

Fall 12-8-2017

## **A Game We Have to Lose: Overcoming the Harm of Coming Into Existence**

Hannah Strang  
hstrang@bgsu.edu

Follow this and additional works at: <https://scholarworks.bgsu.edu/honorsprojects>



Part of the [Ethics and Political Philosophy Commons](#)

---

### **Repository Citation**

Strang, Hannah, "A Game We Have to Lose: Overcoming the Harm of Coming Into Existence" (2017).  
*Honors Projects*. 391.  
<https://scholarworks.bgsu.edu/honorsprojects/391>

This work is brought to you for free and open access by the Honors College at ScholarWorks@BGSU. It has been accepted for inclusion in Honors Projects by an authorized administrator of ScholarWorks@BGSU.

Abstract:

This paper explores the asymmetry of pleasure and pain as expressed in David Benatar's book *Better Never to Have Been*, which is the basis for the argument that it is always an irreparable harm to bring a person into existence, and therefore we are morally obligated to pursue extinction as a species. I will examine Benatar's argument in support of the asymmetry's existence and analyze the strength of his argument for extinction overall, ultimately determining that his conclusion is too strong. I will defend this claim on the grounds that Benatar's asymmetry implies the truth of two claims that must be false according to the nature of objective morality: that a moral universe is unattainable, and that morality is better off without the existence of moral agents. I will conclude by defending Benatar's view about the harm of bringing a person into existence, but revising his secondary argument that we are morally obligated to seek extinction because of this. In its place, I propose that we should adopt a policy of limited procreation, to increase the chance of any person who is brought into existence having a life that creates or experiences enough good to compensate for the initial bad that stems from the harm of coming into existence.

A GAME WE HAVE TO LOSE:  
OVERCOMING THE HARM OF COMING INTO EXISTENCE

HANNAH STRANG

HONORS PROJECT

Submitted to the Honors College  
at Bowling Green State University in partial fulfillment of the  
requirements for graduation with University Honors

DECEMBER 16, 2017

Louis Katzner, Philosophy Department, Advisor

Abigail Cloud, English Department, Advisor

Procreation is a strange philosophical puzzle, incorporating people who are and will be and will never come to be. There is much to consider when it comes to creating new humans—but most people fail to consider the painful cost of that human's existence. In this paper, I will discuss David Benatar's argument against procreation, beginning with the asymmetry of pleasure and pain. I will analyze this argument, use it to provide evidence in support of regulated parenting, and offer my thoughts on other actions we should take based on Benatar's findings, particularly regarding the ideal approach to procreation and the concerns of existent and non-existent persons. Ultimately, I will contend that Benatar is wrong about the harm of coming into existence being insurmountable, and that a life with enough good experiences can have an overall positive value, despite starting out with a negative value from the initial harm of existence.

David Benatar's book *Better Never to Have Been* discusses the ethics of procreation, defending the unpopular view that bringing a person into existence is a harm. Through seven chapters, Benatar addresses five major points: how coming into existence can be a harm and why it is always a harm to be brought into existence; how great a harm is done when a person is brought into existence; anti-natalism and why it is always wrong to have children; abortion from an anti-natal perspective; and goals regarding population and extinction in a world where procreation is immoral. Benatar's argument begins with an explanation of how we can be harmed by being brought into existence, and proceeds to the dramatic conclusion that bringing someone into existence is a wrong because it is always an irreparable harm to the individual in question—a harm that outweighs any benefits occurring in the individual's life. The remainder of the book discusses various implications of this conclusion, particularly those regarding existing persons and procreation. Benatar raises many intriguing questions throughout his book, but the

primary focus of this paper is the starting point of his argument: why bringing a person into existence is always a harm.

