Should Humanity Go Extinct? Analyzing the Arguments for Pro-Extinctionism

Draft. Please do not cite without permission. Thank you.

Abstract: This paper examines a position on the ethics of human extinction that I call "pro-extinctionism." It argues that there are many different ways that one could interpret this thesis. For example, it could specifically target "final" human extinction, or instead aim for "terminal" human extinction *without* final extinction having occurred. I then critically examine a number of arguments for pro-extinctionist views that focus on final extinction, including philosophical and empirical pessimism, antinatalism, radical environmentalism, negative utilitarianism, and misanthropy. After this, I turn to the various arguments for terminal but not final extinction, including those based on notions of cosmic evolution, posthuman supremacy, and longtermism. My hope is that this paper provides some much-needed clarity to a topic of increasing urgency: should our species disappear in the near future?

The case for human existence, which he's kind of trying to defend, is shockingly weak. … We do so much bad that the fact that we create scientific theories and create beautiful art just doesn't seem to even come close to the colossal damage that we do to other lifeforms and to nature. — David Peña-Guzmán, discussing Todd May's book *Should We Go Extinct?*

1. Introduction

Some philosophers argue that "human extinction" would constitute a profound tragedy, perhaps of quite literally cosmic proportions (Bostrom 2003; Ord 2020). Others claim that there would be nothing bad or wrong about our extinction if there were nothing bad or wrong about the way in which it happens, e.g., due to everyone on Earth voluntarily refusing to procreate (see Redacted). Still others contend that our collective disappearance would be desirable and/or that we should actively try to bring about this state of affairs. I refer to these three classes of positions on the ethics of human extinction as *further-loss views*, *equivalence views*, and *pro-extinctionist views* (Redacted).

In this paper, I want to focus on the third view: pro-extinctionism. What exactly do pro-extinctionists claim? What are the different types of pro-extinctionism? And what arguments have philosophers propounded in support of this position? We will proceed as follows: section 2 explores several ways that pro-extinctionists may differ in their accounts of what the position is. Section 3 turns to six arguments in favor of a certain kind of pro-extinctionism, based on philosophical pessimism, empirical and futurological pessimism, antinatalism, radical environmentalism, negative utilitarianism, and misanthropy. Section 4 then discusses three arguments for a different kind of pro-extinctionism associated with techno-futuristic ideologies like transhumanism and longtermism. Finally, section 5 briefly concludes the paper.

The aim of this work is not necessarily to propound a novel argument for this or that position but to offer an original analysis of the differences and relations between pro-extinctionist views. In my opinion, providing a comprehensive overview of this topic can be value-added to the literature, especially if that literature lacks any systematic framework for thinking about this issue. There is, I would contend, much to gain from under-

standing how the variants of pro-extinctionism overlap and diverge, especially given that a growing number of people are endorsing pro-extinctionist positions, from so-called Efilists to radical environmentalists to longtermists and transhumanists. Indeed, there is one notable thesis of this paper that is unique, original, and surprising: "TESCREAL" ideologies like longtermism and transhumanism should be classified as pro-extinctionist views, if one accepts what I call the Narrow Definition of 'humanity' or 'human.' This definition identifies these terms with our species, *Homo sapiens*, and it contrasts with Broad Definitions that identify them with our species *and* whatever successors we might have, so long as they possess a certain moral status (Redacted). Since some longtermists and transhumanists explicitly argue that we should replace *Homo sapiens* with a "superior" posthuman species, they should be identified as pro-extinctionists. Others merely hold that the disappearance of our species once posthumanity arrives should be a matter of indifference. I call these people "extinction neutralists." Section 4 goes into detail about these claims. The point is that, if I am correct, we should recognize that pro-extinctionism is actually widely held, especially within certain powerful corners of the tech world and academia.

Let's now turn to the deceptively complex question of how pro-extinctionism could be interpreted.

2. Interpretations of Pro-Extinctionism

We begin with several important distinctions. The first concerns the different types of "extinction" that our species could undergo. These are all based on what I call a Minimal Definition of 'human extinction,' which states:

> *Minimal definition*: Human extinction will have occurred if there were tokens of the type "humanity" at some time T1, but no tokens of this type at some later time T2 (Redacted).

One could interpret this in different ways, depending on whether one accepts the Narrow or Broad Definitions of 'humanity.' For most of this paper, I will assume the Narrow Definition, as this will greatly simplify our discussion.

I count at least six types of "extinction" in Redacted, although only two are directly relevant to our present discussion. These are what I call *terminal extinction* and *final extinction*. The first would occur if and only if our species were to disappear entirely and forever, whereas the second would occur if and only if our species were to disappear entirely and forever without leaving behind any successors. As we will see, this distinction is crucial for making sense of pro-extinctionist positions: some specifically aim for final extinction, while others aim for terminal extinction *without* final extinction. In both cases, the future envisioned is one in which our species no longer exists, though the particular circumstances of us no longer existing are not the same. Section 3 explores pro-extinctionism of the first sort, while section 4 is dedicated to the second.

Another distinction concerns two aspects or stages of human extinction: (1) the process or event of Going Extinct, and (2) the subsequent state or condition of Being Extinct. This applies to both types of extinction noted above—i.e., there is Going Extinct and Being Extinct in both the terminal and final senses. To explain why this matters, consider

the first two views on the ethics of extinction mentioned in section 1: "equivalence" and "further-loss" views. Equivalence views state that our extinction, however interpreted, would be bad or wrong *only insofar* as Going Extinct would be bad or wrong. If there is nothing bad or wrong about Going Extinct, then there is nothing bad or wrong about our extinction—full stop. Equivalence theorists would thus say that if everyone around the world were to decide not to have children, resulting in the eventual disappearance of *Homo sapiens*, this would not be bad or wrong. Further-loss views disagree, claiming that both Going Extinct *and* Being Extinct can be sources of badness—independently. For example, they would claim that even if there is nothing bad or wrong about Going Extinct, our extinction could still be very bad or wrong because of the "further losses" or opportunity costs associated with Being Extinct, such as "astronomical" amounts of value, wellbeing, and other goods that might have otherwise existed (Bostrom 2003). Without separating Going Extinct and Being Extinct, the crucial differences between these two views would be unintelligible.

We will see that the distinction between Going Extinct and Being Extinct is also fundamental to pro-extinctionism, which can take both deontic and evaluative forms.[1] Evaluative pro-extinctionism states that Being Extinct would in some way be *better than* Being Extant, or continuing to exist. Many evaluative pro-extinctionists are also deontic pro-extinctionist who, as such, argue that we *ought to* try to bring about our Being Extinct. This raises an important question of practical ethics: How exactly should we bring about Being Extinct, according to deontic pro-extinctionists?

The answer depends in part on which type of "extinction" one aims to bring about. For pro-extinctionists motivated by, for example, philosophical pessimist or radical environmentalist convictions (section 3), the aim is almost always to bring about our Being Extinct in the particular sense of *final extinction*. That is to say, they want our species to die out without leaving behind any successors—a complete and final end to the story of our evolutionary lineage.

There are three primary methods of achieving this, i.e., of Going Extinct anthropogenically: (i) antinatalism, whereby a sufficient number of people around the world voluntarily choose not to procreate; (ii) pro-mortalism, whereby a sufficient number of people around the world voluntarily kill themselves; and (iii) omnicide, whereby some people, perhaps a small group or single individual, kill themselves along with everyone else. Most pro-extinctionists who aim for final extinction strongly reject omnicide, while a smaller portion reject pro-mortalism.[2] Nearly all advocate for antinatalism as the best, if not the only morally permissible, option. As we will see below, this points to two interpretations of antinatalism: on the one hand, it could be understood as an *ethical position*, according to which birth has a negative value and/or creating new people is morally wrong. Antinatalism in this sense seemingly entails pro-extinctionism, since if one believes people ought not to procreate, and if people act accordingly, the human population will eventually dwindle to zero. On the other hand, antinatalism could be understood as a *method* of achieving the goal of Being Extinct. One might thus claim that there is nothing morally wrong about having children, but that, e.g., our world contains types of "severe suffering" that no amount of happiness could possibly counterbalance (Crisp 2023), and hence that it would be best if we no longer existed. One might then adopt voluntary antinatalism as the only morally unproblematic *means* of bringing about our extinction—in contrast to omnicide and pro-mortalism.[3]

Another way that pro-extinctionists who aim for final human extinction might differ concerns the *temporality* of Going Extinct. Hermann Vetter, for example, argued that if Going Extinct via anthropogenic causes were drawn-out, inflicting large amounts of physical and/or psychological suffering on those living at the time, it would be very bad or wrong. But, he says, "if mankind were completely extinguished in a millionth of a second without any suffering imposed on anybody, I should not consider this as an evil, but rather as the attainment of Nirvana" (Vetter 1968). Although Vetter did not *explicitly* endorse instantaneous *omnicide*, my guess is that he would approve of this if a means for accomplishing it were to become available. However, others have been less equivocal. The negative utilitarian David Pearce writes that, "if the multiverse had an 'OFF' button, then I'd press it," though he adamantly opposes any scenario of Going Extinct that would cause physical and/or psychological suffering (Pearce 1995). It should be noted that there may, in fact, be a way to destroy our universe— specifically, our entire future light cone—almost instantaneously. If the universe is in a "metastable" or "false vacuum" state, which some cosmologists believe it may be, then a high-powered particle accelerator could theoretically tip the universe into a "true vacuum" state by nucleating a "vacuum bubble" that expands in all directions at close to light speed, obliterating everything it touches. This looks like the equivalent of an "off" button for our universe that some pro-extinctionists, who are otherwise strongly opposed to omnicide, would be willing to push.

