1. Introduction

In his philosophically penetrating, and characteristically bold, book *The Human Predicament*, David Benatar defends a pessimistic view about the human condition. In doing so, he offers a modified defense of anti-natalism, viz. the view that it’s better never to come into existence. One objection raised against anti-natalists is that if it’s better never to have been, then anti-natalists must think it would be prudent to immediately kill themselves rather than to endure living any additional life. I’ll refer to this objection as the suicide objection to anti-natalism. Benatar responds to this objection, in part, by rightly noting that this conclusion simply does not follow. For while coming into existence may be bad for all persons, premature death may generally be even worse for them than continued life. It’s certainly conceptually possible that the badness of death at time \( t \) for agent \( A \) is worse for \( A \) than the additional life \( A \) would have received had \( A \)’s actual death not occurred, even if it’s also true that it would be better for \( A \) to have never existed (and so, supposedly, had no interest in life) in the first place. If this is right, then anti-natalism needn’t entail that it would be prudent for people to commit suicide.

Getting to the heart of this issue about anti-natalism requires answering a fundamental question about death. Can death be bad for the person who dies and, if so, what makes death bad? The success of Benatar’s reply to the suicide objection hinges on whether premature death really is worse for most people than continued life. With respect to this question, Benatar accepts the basic framework of the most popular view in the literature, viz. deprivationism. Roughly, deprivationism is the view that one’s death is bad for them to the extent that it prevents them from living additional good life. Deprivationism itself may be thought to do little to help the anti-natalist respond to the suicide objection. Though, as I’ll illustrate later, I think standard forms of deprivationism can avoid each of the problems discussed. Still, it’s worth noting why deprivationism may be thought to be powerless to handle the suicide objection. First, some people’s death may be overdetermined and so their death deprives them of nothing at all. Some people are very sick at the end of their life, and their death deprives them of a net harm, and so (according to standard versions of deprivationism) is actually good for them. Benatar, however, believes that (i) such people’s death still *seems bad*, something which deprivationism may have a hard time of capturing. One may also worry that deprivationism lacks the explanatory power to (ii) account for the supposed fact that it’s worse for a teenager to

---

1 Benatar provides his original, and highly influential, defense of anti-natalism in his (2006). See also his (1997) and (2013).
2 Rothman (2017) correctly describes this common layperson objection as a “knee-jerk response.”
3 See Benatar (2017, chs. 5 and 7).
4 Benatar (2017, p. 103).
5 Benatar (2017, pp. 100-102).
6 For the seminal defense of deprivationism, see Nagel’s (1970) and for a more recent, and robust, defense, see Bradley’s (2009).
7 Benatar (2017, p. 102 and pp. 131-133).
die than for a fetus to die and (iii) account for why we should fear earlier-than-necessary death, but not later-than-necessary conception (i.e. the asymmetry problem).  

To help solve these problems, Benatar develops and defends the novel and ingenious annihilation view, according to which “death is bad in large part because it annihilates the being who dies.” Annihilation of moral persons is supposed to be bad enough that it’s prudent for most people to delay their annihilation and continue living instead of committing suicide. It also purports to solve the aforementioned problems to which deprivationism is subject. In this paper, I make both a positive and negative argument against the annihilation view. My positive argument consists in showing that the annihilation view generates implausible consequences in cases where one can incur some other (intrinsic) bad to avoid the supposed (intrinsic) bad of annihilation. Avoiding the implausible consequences I identify requires denying that annihilation is itself bad. My negative argument consists is attempting to undercut some motivation for the annihilation view by showing how standard forms of deprivationism can avoid the particular problems Benatar ascribes to them. This paper is structured as follows. In the next section, I review deprivationism, as well as some motivations for (and objections against) standard versions of the view. In the third section, I review Benatar’s annihilation view, as well as his motivations for accepting the annihilation view in more detail. In the fourth section, I make my positive argument against the annihilation view. In the fifth section, I make my negative argument against the annihilation view. If my arguments are successful, they should collectively provide defeasible reason to accept standard forms of deprivationism over the annihilation view.

2. Deprivationism, Its Appeal, and Its Supposed Problems

2.1 Deprivationism

In its most basic form, deprivationism is the view that death is bad because, and to the extent that, it deprives a person of goods she would have received had her actual death not occurred. According to this view, the more net good death prevents the deceased from gaining, the worse death is for that person. To determine how bad one’s death is for the one who dies, the deprivationist needs some way to measure the net level of goodness of which death deprives its victims. Here I will appeal to Ben Bradley’s view because I think it’s the most plausible and well-developed of all the views on offer. Although nothing for the purposes of my argument hinges on accepting Bradley’s specific version of deprivationism. Bradley measures the amount of goodness of which death deprives its victims with the difference making principle.

**The difference making principle**: The value of event $E$, for person $S$, at world $w$, relative to similarity relation $R = \text{the intrinsic value of } w \text{ for } S, \text{ minus the intrinsic value for } S \text{ of the most } R\text{-similar world to } w \text{ where } E \text{ does not occur}$.  

In other words, the value of an event depends on what would have happened in the nearest possible world where that event had not occurred. This is supposed to hold for all events, including the event of death. I’ll refer to the difference making principle’s calculation of the badness of death specifically as death’s deprivational bad. To illustrate the difference making principle, suppose that the 30 year old Emma Stone tragically dies in a car crash and the nearest world where the car crash didn’t occur is one where she lived an additional 60 years of happy life

---

8 Benatar (2017, p. 119).
9 Benatar (2017, p. 102).
10 Bradley (2009, p. 50).
(containing, let’s suppose, 600 additional net utiles). Deprivationism entails that Stone’s death was very bad for her. It deprived her of 600 net utiles over the course of 60 years. This is the intuitively correct result.