Benatar uses the asymmetry of pleasure and pain (hereafter “the asymmetry”) as the basis for his argument, claiming that those who would argue the benefits of existence outweigh its harms are mistaken: “There is a crucial difference between harms (such as pains) and benefits (such as pleasures) which entails that existence has no advantage over, but does have disadvantages relative to, non-existence” (31). This difference, as Benatar explains it, can be observed in commonly held ideas about pleasure and pain. It is uncontroversial that the presence of pain is viewed as bad and the presence of pleasure is viewed as good; the controversial aspect arises when we consider the absence of pain and pleasure.

Benatar argues that pleasure and pain must be viewed in terms of an individual’s interests. The absence of pain can be considered good, even if it’s not experienced by anyone, because it would be in the interest of any person (actual or potential) to avoid feeling pain—if a potential person did exist, the lack of pain would be good for her. The absence of pleasure, on the other hand, is only bad if there is someone who can be deprived of this pleasure; thus, if we consider a potential person (an individual who does not exist) the absence of pleasure is not bad for her, because she cannot be deprived of that pleasure by virtue of her nonexistence.<sup>1</sup> So we see that the absence of pain can be good for potentially existing persons in a way that the absence of pleasure cannot be bad for them, since the absence of pleasure is only bad in specific situations while the absence of pain is good in all cases. As aforementioned, this is the more controversial

---

<sup>1</sup> It should be noted, however, that if this person did exist and we withheld pleasure from her, it could be bad for her. So, we can see Benatar’s asymmetry is intended to apply primarily to cases of potential persons.

of the Benatar's claims, and he proceeds to defend it by grounding the assertion with four commonly held views that appear to rely on the asymmetry.

Benatar first cites the view that we have a duty to avoid bringing suffering people into existence, yet no corresponding duty to bring happy people into existence. This relates to the widely held belief that there is no moral obligation for humans to have children, but there is an obligation not to have a child if she will have a low quality of life. Benatar says we think there is a duty to avoid bringing suffering people into existence because, "the presence of this suffering would be bad (for the sufferers) and the absence of the suffering is good (even though there is nobody to enjoy the absence of suffering)" (33). On the other hand, we think there is no duty to bring happy people into existence; the absence of pleasure would not be bad for them (since no one would be deprived of that pleasure if they never exist) even though the presence of pleasure would be good for them if they did exist.

Benatar foresees a challenge to his claim that the asymmetry provides the best explanation for our belief that there is no obligation to procreate, which emerges if we consider a more limited set of duties. One could argue that we have negative duties (duties that require inaction) to avoid harm, but lack positive duties (duties that require action) to cause happiness, and this would explain our intuitions without using the asymmetry. In other words, it could be that morality only requires us to act in ways that minimize harm without also requiring that we act in ways that increase happiness. To give a mundane example, we would be obligated not to insult others, but not obligated to give them compliments. Benatar responds by conceding that, if positive duties do not exist, our beliefs about procreation being non-obligatory are not necessarily grounded in the asymmetry, and he is down one supporting example for his argument. He pushes back, however, arguing that most sets of duties involve both negative and

positive duties. Though accounts of positive duties rarely include the duty to bring happy people into existence, if there *were* a positive duty to bring happy people into existence, we would still need an explanation for why we aren't obligated to procreate whenever possible (since so many people fail to do this). Benatar suggests that the extent of personal sacrifice involved with a task is one thing that can exempt us from positive duties. If we apply this concept to procreation, it would mean that the pregnancy would have to constitute enough suffering on the woman's behalf to allow her to ignore this duty (and it may be the case that it does in some or all cases). Benatar argues, however, that even if this is the case we don't truly believe the toil of pregnancy is what gives us reason to reject the duty to have children, citing the fact that many would still think there would be no obligation to have children if there was no suffering or inconvenience involved. And if suffering isn't the reason why we can avoid positive duties to procreate, he argues, then we are brought back to the asymmetry, because we see that the interests of potential persons are what determine our moral reasons for procreation rather than the interests of existing persons.