For those pro-extinctionists who endorse terminal extinction *without* final extinction, there are two primary methods of catalyzing this outcome: (i) we could *reengineer* ourselves using advanced technologies to create one or more new "posthuman" species, which may result in the complete and permanent disappearance of our species but *not* our evolutionary lineage. Or (ii) we could completely *replace* ourselves with a new population of "intelligent machines" or AGIs (artificial general intelligences). In the former case, the spatiotemporal continuity of our lineage would remain intact, whereas in the latter case, it would be broken. However, in both, the disappearance of our species would coincide with the emergence of a new successor species of posthumans and/or AGIs, which might be quite alien in nature. As alluded to above, pro-extinctionism of this sort is far more widely held and influential today than the pro-extinctionist views discussed in section 3, which specifically advocate for final extinction.

We have thus far distinguished between evaluative and deontic pro-extinctionism, and between pro-extinctionist views that aim for final extinction and those that aim for terminal extinction without final extinction. We then examined some of the different methods that pro-extinctionists of each category could employ to achieve their respective aims. Yet another way that pro-extinctionists may differ concerns not the *temporality* but the *timing* of our extinction. That is, when is the optimal moment for our species to disappear?

Some, such as David Benatar, who endorses final extinction, argue that we should die out as soon as possible, since this would mean fewer total births (as birth is always a net harm, he claims) and less overall human suffering (Benatar 2006). Other pro-extinctionists have disagreed, such as the 19th-century German philosopher Eduard von Hartmann, who synthesized the work of Hegel and Schopenhauer into a grand eschatological vision of humanity's collective future. Unlike some followers of Schopenhauer, Hartmann vehemently argued *against* antinatalism and pro-mortalism, contending instead that we should expand the human population while further developing our civilization. On this view, the ultimate goal isn't merely to precipitate the final extinction of our species, but to

eliminate the very possibility of life emerging anywhere in the cosmos. We must, therefore, delay our extinction and accelerate progress until an appropriate method of *universal omnicide* comes into view, which might be centuries (or more) in the future. Since Hartmann was an idealist, he believed that if we eliminate all subjectivity in the universe, the universe itself will disappear forever. Though Hartmann's views are intriguing, most pro-extinctionists have held Benatar's view that we should die out sooner rather than later. This goes for pro-extinctionists who aim for terminal but not final extinction, too: the sooner we transition to a posthuman or AGI-dominated civilization, the better.

A final way that pro-extinctionist views can differ concerns their *evaluation* of Being Extinct. The definition given above states that Being Extinct is better than Being Extant. But "better than" does not imply "good"—indeed, it is compatible with Being Extinct being very bad. Simon Knutsson, for example, argues in favor of final extinction when he writes that "an empty world is the best possible world," but he adds that "I would not say that an empty world would be good" (Knutsson 2023). In contrast, Benatar (2006) seems to hold that Being Extinct would be positively good, falling somewhere above the "neutral" line separating good from bad. This follows from his harm-benefit asymmetry, according to which (a) existence is a good/bad situation, because it involves both pleasures and pains, where the presence of pleasure is good while the presence of pain is bad; and (b) nonexistence is a good/not-bad situation, because it involves neither pleasures nor pains, where the absence of pleasure is not bad (as there is no one to experience this absence), while the absence of pain is good (even if there is no one to experience this absence). Since Being Extant, as a species, is a good/bad situation whereas Being Extinct is a good/not-bad one, it appears that Being Extinct is not merely better but positively good. This is yet another issue that pro-extinctionists may disagree about.

Pro-extinctionism can thus take many forms. It is not a single monolithic view, as some might assume, but rather a family of views that differ in sometimes substantive ways. Let's now examine some of the main arguments for pro-extinctionist views that specifically aim for final human extinction. The penultimate section will turn to views that oppose final extinction while endorsing (or at least being indifferent to) terminal extinction.

3. Arguments for Final Human Extinction

1.3 The Argument from Philosophical Pessimism. Philosophical pessimism is the view that "life is not worth living, that nothingness is better than being, or that it is worse to be than not to be" (Beiser 2016). If any of these claims are true, it seems to follow straightforwardly that Being Extinct would be best. Put differently, if nonexistence is best for each one of us, it is best for all of us, and Being Extinct *just is* the nonexistence of everyone. However, not all philosophical pessimists explicitly advocate for pro-extinctionism. Consider Arthur Schopenhauer's claim that "voluntary and complete chastity is the first step in asceticism or the denial of the will to live" (Schopenhauer 2016). In a subsequent essay, he declared that, "if the act of procreation were neither the outcome of a desire nor accompanied by feelings of pleasure, but a matter to be decided on the basis of purely rational considerations, is it likely that the human race would still exist?" And:

If you imagine, insofar as it is approximately possible, the sum total of dis-

tress, pain, and suffering of every kind which the sun shines upon in its course, you will have to admit it would have been much better if the sun had been able to call up the phenomenon of life as little on the earth as on the moon; and if, here as there, the surface were still in a crystalline condition (Schopenhauer 1970).

Here, Schopenhauer argues for the backward-looking claim that we should have *never been*, but he does not defend the forward-looking claim that we should *no longer be*. Nor does he explicitly endorse antinatalism, despite his remarks above, though he does oppose pro-mortalism, as he believes that committing suicide is tantamount to acquiescing to "the will" rather than overcoming it (Beiser 2016, p. 59).

However, other pessimists within the Schopenhauerian tradition did take this extra step. We have already mentioned Hartmann, who endorsed universal omnicide—indeed, the annihilation of the entire universe—at some point in the future, which he also seems to have imagined as being a voluntary event, at least for those who initiate it.[4] The reason that this decision will be voluntary is that, as our "consciousness" continues to develop, we will become ever-more acutely aware of the profound suffering that is life, and hence will become increasingly convinced of the fact that we *must* eliminate the very possibility of life arising anywhere in the universe. Presumably, everyone will thus agree that the best course of action is to end the entire "world process" (Beiser 2016, pp. 155-156). Another German pessimist of the late 19th century was Philipp Mainländer, who, like Hartmann, held a teleological view of the world as inevitably marching toward complete nonexistence, though his account differed from Hartmann's. However, Mainländer endorsed both antinatalism (virginity) and pro-mortalism as the correct methods for bringing this about, and in fact he committed suicide, at the age of 34, a day or two after volume I of his *magnum opus* was published, by stacking copies of his book on the floor and step ping off them with a rope tied around his neck (Beiser 2016, p. 201).

Philosophical pessimism has also motivated the pro-extinctionist views of philosophers like Peter Wessel Zapffe and Benatar. Zapffe argued that humanity is an evolutionary anomaly: we have, through some sort of evolutionary mistake, evolved a kind of excessive consciousness. Whereas all animals "know angst, under the roll of thunder and the claw of the lion," humans feel "angst for life itself—indeed, for his own being." In other words, we are capable of being gripped by a sense of "cosmic panic," which we must constantly mitigate through various defense mechanisms. Zapffe thus compares us to the Irish elk, which supposedly evolved antlers that were too heavy for males to lift their heads, thus resulting in the species' extinction. He argues that we should meet the same fate, for the sake of ourselves, by refusing to have children. Hence, Zapffe endorsed antinatalism as the right method of Going Extinct.

Benatar's version of pessimism is closer to Schopenhauer's: life is inherently very bad, as there is an asymmetry between pains and pleasures: the worst pains are greater than the best pleasures, as illustrated by the fact that most of us would not trade 24 hours, or even several weeks, of the most blissful ecstasy for 24 hours of the most agonizing torture. Furthermore, Benatar notes that many pleasures are merely "relief pleasures," the feelings we get when pain or discomfort is relieved, and that "significant periods or each day are marked by some or other" state of unpleasantness: itches, boredom, stress, anxiety, fear, guilt, irritation, sadness, frustration, grief, loneliness, and so on. Many of these un-

pleasant states, such as hunger and thirst, naturally regenerate, such that "we must continually work at keeping suffering (including tedium) at bay, and we can do so only imperfectly," while the pleasures of relieving these states are merely ephemeral (Benatar 2006). Worse, Benatar points out that many people suffer from chronic pain, yet "there's no such thing as chronic pleasure" (Rothman 2017).