Deprivationism has a lot going for it. It nicely accounts for a variety of intuitive judgements about the badness of death. Consider, for instance, a variation of the case above, but with the 105 year old actor Norman Lloyd in Emma Stone’s place. Lloyd’s death would be bad for him to the extent that it prevented him from missing out on additional good life, which (let’s assume) would be an additional 2 years of happy life containing 20 additional net utiles. Intuitively, Stone’s death seems much worse for Stone that Lloyd’s death seems for Lloyd, even though both may be tragic. Deprivationism generates the intuitively correct verdict here because Stone’s death deprives her of more net good than Lloyd’s death deprives him of (i.e. 600 net utiles compared to 20 net utiles).

The explanatory power of deprivationism also extends to common sense judgements about when death is, all things considered, good for the dying. Death is intuitively thought to be all things considered good for the dying in cases where the dying persons’ quality of life, were they to continue living, would be so low as to make death preferable. To take an unpleasant example, imagine a person who, in the final stages of amyotrophic lateral sclerosis, seeks physician assisted suicide. Intuitively, this person’s death is, all things considered, good for him because it prevents him from experiencing any more of the overwhelming agony caused by amyotrophic lateral sclerosis. This is so even if his death is still tragic in some sense. Deprivationism captures the judgment that his death was, all things considered, good for him and (I’ll argue) can capture the judgment that his death was nevertheless tragic in virtue of the possible good life it precludes.

2.2 Benatar’s Objections to Standard forms of Deprivationism

In spite of its virtues, Benatar argues that standard forms of deprivationism are insufficient to fully account for the badness of death. His first objection is that they entail that overdetermined deaths are not bad for the one who dies.11 To illustrate, consider the following case.

**Assassins:** Two independent assassins, each of whom have been hired to kill the same person, Trevor the Target, shoot and kill Trevor at the exact same moment. Had the first assassin not fired his shot, Trevor would have been killed by the second assassin’s shot.12

According to deprivationism, Trevor’s death is not bad for him because it does not deprive him of any good life. This is a surprising verdict to say the least.

Benatar’s second objection is that standard forms of deprivationism entail that the worst time to die is generally the moment one comes into existence.13 After all, that is generally when death would deprive one of the greatest net amount of good. Many find this result counterintuitive. To illustrate, compare the death of a newborn baby (or a fetus) to that of a twenty year old. Suppose that each would have lived to eighty had they not died

---

11 This issue is discussed frequently in the literature, including McMahan (2002, p. 117), Bradley (2007, p. 126), and Timmerman (2016).
12 This case originally appears in Parfit (1984, p. 70) and is replicated in Timmerman (2015).
13 See McMahan (2002, pp. 162-165), who regards this conclusion as absurd. See also Bradley (2008), who accepts the conclusion.
when they did. According to standard forms of deprivationism, the baby’s death was worse for the baby than the twenty year old’s death was for the twenty year old.

A third objection, albeit one not endorsed by Benatar, is that standard forms of deprivationism supposedly cannot account for the alleged asymmetry between the badness of death and conception. If earlier-than-necessary death is bad because it deprives us of additional good life, then why isn’t later-than-necessary conception bad for the same reason? Perhaps standard forms of deprivationism have to hold that both events (late-conception and early-death) are equally bad for people, yet many find this conclusion absurd.14

3. The Annihilation View

3.1 Formulating the Annihilation View

Benatar develops and defends the annihilation view, which is supposed to be immune from the aforementioned problems that supposedly plague standard forms of deprivationism.15 In an early articulation of the view, he writes that

we should consider is that death is bad in large part because it annihilates the being who dies. Death is bad not merely because it deprives one of the future good that one would otherwise have had, but also because it obliterates one. Put another way, we have an interest not only in the future goods we would have if we continued living, but also an interest in continued existence itself.16

Simply put, Benatar rejects standard forms of deprivationism because there is supposedly “a further explanation for why death is bad—and the further explanation is that one’s annihilation is an independent bad.”17 I will formulate the annihilation view more formally as follows.

**The Annihilation View:** The value of the event of death \( D \) at time \( t \) for person \( S \) is a function of death’s deprivational bad and the non-zero intrinsic badness of \( S \)’s being annihilated at \( t \).

This view may seem a bit odd since death’s deprivational bad is merely extrinsically bad, while the badness of annihilation is, presumably, intrinsically bad. There are, however, a number of ways one may combine intrinsic and extrinsic badness to determine how bad

---

14 See, for instance, Kaufman (1996) and (2016), as well as Draper (2004). Some philosophers allow that conception can be bad for a person in the same way death can, but posit an asymmetry with respect to which attitudes one ought to have towards these harmful events. See, for instance, Ben Bradley’s (2009, pp. 52–60), Anthony Brueckner’s and John Martin Fischer’s (1993), (1998), and (2013).
15 In his (2012), Stephen Blatti defends a similar, but importantly distinct, view that death is a “restriction harm” because it can always thwart of our autonomy. His motivation for positing the existence of restriction harms mirrors some of Benatar’s motivations for accepting the annihilation view. For a critical response to Blatti’s view, see Eric Yang’s (2018).
17 Benatar (2017, p. 103). One complication is that if annihilation is intrinsically bad (or continued existence is intrinsically good), then standard forms of deprivationism could accommodate this claim since death would bring about this intrinsic bad (or deprive one of this intrinsic good). Benatar recognizes this issue, though he suggests that it “makes no difference to [the annihilation] view whether we see the loss of one’s life as the deprivation of an additional good or as a further loss over and above any deprivations it may cause” (2017, p. 110).
one’s death is all things considered. However Benatar wishes to precisify the annihilation view to answer this question is not important for the purposes of my argument. All that matters is that Benatar’s annihilation view holds that the badness of death for a person is determined both by the deprivational badness and the badness of annihilation.