To understand this point, consider a case where a woman and man are considering having a child. There is a drug that has been recently developed, which eliminates the symptoms of pregnancy entirely, there are no aches, no cravings, no nausea, and so on. In addition, the woman is assured she will be provided whatever clothes she needs to accommodate her body's changes throughout and after the pregnancy, and she will be given another drug at the birth that will make the birthing process painless as well. All of this is provided at no cost to her and her husband, and it is all guaranteed to work. Lastly, the man and woman know the child will have no birth defects, and that they are financially and emotionally capable of taking on the process of childrearing without having to sacrifice anything they might want to have in the future. Now our

couple are considering whether to have a child, and we can see there is nothing resembling suffering that can be cited as a reason why they should not go through with the process. In this case, does it seem like the man and woman are *obligated* to have a child? Benatar thinks not, and I agree; it seems coherent to argue that the couple would not be violating any moral duty even if they were entirely capable of raising a child, yet decided not to do so. Those who wish to deny the asymmetry's force in this scenario<sup>2</sup> would have to reply that this couple *are* obligated to have a child, however. This shows us the force of Benatar's reply: he thinks this idea of obligation deviates too far from the common view, whereas the asymmetry aligns with the common view perfectly, when we consider the interests of potential persons versus existing persons.

The second view Benatar discusses is that it is considered strange to have a child on the grounds that the child will be benefited, yet it is not considered strange to avoid having a child on the grounds that it is against the child's best interests to be born into unfavorable circumstances. He asserts, "If absent pleasures were bad irrespective of whether they were bad for anybody, then having children for their own sakes would not be odd" (35). For example, if someone said, "We're having a child because our child will receive free tuition in college if she goes to the college where I work" we would find it odd. Benatar claims this oddness stems from the asymmetry, since we can see how its pain-avoiding variant is not odd at all, as when someone says, "We're not having a child because we won't be able to afford to send her to college." Examples like this demonstrate why "The welfare of potential children is a sound basis for deciding *not* to have that child" (Benatar 35), and this would not be true unless the absence of pain was good for anybody, even non-existing persons. The point he illustrates here is that we

---

<sup>2</sup> Assuming they deny this by claiming a positive duty relating to procreation

cite the potential child's interests as the relevant consideration in cases of potential pain while overlooking them in cases of potential pleasure, which supports the asymmetry.

Benatar's third claim is that we can regret bringing people into existence for the sake of a person who exists in pain because of our decisions, but we cannot regret that a never-existent person never experiences a good (because, as this person never exists, she can't be deprived of this good). "Remorse about not having children is remorse for ourselves..." Benatar asserts, "However, we do regret having brought into existence a child with an unhappy life, and we regret it for the child's sake, even if also for our own sakes" (35-36). To understand this case, we can envision a scenario where a child is born to a couple who are unable to provide for it. This couple were barely able to feed themselves before and, though they try their best to provide for their child, in the end all three are malnourished and sickly. These parents would likely regret having this child—certainly for their sake, as they suffered more sleepless nights and empty dinner plates because of her—but also for her sake, since they're unable to provide her with the things she needs. It would make sense for one of the parents to say, "I'm sorry we had this child, she leads such a miserable life," and we would likely feel sympathy for the plight of this unfortunate family. On the other hand, we can envision a couple who are in an ideal childrearing situation for much of their lives, who eventually grow old and come to regret that they had no children. It would make sense for one of them to say, "I'm sorry we never had a child, I would have liked to raise a family," but it doesn't seem to make sense for one of them to say, "I'm sorry we never had a child, she would have had such a wonderful education." This latter sentence seems odd because it feels less like the primary reason for regretting having a child and more like a secondary regret of an individual whose primary regret is not having children for her own sake. Rather than thinking of what could benefit the children we may have had for their sake, we

think of what could harm them for their sake, and Benatar suggests this discrepancy is best explained by the asymmetry.