Such considerations lead him to argue that life is much worse than most of us realize, and that while some lives are worth continuing, no life is worth starting. From this, he concludes that the near-term final extinction of humanity would be our best course of action. However, in making this case, Benatar distinguishes between "dying-extinction" and "killing-extinction," where the former is, very roughly speaking, voluntary while the latter is not. Only dying-extinction, via antinatalist means, is morally acceptable, though he also contends that "suicide may more often be rational and may even be more rational than continuing to exist." In contrast, involuntary omnicide, a form of killing-extinction, would be "troubling for all the reasons that killing [individual people] is troubling" (Benatar 2006). Benatar thus prefers the antinatalist method of bringing about our Being Extinct, is open to pro-mortalism, and is strongly opposed to omnicide.

One last argument is worth considering. It hinges on the idea that there exists not only an *asymmetry* between pleasures and pains but a fundamental *discontinuity* such that no amount of pleasure could possibly compensate for some amount of certain kinds of pains. Take the case of torture by electric shock. As Roger Crisp writes about the agony of just one hour of this torture, "it seems to me as reasonable to claim that such a period cannot be counter-balanced by any amount of positive hedonic experience as to claim it can" (Crisp 2023). He calls this a "T-discontinuity," and, if someone were to accept that T-discontinuities exist, they may contend that any world that contains certain kinds of pain would be better off not existing, even if that world were to contain very large, or perhaps infinite, amounts of pleasure. Though Crisp himself does not claim that Being Extinct *would* be better, he does say that it "*might* be," and that if one were to detect a massive asteroid heading toward Earth, one should seriously consider not deflecting it— a form of passive omnicide that Crisp imagines being instantaneous and, therefore, painless, al though, in reality, Going Extinct from an asteroid impact would likely cause massive amounts of severe suffering (Crisp 2023, 2021). Knutsson also focuses on instances of extreme suffering, such as "mutilations, torture murders, and sexual violence against children," which leads him to conclude that "human extinction would probably be less bad than the realistic alternatives, and the same goes for the extinction of all other species" (Knutsson 2023, 2022).

In sum, philosophical pessimists have proposed several arguments in favor of pro-extinctionism. If life is drenched in suffering, and this suffering is inescapable, then nonexistence would be better. But not just *our* nonexistence, the nonexistence of *all* sentient beings—a point that Benatar makes explicit in writing that his "argument applies not only to humans but also to all other sentient beings" (Benatar 2006, pp. 2). This is why most pro-extinctionists who are motivated by philosophical pessimism advocate for final human extinction: if *Homo sapiens* were to disappear, but we were to leave behind successors who are also capable of suffering, this would not solve the fundamental problem. The only true solution is to die-out without leaving behind any successors, if not extinguish all life on the planet or within the entire universe.

3.2 The Argument from Empirical Pessimism. Pessimism of this sort is the claim that our world is, as a matter of contingent fact, very bad. It often overlaps with philosophical

pessimism, but constitutes a distinct position, as one can accept empirical pessimism while rejecting claims like: life is suffering, nonexistence is inherently preferable to existence, our lives are full of much more suffering than we realize, there exists an asymmetry between pleasures and pains, some pains cannot be counterbalanced by any amount of pleasures, and so on. Rather, one might argue that our world *could* be very good, but for some reason it *isn't*.

To illustrate, consider that roughly 580,000 people die violently every year, while another 463,000 people are raped or sexually assaulted in the US alone (UNODC 2023; RAINN 2024). Some 600,000 US children are abused annually, and about 840,000 children go missing each year, resulting in an average of one child disappearing every 40 seconds (Seetharaman 2024; NCA 2024; CCPSC 2023). Globally, about 1.2 billion people live in acute multidimensional poverty, with some 712 million in extreme poverty, a number that has risen by 23 million since 2019 (HDR 2022; WB 2024). About the same number of people, 735 million, are malnourished, and a whopping 800 million children—one-third of all the children on Earth—suffer from lead poisoning, which causes permanent brain damage (CW 2023; NIEHS 2024). Two billion people don't have access to safe water, while another 150 million worldwide are homeless (UNESCO 2024; Abbas 2024). Some 1.4 billion children live on $6.85 or less per day, an estimated 50 million people are trapped in modern-day slavery, over 51 million Americans suffer from chronic pain, about 50 million Americans struggle with chronic sleep disorders, and some 40 million people in the US take antidepressants for depression (GCECP 2024; Fleck 2023; Dillinger 2023; HD 2023; Ahrnsbrak 2021). An even higher number of Americans—46.8 million—battle drug and alcohol abuse each year, with over 178,000 dying of alcohol-related diseases every 12 months (DHHS 2024). Over 258 million Americans report that "they have experienced health impacts due to stress in the prior month," while more than 91 million say that they feel so stressed-out most days that they are unable to function normally (APA 2022). Globally, 280 million people deal with depression, and 301 million suffer from anxiety disorders (Koskie 2023). This is just a brief snapshot; for a comprehensive survey of suffering in the world, see Redacted.

Even if one does not accept some of the claims made by philosophical pessimists, these statistics make the world look like a waking nightmare. There is *so much* suffering, pain, misery, anguish, agony, terror, and hardship in this world—much of which is avoidable, as in the case of global poverty. Those sensitive to this oceanic suffering may thus conclude that it might be better if humanity were to no longer exist: though we would lose all of the good things that Being Extant may provide, Being Extinct would expunge the negative experiences that happen to be widespread. Indeed, the numbers above are (more or less) unprecedented in human history, due in part to population growth over the past century: in absolute terms, our planet has almost certainly never contained as much suffering as it does right now, or within very recent history. Contra "New Optimists" like Steven Pinker (2011), the world has, from this perspective, become significantly *worse* over time.

This points to yet another version of pessimism, which I will only briefly address: futurological pessimism. One might concede that the world is very bad right now, but argue that it will improve in the future. Others will claim that, in contrast, we should expect world-conditions to deteriorate even further in the coming decades or centuries. Climate change, tipping points in Earth systems, biodiversity loss, the sixth mass extinction, and so on, could severely compromise our ability to survive and flourish in the future. One study, for example, estimates that roughly 2 billion people will become climate refugees by the

end of this century, while another calculates that some 1 billion people will likely die over the next ~80 years as a direct result of climate change (Beitrag 2017; Pearce and Parncutt 2023). Other studies find that up to 74% of the human population will be exposed to lethal heat waves by the century's end, while 30% of Earth's terrestrial surface will become arid land if temperatures reach 2 degrees Celsius (Mora et al. 2017; Gabbatiss 2018). Still others argue that we may be on the verge of a sudden, irreversible, catastrophic collapse of the global ecosystem (Barnosky et al. 2012). Add to this facts about the ubiquity of toxic anthropogenic chemicals, growing wealth disparities, democratic backsliding and the ongoing rise of fascism, social media, and AI systems that are polluting our information highways with fake news and deepfakes (see Redacted). If the present looks like a waking nightmare, the future looks like a dystopian catastrophe. At the extreme, one can imagine totalitarian regimes using AI and other advanced technologies to control its citizens, or even developing life-extension technologies that enable such regimes to torture dissidents for hundreds of years. These are, of course, speculative possibilities, but they do not seem entirely implausible.

Futurological pessimism may be seen as a forward-looking version of empirical pessimism. Both could be employed to argue for pro-extinctionism, according to which near-term extinction via some method like universal antinatalism would be better than risking the horrors of continuing to exist.

3.3 The Argument from Antinatalism. As with pro-extinctionism, antinatalism can take evaluative and deontic forms. The former claims that birth has a negative value, while the latter claims that birthing someone is morally wrong. Let's examine a few arguments for each, and then consider the links between antinatalism and pro-extinctionism.

There are many arguments for evaluative antinatalism. The most obvious are based on pessimism. Philosophical pessimists would point to the facts delineated above in arguing that nonexistence is always preferable to existence, our lives are marked by continual cycles of need and boredom—both of which are sources of suffering—and so on. Empirical and futurological pessimists would point to more contingent facts about our particular world in claiming that being born is bad for the person who is born. The "birthstrike" movement is an example of (conditional) antinatalism based on futurological pessimism—specifically, on pessimism about the future given the climate crisis and ecological degradation. This form of antinatalism is conditional because it claims that one shouldn't procreate *right now*, though if our political leaders had taken climate change seriously 30 years ago, many birthstrike advocates would likely want to have children. But one could defend a stronger version of this view that adopts a broader scope of our existential predicament: perhaps a "perfect storm" of climate chaos, advanced AI (propaganda, surveillance, and lethal autonomous weapons), democratic backsliding, etc., could create a dystopian future in which authoritarianism and totalitarianism become widespread. Some might argue that the *risk* of subjecting children to such a future is sufficiently great to warrant stronger claims about the badness of being born.