3.2 Motivating the Annihilation View

In support of the annihilation view, Benatar references Frances Kamm’s *Limbo Man*, someone who has the supposedly common preference to put “off a fixed quantity of goods of life by going into a coma and returning to consciousness at a later point to have them.”18 Kamm’s suggestion is that both options are equally prudentially good for Limbo man, though he may still rationally prefer “to postpone things being all over, even if this did not increase the total amount of [net good he] had in [his] life.”19 I find it hard to see how one could rationally prefer (rather than hold an attitude of indifference toward) one option over others that are equally prudentially and morally good. Nevertheless, to the extent that we find *Limbo Man*’s preference rational, we should take it as evidence that earlier-than-necessary annihilation is, in some sense, bad.

In addition to the prima facie intuitive support provided by *Limbo Man*, Benatar defends the annihilation view by arguing that it avoids the aforementioned problems that supposedly plague standard forms of deprivationism. First, it allows that death can be bad for people who die with no (net) prospective good in their future.20 The annihilation view can generate this consequence because, while their death deprived them of no good, it still resulted in their annihilation and this is intrinsically bad for them. Second, the annihilation view can identify a relevant asymmetry between later-than-necessary conception and earlier-than-necessarily death. Later-than-necessary conception does not result in annihilation, while earlier-than-necessary death does, making premature death especially bad for the one who dies. This supposedly allows the annihilation view to avoid the asymmetry problem.21

Finally, the annihilation view avoids the consequence that the worst time to die is typically the instant one comes into existence.22 This is because annihilation is supposedly bad for a person in virtue of the fact that it “thwarts an interest in continued existence that one acquires when one comes into existence.”23 Benatar suggests that one’s interest in continued existence increases as one achieves personhood and so it may be worse to die as a teenager than as a newborn, which is thought to be the intuitively correct verdict. It’s worth noting that this claim seems to potentially be in tension (but not strictly inconsistent) with Benatar’s claim that, *ceteris paribus*, it’s worse to be annihilated earlier rather than later.24 Moreover, to maintain this conclusion, Benatar would also have to maintain that the potential good lost by the increased interest in continued existence is greater than the net good one accrues before achieving personhood. This needn’t be a problem for Benatar,

---

22 On standard forms of deprivationism, this consequence only applies to a subset of people whose lives are (or would be) good, all things considered. Benatar (2017, pp. 130-131). Benatar notes that his annihilation view can generate the same, seemingly correct, verdicts as Jeff McMahan’s time-relative interest account (TRIA) of the badness of death. For more on McMahan’s account, see his (2002, pp. 165-174 and 232-240) and (2019).
23 Benatar (2017, p. 130).
however, as it seems to fit nicely with his anti-natalist views about the objective quality of life. However, this solution may seem less appealing to those who believe that typical humans have lives that are reasonably good before they gain personhood. Now that I have explained the annihilation view, and Benatar’s motivations for accepting it, I am in a position to present my positive argument against it.

4. A Positive Argument Against the Annihilation View

4.1 My Positive Argument

My positive argument against the annihilation view may be seen as a sort of *reductio ad absurdum*. The basic idea is that if annihilation is intrinsically bad for persons, then it would be prudent to deliberately incur some intrinsic bad less bad than annihilation in order to avoid annihilation. However, that commitment leads to an absurd conclusion. It’s not prudent to deliberately incur any amount of something clearly intrinsically bad (e.g. pain) in order to avoid annihilation itself. Thus, annihilation is not intrinsically bad for persons. To illustrate the problem, consider the following morose case.

**Cancer:** Diego is dying of abdominal cancer and his prospective future life contains nothing but immense amounts of pain and absolutely no compensating goods. In other words, if Diego continues to live with cancer, his life will contain much that is intrinsically bad for him and nothing that is intrinsically good for him.

Now, at t₀, Diego is faced with the following two options.

**(A)** Diego can die a painless death immediately at t₁, thereby avoiding any future pain.

**(B)** Diego may be granted an immortal life where he will suffer in misery until t₁₀ and then exist with a well-being level of 0 from t₁₁ through eternity, thereby avoiding annihilation.

There are a multitude of ways that Diego may continue to exist for an eternity with a well-being level of 0. He could be rendered comatose with no conscious experiences from t₁₁ onwards. Alternatively, he could continue to exist for an eternity with conscious experiences, but ones that are neither (intrinsically) good nor bad for him. The important point is that were

---

25 For a seminal defense of anti-natalism, see Benatar’s (2006). See also his (2017) for additional arguments that support anti-natalism. For a distinct defense of the view, see Belshaw (2012).

26 Some might believe that continued life necessarily contains some prospective good, even if it’s outweighed by the bad. I don’t accept this, and I doubt Benatar does either. But even if living necessarily entails that one is accruing some intrinsic good, Cancer may simply be re-described so that Diego’s prospective future contains net bad.

27 Perhaps Benatar would regard this as a form of annihilation. He does write that the “irreversible cessation of consciousness is annihilation of the conscious being.” (2017, p. 109). However, being in a coma is not irreversible and presumably people have not been annihilated during the time they’re temporarily in a coma. So being in a reversible coma for eternity does not, I think, constitute annihilation as Benatar uses the term.

28 If we tweak the thought experiment a bit, a third option allows Diego to have experiences that are intrinsically good after t₁₀, but each good experience is perfectly balanced out by a bad experience, such that Diego’s momentary well-being is constantly 0 for eternity. I set aside this version of the thought experiment, however, in order to avoid begging the question against Benatar’s asymmetry argument. My
Diego to exist in such a state, he would be living a life that is not good for him in any way, apart from the apparent good of avoiding annihilation.