Fourthly, Benatar considers our reactions to stories about places distant from us, and what this says about our views. Oftentimes we are saddened to hear that distant people are suffering, from war, natural disaster, or some other tragedy; in contrast, we are rarely saddened to hear about idyllic, uninhabited areas of the world. Benatar uses this example to demonstrate that we regret suffering, but don't regret unactualized pleasure, since we should be sad to hear of uninhabited places where people could live pleasurable lives if the absence of pleasure was bad, but we don't tend to have any strong feelings about these places at all.

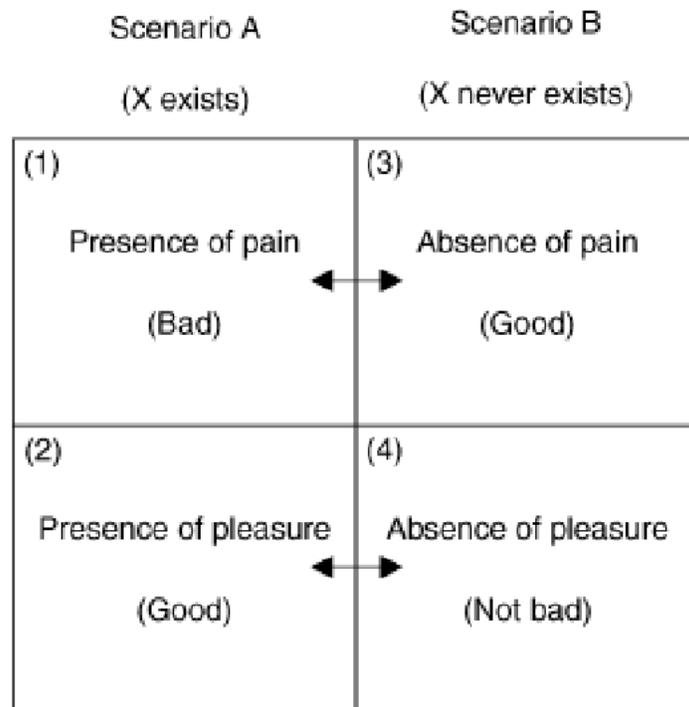
Benatar anticipates an objection to his own argument that uses this same example, which emphasizes that we don't feel joy for the pain avoided by non-existent persons. There are so few actual persons relative to the number of potential persons that the amount of pain avoided should be utterly joyous to us if the absence of pain is truly as good as Benatar claims, yet we don't feel anything in particular about non-existent persons. Benatar responds that this parallel fails because joy is not the proper opposite of regret. Melancholy is a better opposite of joy; however, melancholy suggests being overwhelmed with sadness, and we are rarely overwhelmed by news of any people distant to and incapable of affecting us. So, he argues, we are not made joyful or melancholic by distant people, but we *can* regret their situation or find it "welcome" or "good" (37). Benatar goes on to suggest that if absent pain is "the opposite of regrettable," then we must accept that absent suffering is a good part of not bringing someone into existence. Our attitudes toward distant places show that we prefer them to be absent of pain and view this as good, yet the absence of pleasure doesn't register as something particularly good or bad. The asymmetry

can provide an explanation for this discrepancy as well, giving Benatar his last bit of ammunition in support of the asymmetry's existence.

Before ending with his main argument, Benatar acknowledges that these four claims do nothing more than express the broad appeal of the asymmetry. He has not given evidence for its truth, but has merely expressed the scope of its appeal. Benatar suggests that the asymmetry's existence as a grounding force for so many widely accepted claims should give us reason to believe in the force of his argument: that so many other commonly held beliefs can be explained with the asymmetry his argument hinges on makes it a strong starting point for his main argument—that it is always a harm to bring someone into existence. He also acknowledges that the judgments spoken of in the four claims above are not universally shared, and demonstrates this by entertaining a positive utilitarian objection to his claims.

