Can Anti-Natalists Oppose Human Extinction?

Abstract: This article outlines a novel philosophical position according to which people should (a) value the long-term survival of humanity, and (b) oppose procreation on moral grounds. While these two propositions may appear to be contradictory, they need not be: Future life-extension technologies could enable members of a "final generation" to live indefinitely long lives and, therefore, to avoid human extinction. After examining a range of arguments for (a) and (b), I turn to a number of interesting complications for this position, which I call "no-extinction anti-natalism."

1. Introduction

This paper argues that one can accept, without contradiction, the following three propositions: (i) it would be better if no people had ever existed, (ii) it would be better if there are no more people, and (iii) human extinction would be a terrible tragedy that we should work to avoid. The most influential contemporary anti-natalist, David Benatar, endorses what he calls "dying-extinction," whereby the voluntary cessation of procreative activities results in the global population dwindling to zero, and he seems to believe that a pro-extinction view necessarily follows from his central thesis that coming into existence is always a net harm. But this is mistaken: Anti-natalists need not also favor human extinction. The argument presented in this paper goes as follows:

(1) Coming into existence is always a net harm.
(2) There are some lives worth continuing even though they weren't worth starting; furthermore, lives in the future will likely become even more worth continuing than lives today are.
(3) There are strong reasons for believing that most (but not all) instances of human extinction would constitute a terrible tragedy.
(4) Life-extension technologies could enable present or future people to live indefinitely long lives. Thus, such technologies could enable humanity to survive for an indefinitely long time.

(5) It follows that one can accept anti-natalism without favoring human extinction, as Benatar does. Insofar as one finds the moral intuitions behind anti-natalism compelling, one can espouse (1) and reject Benatar's pro-extinction view if one endorses the development and use of safe/effective life-extension technologies. Call the resulting position "no-extinction anti-natalism."

The following sections provide arguments for these premises. Section 2 recapitulates the basic anti-natalist argument that Benatar puts forward. Section 3 explains how one could oppose procreation while advocating for life-extension technologies that would enable humans to live indefinitely long lives, and section 4 explores a number of interesting implications and complications of no-extinction anti-natalism. I myself am not entirely convinced of these conclusions; thus, the primary aim of this paper is to fill-in a lacuna in the literature on anti-natalism. Indeed, I accept a distinction between *ideas worth considering* and *ideas worth accepting*. The position articulated in (5) above constitutes, at the very least, the former.

Finally, it is worth noting that the thesis of this paper could potentially increase the appeal of anti-natalism. As Benatar notes, many people intuitively accept the "starting point" of his argument, i.e., the harm-benefit asymmetry, yet they ultimately reject anti-natalism because its (apparent) conclusions—one being that humanity should go extinct sooner rather than later—are unsavory. For individuals who, like myself, are sympathetic with anti-natalism but also believe that human extinction would be among the worst things to ever happen, the present paper offers a middle path. One can, as the cliche goes, have one's cake and eat it too.

2. Benatar's Pro-Anti-Natalism Arguments

Benatar provides both philanthropic[1] and misanthropic arguments for anti-natalism. Within the former category, he offers a philosophical and empirical argument. Taking these in turn: the central philosophical case for anti-natalism hinges upon an asymmetry between pain (harms) and pleasure (benefits).

---

[1] Or, more accurately, "zoophilic" arguments, because Benatar's contentions apply more broadly to all sentient life. Given the prior meaning of this term as "having an attraction to or preference for animals," I suggest the term "sentiophilic" instead.

According to Benatar, the presence of pain is uncontroversially bad and the presence of pleasure is uncontroversially good. Less obvious is the additional assertion that "the absence of pain is good, even if that good is not enjoyed by anyone, whereas the absence of pleasure is not bad unless there is somebody for whom this absence is a deprivation" (Benatar 2006). Thus, a lack of pleasure is bad only if an existing person is actually deprived of that pleasure; otherwise this lack is "not good, but not bad either." In contrast, a lack of pain is positively good whether or not there exists a person to experience this state of lacking. I do not believe that the present paper requires a detailed reconstruction of Benatar's many arguments for this asymmetry, so the following will be somewhat cursory. I hope only to convince readers not especially conversant with Benatarian anti-natalism that this view is plausible.

To support the harm-benefit asymmetry, Benatar claims that it can explain four additional asymmetries that are "widely endorsed." He thus argues that his asymmetry may already be "widely accepted," even if people do not realize this. First, consider the common intuition known as the "procreation asymmetry." This arises from the belief that we have a duty not to bring into existence people who will suffer, but no corresponding duty to create new people who will be happy. According to Benatar, we have this intuition because a life worth living is good for those who exist but not bad for those who don't exist—since, once again, in the latter case there is no one who is deprived of this worthwhile life. But the presence of a miserable life is bad, while the absence of a miserable life is good—that is, even though there is no person who exists to experience the goodness of no misery.

Other intuitions that the harm-benefit asymmetry purportedly accounts for include the claims that (a) "it is strange (if not incoherent) to give as a reason for having a child that the child one has will thereby be benefited," but "it is not strange to cite a potential child's interests as a basis for avoiding bringing a child into existence," (b) "only bringing people into existence can be regretted *for* the sake of the person whose existence was contingent on our decision," and (c) discovering that a habitable island or exoplanet is uninhabited by happy people does not elicit the same degree of sadness as discovering that this island or exoplanet is populated by people who are suffering greatly (Benatar 2006).

Insofar as the harm-benefit asymmetry explains these common intuitions, Benatar argues that it should "have broad appeal" (Benatar 2006). Yet if one accepts this asymmetry, then one must also accept the anti-natalist proposition that bringing someone into existence is always a net harm. Consider the fol-

lowing two scenarios called “Existing” (X) and “Not-Existing” (~X). In X, the existence of an individual entails the presence of both pain and pleasure, the former being bad and the latter being good. In ~X, the non-existence of an individual entails the absence of both pain and pleasure, the former being good and the latter being not bad (or good). Now ask which of these scenarios is morally better. Since X yields a situation that is *good and bad* while ~X yields one that is *good and not bad*, it appears that ~X is better, from which it follows that non-existence is always preferable to existence.

The second, more empirical argument that Benatar presents aims to show not that coming into existence is always harmful, but that our lives are full of (a) great suffering, and (b) more suffering than most of us realize. As Benatar writes, “if people realized just how bad their lives were, they might grant that their coming into existence was a harm even if they deny that coming into existence would have been a harm had their lives contained but the smallest amount of bad” (Benatar 2006). Consider that there have existed roughly 100 billion humans so far, meaning that some 93 billion people have already died on planet Earth. Such deaths have been caused by murders, suicides, genocides, wars, disease, accidents, and aging. Indeed, over 15 million people may have perished in natural disasters in the last millennium; some 840 million people suffer from malnutrition and hunger, and roughly 20,000 die *every single day* from the latter; perhaps 133 million people died in mass killings before the twentieth century; some 110 million humans died in the hemoclysmic conflicts of the twentieth century; and diseases continue to trip millions of people into the eternal grave every year (see Benatar 2006).