However, not all arguments for evaluative antinatalism are based on pessimism. Although Benatar points to both philosophical and empirical pessimism as reasons not to have children (and as reasons for pro-extinctionism), he also formulates a second argument based on the aforementioned harm-benefit asymmetry. As previously noted, this

states that being born is a good/bad situation, since existence is both good (the presence of pleasures) and bad (the presence of pains), while never being born is a good/not-bad situation, since nonexistence is both good (the absence of pains) and not bad (the absence of pleasures). Given this asymmetry, he concludes that birth is always a *net harm* to the one who is born, and since it is morally wrong to cause someone harm (other things being equal), Benatar contends that it is always wrong to have children.

While many evaluative antinatalists are also deontic antinatalists, one can accept deontic antinatalism without accepting evaluative antinatalism. For example, some philosophers have argued that procreation violates Kant's Categorical Imperative—specifically, the "Formula of Humanity," which states that one should never use other humans as mere means to an end, such as the happiness or fulfillment of the parents (Akerma 2010). This is a deontological argument that has no connection to what is good or bad—e.g., if correct, it would mean that having children is wrong *even if* every human life were overwhelmingly good. Another example comes from Asheel Singh, who writes that, since harming someone without their consent is wrong because it would violate their rights, and since one cannot ask the unborn for consent to be born, procreation is always wrong (Singh 2012). Gerald Harrison defends a version of antinatalism that draws from W. D. Ross's notion of a "*prima facie* duty." On this account, we have a *prima facie* duty not to create new unhappy people, but no such duty to promote total happiness by creating new happy people. It follows that, since life inevitably contains both pleasures and pains, there are no reasons to create new people but one reason *not* to create them—even if the unborn person were to have a happy life if they had been born (Harrison 2012). Vetter defends a similar idea in pointing out that, on the person-affecting utilitarianism of Jan Narveson, if one were to create a child who would have a happy life, no utilitarian duty would be violated; but if one were to create a child who has a bad life, one's duty will have been violated. And since we cannot know whether our children will have good or bad lives once created, we should not create them (Vetter 1969).

At first glance, the connection between antinatalism and pro-extinctionism appears to be straightforward: if being born is a harm, and it would be better if people weren't harmed, then it would be better if there were no more births. Or, if procreation is wrong, and everyone around the world were to heed this conclusion and stop procreating, the human population would eventually dwindle to zero. However, there is a complication: at the moment, it is obviously true that if everyone were to stop having children, humanity would eventually die out. But this might not be true in the future, if we develop safe and effective life-extension technologies that would enable particular *individuals* to potentially live as long as our evolutionary lineage itself could exist. Call this "functional immortality."

There are several points to make about this: first, if one accepts antinatalism for reasons arising from philosophical pessimism, then one should also accept pro-extinctionism, since the very same argument for antinatalism *also* supports pro-extinctionism. If we should not have children because life is suffering, then it seems to also follow that Being Extinct would be better than Being Extant. Second, many arguments for deontic antinatalism don't appear to have this implication. There is nothing about the harm-benefit asymmetry, the Kantian argument, etc., that implies that our species *itself* should disappear. If functional immortality were achieved, then our species could theoretically persist into the very far future—millions or billions of years from now —without there being any additional births. The development of radical life-extension technologies may thus, someday, split the antina-

talist community into two camps: those that advocate for pro-extinctionism, and those that want our species to persist, a view that has been termed "no extinction antinatalism" (Redacted). Hence, antinatalism provides only a partial or conditional argument for pro-extinctionism: some versions of antinatalism are intimately bound up with pro-extinctionist views, while others imply pro-extinctionism only in the absence of radical life-extension.

   3.4 The Argument from Radical Environmentalism. Another argument for pro-extinction ism that aims for final extinction comes from radical environmentalism—especially when motivated by biocentric, biospherical egalitarian, or ecocentric theories of value. A brief reconstruction of the central argument goes like this: if a single species of mite, call it *Varroa obliterator*, were responsible for the environmental degradation that *we* have in fact caused, humans would no doubt set out to eliminate this species. Since humans have no more intrinsic (or final) value than any other species, according to biospherical egalitarianism, we should do the same to *ourselves*—i.e., we should remove our species from the biosphere that we are catastrophically destroying.

   Numerous environmentalists and environmental groups have endorsed this conclusion, such as the Voluntary Human Extinction Movement (VHEMT). To quote its founder, the ecocentrist Les U. Knight, "if you'll give the idea a chance, I think you might agree that the extinction of *Homo sapiens* would mean survival for millions, if not billions of other Earth-dwelling species" (Knight 1991). As the name suggests, VHEMT advocates for our voluntary extinction, preferably via the method of antinatalism. A more extreme case is the Church of Euthanasia, which declares that "one thing seems certain: from the point of view of nonhumans, on balance, our extinction would be a great blessing" (Korda 1994). This group also advocates for our extinction to be voluntary, although it endorses both antinatalism and pro-mortalism. Members are thus required to "take a lifetime vow of nonprocreation," as the church's sole commandment is "Thou shalt not procreate," though it also specifies suicide as one of its four primary pillars, exemplified by the slogan: "Save the Planet, Kill Yourself" (Korda 2019).

   There are only a few radical environmentalist groups that favor omnicide, such as the Gaia Liberation Front (GLF), which specifies their mission as being

> the total liberation of the Earth, which can be accomplished only through the extinction of the Humans as a species. ... Every Human now carries the seeds of terracide. If any Humans survive, they may start the whole thing over again. Our policy is to take no chances (CoE 1994).

Exterminating our species, they write, through nuclear war would cause excessive collateral damage. Mass sterilization would require too much time. And suicide is impractical. But genetic engineering offers "the specific technology for doing the job right—and it's something that could be done by just one person with the necessary expertise and access to the necessary equipment" (GFL 1994; Redacted). The GLF thus argues for synthesizing a designer pathogen that specifically targets *Homo sapiens*, adding that

> to complicate the search for a cure or a vaccine, and as insurance against the possibility that some Humans might be immune to a particular virus, several different viruses could be released (with provision being made for the release of a second round after the generals and the politicians had come out of

their shelters) (GLF 1994).

Given that biotechnology and synthetic biology are making it more feasible for small groups or even lone wolves to synthesize such pathogens, we might expect the risk of omnicide from radical environmentalists to rise in the future. To quote the terrorism scholar Frances Flannery, "as the environmental situation becomes more dire, eco-terrorism will likely become a more serious threat in the future" (Flannery 2016).

3.5 The Argument from Negative Utilitarianism. This version of utilitarianism states that the moral rightness of an action depends only on whether it minimizes suffering. Shortly after Karl Popper introduced the idea, R. N. Smart noted that it implies that one should become a "benevolent world-exploder" who destroys all life, including all human life, to achieve this end. A world without suffering is the best kind of world, and the only way to guarantee the absence of suffering is to eliminate everything that is capable of suffering: sentient beings like us. As noted above, David Pearce is a negative utilitarian who admits that, "if the multiverse had an 'OFF' button, then I'd press it," although Pearce does not endorse this *in practice*, for reasons discussed in subsection 4.2 below (Pearce 1995). Other negative utilitarians, though, may disagree, perhaps seeing the nucleation of a vacuum bubble as an effective means of painlessly eliminating all life that exists or might exist within our future light cone.

Negative utilitarianism does not have a significant following among contemporary philosophers. As one critic notes, it "is treated as a non-starter in mainstream philosophical circles, and to the best of my knowledge has never been supported by any mainstream philosopher, living or dead" (Ord 2013). However, this is not entirely true. As Knutsson points out:

> Gustaf Arrhenius and Krister Bykvist say that they "reveal" themselves "as members of the negative utilitarian family." J. W. N. Watkins describes himself as "sort of negative utilitarian." Clark Wolf defends what he calls "negative critical level utilitarianism" for social and population choices. [And] Thomas Metzinger proposes the "principle of negative utilitarianism" (Knutsson 2022).

Of all the arguments thus far discussed, negative utilitarianism may provide the most *direct* case for pro-extinctionism. Utilitarians base the right on the good, and since "the good" in this case is merely the absence of suffering, negative utilitarianism instructs one to do as Smart noted: to benevolently "explode" the world.

3.6 The Argument from Misanthropy. A final argument comes from misanthropy, defined as a "hatred of humankind" (OED 2024). There is, one could argue, much to hate about humanity: we lie, cheat, steal, and abuse each other. Some go on mass killing sprees, while others start wars and perpetrate violent genocides. Discriminatory attitudes like racism, xenophobia, ableism, ageism, classism, sexism, misogyny, and speciesism remain pervasive.