According to positive utilitarianism, the morally right action in all situations is the action that maximizes pleasure and minimizes pain. This means the positive utilitarian can argue there is always a duty to create more people if it will bring more happiness into the world, because this will lead to maximal pleasure. Benatar responds that people care about their own happiness, meaning happiness is valuable because it is good for people, rather than people being valuable only because they are “vessels of happiness.” What he means is that we shouldn't be looking to maximize happiness for its own sake, but for the sake of people: it's not as though people are irrelevant to happiness—happiness is supposed to be a good thing because it is good for those who experience it since it causes pleasure. Thus, Benatar claims the emphasis should be placed on people rather than an abstract “happiness value,” since happiness can only be good in virtue of how good it is for those experiencing it.

This section of the book concludes with the argument that drives Benatar’s thesis, demonstrating how the asymmetry he’s fought for thus far demonstrates why it is always a harm to come into existence. He uses a square chart containing four cases involving pleasure and pain, and a person, X, who either exists or does not exist, to convey this point. The cases where X exists are simple, as it is uncontroversial that the overall state is bad when X exists and there is pain, and good when X exists and there is pleasure; the cases where X does not exist are where Benatar’s argument unfolds. According to the asymmetry, where X does not exist and there is no pain the overall state is good, but where X does not exist and there is no pleasure the overall state of things is “not bad” instead of the symmetrical “bad.”



3

Benatar addresses multiple ways one might try to preserve the symmetry thereafter, and concludes that each one fails. Trying to preserve the symmetry by calling the absence of pleasure

<sup>3</sup> Figure 2.1 from Chapter 2 of *Better Never to Have Been*

“bad” makes the claim too strong, since this means we’d be obligated to regret that X never existed for X’s sake, though this is not regrettable according to our intuitions (as seen in the four claims above). Alternatively, if we try to preserve symmetry by calling the absence of pain “not bad” as well, we see one of two results. The first is that “not good” and “not bad” can be interpreted as simply equivalent to their opposite terms (“not good” is bad and “not bad” is good); if that’s the case the previous argument applies, since we already established that “bad” is too strong a term to describe the absence of pleasure. The second is that “not bad” can include the added stipulation of being “not good” as well (so “neither good nor bad”), but that makes the label too weak to capture how good it is to avoid the pains of human existence.<sup>4</sup> This brings out our need for the asymmetry again, because we cannot change a non-existent person’s failure to experience pleasure into something bad without forfeiting the previously discussed views, but if we say that a non-existent person being robbed of pleasure is anything less than bad, we are acknowledging and accepting the asymmetry.

Now that we’ve examined Benatar’s argument for the asymmetry of pleasure and pain, we can better consider its strengths and weaknesses. I will discuss the problem of basing an argument off common intuition, a concern about internal bias in the four claims Benatar uses to support his argument, and a similar issue in Benatar’s final claim that the depth of pain in human existence outweighs any good that comes from it. By examining the problems with Benatar’s argument here, I will build the foundation for my own argument in the final section of this paper.

---

<sup>4</sup> The depth of these pains is discussed in greater detail in Chapter 3, but to understand Benatar’s argument one must know that Benatar believes them to be bad enough that no good in a person’s existence can compensate for their harm.

To simplify matters, I will begin by restating Benatar's argument in premise-conclusion form.

- 1) Uncontroversially, the presence of pain is bad and the presence of pleasure is good.  
(From common perception)
- 2) There is an asymmetry of pain and pleasure, such that the presence of pain is bad and the absence of pain is good in all circumstances, whereas the presence of pleasure is good and the absence of pleasure is bad only in circumstances involving currently existing persons. (From the four asymmetries and common perception)
- 3) The absence of pain is good even when there is no one to enjoy that good. (From P2)
- 4) The absence of pleasure is not bad in circumstances where no one is deprived of that pleasure's benefit. (From P2)
- 5) When considering non-existent individuals, it is good that they experience no pain and merely not bad that they experience no pleasure. (From P3 & P4)
- 6) Non-existent individuals have a net positive "goodness value."<sup>5</sup> (From P5)
- 7) No amount of pleasure an individual experiences in life can outweigh the harm of being brought into existence. (From Benatar's Harm of Being Brought into Existence Argument<sup>6</sup>)
- 8) Any individual who exists will have a net negative "goodness value." (From P7)
- 9) It is better, if we aim to maximize the good, not to bring individuals into existence. (From P6 & P8)