In fact, as [redacted], the Black Death may have killed more people than World War II, the Taiping Rebellion, Mongol conquests, World War I, Napoleonic Wars, Vietnam War, American Civil War, 2003 Iraq War, and the War of 1812 *combined*—and public health experts claim that “we are at greater risk than ever of experiencing large-scale outbreaks and global pandemics,” where “the next outbreak contender will most likely be a surprise” (see redacted). The point is that the amount of human suffering throughout history is utterly unfathomable.[2] But so does the amount of misery caused by factory farming and, in the wild, the Darwinian struggle for existence (see Tomasik 2016, 2017). For example, consider Richard Dawkins’ (1995) observation that

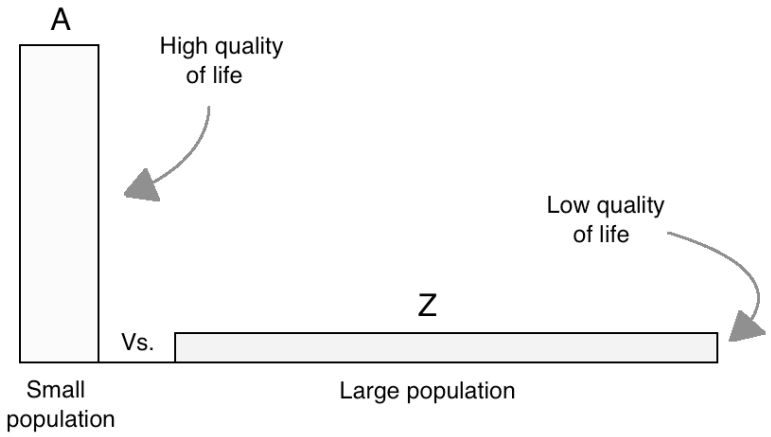
---

[2] Thanks in part to cognitive biases like “psychophysical numbing” that limit our ability to emotionally and cognitively register the true enormity of the harm

> the total amount of suffering per year in the natural world is beyond all decent contemplation. During the minute it takes me to compose this sentence, thousands of animals are being eaten alive; others are running for their lives, whimpering with fear; others are being slowly devoured from within by rasping parasites; thousands of all kinds are dying of starvation, thirst, and disease.

Benatar also argues that his empirical assessment that the proposition "life is very bad" is accurate independent of whether one accepts a hedonistic, desire-fulfillment, or objective list theory of well-being; and he claims that most people are self-deceived about the true quality of their lives, which are worse than we generally assume. This arises, Benatar argues, from psychological phenomena like the optimism bias and habituation, as well as our tendency to judge our lives by relative rather than absolute criteria—i.e., "A's life is *better than* B's; therefore, A's life must be *good*." A full exposition of these interesting points is beyond the scope of this paper, as mentioned above. Suffice it to say that there is a good case to make that human suffering—or, more generally, the suffering of sentient organisms on Earth—is enormous and self-reports about the goodness of one's life are often unreliable.

Before moving on to the next section, it is worth noting that Benatar's anti-natalism appears to solve a number of recalcitrant problems associated with population axiology. For example, it seems to imply that a "narrow" version of the person-affecting view could, in fact, solve the "non-identity problem." Parfit (1984) delineates this view as follows: "Suppose that we are comparing outcomes *X* and *Y*. Call the people who will exist in outcome X *the X-people*. Of these two outcomes, call X 'worse for people' in the narrow sense if the occurrence of X rather than Y would be either worse for, or bad for, the X-people." If we stipulate that X is "coming into existence" and Y is "not coming into existence," then X is "worse for people" than Y, given the harm-benefit asymmetry. This solves the non-identity problem by supporting a version of Parfit's "no-difference view," according to which "if A and B are two situations, and if the only difference between them is that A is a different people choice and B is a same people choice, then there is no moral difference whatsoever between A and B" (Mulgan 2009). That is to say, the identity of the relevant individuals in A and B is immaterial; moral intuitions about A should apply no less



to B and *vice versa*. Even more, Benatar's position can also avoid the "repugnant conclusion" and "mere addition principle" that undermine total and average utilitarianism (see Greaves 2017). Since smaller populations are always preferable to larger populations, this (a) makes Parfit's "population Z" (consisting of many people with very low-quality lives) worse than "population A" (consisting of few people with very high-quality lives), and (b) renders *any* addition of people—whether their quality of life is higher *or* lower than the average—morally wrong. (See figure 1.)

Thus, there are multiple quite cogent, in my view, arguments for propositions (i) and (ii) in section 1: It would have been better if there were no people at all, and indeed it would be good if humanity desists from bringing new people into the world. In addition, Benatarian anti-natalism appears to circumvent some of the most intractable conundrums in contemporary ethics. Yet few people would self-identify as "Benatarian anti-natalists," largely because of what this position *seems* to imply, as discussed in the next section.

3. Until Entropy Death Do Us Part

Among the more unsavory implications of Benatar's view is that it would be best if human extinction were to occur sooner rather than later, that is, as the result of people voluntarily refraining from procreation. Benatar refers to this strategy for ending all human life a "dying-extinction" scenario and contrasts it with "killing-extinction" scenarios that are brought about by natural or anthropogenic disas-

ters. Crucially, Benatar seems to believe that his pro-extinction view follows necessarily from his central thesis "that there should, ideally, be no (more) people." Or, as he elaborates,

> my answer to the question "How many people should there be?" is "zero." That is to say, I do not think that there should ever have been any people. Given that there have been people, I do not think that there should be any more (Benatar 2006).

But this doesn't actually entail that humanity must go extinct.[3] The reason has already been stated: Future technologies could enable people to live indefinitely long lives, thus achieving what I will call "functional immortality." Thus, a dual policy of "stop procreating entirely" (anti-natalism) and "develop safe and effective life-extension technologies" (immortalism) could satisfy the philosophical conclusions that Benatar defends while also avoiding the dismal prescription that humanity should voluntarily bring about its own demise.

This "dual policy" position is the heart of no-extinction anti-natalism, a view that should be appealing to Benatar himself given his anti-Epicurean view that death is (typically) bad and, as such, one ought to (typically) avoid it. The reason is that, according to Benatar, the avoidance of *any* harm "will be decisive." Thus, since "we (usually) have an interest in continuing to exist … death may be thought of as a harm, even though coming into existence is also a harm" (Benatar 2006). This gestures at an important distinction between "lives worth starting" and "lives worth continuing." Although Benatar argues that life is much worse for us than we realize, he nonetheless concedes that some lives aren't *so bad* as to warrant suicide. In this sense, they are worth continuing even though they weren't worth starting. It should be emphasized here that not only are many contemporary lives worth continuing, but advanced "person-engineering" and "world-engineering" technologies could make lives in the future *far more* worth continuing than lives currently are. I won't explore this point in detail here—an idea that [redacted] dubbed the "quality contention" (redacted)—but suffice it to say that there are *posthuman modes of being* that are

---

[3] In his words, "my arguments in [chapter 6] and previous ones imply that it would be better if humans (and other species) became extinct" and, elsewhere, "extinction … would result from universal acceptance of my view" (Benatar 2006).

almost certainly much more valuable, in many senses of that term, than our current mode of being, haphazardly shaped and molded as it was by millions of years of blind, non-teleological evolution. Using Benatar's phraseology, the "magnitude" of harm in the world could decrease significantly as advanced technologies improve the human condition and the "moral Flynn effect" continues to augment our circles of moral concern for sentient life (see Pinker 2011).[4]

Furthermore, there are a number of strong arguments *for* ensuring the perpetuation of our evolutionary lineage that are compatible with Benatar's anti-natalist utilitarianism. For example, Samuel Scheffler (2016, 2018) argues that most people embody a "love of humanity," and that this love is *revealed* by the putative fact that most of us would become despondent if we were to discover that human extinction were imminent, or even that human extinction were going to occur for certain, say, 100 years after our death. As Scheffler writes, referring specifically to future generations—because he doesn't consider the possibility of technological life-extension—"we have an interest in their survival in part because they matter to us; they do not matter to us solely because we have an interest in their survival" (Scheffler 2018). The key idea is that most of us care about the survival of our lineage far more than we realize, and indeed the assumption that we will persist into the future under relatively good conditions is an integral component of what gives us "value-laden lives."