Now, suppose that the pain Diego experiences from $t_1$-$t_{10}$ is quite (intrinsically) bad for him, but not quite as (intrinsically) bad as annihilation happens to be. Benatar does not specify how bad annihilation is for the person who dies, only suggesting that it’s very bad for everyone. This is not a problem for my argument though. However bad Benatar believes annihilation is, I will stipulate that the net pain experienced from $t_1$-$t_{10}$ is ever so slightly less bad than annihilation. Here is the problem. Given options (A) or (B), (A) seems to be the more prudent choice. At least, this seems to prudent choice to me and I suspect (but am not certain) it will seem to be the prudent choice for others who lack pre-existing theoretical commitments in this debate. Here’s the problem. If annihilation is intrinsically bad, then (B) would actually be the prudent choice. This seems quite counterintuitive, so we should infer that annihilation is not intrinsically bad. Choosing (B) for self-interested reasons seems to be a simple prudential mistake. It seems to me that there could be no good reason to endure suffering in order to gain additional life that is stipulated to contain no intrinsic goods and avoid no extrinsic bads apart from simply remaining in existence.

In response, Benatar might reiterate that he takes annihilation itself to be (intrinsically) bad, so avoiding annihilation would be extrinsically good and it can be prudent to deliberately incur some intrinsic bad in order to gain an extrinsic good. That’s a fair response, but it’s one that I believe many will find unconvincing. Here’s why. I predict that, pre-theoretically, many would find it extremely counterintuitive to hold that it would be prudent to endure any amount of physical pain in order to avoid annihilation. At least, I think this is true in cases where avoiding annihilation does not allow one to accrue any additional good life. This should provide strong, yet defeasible, reason to believe that annihilation itself cannot be intrinsically bad.

There is even stronger reason to reject the annihilation view if we hold, as Benatar does, that, ceteris paribus, it’s less bad to be annihilated later rather than sooner. If, ceteris paribus, it’s in one’s prudential interest to delay their annihilation (even when a later annihilation would not make a difference to the other goods of which one’s death deprives them), then it would be prudent for a person to incur some intrinsic bad, such as a small amount of pain (e.g. brief, but intense or mild, but less brief) in order to delay their annihilation. For the purposes of my argument, it doesn’t matter what the exact ratio is, so long as it’s worth incurring some intrinsic bad to delay annihilation. Suppose, for the sake of argument, that it’s prudent to endure one millisecond of intense pain in order to have one’s annihilation delayed one week.\(^{29}\) If this is right, then it would be in one’s prudential interest to incur an infinite amount of suffering in order to avoid annihilation. To illustrate why the annihilation view generates this conclusion, imagine that each week Diego is offered the choice to have his annihilation delayed one week if he incurs one millisecond of pain and each week he makes the supposedly rational choice to delay his annihilation. If he always makes this choice, then he will delay his annihilation indefinitely and, in the process of doing so, incur an infinite number of millisecond long pains.

---

\(^{29}\) Benatar (2017, p. 133). It’s unclear why, ceteris paribus, it would be less bad to be annihilated later rather than sooner. Why would the intrinsic badness of annihilation decrease as time goes on, rather than remain constant? Most people do not regard other intrinsic bads in this way. Breaking my arm and suffering 10 units of pain is, ceteris paribus, equally bad for me whether it happens tomorrow or fifty years from now.

\(^{30}\) Assume also that the delayed annihilation never results in the agent receiving any other intrinsic goods.
This would result in Diego incurring an infinite amount of pain in order to avoid annihilation. But it’s utterly absurd to hold that it’s better for one to suffer an infinite amount of pain than to avoid going out of existence. So something has gone wrong.

4.2 Objections to My Positive Argument, and Responses to Those Objections

Benatar could respond to my objection in a few different ways. One option is to hold that the value of delaying annihilation decreases exponentially as time goes on, such that the amount of pain it’s worth enduring to delay annihilation has an upper limit. Such a response strikes me as objectionably ad hoc. Setting that worry aside, however, the bigger problem is that this would still allow that it would be prudent to incur some finite amount of pain to avoid annihilation. But even that more moderate claims strikes me as absurd.

A second option is for Benatar to retract his claim that, ceteris paribus, it’s better to delay annihilation, instead holding that annihilation is equally bad for all full-fledged moral persons. The primary problem (at least for Benatar) with this response is that it undermines the annihilation view’s force against the suicide objection. If, ceteris paribus, annihilation is equally bad independent of when it occurs, then one would have no interest in delaying annihilation, only in avoiding it altogether. Since choosing to continue living instead of killing oneself would, in every actual case, only delay one’s annihilation, the supposed badness of annihilation would provide people with no reason to continue living at any given time. The second problem with this response is that it too allows that it can be prudent to incur some finite amount of pain to avoid annihilation.

A third option is for Benatar to suggest that annihilation is merely extrinsically bad and not, as I have assumed, intrinsically bad. Unfortunately, this move would not help the annihilation view since extrinsic badness matters prudentially only insofar as it affects the net intrinsic good one accrues over the course of their life.31 Since Benatar was attacking standard forms of deprivationism, he was focusing on cases where avoiding annihilation made no difference to the net receipt of intrinsic goods. As such, it would make no sense to hold that annihilation is merely extrinsically bad for people. Moreover, even if annihilation were merely extrinsically bad, its extrinsic badness would be prudentially irrelevant in the cases that matter for Benatar’s argument. Given these considerations, I believe that Benatar is unlikely to hold that annihilation is merely extrinsically bad.

