of the catastrophes we face which are naturally occurring. For example, we have good evidence to think that the dinosaurs died out because of a devastating meteor impact with the Earth. There is no reason why some future meteor impact could not do the same to us humans. In fact given the frequency of close encounters with outer-space objects this is not a wholly unlikely occurrence. Instead let us focus on the catastrophes which may be due to humans and their activities. Individuals can perform only quite limited devastation with just their bare hands; but they can perform acts of mega-terror with their hands when assisted by some technologies.

1. *Nuclear Energy*. Nuclear warfare is something the horrors of which we have been familiar with since the end of world War II. Such wars may be deliberately caused or accidental; and they are taken to include the effects of any "nuclear winter" which might follow. Despite what has been said about the balance between hostile parties achieved through nuclear deterrence, many of us remain unconvinced by the deterrence doctrine and the threat of nuclear disaster still looms large. ☒☒Here people have in mind the fact that nuclear war can be accidentally started. Luckily there are those such as the Soviet Colonel Stanislav Petrov who, on 26 September, 1983, noted that his early warning system indicated that 5 USA missiles were about to attack the Soviet Union; but instead he realised that the warning was a false one and he did not order a retaliatory strike. And he was right; what the system detected was a rare alignment of the sun's reflection off clouds and orbiting satellites. Events such as these do not fill people with any confidence about warning systems and suggest that we have merely been lucky in avoiding nuclear war. Our deterrence systems must always be correct about whether one is under attack or not; one needs to be wrong only once to unleash a

nuclear disaster.

Nuclear energy can lead to other catastrophes. Domestic nuclear reactors have malfunctioned producing much devastation. For example, the consequences of the Chernobyl and Fukushima nuclear disasters are still with us. Finally, there is the possibility of "dirty bombs" being unleashed by terrorists, or nuclear power plants being attacked by terrorists. Rees discusses these possibilities in a section called 'Nuclear Mega-terror" (pp 43-7). There is already enough fiction (books, films, TV programmes) depicting these grim, but local, catastrophes. Though some of our nuclear technology has been beneficial, other aspects of it can be a threat to us humans, and clearly an existential threat. An extreme threat would be the total extinction of mankind. Slightly less extreme would be the survival of humans in a collapsed civilisation (perhaps akin to a pre-agricultural human existence). It is important to recognise the different kinds of threat and different kinds of subsequent collapse which envisaged.

2. *Biological Risks* (Rees, pp. 47-60). We are acquainted with naturally occurring pandemics throughout human history and in recent times pandemics of smallpox, polio, measles, SARS, and these days COVID-19. Even when naturally occurring, our technology has helped these pandemics thrive. It is no accident that with the pandemic of COVID-19 the world's airline industry has collapsed; it was one of the main ways in which the infection was spread around the world. In a different vein is biodiversity loss with the growth in the number of humans and human technology. In an alarming fashion it has recently been claimed that the world's insect population is in decline – surely a by-product of human agricultural activity. And it is often pointed out that owing to human activity we are currently in the middle of the 6th mass extinction of life on this Earth – the Anthropocene extinction.

Finally, there is biological terrorism. Rees claims that it is only a short time before terrorists turn from the gun to biological weapons which they can make after a little study of the necessary biology (perhaps undergraduate university level). We have already seen attacks involving the nerve gas sarin, and various attempts to spread anthrax. There is no reason why some "loner" cannot develop such agents in their own laboratory and then find some effective way to spread them.

The world has a Biological Weapons Convention (BWC) which needs to be funded and upheld as well as updated and checked for compliance. However due to international political disputes the BWC has serious shortcomings which need to be addressed. (Toby Ord in his book *Precipice* (2020) estimates that BWC's funding is smaller than the average McDonalds.) The COVID-19 pandemic reveals the very tangible risks the world faces if the BWC is not strengthened. One need only consider issues around whether Saddam Hussain of Iraq had weapons of mass destruction to show how difficult these matters can become politically. Biological terrorism can yield high levels of existential threat in which much of the human population can be wiped out.

3. *Climate change* (Rees, chapter 8). Hopefully by now arguments that climate change does not have some anthropogenic origin are receding fast and it is recognised that human activity has made a large contribution to it. Already during the current COVID-19 epidemic, even the casual observer can note that with humans under some form of "lockdown", so they are much less active, the skies are much clearer of pollution (e.g., Beijing, Los Angeles, etc). Once we cease to be preoccupied

with the current pandemic, we will have to return to the issue of climate change. The pandemic may go away but climate change will not. It remains to be seen if we could survive radical climate change but in a collapsed civilisation (i.e., something akin to humans in previous ice ages).

4. *Cyberterrorism*. We are all aware that our individual computers can be subject to various "viruses" and to cyberattack. However such attacks can be extended to systems of computers involved in critical national infrastructures such as a nation's banking systems, its government, its policing organisations, its businesses, its stock markets, its military and defence organisations and its voting systems (if they are computerized). For such reasons organisations go to great lengths to protect their computer systems from attack. They know how debilitating and destructive such attacks can be. Of equal concern is when computer systems are secretly infiltrated for the information they store, thereby the undermining privacy of citizens.

However, one can ask if this is really an existential threat? I am old enough to have been a student when no computers existed. But now I can appreciate the way in which they have enhanced human activity. To suddenly find that computers cannot be used (due to virus infection or whatever) would be vastly inconvenient and a threat to our current style of life; but this would not be a threat to our very existence. Not all threats to some aspect of our existence (such as our use of computers) are threats to existence itself. Here we also need to distinguish not only the different kinds of threat but the level of the threat (high, middling or low) and the probability of occurrence of the threat. These are matters for further research into claims about the threat of cyberterrorism.