Johann Frick (2017) propounds a similar argument, which he calls the "argument from final value." In brief, Frick observes that people commonly attribute "final value" to a range of phenomena, such as languages, species, and cultures, which suggests to him "that humanity too, with its unique capacities for complex language use and rational thought, its sensitivity to moral reasons, its ability to produce and appreciate art, music, and scientific knowledge, its sense of history, and so on, should be deemed to possess final value." If this is the case, then one should strive to ensure our continued survival since, quoting earlier work by Scheffler (2007), "what would it mean to value things but, in general, to see no reason of any kind to sustain them or retain them or preserve them or extend them into the future?" It follows that we should take actions to perpetuate humanity indefinitely into the future, given that "it would be very

---

[4] Notice here that I have, throughout this paper, habitually referred to "lives/people in the future" rather than "future lives/people." This was done to emphasize the possibility that future people aren't the result of a new generation being spawned by an older generation, but have themselves survived across transgenerational time periods.

bad, indeed one of the worst things that could possibly happen, if, for preventable reasons, the end came much sooner rather than later" (Frick 2017).

Other arguments for ensuring our survival include:

(1) We may be the only intelligent lifeforms in the galaxy or perhaps universe (see Sandberg et al. 2018). If so, one could argue that it would be a shame for such a special, unique feature of the furniture of the cosmos to disappear.

(2) Civilization could be understood as a "partnership" across time, where past generations have sacrificed for the sake of creating a better, happier, more prosperous future. Thus, it would be too bad—a real let down—if this partnership were to terminate due to the actions or inactions of those alive at the time.

(3) Many people, both past and present, have been motivated to create art, make scientific discoveries, and improve the world in various ways to achieve a kind of "vicarious immortality." It would, therefore, be "disrespectful" to their posthumous wishes and interests if we were to allow humanity to go extinct.

(4) As the negative utilitarian David Pearce (2007) argues, "*Homo sapiens* is the only species capable systematically of preventing harm to the rest of the living world—despite the frightfulness of what we do to non-human animals today. So it is vital that humans survive in the face of existential risks to bring the abolitionist project to completion," where the "abolitionist project" aims to used advanced biotechnology to eliminate as much suffering in the universe as possible.

And (5) moral uncertainty suggests that we should aim to keep our options open; since the ultimate form of closing our options would be to succumb to extinction, we should prioritize the avoidance of extinction (see Tonn 2009; Bostrom 2013). As Will MacAskill (2014) puts this point,

> we should expect ourselves to progress, morally, over the next few centuries, as we have progressed in the past. So we should expect that in a few centuries' time we will have better evidence about how to evaluate human extinction than we currently have. … In general, when one has the

> choice between two options, one of which is irreversible, and one expects to make moral progress, then option value gives one additional reason in favour of choosing the reversible option.

These are just a few reasons for maintaining that human extinction would be bad, and that are also compatible with Benatar's anti-natalism. To quote Wendell Bell (1993), writing about the importance of safeguarding our existence: "Without the possibility of a future, there is nothing left but despair. Thus, if we give up on the future, we give up on ourselves."

Section 4: Could No-Extinction Anti-Natalism Actually Work?

For the present purposes, I will not examine in detail the many *generic* problems associated with the development of effective life-extension technologies. Suffice it to say that there are serious questions about how society will have to restructure itself to accommodate people who will never retire, or whether people with indefinitely long lives will suffer from crushingly oppressive ennui, as some authors have argued.[5] There are also concerns that if anti-natalist policies are *not* implemented once humanity attains "actuarial escape velocity" (AEV), then overpopulation could bring about a Malthusian disaster, assuming that we remain sequestered on Earth (rather than having spread elsewhere in the universe) (see Häggström 2016; redacted).[6] By "AEV," I refer to scenarios in which, for example, anti-aging technologies can extend one's life, at T1, for another 20 years to T2, at which point, given the exponential pace of biomedical innovation, new anti-aging technologies can extend one's life, at T2, another 30 years to T3, *and so on*. As Aubrey de Grey (2004) explicates this idea,

> I term this rate of reduction of age-specific mortality risk "actuarial escape velocity" (AEV), because an individual's remaining life expectancy is affected by aging and by improvements in life-

---

[5] Recall here Soren Kierkegaard's quip that boredom is the root of all evil; see also, e.g., Williams 1973.
[6] See [redacted] for reasons why we *should* remain on Earth, at least for the foreseeable future.

> extending therapy in a way qualitatively very similar to how the remaining life expectancy of someone jumping off a cliff is affected by, respectively, gravity and upward jet propulsion.

He later notes that

> those who get first-generation therapies only just in time will in fact be unlikely to live more than 20-30 years more than their parents, because they will spend many frail years with a short remaining life expectancy (i.e., a high risk of imminent death), whereas those only a little younger will never get that frail and will spend rather few years even in biological middle age. Quantitatively, what this means is that if a 10% per year decline of mortality rates at all ages is achieved and sustained indefinitely, then the first 1,000-year-old is probably only 5-10 years younger than the first 150-year-old.

Thus, the threshold for "eternal life" *isn't* the realization of a technological strategy that can extend one's life forever, it is the moment at which each new advancement extends one's life *long enough* to benefit from the next advancement, until one achieves functional immortality. On the other hand, though, if anti-natalist policies *are* successfully implemented once we attain AEV, one might worry about how the lack of new generations could affect the life-quality of those in the long-lived final generation. As Larry Temkin (2008) writes, "I think if the cost of immortality would be a world without infants and children, without regeneration and rejuvenation, it wouldn't be worth it." Even more, consider that, at the behest of Joseph Stalin, "life-extension became a central subject of Soviet medical research" (Medvedev and Aleksandrovich 2006). One wonders how history might have differently unfolded if this research had yielded effective life-extension technologies and if Stalin had reached AEV. Would the Soviet Union still exist? Would Stalin still be in power? How might the Cold War have played out? The issue here can be rephrased as follows: Who gets to be included, and excluded, in the final generation, given that certain bad elements of humanity, as it were, could get "locked in" after reaching AEV?