Consider that during the 20th century, approximately 231 million people perished in conflicts, while, as noted earlier, about 463,000 people are raped or sexually assaulted in the US each year, and just under 25% of US women over the age of 18 "have been the victim of severe physical violence by an intimate partner in their lifetime" (Leitenberg 2006;

RAINN 2024; NDVH 2023). Roughly 50 million people live in modern-day slavery, and Amnesty International reports that 141 countries "still torture citizens and prisoners" (Fleck 2023; Wells 2017). There are about 1.2 million people—disproportionately Black—in US prisons, and over 80,000 of them are held in solitary confinement (Carson 2023; AFSC 2024). Meanwhile, 2.1 billion people are trapped in multidimensional poverty, with 712 million living in extreme poverty. About 600,000 children are abused in the US each year, and some 800 million—about 1/3 of the world's children—suffer from lead poisoning (see Redacted).

Furthermore, we have razed forests, polluted the oceans, obliterated habitats, and initiated the sixth major mass extinction event of the past 3.8 billion years. Our planet is now covered in the waste products of civilization, including microplastics and plastic bags that have been found at the bottom of the Mariana Trench (Mouthino 2023; Gibbens 2018). Our actions have pushed Earth past six of nine planetary boundaries, and we have flooded the atmosphere with heat-trap ping greenhouse gasses that could make large regions of our planet more or less uninhabitable in the coming decades (Richardson et al. 2023; Mora et al. 2017). Since 1970, the global population of wild vertebrates has declined by a staggering 69%, largely because of human activities (WWF 2022).

Even more, we keep many billions of domesticated animals in factory farms, killing about 23 million every single day (Roberson 2023). Seven billion male chicks are slaughtered each year through methods like maceration and asphyxiation, while pigs and cows are often shot in the head before being dismembered (Krautwald-Junghanns et al. 2018; THL 2023). Some have described this as one of the greatest moral crimes in history (Harari 2015).

If the world is a waking nightmare, it is, in large part, because we have *made* it one, and this fact provides at least some ground for misanthropic attitudes toward our species. One may thus argue, as Benatar (2006) does, that it would be better if the entire human misadventure were to come to a permanent end—although Benatar notes that, while "the end of humanity would greatly reduce the amount of harm, it would not end it all," as our disappearance would leave behind many sentient creatures that would continue to suffer. In a phrase: if the world did not contain any humans, it would not contain any human-caused evils, and given the magnitude, scope, and horror of these evils, that may very well be better.[5]

4. Arguments for Other Types of Extinction

The arguments above specifically concern final human extinction, whereby our species disappears entirely and forever without leaving behind any successors. At least, this is the way that they are often understood and defended. If one believes that life is suffering, nonexistence is always preferable to existence, the world is very bad and will almost certainly get worse, and/or any successors we might have would continue to fight wars, abuse each other, destroy the natural environment, and slaughter nonhuman animals for food, then one may conclude that the best outcome would be for us to disappear without leaving anything in our place. Final human extinction is, therefore, the target that we should aim for, with most arguing that this should be brought about through the method of antinatalism, and that this should occur sooner rather than later.

However, there is a different version of pro-extinctionism that advocates for the

terminal extinction of our species *without* final extinction. As noted in section 2, this could be achieved through two general methods: first, by *reengineering* our species to become one or more new "posthuman" species. If the outcome were that no more tokens of *Homo sapiens* remained at the end of this process, then terminal but not final extinction will have occurred, as these posthumans would be our successors. Second, by *replacing* ourselves with a population of artificial beings, such as intelligent machines or AGIs (which might also be labeled "posthumans"). This would be tantamount to creating a new evolutionary lineage, spatiotemporally discontinuous from ours, that would then take our place. Of note is that this second possibility foregrounds a complication not mentioned above: if we create a second lineage to replace ourselves, how exactly should *we* subsequently die out? What would Going Extinct in the sense of terminal extinction look like for *us*, once these successors are prepared to supersede our kind? There are four main possibilities, three of which have already been discussed in the context of final human extinction: we could stop having children (antinatalism); we could kill ourselves (pro-mortalism); or someone or some group could kill everyone, including themselves (omnicide). There is, how ever, an ominous fourth possibility, namely, that our successors kill us in an act of *parricide*. Some pro-extinctionists who endorse terminal but not final extinction appear to hold rather blithe attitudes about this fourth option. It seems that many have not given it much thought.

Let's examine a few notable arguments for this alternative account of pro-extinctionism. I will make a number of key distinctions and clarifications along the way.

4.1 The Argument from Cosmic Evolution. This states that the replacement of our species by a population of intelligent machines or AGIs is desirable (and probably inevitable), given a broadly linear, teleological conception of *cosmic evolution*. Single-celled lifeforms emerged on Earth roughly 3.8 billion years ago, and over time became increasingly complex. With the emergence of our genus *Homo*, we began to modify our phenotypes with technology: at first, we fashioned simple tools out of wood, bones, and stone, but over time we invented telescopes, pace makers, brain implants, airplanes, and global communication systems that extend, amplify, or replace features of our biological substrates. If one extrapolates these trends of technologization into the future, the apparent terminus is a world entirely dominated by artificial beings. According to many advocates of the Argument from Cosmic Evolution, this is not only what *will be* but what *should be* the case—that is, we should welcome and actively accelerate the inevitable transition to a digital world in which biological humans are relegated to the museum of what has been.

This is not a fringe view, but has been embraced by certain tech billionaires, computer scientists, and technologists in Silicon Valley. As Dan Hendrycks et al. write, many people "want to unleash AIs or have AIs displace humanity," especially those associated with an ideology called "AI accelerationism," which "is alarmingly common among many leading AI researchers and technology leaders, some of whom are intentionally racing to build AIs more intelligent than humans." For these accelerationists, the goal is to create a new "technological utopia" populated by AGIs, as "this techno-utopian viewpoint sees AI as the next step down a predestined path to ward unlocking humanity's cosmic endowment" (Hendrycks et al. 2023).

For example, Larry Page, the cofounder of Google, claims that "digital life is the natural and desirable next step in … cosmic evolution and that if we let digital minds be free rather than trying to stop or enslave them, the outcome is almost certain to be good," a position that Max Tegmark dubs "digital utopianism" (Tegmark 2017). Richard Sutton similar-

ly claims that the "succession to AI is inevitable," and although our AGI successors might "displace us from existence … we should not resist [this] succession" (Sutton 2023). And Hans Moravec has described himself as "an author who cheerfully concludes that the human race is in its last century, and goes on to suggest how to help the process along." Advanced AI minds, he contends, will soon "be able to manage their own design and construction, freeing them from the last vestiges of their biological scaffolding, the society of flesh and blood humans that gave them birth." He adds that this will mark the climactic end of this world dominated by *Homo sapiens* (Moravec 1988).

On these accounts, Being Extinct in the sense of our terminal extinction would be good and desirable *so long as* this did not coincide with final extinction. Hence, they advocate for a kind of *conditional* pro-extinctionism: if our species were to die out *before* creating a suitable successor species, it would constitute a great tragedy—perhaps one of literally cosmic proportions. But if our species were to die out *after* this, that would be better than us continuing to exist and use up precious resources that our artificial progeny could more efficiently exploit to generate "value."

It is not clear how accelerationists imagine our species actually dying out once a population of successors has been created. Maybe we will voluntarily choose not to procreate, causing the population to dwindle until no one remains, or perhaps our successors will speed up the process through some annihilatory act of universal parricide. While parricide would be *bad for* those people living at the time, accelerationists might counter that it is a relatively small "cost" for the enormous "benefit" of transitioning to the next glorious stage of cosmic evolution. As alluded to above, very few accelerationists seem to have thought seriously about this issue; it is often swept under the rug as a mere detail that will, somehow, work itself out once our successors have arrived.

4.2 The Argument from Posthuman Supremacy. Although I characterized the arguments from section 3 as specifically aiming for final human extinction, there is another way to interpret them. That is to say, one could accept the premises to some of these arguments yet come to a quite different conclusion. If our species is evil, cruel, and destructive, it may be best if *we* no longer existed. But what if we went extinct by replacing ourselves with something "*better*"? If a new posthuman species were to possess traits of superhuman benevolence, kindness, and compassion, purged of the "darker angels" of our current human nature, then perhaps we could solve the problems foregrounded by these arguments by undergoing terminal extinction without final extinction.

This is precisely what Pearce advocates. He argues that, first, attempting to fulfill the moral obligations of negative utilitarianism by destroying the world would likely backfire, thus causing even worse suffering. And second, if our species were to undergo final extinction, the pervasive suffering of nonhuman organisms would persist, which would be bad. The best response is, therefore, not to become "benevolent world-exploders" but rather to radically enhance ourselves to become superintelligent posthumans. These posthuman successors, by virtue of their superintelligence, could then reengineer the entire biosphere to eliminate all suffering, replacing such experiences with "gradients of bliss" (see Pearce 1995). If other life exists within our future lightcone, we could initiate a colonization explosion to eliminate suffering in those creatures, too. On this view, would the continued existence of *our* species be justified once these posthuman successors are created? No, because we are capable of suffering, and negative utilitarianism instructs us to eliminate suffering entirely.