---

<sup>5</sup> By "goodness value," I mean whether the good outweighs the bad in a life (net positive goodness value) or the bad outweighs the good in a life (net negative goodness value).

<sup>6</sup> See *Better Never to Have Been* Chapter 3

As you can see, much of this argument stems from the controversial premises 2 and 7; thus, these will be the premises I focus on in my critique.

The first thing to discuss about premise 2 is that it appeals to widely held beliefs for its basis. The strategy of appealing to common intuition is a frequently used tactic, often referred to as appeal to belief, and it is well known for being a logical fallacy. While Benatar's use of these claims is honest, and he does not overstate the strength of appeal to belief, we must consider this form of fallacy further, before we continue.

Appeals to belief take the form:

- A) Many people believe X
- B) Therefore, X is true

This reasoning is fallacious because, as has been demonstrated countless times throughout human experience, the majority can be wrong. To give a familiar example, mere centuries ago, the belief that the Earth was flat was widely accepted as true, yet now it is a widely accepted belief that the Earth is round. This and similar cases serve to demonstrate that the belief of the many does not always align with the truth.

In the case of Benatar's argument, he never claims that the views he discusses are true, thus his argument does not fall into the realm of fallacy.<sup>7</sup> Instead, he explicitly states that he has only expressed the scope of asymmetry's appeal using his four examples. Given what I've just stated about appeals to belief, however, I would like to consider Benatar's claim about the broad appeal of his argument, and the effectiveness of that reasoning.

---

<sup>7</sup> In addition, some individuals would argue that appeals to belief can be effective in moral discourse, because we have little more than our intuitions to guide us.

The four asymmetries Benatar discusses involve intuitions about bringing people into existence, reasons for having children, reasons to regret having brought a person into existence, and reactions to knowledge of distant places and persons. Benatar demonstrates how we can use the asymmetry of pleasure and pain as a means of grounding these beliefs, providing explanations for why we respond to the potential for or presence of pain while ignoring the potential for or absence of pleasure in these situations. These intuitions are widely occurring, making his argument compelling; however, Benatar neglects to properly consider that these asymmetrical beliefs could be misguided.

Consider this: Humans have a difficult time looking at their lives from an outside perspective. In the chapter following the one discussed in this paper, Benatar considers a phenomenon called Pollyannaism—a psychological tendency humans possess that makes them view their lives with a bias of optimism—and how this might create a discrepancy between the actual and perceived quality of our lives. We will return to this idea later, when discussing premise 7, but I reference it here to illustrate an important point: It seems humans are poor at judging quality of life from an internal perspective. If this is the case, then revisiting Benatar's asymmetries prompts an interesting question—are our asymmetrical intuitions a result of this biased, internally limited perspective?

To illustrate this point, let us consider the second view, where Benatar examines the reasons individuals might pursue or avoid childbirth. He claims that we cite potential suffering as a reason to avoid bringing a person into existence, yet we do not cite potential pleasure as a reason to pursue bringing a person into existence. If we consult the pain-pleasure asymmetry, we observe that this is because we don't see the loss of potential pleasure as bad, but we do see the loss of potential pain as good. I wish to suggest that perhaps we don't see the loss of potential

pleasure as a bad thing because we cannot examine life from an external perspective. Being internally biased, we consider our perceptions of our happiness to be relevant, but this may not be the case; perhaps our lives can improve even without us perceiving them as being better. As an example, consider an individual. In high school, this individual had people who pretended to be her friends but secretly didn't care for her at all; after going to college, the same individual acquired new, true friends. This individual can't perceive the difference between the false friends from high school and the true friends from college, yet it seems possible that she could be said to be living a better life when she acquires true friends, even though she can't perceive that improvement. In this same way, perhaps an objective perspective is needed to determine the quality of a person's life.