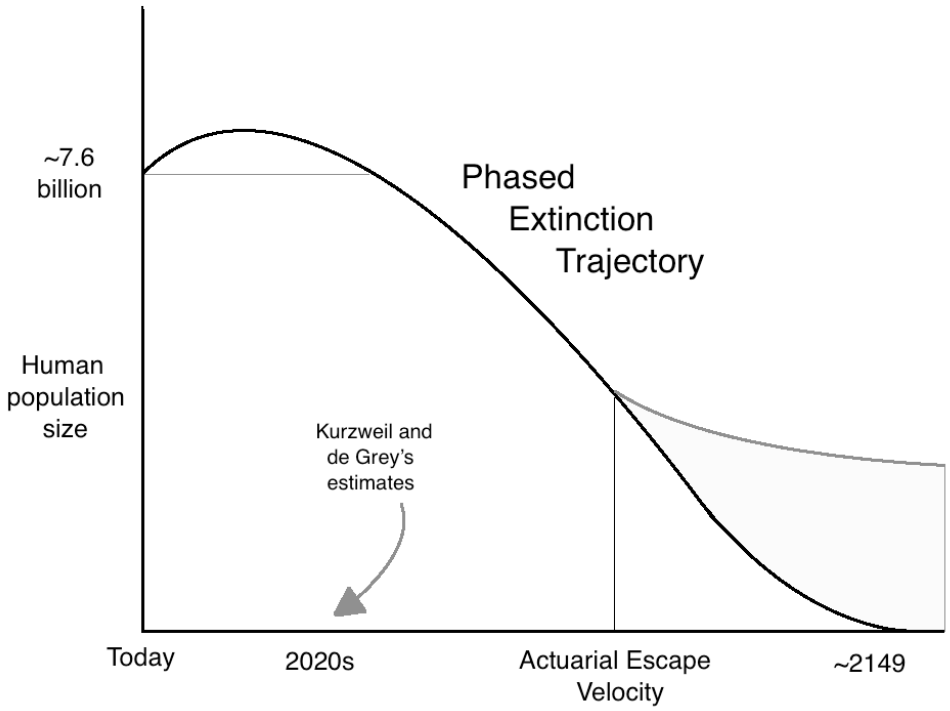
Along these lines, the question arises: Is it possible to justify continued procreation until moral progress yields a more ethical population of human beings that, as such, would increase the probability of

better outcomes in the long run? Relatedly, if one values the avoidance of human extinction and believes that death is bad, but also accepts the harm-benefit asymmetry, then are there resources within the Benatarian framework for permitting the creation of new people until we reach AEV? In fact, Benatar argues that there *are* circumstances under which further procreation could be justified. He claims that there are two moral positions that appear compatible with procreating to "phase-out" human extinction: (a) A "negative total view" according to which "we may create new people where the total amount of harm in doing so is equivalent to, or less than, the harm that would be suffered by existing people if the new people were not created," and (b) a "less stringent rights or deontological view" according to which "creating new people cannot be justified by mere reduction in total harm," but it "may be justified by substantial (but not mere) reduction in total harm" (Benatar 2006). (The narrow person-affecting view that Benatar previously endorses is unable to confer moral license to procreate, ever.)

Thus, can either of these views warrant the creation of new people to avoid extinction? First, if already existing individuals are going to live long enough to attain AEV, then there's no need to procreate, because this would enable humanity to satisfy the "no-extinction" desideratum of the present view, and thus there would not be a *dying* last generation whose plight is severely exacerbated by the absence of a younger generation that can provide them company, support, palliative care, and so on. Second, if existing individuals are not going to live long enough to attain AEV, neither the negative total view nor the rights or deontological view seems to offer any reasons for continuing to procreate until the AEV threshold, since both focus entirely on reducing harm to, and only to, *extant* individuals. That is to say, one is permitted to bring new people into existence only insofar as doing so could mitigate the misery of those already alive.

From a practical perspective, though, it may not matter whether one can morally justify procreation until the attainment of AEV. Consider the following statement from Benatar (2006):

> Whether the number of people could be reduced fast enough without the costs of rapid population decline, to a level where the number of final people was small enough to offset the harm to intervening generations, is a difficult one to answer. Whatever the answer, we can say that extinction



within a few generations is to be preferred to extinction only after innumerably more generations (italics added).

Here Benatar seems to allow for continuing procreation such that there are a "few" more generations. The length of a contemporary generation is ~25.5 years. Thus, if we interpret "few" as denoting "3" and assume that the average life expectancy in the future remains around 80 years, then the last humans alive would perish in about 131 years, or circa 2149, if a global policy of phased extinction were implemented today. In other words, the generation born today would have children in 25.5 years, and these children would grow up to have children 25.5 years later. These children of children of children would live another 80 years or so, which yields the twenty-second century date above.[7] Now consider that Ray Kurzweil, de Grey, and other "in-the-know" futurists conjecture that we could attain AEV by the 2020s. If these estimates appear overly optimistic, one can perhaps fall back on Nicholas Wade's historical observation that major advances in biotechnology—e.g., the human genome being sequenced—have often occurred *soon-*

[7] Of course, some humans could naturally live longer than 80 years—up to at least 122 years—but for the present purposes, we can bracket this.

er than experts anticipate. Thus, he writes that "longevity increases might be one of those big steps that arrive much sooner than expected" (Wade 2004). The point is simply that there is a good chance that *even if* Benatar's phased extinction proposal were implemented, humanity could still develop life-extension technologies *before* the final generation dies out, thus giving this generation the option of evading individual death and collective extinction. One might rejoin to this idea that a dwindling population would slow down scientific progress, thereby pushing back the AEV threshold. Yet this is unlikely to actually matter, since even if every human currently alive today were to cease all procreative activities starting *now*, there would still probably be (plenty) enough time for the scientific community (in particular, the currently growing demographic of scientists working on life-extension technologies) to prolong current human lives long enough to reach AEV.[8]

Yet another issue that problematizes no-extinction anti-natalism concerns the troublesome metaphysical topic of personal identity. First, consider one of the primary strategies for obtaining functional immortality, namely, "mind-uploading" or "whole-brain emulation." This can take three general forms: (a) *Destructive* (the original brain is destroyed either instantaneously or gradually), (b) *non-destructive* (the original brain remains fully intact while a hardware-substrate copy is made), and (c) *reconstructive* (the original brain dies but the person is recreated based on historical records of them) (Chalmers 2010; see also Sandberg and Bostrom 2008). For example, the "microtome procedure" involves freezing a brain to liquid nitrogen temperatures, slicing it into small sections, scanning these slices, transferring the gathered information about the brain's microstructural features to a computer, and then simulating the scanned brain. Another possibility is to scan the brain "from within … using nanobots," which Ray Kurzweil anticipates being "available by the late 2020s" (Kurzweil 2005). Whereas the former is destructive, the latter need not be. With respect to reconstructive uploading, it could become possible for a future superintelligence to collect enough data about past people to resurrect them from the grave. The idea here is predi-

---

[8] Note that, when discussing phased extinction, Benatar writes, "I am under no illusions. Although humans may voluntarily seek to reduce their number, they will never, under current circumstances, do so with the intention of moving towards extinction. Thus, in considering the question of phased extinction from a large population base, I am not discussing what will ever happen but only what should happen or what it would be best to have happen. Put another way, I am discussing the theoretical implications and applications of my views." I very much share this caveat: I am discussing AEV, population decline, and so on, from a purely theoretical perspective. I have little doubt that what is said in these pages will have almost zero impact on how future human history unfolds.