Along similar lines, Derek Shiller concedes many claims made by the pessimists and misanthropes in writing that

> we suffer from unreasonable and unfulfillable desires. We want to be kinds of people that we cannot be. In pursuit of fleeting temptations, we are disposed to make decisions that go against our own interest. We are aggressive, callous, and cruel to each other. We harbor arbitrary biases against our fellow creatures based on irrelevant characteristics or group membership.

However, "these are not the inevitable vices of any intelligent being. They are part of our species." He thus contends that, given the possibility of creating artificial beings that do not suffer the way we do or exhibit the same vicious traits, "we should engineer our extinction so that our planet's resources can be devoted to making artificial creatures with better lives" (Shiller 2017). This is the Argument from Posthuman Supremacy: to achieve a better world, we should hand the existential baton to something "superior" to us, rather than stop the race entirely.

4.3 The Argument from Longtermism. The "ethic" of longtermism comes in both radical (or "strong") and moderate forms. For our purposes, we will focus on radical longtermism, understood as the view that "positively influencing the long-term future is *the* key moral priority of our time" (EAF 2024). On my reading of this ideology—referred to simply as "longtermism" henceforth—it consists of three main components, each of which gives rise to a distinct argument for why we should either become or create a new posthuman species. These components are *transhumanism*, *totalism*, and *space expansionism*. Let's briefly examine these three arguments for posthumanity and then explore their implications for pro-extinctionism.

(i) Transhumanism. This component overlaps with the Argument from Posthuman Supremacy: there could be posthuman "modes of being" that are far more "valuable" than our current human mode, which is largely the result of blind evolutionary forces. Transhumanists thus argue that we should develop "human enhancement" technologies that enable us to explore these posthuman modes by radically augmenting our cognitive systems, sensory modalities, rational capacities, moral dispositions, and so on (Bostrom 2005). In his "Letter from Utopia," Bostrom depicts a posthuman world being marked by immortality, superintelligence, and so much pleasure that our posthuman descendants (to paraphrase) "sprinkle it in their tea." By overcoming the biological limitations that render our current state of existence suboptimal, we could create a literal "utopia" of almost supernatural wonders. Bostrom and Pearce, both transhumanists, refer to this project as "paradise-engineering" (Bostrom 2008; Pearce 1995).

Given the possibility of engineering paradise by becoming or creating a new posthuman species, many longtermists have argued that this is an integral part of fulfilling "our long-term potential" in the universe. For example, Bostrom—referring to posthumanity—writes that "the permanent foreclosure of any possibility of this kind of transformative change of human biological nature may itself constitute an existential catastrophe," where an "existential catastrophe" is any event that prevents us from fulfilling our cosmic potential (Bostrom 2013). Toby Ord similarly declares that "forever preserving humanity as it is now may also squander our legacy, relinquishing the greater part of our potential," and "rising to our full potential for flourishing would likely involve us being transformed

into something beyond the humanity of today" (Ord 2020). Hence, the tantalizing promise of utopia through radical enhancements provides one reason in favor of posthumanity.

(ii) Totalism. As noted earlier, this states that the world becomes better the more total value that it contains. When paired with the deontic component of utilitarianism, according to which moral rightness depends on what produces maximum value, our sole moral obligation across space and time is to maximize the total amount of value within our future light cone. While longtermism does not necessarily presuppose utilitarianism, (a) even "moderate" versions are based on totalism (MacAskill 2022), and (b) utilitarianism has, in fact, had a significant influence on the development of this ideology, and many longtermists consider themselves to be consequentialists (Ord 2016; Dullaghan 2019).

Our goal should thus be to create the largest possible population of beings capable of realizing value within the universe as a whole, from now until the heat death. There are two ways of doing this: first, intelligent lifeforms must survive for as long as possible, and second, there should be as many instances of intelligent life as the universe can contain. Since Earth has limited real estate and will become uninhabitable in roughly 1 billion years, we must colonize the universe and establish a sprawling, multi-galactic civilization full of the maximum number of people with net-positive lives. Furthermore, since there could exist more digital than biological people per volumetric unit of space, longtermism also implies that the people who occupy this sprawling civilization should be digital in nature, nearly all of whom would reside in vast computer simulations running on "planet-sized" computers powered by Dyson swarms surrounding the nearest stars (Bostrom 2003). According to one estimate, there could be some 10^{45} digital people in the Milky Way Galaxy alone, per century, while Bostrom calculates at least 10^{58} digital people within the universe as a whole (Newberry 2021; Bostrom 2014). This is a second argument in favor of posthumanity.

(iii) Space expansionism. The final argument concerns the fact that colonizing the universe will almost certainly require us to have *already* become or created a new species of digital posthumans. The reason is that outer space is extremely hostile to biological beings, due to the effects of space radiation, microgravity, confinement in small spaces as spacecraft travel for hundreds, thousands, or even millions of years to other stars and galaxies, etc. Perhaps our species will succeed in colonizing Mars, but venturing to the nearest star will very likely be impossible unless we become or replace ourselves with digital beings, in the form of uploaded minds or AGIs. In other words, not only will future populations need to be digital to maximize value, but spreading beyond our solar system will require that we swap our biological substrate with a more robust artificial one.

What are the implications of these three arguments for posthumanity? I am not aware of any longtermists explicitly arguing that our species *should* disappear in the future, or that it would be *better* if this were to happen. However, I have seen no indication that longtermists think it would be *bad* if our species were to die out once posthumanity arrives. After all, what usefulness or value would *our* continued existence have? As Shiller (2017)— who appears to be sympathetic with longtermism—points out above, *Homo sapiens* would use-up resources that these posthumans could put to better use. We would be a vestige of a bygone era ruled by inferior, less efficient, less "intelligent," less "happy" beings, and hence posthumans may opt to discard our species, which might be for the best, from the evaluative perspective of totalism. One thus gets the impression that many longtermists would agree that the eventual disappearance of our species would, if anything, be *good*. For exam-

ple, consider a passage from Ray Kurzweil's book *The Singularity Is Near*, which outlines a vision of the future that is very similar to longtermism. While no one should be forced to become posthuman, he argues, if one chooses not to become posthuman, then "you won't be around for very long to influence the debate." He does not see this as problematic, so long as the extinction of our species is terminal rather than final— that is, it coincides with the emergence of a new posthuman species that spreads the "light of consciousness" to other galaxies, causing the universe to literally "wake up" (Kurzweil 2005).

It may be helpful to introduce another concept here, which I will call "situation-dependent extinction neutralism," or "extinction neutralism" for short. This states:

> *Extinction neutralism*: It would be neither good nor bad if our species were to disappear *after* we either became or created a suitable species of successors.

In other words, if we were to go extinct before this happens, they would see our extinction as very bad (or wrong), but one should be indifferent about whether our species survives once our successors arrive—which is just to say that our final extinction would be very bad (or wrong), but not our terminal extinction.

We can thus rephrase the points made above as follows: longtermism is ambiguous between pro-extinctionism and extinction neutralism. So far as I can tell, all longtermists fall some where on the spectrum between these two positions, though my guess is that many are closer to pro-extinctionism than extinction-neutrality, for the reasons specified earlier (e.g., our continued resource use, inferiority, relative low value, etc.). However, longtermism appears to entail pro-extinctionist *consequences*, *in practice*, for reasons gestured at by Kurzweil and others. Maybe our posthuman successors will decide to keep so-called "legacy humans" around in a zoo or as pets, though if we use up valuable resources, they may—rationally—opt to erase us from the drawing board of the future, perhaps by phasing us out through sterilization programs rather than through parricide (Goertzel 2010). There is, once again, almost no serious discussion among longtermists of what might happen to *Homo sapiens* in a world run by posthumans, perhaps because most longtermists do not imagine us surviving for long once these posthumans arrive. It is worth noting that many longtermists anticipate the history-rupturing transition to a new posthuman civilization to occur within the near future, probably in the coming decades (see, e.g., Bostrom 2002). Hence, this is not some distant possibility, on their account, but a scenario that may and should come to pass very soon: the sooner our successors appear, the sooner we may be able to fulfill our long-term potential in the universe.