A fourth option would rely heavily on the ceteris paribus clause, and posit a lexical priority between the badness of pain and the badness of annihilation. In other words, Benatar could hold that while annihilation is indeed intrinsically bad, it’s worse to incur any amount of other intrinsic bads (e.g. pain) than to be annihilated. Kamm’s intuition that the extra time one receives in Limbo Man may be a “tie breaker” suggests that there is intuitive support for this position.32 My response is twofold. First, I am skeptical that there are lexical relationships

31 See Bradley (2009, chs. 2 and 5) and Timmerman (2016, pp. 20-21).
32 Kamm (1993, pp. 19 and 49-54). It’s worth noting that intuitions in the “Limbo Man” case are susceptible to being influenced by background considerations that are not relevant to the question at hand. For instance, people might prefer to live longer because they think it would be less painful for others were their life spread out, or they may think they will get specific goods that will improve their life (e.g. seeing a child graduate from college) if they live longer. However, a deprivationist could allow that such considerations count in favor of choosing to live longer. If we really imagine that all other things are equal, it’s far from clear that it’s better for one to have their good life spread out over a longer period of time.
between any two goods or bads because of the variety of problems such views incur.\textsuperscript{33} Second, and more importantly for the purposes of my argument, this position would largely undermine the motivation for Benatar’s annihilation view. This is because positing such a lexical relationship renders annihilation less bad than anything else and, as such, would make it barely trivially bad for anyone. In other words, holding that annihilation is less bad than any other bad would prevent it from doing the explanatory work Benatar wants it to do. Annihilation’s supposed badness could not explain why someone’s death is genuinely and seriously bad for them if they have no prospective goods in the future. It could only explain why it would be trivially bad for them, though not even as trivially bad as a pinprick. Furthermore, it could not justify asymmetric attitudes toward prenatal and postmortem non-existence, as prenatal deprivations would be worse than any overdetermined death. Nor could it explain why death is worse for an adult than for a baby, as the additional deprivational bads would, on the lexical view, be worse for the baby than the annihilation would be for the adult.

Finally, a fifth option is for Benatar to avoid the issues that arise for lexical views by instead holding that annihilation and other intrinsic bads are incommensurable. Perhaps these bads are so disparate that there is no way, even in principle, to compare how bad they are relative to one another. While this avoids the issues with positing lexical priority, it’s not a tenable solution for the annihilation view. The fatal problem for this option is that the supposed badness of annihilation and other bads can’t be incommensurable for reasons to which I’ve already alluded. It is, I take it, determinately better for you to be annihilated than to endure an eternity of agony. If so, then the badness of pain really can be weighed against the badness of annihilation. If they were incommensurable, then there would be no fact of the matter about which fate is worse. Clearly, however, infinite agony is worse than annihilation.\textsuperscript{34} Now, to avoid this problem, Benatar could hold that some quantities of pain (or other intrinsic bads) are on a par\textsuperscript{35} with annihilation, while others are determinately worse. That is perhaps the best response available to Benatar. Unfortunately, this response raises the question of how much of each intrinsic bad must be accrued before it becomes determinately worse than annihilation. I see no way of answering that question without relying on completely arbitrary cutoff points. This should make us skeptical of that there is incommensurability at any level of comparison. The second problem with this move is that its proponents have to deny that it’s prudentially better to be annihilated than to incur some amount of pain (assuming, of course, that the pain would not be followed by some compensating good). As I’ve stated earlier, this seems quite implausible and, as such, I am skeptical of the viability of this move.

These considerations collectively provide good reason to be skeptical of the annihilation view, specifically of the claim that annihilation itself is bad for anyone over and above the goods of which annihilation deprives a person. In the next section, I make my negative argument against the annihilation view by attempting to demonstrate that standard forms of deprivationism can avoid the problems Benatar ascribes to them.

\textsuperscript{33} See, for instance, Huemer (2010) and Parfit (2016).
\textsuperscript{34} I think Benatar would likely already accept this claim since he accepts that suicide can be rational. See his (2017, ch. 7) and his (2021).
\textsuperscript{35} For a defense of a framework that would allow Benatar to make this move, see Chang (2002).
5. Three Negative Arguments Against the Annihilation View

My negative case against the Annihilation View consists in illustrating how standard forms of deprivationism can solve the problems that Benatar claims they cannot. I’ll respond to each in the order in which they were raised in section three.

5.1 How Deprivationists Can Handle the Overdetermination and the Good Death Problems

Recall that one advantage of the annihilation view is that it still allows that overdetermined deaths are bad for people since such deaths still result in the annihilation of the person. It also allows that deaths can still be bad, in some sense, for people who have no prospective goods in their future (e.g. a terminally ill person suffering in constant agony). According to Benatar, the correct account of the badness of death should be able to hold that it’s “entirely appropriate for a person facing unbearable suffering to welcome death, all things considered, even while deeply regretting that annihilation is the only means to avoiding this fate worse than death.”

I wholeheartedly agree with Benatar that the correct account of death should allow that it’s perfectly appropriate for one have a non-trivial negative attitude toward their death, even in cases in which they should prefer death to any available alternative. However, standard forms of deprivationism can easily accommodate this judgment and, as such, positing annihilation as an additional intrinsic bad does not provide any explanatory advantage for the annihilation view over deprivationism. Standard forms of deprivationism can account for this intuitive judgment in multiple, consistent, ways. Here are the three primary ways the deprivationist can do this.

First, the deprivationist can point out other comparative claims that are relevant to the badness of (and our attitudes towards) death. It may be true that a terminally ill person’s death is (all things considered) good for them because, had their actual death not occurred, they would not have continued to live a life worth living. But, and this is the crucial point, it’s also true that this person would be better off in a world where they did not get sick in the first place and were able to continue to live a long and happy life. So, this person is better off in the actual world relative to the world where they still got sick in the first place, living an even longer, happier life than the one they actually lived. The deprivationist can (and should) hold that it’s bad for this person to miss out on living in worlds where they would live a longer, happier life. Moreover, they can (and should) hold that it would be entirely appropriate for a person facing this unbearable suffering to regret that they’ll die now rather than being able to live in such a world.