cated on a distinction between *clinical death* and *information-theoretic death*, where the latter refers to the point at which no future technologies could in principle retrieve sufficient information from one's nervous system to recreate the deceased individual (Merkle 2018). Although this is highly speculative, it could be that many people who have undergone clinical death (and ultimately sunk into thermodynamic equilibrium) haven't undergone information-theoretic death, and thus could be "magically" brought back to life at some point in the future.[9]

The point is that mind-uploading, if it "can be made to work," would constitute "the ultimate life-extension technology" (Häggström 2016) given that "uploads would not be subject to biological senescence" and "back-up copies of uploads could be created regularly so that you could reboot if something bad happened" (Bostrom 2003a). But would your uploaded mind be *you* in the metaphysical sense? There is a good case to make that minds are "organizational invariants" that, as such, are multiply realizable by any substrate with the right sort of functional organization, but is this true of selves as well? In other words, *does mind-uploading always or ever entail self-uploading*?

The most obviously problematic case involves non-destructive uploading, since this would result in numerically distinct conscious beings who are both psychologically continuous with the original. This entails that, if one accepts the ontological singularity of selves, then selves are not organizational invariants, because the uploaded mind would be (by definition) a functional isomorph. In this particular case, though, the personal identity issue is ultimately irrelevant to the ethical question of whether uploading one's mind in this manner would be morally unacceptable. The reason is this: (a) Independent of whether the uploaded mind is you or not, the metaphysical fact is that, after the upload event, there exists two distinct conscious beings with different experiences from that moment forward; (b) this implies that non-destructive uploading will result in the creation of an *extra* conscious being; thus, (c) since it is always wrong to create an extra conscious being, it would always be wrong to upload one's mind in this manner. Anti-natalists should therefore strongly oppose non-destructive mind-uploading no matter how one answers the personal identity question. A similar conclusion follows with respect to reconstructive uploading: Whether or not a reconstructed functional isomorph (say, by a machine superintelligence in 300 years

---

[9] See Arthur Clarke's "third law," which states that "Any sufficiently advanced technology is indistinguishable from magic."

that combs through all of my publications, social media accounts, video and podcast interviews, etc.) is *me* in the metaphysical sense, if the resulting mind is conscious, then Benatarians should object to this form of uploading.

The much more complicated case involves destructive uploading. Consider a piecemeal process of destructive uploading whereby each cell in one's brain is replaced by a cellular functional isomorph that receives and transfers information in exactly the same way as biological neurons. Mark Walker (2008) calls this approach "gradualism," and it would presumably preserve continuity of consciousness from start (T1) to finish (T2), at which point the wetware of the brain would be entirely supplanted by non-biological hardware that could enable a digital form of functional immortality. This being said, if the person at T2 is the same as the person at T1, then Benatarians should applaud this form of uploading, since it would enable the uploader to evade death, which is (usually) bad. However, if the person at T2 is not the same as the person at T1, then we have a *doubly bad* situation. The reason is that the process of P1 (at T1) becoming P2 (at T2) would not only entail the creation of a new person, P2, which would be wrong, but the death of P1, which would also be wrong. Thus, from the anti-natalist perspective here defended, the metaphysical question of whether gradual destructive uploading preserves personal identity across time is absolutely crucial. So, does it?

I agree with Chalmers (2010) that questions of personal identity are deeply confounding; perhaps they are, in the sense of Colin McGinn (1993), permanently unknowable "mysteries" relative to the concept-generating mechanisms available to us. Nonetheless, some philosophers have defended various positions with respect to this question. For example, Susan Schneider and Joe Corabi (2014) have put forth a rather clever argument (not here recapitulated) for why destructive uploading of this sort would not preserve identity. In their words, if

> all that is required for this continuity [of consciousness] is that later mental states be caused by or be qualitatively similar to earlier ones (e.g., that qualitatively similar thoughts be entertained, or that later sensory "memories" be qualitatively similar to the original experiences), then plainly the sort of continuity in question could be shared by numerous individuals at later times; what if the

> information were sent to two computers instead of just one, after all? (Schneider and Corabi 2014)

In other words, since selves are singular entities, and since (gradually, destructively) uploaded minds could be duplicated, an uploaded mind couldn't be the same person as the original. Along somewhat different lines, Massimo Pigliucci borrows from John Searle's biological naturalism to argue that conscious minds are *not*, in fact, organizational invariants. Thus, mind-uploading is nothing more than "a very technologically sophisticated (and likely very, very expensive) form of suicide" (Pigliucci 2014). Yet another pessimistic view comes from Nicholas Agar (2016), who contends that mind-uploading "does not satisfy a necessary condition for the transmission of human identities," since it constitutes a "significant interference in the properties and processes that normally accompany our survival." It follows that, on Agar's view, "the replacement of human biology by machines may result in significant enhancement, but it will not result in significant enhancement for us" (Agar 2016). The point is that if one accepts (a) the harm-benefit asymmetry, and (b) that death is (usually) bad, then the lack of a strong consensus that piecemeal destructive uploading would preserve personal identity yields an extremely good reason to reject this form of uploading. That is to say, given the stakes—suicide or murder paired with a kind of procreation—one should oppose uploading of this sort until we can be *very sure* that the same person would persist from T1 to T2. Furthermore, for reasons discussed in Chalmers (2010), if one is uncertain about gradualistic uploading, then one should be even more uncertain about *instantaneous* strategies for replacing the neurons in one's head with functional isomorphs. I will not here elaborate this point because of space constraints.

Finally, it is worth adding that attempts to emulate entire nervous systems could themselves pose some serious ethical hazards. First, as Bostrom notes, "before we would get things to work perfectly, we would probably get things to work imperfectly" (Bostrom 2014). The result could be that imperfectly simulated brains experience moments of truly intense suffering, perhaps in the form of psychotic hallucinations or delusions, grand mal seizures, and so on, before a normal state of consciousness and mentality is established. Second, since emulating parts of the human brain will likely antedate emulating a whole human brain, some form of "neuromorphic AI"—that is, a system that combines simulations of human brain regions with synthetic AI patches—will probably occur before whole-brain emulation. Insofar as

this AI system is (i) superintelligent, and (ii) has a value system that is not sufficiently aligned with our "human values," then we have reason to worry about it bringing about a killing-extinction event (that is, for reasons pertaining to the "instrumental convergence thesis"). In other words, the road to mind-uploading could itself pose grave threats to our survival (see Bostrom 2014).

Now consider an alternative route to achieving functional immortality, namely, "strategies for engineered negligible senescence" (SENS). This theoretical framework, which was pioneered by de Grey, identifies nine primary types of deleterious, cumulative change associated with aging: Cell loss (without replacement); oncogenic nuclear mutations and epimutations; cell senescence; mitochondrial mutations; lysosomal aggregates; extracellular aggregates; random extracellular protein cross-linking; immune system decline; and endocrine changes (de Grey 2003). It follows that interventions to halt or reverse these changes could halt or reverse aging itself. Such interventions might employ biotechnology in the relative near future, but eventually the relevant tasks could be taken over by advanced nanomedicine techniques that we can barely glimpse from our current technological vantage point. Less radically, "geroprotectors" like metformin, an anti-diabetes drug that has been shown to reduce "all-cause mortality and diseases of ageing independent of its effect on diabetes control," could extend one's life by years (Campbell 2017). This is notable because a few extra years could enable a potentially large number of currently living individuals to reach AEV, and thus attain functional immortality.