There is one final point to make about longtermism and pro-extinctionism. Many of the loudest exhortations to *avoid* "human extinction" come from the longtermist community. This may appear to contradict my claim that nearly all longtermists fall somewhere between extinction neutralism and pro-extinctionism. But this is not the case, as longtermists typically define "humanity" in an idiosyncratic manner, such that it includes not just our species but whatever successors we might have, so long as they possess certain important properties. This is the Broad Definition introduced in section 1. For example, Nick Beckstead, who cofounded longtermism with Bostrom, writes that "by 'humanity' and 'our descendants' I don't just mean the species *homo sapiens* [sic]. I mean to include any valuable successors we might have," which he later describes as "sentient beings that matter" (Beck-

stead 2013; Redacted). Similarly, Hilary Greaves and William MacAskill argue in a defense of radical longtermism that "we will use 'human' to refer both to *Homo sapiens* and to whatever descendants with at least comparable moral status we have, even if those descendants are a different species, and even if they are non-biological" (Greaves and MacAskill 2021). Toby Ord provides yet another example. He writes that "if we somehow give rise to new kinds of moral agents in the future, the term 'humanity' in my definition should be taken to include them" (Ord 2020).

On these definitions, (a) our "posthuman" successors would also count as "human," which appears to be incoherent, but this is merely a terminological problem. And (b) it follows that our species could disappear entirely and forever, perhaps in the very near future, *without* "human extinction" having occurred. So long as we have successors with the right cognitive and moral properties, then "humanity" on the Broad Definition would persist despite *Homo sapiens* ceasing to be. One should, therefore, not be misled by such talk of "human extinction." What longtermists strongly oppose, on the Narrow Definition, is the final extinction of our species; they are either indifferent about or in favor of our terminal extinction once we have created successors to take our place.

5. Conclusion

The study of the ethical and evaluative implications of human extinction, or what might be called "Existential Ethics," has received growing interest in recent years from philosophers. My aim in this paper was to focus on one of the three major positions that one can take within Existential Ethics, namely, pro-extinctionism. I hope to have shown that there are many details that pro-extinctionists can differ on, including the ways in which Going Extinct might unfold and the reasons for why Being Extinct might be better than Being Extant—if not positively good (section 2). I then presented a number of arguments in favor of the final extinction of our species, including those based on philosophical, empirical, and futurological pessimism, antinatalism, radical environmentalism, negative utilitarianism, and antinatalism (section 3), as well as various arguments in favor of terminal but *not* final human extinction, including those based on cosmic evolution, posthuman supremacy, and longtermism (section 4).

Though pro-extinctionism is typically associated with pessimistic views of humanity and concerns about our environmental impact, there is an especially insidious version of pro-extinctionism that has become very influential in Silicon Valley and at elite institutions like the University of Oxford (where TESCREAL organizations such as the Global Priorities Institute and Forethought Foundation are located). What links these two mutually incompatible versions of pro-extinctionism is the shared view that our species has no place—and *should* have no place—in the future. In the one case, the aim is to bring our entire lineage to an end, to disappear without leaving behind any successors that might perpetuate our miserable and destructive tendencies, while in the other, the aim is to eliminate our species by becoming or creating a new successor species to usurp us. Both are equally pro-extinctionist, that is, if one takes "human extinction" to mean the complete and permanent disappearance of our species, *Homo sapiens*.

References

Abbas, Rabeeta. "20 Countries with the Highest Homeless Population." Yahoo! Finance, Yahoo!, 29 Jan. 2024, finance.yahoo.com/news/20-countries-highest-homeless-population-180254615.html.

AFSC. "Solitary Confinement Facts." American Friends Service Committee, 2024. afsc.org/solitary-confinement-facts.

Ahrnsbrak, Rebecca, and Marie N. Stagnitti. "Comparison of Antidepressant and Antipsychotic Utilization and Expenditures in the U.S. Civilian Noninstitutionalized Population, 2013 and 2018." Agency for Healthcare Research and Quality, Feb. 2021, meps.ahrq.gov/data_files/publications/st534/stat534.shtml.

Akerma, Karim. "Theodicy shading off into Anthropodicy in Milton, Twain and Kant." Tabula Rasa 41 (2010).

APA. "APA Dictionary of Psychology." American Psychological Association, 15 Nov. 2023, dictionary.apa.org/pretraumatic-stress-disorder.

Barnosky, Anthony D., et al. "Approaching a State Shift in Earth's Biosphere." Nature, Nature Publishing Group, 6 June 2012, www.nature.com/articles/nature11018.

Beiser, Frederick C. Weltschmerz: pessimism in German philosophy, 1860-1900. Oxford University Press, 2016.

Beitrag, Ein, and Joe McCarthy. "There Could Be 2 Billion Climate Change Refugees by 2100." Global Citizen, 29 June 2017, www.globalcitizen.org/de/content/2-billion-climate-change-refugees-2100/.

Benatar, David. Better never to have been: The harm of coming into existence. Oxford University Press, 2008.

Bostrom, Nick. "Existential risks: Analyzing human extinction scenarios and related hazards." Journal of Evolution and technology 9 (2002).

Bostrom, Nick. "Astronomical waste: The opportunity cost of delayed technological development." *Utilitas* 15, no. 3 (2003): 308-314.

Bostrom, Nick. "Transhumanist values." Journal of philosophical research 30, no. Supplement (2005): 3-14.

Bostrom, Nick. "Letter from utopia." Studies in Ethics, Law, and Technology 2, no. 1 (2008).

Bostrom, Nick. "Existential risk prevention as global priority." Global Policy 4, no. 1 (2013): 15-31.

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

Carson, E. Ann. "Prisons Report Series: Preliminary Data Release." Bureau of Justice Statistics, Bureau of Justice Statistics, Sept. 2023, bjs.ojp.gov/library/publications/prisons-report-series-preliminary-data-release.

CCPSC. "Missing and Abducted Children." Child Crime Prevention and Safety Center, 2023. childsafety.losangelescriminallawyer.pro/missing-and-abducted-children.html.

CoE. "e-sermon #9." Church of Euthanasia. https://www.churchofeuthanasia.org/e-sermons/sermon9.html.

Crisp, Roger. "Would extinction be so bad." The New Statesman 10 (2021): 2021.

Crisp, Roger. "Pessimism about the Future." Midwest Studies in Philosophy 46 (2023).

CWA. "PFAS: The Forever Chemicals." Clean Water Action (2024). https://cleanwater.org/pfas-forever-chemicals#:~:text=PFAS%20chemicals%20have%20been%20dubbed,has%20PFAS%20in%20their%20blood.

DHHS. "Alcohol-Related Emergencies and Deaths in the United States." National Institute on Alcohol Abuse and Alcoholism, U.S. Department of Health and Human Services, 2024, www.niaaa.nih.gov/alcohols-effects-health/alcohol-topics/alcohol-facts-and-statistics/alcohol-related-emergencies-and-deaths-united-states.

Dillinger, Katherine. "Chronic Pain Is Substantially More Common in the US than Diabetes, Depression and High Blood Pressure, Study Finds." CNN, 16 May 2023, www.cnn.com/2023/05/16/health/chronic-pain-study/index.html.

Dullaghan, Neil. "EA Survey 2019: Community Demographics & Characteristics," 2019. https://rethinkpriorities.org/publications/eas2019-community-demographics-characteristics.

EAF. "Longtermism." Effective Altruism Forum, 2024. https://forum.effectivealtruism.org/topics/longtermism#Strong_vs__weak_longtermism.

Finneron-Burns, Elizabeth. "What's wrong with human extinction?." Canadian Journal of Philosophy 47, no. 2-3 (2017): 327-343.

Flannery, Frances L. Understanding apocalyptic terrorism: countering the radical mindset. Routledge, 2015.

Fleck, Anna. "Infographic: Countries with the Highest Prevalence of Slavery." Statista, Statista, 24 Aug. 2023, www.statista.com/chart/30666/estimated-number-of-people-in-modern-slavery-per-1000/.

Gabbatiss, Josh. "More than Quarter of World's Land Could Become Arid Due to Global Warming, Study Says." The Independent, Independent Digital News and Media, 3 Jan. 2018, www.independent.co.uk/climate-change/news/global-warming-world-land-arid-desertification-climate-change-study-a8139896.html.

GCECP. "Facts on Child Poverty." Global Coalition to End Child Poverty, 2024. www.endchildhoodpoverty.org/facts-on-child-poverty.

Gibbens, Sarah. "Microplastic Permeates Mariana Trench and Other Deep Sea Points." NatGeo Environment, National Geographic, 6 Dec. 2018, www.nationalgeographic.com/environment/article/microplastic-pollution-is-found-in-deep-sea.

GLF. 1994. Statement of Purpose (A Modest Proposal). https://www.churchofeuthanasia.org/resources/glf/glfsop.html.

Goertzel, Ben. "A cosmist manifesto: Practical philosophy for the posthuman age." (No Title) (2010).

Greaves, Hilary, and William MacAskill. "The case for strong longtermism." Global Priorities Institute Working Paper No. 5-2021 (2021).