Second, along the same lines, the deprivationist can (and should) posit plural harms, which allows them to hold that such deaths are part of an event that is bad for the person dying, even if the event of death itself isn’t. To illustrate, recall the assassins case from section three.

**Assassins:** Two independent assassins, each of whom have been hired to kill the same person, Trevor the Target, shoot and kill Trevor at the exact same moment.

---

37 For more on this kind of response, see Bradley (2015). See also McMahan (2002, ch. 2).
Had the first assassin not fired his shot, Trevor would have been killed by the second assassin’s shot.

According to standard forms of deprivationism, <Trevor’s dying from the first assassin’s shot> was not bad for him because, had he not died from the first shot, he would have died one second later from <the second assassin shooting him>. However, deprivationists can hold that Trevor’s death from the plural event of <both assassins shooting at the times they did> was bad for him because, had neither shot him, he would have lived a much longer and happier life.\(^{38}\) Now, here’s the kicker. All actual deaths are likely part of some larger plural event of which that person’s death is a part. Dying of cancer, for instance, is caused by some set of single events, which we can group together and think of as a plural event. Had this plural event that caused cancer not occurred, then one would not get cancer in the first place, instead going on to live a long happy life. In short, standard forms of deprivationism can admit that the non-plural event that caused one’s death was not bad for the person in question, but avoid counterintuitive implications by noting that it was part of a plural event that was bad for the person in question. It would, presumably, be entirely appropriate for someone who sees death as their best available option to have a negative attitude toward the bad plural event that is preventing them from living a long and happy life.

Third, Benatar seems to be operating under the common, though I believe mistaken, assumption that one’s attitudes toward their death should track how good or bad their death is for them. But, as I’ve argued elsewhere, the deprivationist can (and should) reject this assumption.\(^{39}\) This is because agents can act in ways that make one’s death good or bad for a person without providing them any reason to change their attitudes toward death. To illustrate, consider the case of Sick Suzy.

**Sick Suzy:** Suzy has a heart defect and is going to die within a month. She might think that the nearest possible world in which she does not die is one where she receives a heart transplant and lives a long happy life. However, Naïve Ned is Suzy’s friend and he wants to minimize the badness of Suzy’s death. Knowing that Suzy will die within a month, Ned hires a professional torturer to immediately torture Suzy if and only if she recovers from her heart defect. Now, the nearest possible world in which Suzy survives her illness is one where she is tortured for the remainder of her natural life.\(^{40}\)

Here’s the takeaway from the case. Ned’s actions do prevent Suzy’s death from being bad for her, but they do not give her any reason to change her negative attitude toward missing out on additional good life that is, in some sense, a relevant possibility for her. To further illustrate this point, consider the case of Betty and the Billionaire.

**Betty and the Billionaire:** Betty has the same heart defect as Suzy and she knows that she will die within a month. Betty’s enemy is a naive billionaire who, in a ludicrous attempt to exact revenge on Betty, resolves to make Betty’s death as bad for her as possible by entering into a legally binding agreement to give Betty all her money if and only if Betty recovers from her heart defect.\(^{41}\)

---

\(^{38}\) Parfit (1984, ch 3) was the original proponent of the plural harm solution. However, Feit (2015) provides the best defense of plural harm, as well as the most plausible formulation of plural harms.

\(^{39}\) See my (2015) and (2016).

\(^{40}\) This example is drawn from Timmerman (2016, p. 20).

\(^{41}\) Timmerman (2016, p. 21).
In this case, the takeaway is that Betty’s enemy does manage to make Betty’s death extremely bad for her, but this does not make Betty’s death more lamentable than it was before it became (extrinsically) worse. It was never a relevant possibility for Betty to live to inherit billions of dollars.

These cases should help illustrate that one’s attitudes toward death should track something other than how good or bad death is for the person in question. Here is how this helps standard forms of deprivationism avoid the counterintuitive conclusion. They can hold that, even in cases in which someone’s death is not bad for them it’s perfectly appropriate to have a negative attitude toward their death. Of course, this raises the question of what our attitudes toward death should track. If they shouldn’t track death’s goodness or badness, what should they track? Elsewhere, I have argued that our attitudes toward death should track at least two things. First, they should track the difference between how good one reasonably expected their life to be and how good their life actually was. Second, they should track the difference between how good one’s life actually was and how much good one reasonably believed was (metaphysically) possible for them to have gained in their life. More precisely, I defended the following two principles.

**The Expectation Principle (EP):** The degree to which a person S ought to lament her death is partly determined by how well S fares relative to her justified beliefs about her expected quality and quantity of life. The better S fares relative to her justified expectations, the less she should lament her death and vice versa.42

**Metaphysical Possibility (MP):** The degree to which a person S ought to lament her death is partly determined by the amount of good S is justified in believing was metaphysically possible for her to have obtained had S not died at the time she did.43

I am not sure how confident I am in these two principles, but the general point stands. Deprivationists can hold that one’s death is lamentable even if their death is not bad for them, thereby allowing them to avoid Benatar’s worry. The three responses I’ve reviewed in this section may be consistently adopted by standard forms of deprivationism. Adopting them should provide deprivationism with the same explanatory power as the annihilation view, at least with respect to cases of death where one had no prospective good in their future.

### 5.2 How Deprivationists Can Solve the Asymmetry Problem

The second supposed advantage of the annihilation view is that it can help solve the asymmetry problem by explaining why earlier-than-necessary death is bad for people in a way that later-than-necessary conception is not. Later-than-necessary conception does not result in annihilation, while earlier-than-necessary death does, making premature death especially bad for the one who dies. I see the appeal of such a claim. But, unfortunately, I do not think this provides the annihilation view any substantive explanatory advantage for the following reasons.