The point is that, while mind-uploading would entirely replace the biological substrate of *Homo sapiens* with non-biological hardware, SENS would result in biology-based beings that do not senesce. In either case, though, I propose the following (uncontroversial) descriptive hypothesis: If humanity becomes functionally immortal, it will almost certainly undergo evolutionary changes at some point by modifying its own phenotypic traits, thus resulting in one or more *posthuman species* that are either fully biological, technobiologically hybrid, or wholly artificial in nature. Indeed, Milan Ćirković and Robert Bradbury (2006) identify the transition from humanity to posthumanity as the next "mega-trajectory" in the 3.8-billion-year history of terrestrial life. With respect to humans who attain functional immortality through SENS (or something like it), this could occur via genetic interventions, nootropics, brain-computer interfaces (BCIs), and so on; with respect to uploaded minds, enhancement could be even easier to achieve given that emulated brains don't bleed and errors could be easily reversed by reverting to an ear-

lier saved copy. If this hypothesis is true, then it introduces questions of personal identity that are similar to those above: Does the transformation of a human (including uploaded humans), H, into a posthuman, H+, preserve one's personhood? In other words, is Joe the same person, in the metaphysical sense, after he gains 300 IQ points, can detect infrared light with bionic eyes, has access to completely novel emotional experiences, and so on, as we was before these changes? Note that these questions are additional to whether the process of mind-uploading preserves identity. That is to say, whether or not piecemeal destructive uploading preserves identity, there remains a further question of whether the uploaded mind remains the same person over time if it undergoes radical cognitive, emotional, moral, etc. enhancements.

Thus, paralleling the case above, (a) if the transformation from H to H+ preserves identity, then this could render life far more *worth continuing* than before by increasing the individual's moral well-being, life opportunities, knowledge, and so on. But (b) if the transformation from H to H+ does not preserve identity, then this could be, once again, *doubly bad* given the harm-benefit asymmetry and Benatar's anti-Epicurean view of death. The reason for the latter is that H becoming H+ would entail that H no longer exists (a form of death) while also introducing a new, in the metaphysical sense, conscious entity (a form of procreation, albeit *developmental* rather than *generational*). In fact, numerous philosophers—including some transhumanists who are sympathetic with the radical human enhancement project—have argued that becoming posthuman would not in fact preserve one's identity. For example, Mark Walker (2008) has outlined an "Aristotelian identity argument" according to which "some changes may be so drastic that they will mean that I cease to exist." Or, to use Aristotle's own example, if a "man wishes his friend's good for his friend's sake, the friend would have to remain the man he was. Consequently, one will wish the greatest good for his friend as a human being" rather than him becoming a superhuman "god" (Walker 2008). A stronger view has been defended by Susan Schneider, who contends that "even mild enhancements are death inducing." In her words,

> for radical enhancement to be a worthwhile option for you, it has to represent a desirable form of personal development; at bare minimum, even if enhancement brings such goodies as superhuman intelligence and radical life extension, it must not involve the elimination of one of your essential properties. *For in this case, the sharper mind and fitter body would not be experienced by*

> *you—they would be experienced by someone else*. For even if you would like to become superintelligent, knowingly embarking upon a path that trades away one or more of your essential properties would be tantamount to suicide—that is, to your intentionally causing yourself to cease to exist (Schneider 2008).

Thus, from the perspective of no-extinction anti-natalism, the *possibility* that becoming posthuman could fail to preserve personal identity poses a major theoretical and practical problem, given the hypothesis above that becoming posthuman would almost certainly occur if humanity achieves functional immortality. My own intuitions, as it happens, perhaps align best with the "no-self" position defended by James Hughes (2006).[10]

One final related point ought to be considered as well. Independent of whether some form of mind-uploading could preserve personal identity, once a mind M is uploaded, M could be duplicated an infinite number of times. On the one hand, this reintroduces problems of personal identity: If M is copied to produce M', then which would be personally identical to the original? On the other hand, there are anti-natalist reasons for maintaining that duplication would be deeply unethical, given the harm-benefit asymmetry. Notice the difference between this and the case of "H —> H+": Whereas the transformation from human to posthuman involved two qualitatively non-identical but numerically identical beings (that are spatiotemporally and perhaps consciously continuous over time), the duplication of already uploaded minds would involve creating two numerically non-identical but qualitatively identical beings (that is, at the moment of duplication).[11] Once more, anti-natalists ought to oppose the duplication of M to yield M', since the non-existence of M' would result in an absence of pain, which is good, and pleasure, which is not bad, whereas the existence of M' would result in the presence of pain, which is bad, and the presence

---

[10] Indeed, Hughes imagines that mind-uploading will usher in a "post-individual age" whereby "we will be able to copy, share and sell our memories, beliefs, skills and experiences. We will selectively adopt personalities for specific purposes—Machiavelli for politics, Cyrano for love. … Some people will live broadcast VoyeurLives, just as some now put VoyeurCams in their homes, and others will choose to spend a lot of time in someone else's life—like climbing into John Malkovich's head for weeks instead of 15 minutes at a time. … Personalities will begin to bleed and blur. We will write copious reams about the decline of the old discrete, continuous self, and the rise of the new creative, collaborative self-process. … The most dramatic challenges to our social and philosophic world will probably come from hive minds and distributed selves (Hughes 2006)."

[11] Thus, the former result in *replacement* whereas the latter result in *addition* (like procreation). The addition phenomenon also applies to non-destructive uploading, of course.

of pleasure, which is good. Again, a *good and not bad* situation is preferable to a *good and bad* situation, meaning that duplication would be morally unacceptable.

This point is directly relevant to various futurological speculations that uploaded minds—i.e., "ems" for short—could create short-lived duplicates, or "spurs," to complete specific tasks, after which they would be terminated. As Robin Hanson (2016) speculates,

> most ems … are comfortable with often splitting off a "spur" copy to do a several hour task and then end, or perhaps retire to a far slower speed. They see the choice to end a spur not as "Should I die?" but instead as "Do I want to remember this?" At any one time, most ems are spurs.

This is, in my view, both unconvincing and disturbing. If spurs are conscious persons in their own right, where a person is "a thinking intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking thing, in different times and places" (Locke 1689), then terminating them would be tantamount to murder. In Sandberg's words, "if ending the identifiable life of [a spur] is a wrong, then it might be possible to produce a large number of wrongs by repeatedly running and deleting instances of an emulation even if the experiences during the run are neutral or identical" (Sandberg 2014).[12] Even more, the anti-natalist perspective entails that *any* instance of duplicating already uploaded minds would instantiate a "mind crime"—that is, whether or not the resulting copy were treated well, allowed to survive indefinitely, and so on (see Bostrom 2016).