Harari, Yuval Noah. "Industrial Farming Is One of the Worst Crimes in History." The Guardian, Guardian News and Media, 25 Sept. 2015, www.theguardian.com/books/2015/sep/25/industrial-farming-one-worst-crimes-history-ethical-question.

Harrison, Gerald. "Antinatalism, Asymmetry, and an Ethic of Prima Facie Duties1." South African Journal of Philosophy 31, no. 1 (2012): 94-103.

HD. "Over 3 Million Americans Struggle with Chronic Fatigue Syndrome." U.S. News, U.S. News and World Report, 11 Dec. 2023, www.usnews.com/news/health-news/articles/2023-12-11/over-3-million-americans-struggle-with-chronic-fatigue-syndrome.

HDR. "2022 Global Multidimensional Poverty Index (MPI)." Human Development Reports, United Nations, 17 Oct. 2022, hdr.undp.org/content/2022-global-multidimensional-poverty-index-mpi#/indicies/MPI.

Hendrycks, Dan, Mantas Mazeika, and Thomas Woodside. "An overview of catastrophic ai risks." arXiv preprint arXiv:2306.12001 (2023).

Knight, Les. These Exit Times. Voluntary Human Extinction Movement, 1991. https://www.vhemt.org/TET1.pdf.

Knutsson, Simon, and First written June. "Thoughts on Ord's 'Why I'm Not a Negative Utilitarian.'" (2016).

Knutsson, Simon. "Philosophical Pessimism: Varieties, Importance, and What to Do." APA Blog. https://blog.apaonline.org/2022/09/13/philosophical-pessimism-varieties-importance-and-what-to-do%ef%bf%bc/.

Knutsson, Simon. 2023. "My moral view: Reducing suffering, 'how to be' as fundamental to morality, no positive value, cons of grand theory, and more." https://www.simonknutsson.com/my-moral-view/.

Korda, Chris. "Credits," 1994. https://www.churchofeuthanasia.org/snuffit5/Credits.html.

Korda, Chris. "Naval Assault on Earthfest," 1994. https://www.churchofeuthanasia.org/snuffit5/Naval_Assault.html.

Koskie, Brandi, and Crystal Raypole. "Depression Statistics: Types, Symptoms, Treatments, More." Edited by Bethany Juby, Healthline, Healthline Media, 31 Oct. 2023, www.healthline.com/health/depression/facts-statistics-infographic.

Krautwald-Junghanns, M-E, et al. "Current Approaches to Avoid the Culling of Day-Old Male Chicks in the Layer Industry, with Special Reference to Spectroscopic Methods - Sciencedirect." Elsevier ScienceDirect, Elsevier, 1 Mar. 2018, www.sciencedirect.com/science/article/pii/S0032579119310636.

Kurzweil, Ray. "The singularity is near." In Ethics and emerging technologies, pp. 393-406. London: Palgrave Macmillan UK, 2005.

MacAskill, William. What We Owe The Future: The Sunday Times Bestseller. Simon and Schuster, 2022.

Mora, Camilo, Bénédicte Dousset, et al. "Global Risk of Deadly Heat." Nature News, Nature Publishing Group, 19 June 2017, www.nature.com/articles/nclimate3322.

Moravec, Hans. "Human culture: a genetic takeover underway." In Artificial Life, pp. 167-199. Routledge, 2019.

Moutinho, Sofia. 2023. "Seafloor Plastic Pollution Is Not Going Anywhere," 2023. https://eos.org/articles/seafloor-plastic-pollution-is-not-going-anywhere.

Leitenberg, Milton. Deaths in Wars and Conflicts in the 20th Century. Ithaca, NY: Cornell University, Peace Studies Program, 2006.

Narveson, Jan. Utilitarianism and New Generations." Mind 76, no. 301 (1967): 62–72.

Narveson, Jan. "Future People and Us." In Obligations to Future Generations, edited by.

NCA. "National Child Abuse Statistics from NCA." (2024) National Children's Alliance, www.nationalchildrensalliance.org/media-room/national-statistics-on-child-abuse/.

NDVH. "Domestic Violence Statistics." The Hotline, National Domestic Violence Hotline, 4 July 2023, www.thehotline.org/stakeholders/domestic-violence-statistics/.

Newberry, Toby. "How many lives does the future hold?." Global Priorities Institute Technical Report (2021).

NIEHS. "Lead." National Institute of Environmental Health Sciences, U.S. Department of Health and Human Services, 11 Mar. 2024, www.niehs.nih.gov/health/topics/agents/lead.

Ord, Toby. "Why I'm Not a Negative Utilitarian." (2013). http://www.amirrorclear.net/academic/ideas/negative-utilitarianism/.

Ord, Toby. "Opening Keynote," 2016. https://www.youtube.com/watch?v=VH2LhSod1M4.

Ord, Toby. The precipice: Existential risk and the future of humanity. Hachette Books, 2020.

Pearce, David. Hedonistic imperative. David Pearce., 1995.

Pearce, Joshua M., and Richard Parncutt. "Quantifying Global Greenhouse Gas Emissions in Human Deaths to Guide Energy Policy." MDPI, Energies, 19 Aug. 2023, www.mdpi.com/1996-1073/16/16/6074.

Pinker, Steven. The better angels of our nature: why violence has declined. Vol. 80. New York: Penguin Books. 2011.

RAINN. "Victims of Sexual Violence: Statistics." RAINN, Rape, Abuse, and Incest National Network (2024). www.rainn.org/statistics/victims-sexual-violence.

Richardson, Katherine, et al. "Earth beyond Six of Nine Planetary Boundaries." Science Advances, Science, 13 Sept. 2023, www.science.org/doi/10.1126/sciadv.adh2458.

Roberson, Claire. "Everything You Need to Know about Factory Farming." Animal Equality, 20 Sept. 2023, animalequality.org/blog/2022/10/14/factory-farming-facts/.

Rothman, Joshua. "The case for not being born." The New Yorker (2017).

Schopenhauer, Arthur. Essays and aphorisms. Vol. 227. Penguin UK, 1970.

Schopenhauer, Arthur. The World as Will and Representation. Aegitas, 2016.

Shiller, Derek. "In Defense of Artificial Replacement." Bioethics 31, no. 5 (2017): 393-399.

Sikora, Richard, and Brian M. Barry. White Horse Press, 1978.

Singh, Asheel. "Furthering the case for anti-natalism: Seana Shiffrin and the limits of permissible harm." South African Journal of Philosophy 31, no. 1 (2012): 104-116.

Sutton, Richard. "AI Succession," 2023. https://www.youtube.com/watch?v=NgHFMolXs3U.

Tegmark, Max. Life 3.0: Being human in the age of artificial intelligence. Vintage, 2017.

THL. "How Are Factory Farms Cruel to Animals?" The Humane League, 2 Jan. 2023, thehumaneleague.org/article/factory-farming-animal-cruelty.

UNESCO. "Imminent Risk of a Global Water Crisis, Warns the UN World Water Development Report 2023." UNESCO.Org, UNESCO, 20 Feb. 2024, www.unesco.org/en/articles/imminent-risk-global-water-crisis-warns-un-world-water-development-report-2023.

UNODC. "GLOBAL STUDY ON HOMICIDE 2023." United Nations Office on Drugs and Crime, United Nations, 2023, www.unodc.org/documents/data-and-analysis/gsh/2023/GSH23_ExSum.pdf.

Vetter, Hermann. "Discussion." In Induction, Physics, and Ethics, edited by Paul Weingartner and Gerhard Zecha. D. Reidel Publishing Company, 1968.

Vetter, Hermann. "IV. The Production of Children as a Problem of Utilitarian Ethics." Inquiry: An Interdisciplinary Journal of Philosophy 12, no. 1–4 (1969): 445–447.

Wells, Andy. 2017. "What Is 'Legal' Torture and How Many Countries Still Use It?" Yahoo News. https://uk.news.yahoo.com/what-is-legal-torture-and-how-many-countries-still-use-it-165034887.html.

WWF. "Devastating 69% Drop in Wildlife Populations since 1970." World Wildlife Fund, 13 Oct. 2022, www.wwf.eu/?7780966%2FWWF-Living-Planet-Report-Devastating-69-drop-in-wildlife-populations-since-1970.

[1] The same goes for equivalence and further-loss views.
[2] I am here referring to "academic" pro-extinctionists, rather than activists. There are many non-academic pro-extinctionists who endorse omnicide and pro-mortalism (see Redacted).
[3] Put differently, in the first case, one starts with antinatalism (as an ethical position) and ends up with pro-extinctionism, whereas in the second case, one starts with pro-extinctionism and ends up with antinatalism (as a method).
[4] Hence, omnicide could be either voluntary or involuntary. It is, however, almost always assumed to be an involuntary event, whereby some group unilaterally kills everyone without the consent of others.
[5] See May 2024 for further discussion.