First, recall the problems for the annihilation view in the previous section. If annihilation is less bad than other intrinsic bads (such as pain), then *ceteris paribus* there would be little difference between the badness of a later-than-necessary death and an earlier-than-necessary conception. In fact, if one’s conception deprived them of enough

---

42 Timmerman (2016, p. 31).
43 Timmerman (2016, p. 32).
pleasure, the annihilation view might have to allow that some conceptions are worse than some deaths. But that is at odds with the standard intuitive judgements about the case, which is that there is a very large difference between the badness of conception and the badness of death. More precisely, most seem to think it’s not bad at all to have a late conception, but is very bad to have an early death. Second, along the same lines, even if earlier-than-necessary death were necessarily worse for people than later-than-necessary conception, this wouldn’t completely solve the asymmetry problem since this solution allows that it can still be bad for one to have a late conception. That, itself, is a judgment that people who want to posit an asymmetry will reject. The standard view is that later-than-necessary conception is not bad at all for anyone.

Third, standard forms of deprivationism may have the explanatory resources to avoid the asymmetry problem without having to hold that annihilation is itself bad. It would, I believe, be more parsimonious for deprivationists to adopt one of these other solutions, solutions of which Benatar is aware.\textsuperscript{44} For instance, deprivationists may hold that it simply is not possible in the relevant sense for one to have been conceived earlier than they were, though it’s possible for one to die later than they will.\textsuperscript{45} If this is right, then our deaths could be bad for us, though our conceptions couldn’t be bad for us. Alternatively, they may try to explain the asymmetry in terms of fitting attitudes, suggesting that we’re rationally temporally biased. We rationally care much more about (hedonic) goods and bads and in the future and than (hedonic) good or bads in the past. Consequently, we should care about our earlier-than-necessary deaths though not at all about our later-than-necessary conceptions.\textsuperscript{46} There is also a third option for the deprivationist, one that I have defended. They may hold that the badness of our deaths should be salient to us in a way that the potential badness of our conceptions aren’t.\textsuperscript{47} This may be (a) because the choices we make could affect when we will die, but not when we were conceived, (b) because death is a threat to us while we’re alive, though later-than-necessary conception isn’t, or (c) because the time that one’s death will occur is generally unknown, and so anxiety-inducing, while the time of one’s conception is generally known and not anxiety-inducing.

If these solutions are viable options, and if just one of them succeeds, then positing an additional harm that only applies to death seems superfluous. Moreover, given the theoretical problems with the annihilation view discussed in section four, adopting it may create more problems than it could help solve.

5.3 How Deprivationists Can Account for Intuitions about the Worst Time to Die

The final supposed explanatory advantage of the annihilation view is that it entails that it’s generally worse for a teenager to die than for an infant to die, which is supposed to be the intuitively correct result. Standard forms of deprivationism supposedly cannot capture this judgment because, in typical cases, an infant’s death deprives the infant of more good life than a teenager’s death deprives the teenager. In those cases, standard forms

\textsuperscript{44} Benatar (2017, pp. 117-122).

\textsuperscript{45} This solution was originally defended by Nagel in his (1970). However, Kaufman has done the most to develop and defend this view, having done so in numerous papers including his (1996) and (2016).

\textsuperscript{46} See Fischer’s (2009, ch. 5) and, again, Anthony Brueckner’s and John Martin Fischer’s (1993), (1998), and (2013). For good arguments against the rationality of time bias, see Greene and Sullivan (2015) and Sullivan (2018).

\textsuperscript{47} See Timmerman (2018).
of deprivationism entail that the infant’s death is worse for the infant than the teenager’s death is for the teenager.

I understand that this verdict seems prima facie implausible. But what seems counterintuitive about the conclusion is, I believe, the result of people conflating different senses of badness. After distinguishing the different senses of badness from one another and reviewing what deprivationists can say about each sense of badness, the seemingly counterintuitive implications of deprivationism should disappear. The first sense of badness concerns how bad one’s death is for the person who died. I’ll refer to this sense of badness as the prudential badness of death. The second sense of badness concerns how bad one’s death is from the perspective of the universe, which picks out how impartially bad one’s death is. I’ll refer to this sense of badness as the impartial badness of death.

These two senses of badness can come apart. The death of Stalin may have been prudentially bad for Stalin since he would have lived a longer happier life had he not died when he did. But his death may have been impartially good since it resulted in less undeserved suffering than there otherwise would have been (plus the masterful satirical film The Death of Stalin by Armando Iannucci). The death of Norman Lloyd may not be very prudentially bad, yet it will surely be terrible in the impartially bad sense since his death will be very bad for his family, friends, and cinephiles. The death of a 25 year old tortoise may be prudentially very bad for the tortoise, depriving it of over 100 years of life, but not impartially very bad assuming that it’s death will not be very bad for other sentient beings (and perhaps because it’s less tragic for a being with a lower moral status to go out of existence).

Now, here’s the important point. Judgments about it being worse for a teenager to die than for an infant to die may be tracking impartial badness as opposed to prudential badness. If so, then the deprivationist has no trouble accounting for the intuitively correct judgments about such cases. All the deprivationist is committed to holding is that death is usually prudentially worse for an infant than for a teenager and that seems true. After all, an infant’s death typically deprives it of more net good than a teenager’s death deprives the teenager. Nothing prevents the deprivationist from making the further claim that it’s generally impartially worse for a teenager to die than an infant because, perhaps, typical teenagers are full moral agents who have complex interests, categorical desires, and goals. The deprivationist may hold that the mere fact that a moral agent of this kind is going out of existence makes their death impartially bad, even tragic. Furthermore, young adults may typically have their lives more deeply interconnected with other people, and so, their death may typically harm other people more than the death of newborns.