Lastly, we should note that the no-extinction anti-natalist view does not address the various misanthropic arguments that Benatar propounds. Here one could retort that, if those living in the future become posthumans (or post-*anthropos*), then the reasons for disliking humans (or *anthropos*) might not transfer to this new species. It could indeed be possible for some combination of cognitive enhancements and mostropics—along the lines of Persson and Savulescu's (2012) proposal—to enable humanity to overcome the tragedy of the commons that constitutes the primary social cause of climate change. In oth-

---

[12] Note that if we someday upload, duplicate, or simulate a *large number of minds like ours*, then this would constitute evidence that posthuman civilizations tend to simulate a large number of minds, and this would constitute evidence that we are probably inside a simulation (Bostrom 2003b). Yet another issue worth considering.

er words, it could be possible to solve the global-scale coordination issues that are responsible for the changing climate. Along these lines, our expanding circles of moral concern, driven by the aforementioned moral Flynn effect (Pinker 2011), could lead humanity to care more about biodiversity loss and other environmental harms. Our species could, through attitudinal maturation and/or directed artificial evolution, become much wiser, less violent, more peaceable, more compassionate, kinder, happier, and so forth. If this occurs, then the "superb misanthropic argument against having children and in favour of human extinction" that "rests on the indisputable premiss that humans cause colossal amounts of suffering" could become irrelevant (Benatar 2006). In this sense, the present view offers at least some hope of overcoming the misanthropic argument for favoring human extinction.

5. Conclusion

Benatar endorses neither pro-mortalism nor pro-immortalism, as it were, although he could endorse the latter without giving up the harm-benefit asymmetry. If Benatar were to do this, then he could continue to advocate for anti-natalism and also hold that humanity should not go extinct—the second issue being a primary reason that people do not call themselves "Benatarian anti-natalists." Indeed, not only could one simultaneously (i) accept a proscription against procreation, and (ii) embrace the hope that "eternal life" has to offer, but there are reasons for believing that the human (or posthuman) condition could be immensely better in the future, thus yielding lives that are far more worth continuing than worth-continuing lives are today. To put this point in terms of E.E. Cummings's short poem "Into the Strenuous Briefness," the lives of people in the future need not be either brief or strenuous. On a personal note, I find that I care deeply about the continued existence of humanity; yet the harm-benefit asymmetry in particular also strikes me as extremely plausible. Thus, the present paper has attempted to provide a mechanism for relieving this cognitive dissonance. I am, once again, not fully convinced that no-extinction anti-natalism is the correct moral interpretation of the relevant issues—indeed, some of the complications above are formidable—but, at the very least, I believe that this position deserves to be taken seriously in the philosophical marketplace of ideas.

References:

Adams, Fred. 2008. Long-Term Astrophysical Processes. In Nick Bostrom and Milan Ćirković (eds.), *Global Catastrophic Risks*. Oxford: Oxford University Press.

Adams, Robert. 1989. Should ethics be more impersonal? A critical notice of Derek Parfit, Reasons and persons, *Philosophical Review*. 98 4): 439-484.

Agar, Nicholas. 2016. Enhancement, Mind-Uploading, and Personal Identity. In Steve Clarke, Julian Savulescu, C.A.J. Coady, Alberto Giubilini, and Sagar Sanyal (eds.) *The Ethics of Human Enhancement: Understanding the Debate*. Oxford: Oxford University Press.

Althaus, David, and Lukas Gloor. 2018. Reducing Risks of Astronomical Suffering: A Neglected Priority. Foundational Research Institute. https://foundational-research.org/reducing-risks-of-astronomical-suffering-a-neglected-priority/.

Bell, Wendell. 1993. Why Should We Care About Future Generations? In H.F. Didsbury, Jr. (ed.) *The Years Ahead: Perils, Progress, and Promises*. Bethesda, MD: The World Future Society.

Benatar, David. 2006. *Better Never To Have Been: The Harm of Coming Into Existence*. Oxford: Oxford University Press.

Bostrom, Nick. 2003a. The Transhumanist FAQ: A General Introduction, Version 2.1. https://nick-bostrom.com/views/transhumanist.pdf.

Bostrom, Nick. 2003b. Are You Living in a Computer Simulation? *Philosophical Quarterly*. 53(211): 243-255.

Bostrom, Nick. 2008a. Three Ways to Advance Science. *Nature* Podcast. https://nickbostrom.com/views/science.pdf.

Bostrom, Nick. 2008b. Why I Want to be a Posthuman When I Grow Up. In Bert Gordijn and Ruth Chadwick (eds.) *Medical Enhancement and Posthumanity*. New York, NY: Springer.

Bostrom, Nick. 2013. Existential Risk Prevention as Global Priority. *Global Policy*. 4(1): 15-31.

Bostrom, Nick. 2014. *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press.

Bratsberg, Bernt, and Ole Rogeberg. 2018. Flynn Effect and Its Reversal Are Both Environmentally Caused. *Proceedings of the National Academy of Sciences*. Epub ahead of print.

Burke, Edmund. 1834. Reflections on the Revolution in France. In *The Works of Edmund Burke*. London: Holdsworth and Ball.

Campbell, Jared, Susan Bellman, Matthew Stephenson, and Karolina Lisy. 2017. Metformin Reduces All-Cause Mortality and Diseases of Ageing Independent of Its Effect on Diabetes Control: A Systematic Review and Meta-Analysis. *Ageing Research Reviews*. 40: 31-44.

Chalmers, David. 2010. The Singularity: A Philosophical Analysis. *Journal of Consciousness Studies*. 17(9-10): 7-65.

Ćirković, Milan, and Robert Bradbury. 2006. Galactic Gradients, Postbiological Evolution, and the Apparent Failure of SETI. *New Astronomy*. 11(8): 628-639.

Collings, David. 2014. *Stolen Future, Broken Present: The Human Significance of Climate Change*. Ann Arbor, MI: Open Humanities Press.

Dawkins, Richard. 1995. *River Out of Eden: A Darwinian View of Life*. New York, NY: Basic Books.

de Grey, Aubrey. 2003. An Engineer's Approach to the Development of Real Anti-Aging Medicine. *Science of Aging Knowledge Environment*. 2003(1).

de Grey, Aubrey. 2004. Escape Velocity: Why the Prospect of Extreme Human Life Extension Matters Now. *PLOS*. 2(6).

Falk, an. 2018. Why Some Scientists Say Physics Has Gone Off the Rails. NBC. https://www.nbcnews.com/mach/science/why-some-scientists-say-physics-has-gone-rails-ncna879346.

Feynman, Richard. 1965. *The Character of Physical Law*. Cambridge, MA: The MIT Press.

Frick, Johann. 2017. On the Survival of Humanity. *Canadian Journal of Philosophy*. 47(2-3): 344-367.

Greaves, Hilary. 2017. Population Axiology. *Philosophy Compass*. 12(11): 1-21.

Häggström, Olle. 2016. *Here Be Dragons: Science, Technology, and the Future of Humanity*. Oxford: Oxford University Press.

Hanson, Robin. 1998. The Great Filter: Are We Almost Past It? http://mason.gmu.edu/~rhanson/greatfilter.html.

Hanson, Robin. 2016. *The Age of Em: Work, Love, and Life When Robots Rule the Earth*. Oxford: Oxford University Press.