5. *Risks from Artificial intelligence*. (Rees, pp. 18-21) Though this is more in the realm of futurology, Rees does consider the possibility that the machines we make may one day surpass the intelligence of us humans. The point at which this occurs more ardent futurologists call the "singularity". After the singularity, do we humans then become subordinate to machines? Or do we lose our humanity? Or do we need to expand our notion of what counts as human to include the new machines?

Whether or not there is a singularity, many researchers agree that as we progress in investigating AI, new possibilities are opened for us humans in our interaction with the non-biological intelligent machines. It has been argued that we have little idea what these future interactions will be like, and here is said to lie a serious problem for us humans due to our current ignorance. What we ought to be doing now is developing research to ensure that future progress in AI is safe and beneficial for us. As Huw Price (a co-founder with Rees of CSER) puts the matter: "We need the best of human intelligence to make the best of artificial intelligence."

Popular fears over these matters are widespread. Recently no less than Henry Kissinger has voiced fears concerning a burgeoning AI which we have to get under control least it overwhelms us. His 2018 article in *The Atlantic* is ominously entitled 'How the Enlightenment Ends'! (*The Atlantic* June 2018.) A good survey of the issues here can be found in Müller, 'Ethics of Artificial Intelligence and Robotics', *Stanford Encyclopedia of Philosophy*.)

Again, we can ask: is this some kind of threat to our existence or some kind of non-fatal threat to some aspect of our existence? It is hard to answer this given that even those who propose its possib-

ility are not sure how the AI-based threats might come about. Those who advocate a future singularity tend to exaggerate the threat to humans on at least one scenario in which the future machines are said to enslave us rather than be nice to us. The level of threat and its probability are unclear. However Toby Ord in his excellent book on existential threats, *Precipice* (2020), estimates that 'humanity spends more on ice cream every year than on ensuring that the technologies we develop do not destroy us' (Chapter 2, 'Our Neglect of Existential Risks'). (Importantly in chapter 6 Ord gives his reasons for rating highly the risks we face due to artificial intelligence as 1 in 10 over our century! These are Ord's judgements of degree of belief which he says have some evidential basis.)

Research on Risks

The above are some of the important technological crises Rees alleges faces humans. But we can agree that for most of us fears of such impending catastrophes need not always be rationally well justified fears. As a result, there have emerged several well-funded centres of research set up to investigate these issues. This is all to the good as the matters raised do need careful exploration before we can live happily with them. In this area quite often claims of catastrophe are made without much, or no, evidence for them. We need to know, at least for some possible outcome, the level of the threat (high, low, etc) and the probability of it. But quite often the probabilities are subjective best guesses and not, as they ought to be, probabilities based on evidence or calculation. To resolve these problems research is vital. Finally, some catastrophe may have a very small probability in occurring but would have a huge disvalue if it were to occur. Here serious problems are raised if we were to make important applications of Decision Theory to any catastrophe we might face.

One of the research centres already mentioned is the Cambridge Centre for the Study of Existential Risk (CSER) with which Rees is involved. Also important is The Oxford Martin School and its various programmes. (Similar research groups have been set up around the world at other universities including the University of Auckland). Importantly the Oxford Martin School includes the Programme on Misinformation, Science and Media. The concerns of this programme are not listed by Rees, but they ought to be. It attempts to address issues concerning the public discussion of science, countering misinformation about science, confronting non-scientific views which clash with science and decision-making involving science (usually under conditions of uncertainty).

That these issues have become pertinent is both surprising and disappointing. We will not get far in public discussion addressing Rees' concerns about technology if the underlying science is not better understood. Given elementary misunderstandings about, say, vaccination and a general dismissal of science by some politicians, we remain in a parlous position regarding an understanding of the very science which underlies the technology that Rees finds challenging. Worse than this is the low level of human understanding of probability, as much research in psychology into us as probabilistic reasoners has shown. (Note that to address some of these problems Passmore wrote a book *Science and its Critics*, 1978.)

A Note on Pessimism

The positions of Passmore and Rees have been called "pessimistic". What does this mean? Neither are philosophical pessimists, a stance in

philosophy from ancient times to contemporary philosophy which need not concern us. Pessimism has also been characterised as a mental attitude, usually of the depressed; but this does not fit either Rees or Passmore. Moreover, Rees and Passmore remain activists in attempting to combat those matters of which one might be pessimistic; so their pessimism is to be contrasted with fatalism or some other non-activist stance. What their brand of pessimism invites us to do is to face up to certain very unpalatable realities which we ought to redress in some way. So, their pessimism is not global but local in that it is directed to particular aspects of the world and our existence in it. For example, Passmore directs us to our flawed morality and flawed epistemic capabilities; and Rees directs us to the risks our technology poses for the inventors of it. And each of these can have different kinds of threat to our existence or the quality of it.

Pessimism can also be directed towards the belief in progress or the efficacy of religious faith. Passmore is strongly pessimistic about religion and rejects it while Rees is much less pessimistic about religion and attempts a reconciliation with it. More difficult to determine is whether we ought to be pessimistic about human progress. In general, both Passmore and Rees think that if we humans are capable of exercising our intellect then we can overcome some of those situations about which we are currently pessimistic. It remains to be seen if Passmore's book *The Perfectibility of Man* allows for a limited human progress. But these are issues which would take us beyond the present discussion. (To pursue them we would have to at least consider Passmore and Rees in the light of Steven Pinker's case for human progress made out in his recent book *Enlightenment Now* (2018), and in particular chapter 19, "Existential Threats" for Pinker's critique of a position like Rees's.)