Perhaps Benatar will respond that he simply has different intuitive judgments than I do. He may want to maintain that death really is prudentially worse for a teenager than an infant. I am skeptical. To illustrate why, imagine that my parents were given one of two choices after I was born. They could have let me have been painlessly killed that day or they could let me live for another sixteen years before being painlessly killed without warning. If forced to choose between those two options, I would hope that they would have let me live for another sixteen years. After all, allowing me to live into young adulthood would have resulted in my having a higher total well-being than I would have had had I been killed as a newborn. Thus, my dying as a baby would have been prudentially worse for me than my dying as a teenager. If one were to actually hold the view that death is typically prudentially worse for a teenager than for a newborn, they would seem to commit themselves to the view that it would have been prudentially best for me to have been killed as a baby. I find that conclusion implausible.
Perhaps my judgments in such cases are mistaken. I believe that Benatar will maintain that it’s prudentially better to die as an infant than as a teenager. According to Benatar, we should think this, in part, because there will be very little (to no) psychological unity between my infant self and my counterfactual teenage self and we should think that infants have very little interest in continued life. On the other hand, there would be a great deal of psychological unity between my teenage self and my counterfactual adult self. This may mean that my infant self had little interest in continued life, while my teenage self had a great deal of interest in continued life. Appealing to the supposed relevance of psychological relations allows Benatar to hold that death is typically prudentially worse for teenagers than for infants in spite of the fact that teenagers typically have higher total well-being levels.\(^{48}\)

I see the appeal of Benatar’s view, even though I am ultimately more confident in standard forms of deprivationism. I also grant that his view may have an explanatory advantage over standard forms of deprivationism if most people’s judgments of the badness of death for teenagers and infants really is tracking prudential badness, and not impartial badness. Unfortunately, even assuming Benatar is right in his judgment about these cases, he still has not provided good reason to accept the annihilation view. This is simply because standard forms of deprivationism can be amended to accommodate the supposed relevance of psychological unity relations, thereby generating the verdicts Benatar wants. Crucially, they can do this without having to posit that annihilation is itself bad. This is exactly what Jeff McMahan’s time-relative interest account does, as Benatar recognizes.\(^{49}\) Adopting the time-relative interest account would allow the deprivationist to accommodate Benatar’s judgments about the relative prudential badness of infants’ and teenagers’ death without incurring the problems associated with the annihilation view.

To recap, I have provided a sort of debunking argument of the intuitive judgment that it’s worse for a teenager to die than for an infant to die. Such judgments are really tracking impartial badness, not prudential badness, or so I suggested. I have also conceded that there is a plausible principled reason, which Benatar appeals to, that supports the judgment that death really is generally prudentially worse for teenagers than infants. This, however, does not support the annihilation view since deprivationists can accommodate Benatar’s claims about the relevance of psychological unity relations without having to hold that annihilation is itself bad.

If my negative arguments are successful, they should undercut some of the motivation for the annihilation view by showing how standard forms of deprivationism can avoid the problems that Benatar ascribes to them.

6. Conclusion

In this paper, I have made a positive and negative argument against the annihilation view. My positive argument consisted in highlighting the theoretical problems generated by positing that annihilation is itself bad. The primary problem is that this suggests it would be prudent for someone to incur some (intrinsic) bad, such as some quantity of pain, in order to avoid the (intrinsic) bad of annihilation. Yet, it seems imprudent to incur any amount of pain to avoid annihilation, so annihilation itself is not (intrinsically) bad.

\(^{48}\) Benatar (2017, pp. 129-130).
\(^{49}\) Benatar notes that his annihilation view can support the verdicts of the time-relative interest account (2017, pp. 129-130). See McMahan (2002, pp. 165-174).
My negative argument consisted in trying to demonstrate that standard forms of deprivalism can avoid three of the major problems that Benatar attributes to them without having to posit that annihilation itself is bad. They can handle overdetermination cases by relying on comparisons between the actual world and more distant worlds where the agent lived a better life, appealing to the notion of plural harm, and by divorcing fitting attitudes toward death from how good or bad death is for the person. They can handle the asymmetry problem by holding that it’s impossible to be conceived substantially earlier in time, or by appealing to time bias to justify an attitudinal asymmetry between earlier-than-necessary death and later-than-necessary conception, or they may appeal to contingent facts about our lives to justify an asymmetrical salience with respect to the badness of each event. Finally, they can try to debunk the intuitive verdicts about the worst time to die by relying on the distinction between prudential and impartial badness or they can accommodate those verdicts by adopting McMahan’s time-relative interest account of the badness of death.

If my arguments are successful, they should collectively provide reason to accept standard forms of deprivalism over the annihilation view. That being noted, my arguments are certainly not conclusive. The annihilation view may be able to better handle these alleged problems than standard forms of deprivalism for reasons I have overlooked. Moreover, there are other considerations in favor of the annihilation view that I have not been able to address in this paper, and perhaps there are some that are yet to be discovered. Fully answering these questions necessitates further engaging with Benatar’s rich and nuanced view.

As one example, Benatar suggests that one consideration in favor of the annihilation view is that it’s “consistent with (but not the same as) judgments about the destruction of objects that lack prudential value but have value of another kind” (2017, p. 106). His thought is that “if damaging an object of value is bad, then annihilating it—an extreme form of damage—is also bad.” The examples he gives concern damaging the Grand Canyon or the Mona Lisa. I am not convinced by such examples because I deny that nature and works of art have any intrinsic value. On my view, they’re only extrinsically valuable insofar as their destruction would harm sentient beings. So, I’d want to explain the badness of their destruction in terms of the bad effects it has on well-being and that is perfectly consistent with standard forms of deprivalism.

I would like to thank David Benatar, James DiGiovanna, David Kaspar, Felipe Pereira, Rhys Southan, and Jacob Sparks for very helpful comments on earlier drafts of this paper.
References


Parfit, Derek. 2016. “Can We Avoid the Repugnant Conclusion?” *Theoria* 82 (2): 110-127.