Hughes, James. 2006. The Illusiveness of Immortality. In Charles Tandy (ed.) *Death And Anti-Death, Volume 3: Fifty Years After Einstein, One Hundred Fifty Years After Kierkegaard*. Ann Arbor, MI: Ria University Press.

Kaku, Michio. 2005. *Parallel Worlds: A Journey Through Creation, Higher Dimensions, and the Future of the Cosmos*. New York, NY: Doubleday.

Knoll, Andrew, and Richard Bambach. 2000. Directionality in the History of Life: Diffusion from the Left Wall or Repeated Scaling of the Right? *Paleobiology*. 26(sp4): 1-14.

Kurzweil, Ray. 2005. *The Singularity is Near: When Humans Transcend Biology*. New York, NY: Viking Penguin.

Locke, John. 1689. *An Essay Concerning Human Understanding*. London: William Tegg & Co., Cheapside.

MacAskill, Will. 2014. *Normative Uncertainty*. Dissertation. http://commonsenseatheism.com/wp-content/uploads/2014/03/MacAskill-Normative-Uncertainty.pdf.

Metzinger, Thomas. 2017. Benevolent Artificial Anti-Natalism (BAAN). *Edge.org*. https://www.edge.org/conversation/thomas_metzinger-benevolent-artificial-anti-natalism-baan.

McGinn, Colin. 1993. *Problems in Philosophy: The Limits of Inquiry*. Oxford: Blackwell Publishers Ltd.

Medvedev, Zhores, and Roy Medvedev. 2006. The Unknown Stalin. New York, NY: I.B. Tauris & Co. Ltd.

Merkle, Ralph. 2018. Information-Theoretic Death. Unpublished manuscript. http://www.merkle.com/definitions/infodeath.html.

Miles, Robert. 2017. What Can AGI Do? I/O and Speed. YouTube. https://www.youtube.com/watch?v=gP4ZNUHdwp8.

Mulgan, Tim. 2009. Rule Consequentialism and Non-Identity. In Melinda Roberts and David Wasserman (eds.), *Harming Future Persons: Ethics, Genetics, and the Nonidentity Problem*. New York, NY: Springer.

Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Clarendon Press.

Parfit, Derek. 2017. *On What Matters: Volume Three*. Oxford: Oxford University Press.

Paul, L.A. 2014. *Transformative Experience*. Oxford: Oxford University Press.

Pearce, David. 1995. *The Hedonistic Imperative*. https://www.hedweb.com/hedethic/tabconhi.htm.

Pearce, David. 2007. Selection Pressure an Radical Anti-Natalism. https://www.abolitionist.com/anti-natalism.html.

Persson, Ingmar, and Julian Savulescu. 2012. *Unfit for the Future: The Need for Moral Enhancement*. Oxford: Oxford University Press.

Pigliucci, Massimo. 2014. Uploading: A Philosophical Counter-Analysis. In Russell Black ford and Damien Broderick (eds.) *Intelligence Unbound: The Future of Uploaded and Machine Minds*. New York, NY: Wiley-Blackwell.

Pinker, Steven. 2011. *The Better Angels of Our Nature: Why Violence Has Declined*. New York, NY: Penguin Books.

Russell, Bertrand. 1954. *Human Society in Ethics and Politics*. New York, NY: Routledge.

Sandberg, Anders. 2014. Being Nice to Software Animals and Babies. In Russell Blackford and Damien Broderick (eds.) *Intelligence Unbound: The Future of Uploaded and Machine Minds*. New York, NY: Wiley-Blackwell.

Sandberg, Anders, and Nick Bostrom. 2008. Whole Brain Emulation: A Roadmap. Future of Humanity Institute Technical Report #2008-3. https://www.fhi.ox.ac.uk/brain-emulation-roadmap-report.pdf.

Sandberg, Anders, Eric Drexler, and Toby Ord. 2018. Dissolving the Fermi Paradox. arXiv. https://arxiv.org/pdf/1806.02404.pdf.

Scharf, Caleb. 2016. Where Do Minds Belong? *Aeon*. https://aeon.co/essays/intelligent-machines-might-want-to-become-biological-again.

Scheffler, Samuel. 2007. Immigration and the Significance of Culture. *Philosophy & Public Affairs*. 35 (2): 93-125.

Scheffler, Samuel. 2016. *Death and the Afterlife*. Oxford: Oxford University Press.

Scheffler, Samuel. 2018. *Why Worry About Future Generations?* Oxford: Oxford University Press.

Schneider, Susan. 2008. Future Minds: Transhumanism, Cognitive Enhancement, and the Nature of Persons. *Neuroethics Publications*. https://repository.upenn.edu/cgi/viewcontent.cgi?article=1037&context=neuroethics_pubs.

Schneider, Susan. 2016. It May Not Feel Like Anything To Be An Alien. *Nautilus*. http://cosmos.nautil.us/feature/72/it-may-not-feel-like-anything-to-be-an-alien.

Schneider, Susan, and Joe Corabi. 2014. The Metaphysics of Uploading. In Russell Blackford (ed.) *Uploaded Minds*. New York, NY: Wiley-Blackwell.

Temkin, Larry. 2008. Is Living Longer Living Better? *Journal of Applied Philosophy*. 25(3): 193-210.

Tomasik, Brian. 2016. Should We Intervene in Nature? Reducing Suffering. http://reducing-suffering.org/should-we-intervene-in-nature/.

Tomasik, Brian. 2017. The Importance of Wild-Animal Suffering. Foundational Research Institute. https://foundational-research.org/the-importance-of-wild-animal-suffering/.

Tonn, Bruce. 2009. Obligations to Future Generations and Acceptable Risks of Human Extinction. *Futures*. 41(7): 427-435.

Turchin, Alexey, and Maxim Chernyakov. 2018. Classification of Approaches to Technological Resurrection. Unpublished manuscript. https://philarchive.org/archive/TURCOA-3.

Verdoux, Philippe. 2011. Emerging Technologies and the Future of Philosophy. *Metaphilosophy*. 42(5): 682-707.

Wade, Nicholas. 2004. Comment on "The Impact of the Revolution in Biomedical Research on Life Expectancy by 2050. In Henry Arron and William Schwartz (eds.) *Coping With Methuselah: The Impact of Molecular Biology on Medicine and Society*. Washington D.C.: Brookings Institution Press.

Walker, Mark. 2002. Prolegomena to Any Future Philosophy. *Journal of Evolution and Technology*. 10. https://www.jetpress.org/volume10/prolegomena.html.

Walker, Mark. 2008. Cognitive Enhancement and the Identity Objection. *Journal of Evolution and Technology*. 18(1): 108-115.

Ward, Peter, and Donald Brownlee. *Rare Earth: Why Complex Life is Uncommon in the Universe*. New York, NY: Copernicus Books.

Wiblin, Robert. 2017. Why the Long-Term Future of Humanity Matters More Than Anything Else, and What We Should Do About It. 80,000 Hours. https://80000hours.org/podcast/episodes/why-the-long-run-future-matters-more-than-anything-else-and-what-we-should-do-about-it/.

Williams, Bernard. 1973. *Problems of the Self: Philosophical Papers, 1956-1972*. Cambridge: Cambridge University Press.