Benatar responds to this objection in utilitarian terms, claiming that this claim suggests humans are nothing more than vessels of happiness. He objects to the idea that people are a mere means to increase the production of happiness, claiming that our experiences of happiness matter, since happiness should be good for us. The trouble is that this belief stems from an internal bias in favor of our own thoughts and emotions, that cannot expand to discuss the potential thoughts and emotions of non-existent people. If morality is an ultimate goal of the objective universe<sup>8</sup>, then it seems perfectly plausible that human thoughts and emotions don't impact this goal: rather, morality is only a matter of the sum of goodness being greater than the sum of badness.

As Benatar bases a large portion of his final argument off this internal bias, it seems odd that he neglects it when developing his premises. If individuals' personal experiences of

---

<sup>8</sup> The discussion about whether morality is objective or subjective is a matter for another paper; however, it is worth noting that a subjectivist about morality can avoid this difficulty entirely.

happiness are relevant when considering how to maximize happiness, yet aren't relevant when considering the value of their lives after they are brought into existence, it is important that we understand the relevant distinction between these two scenarios. As Benatar's argument currently stands, this distinction is notably absent, yet all four asymmetries Benatar references can be called into question because of internal bias, making the need for this distinction crucial to the development of his argument. For instance, our asymmetrical beliefs about procreation could be a result of an internally biased belief that childrearing requires a good amount of self-sacrifice, which could detract from the parents' happiness. It may be true that, from an objective standpoint, more happiness would enter the parents' lives and the world as a whole if the child were born, but the parents cannot see this, because they are caught in their own subjective worlds as all humans inevitably must be.

Now that we've addressed the preliminary difficulties with premise 2 and the existence of pain-pleasure asymmetry, we must also consider premise 7, why it is always irreparably bad to bring a person into existence. Assuming the truth of the pain-pleasure asymmetry leads us to the conclusion that all non-existent individuals have a net positive "goodness value," because they represent pain and pleasure that will never be experienced; however, Benatar cannot arrive at his claim that it is better to avoid bringing individuals into existence without a second claim, stating that those individuals who are brought into existence have a net negative "goodness value." He makes this claim in the section of Chapter 2 immediately after the section we've been discussing: "Because there is nothing bad about never coming into existence, but there is something bad about coming into existence, it seems that all things considered non-existence is preferable" (45).

Considering this claim at face value, it appears to hold some clout. Benatar is arguing that, because we can compare an existent person's life to a non-existent person's life, we can see

that the non-existent person has greater “goodness value” than the person who does exist.<sup>9</sup> The immediate response is that there are many other goods that will happen to the existent person, so many that it seems like the good made possible by existence could compensate for the harm of being brought into existence, and even extend beyond that point to create a greater net positive than the non-existent person’s value. For example, if a person discovered the cure to cancer, giving us the ability to save countless lives, it seems this person may have created enough good to counteract the harm of being brought into existence. Perhaps leading a fulfilling life with a successful career, loving family, engaging hobbies, and enriching friendships is enough to even out the scales if that’s all the individual needs to be happiest.<sup>10</sup>

Benatar responds to this concern with the claim that the goods of existing persons and the goods of non-existing persons are of two different categories that are incomparable: Any good that happens to a person after they come into existence is a lesser sort of good, because it would be unnecessary if that person hadn’t existed to begin with. He uses an analogy featuring two individuals, one sick, one healthy, to prove his point. If the sick person has the ability to recover from illness quickly and the healthy person does not have this ability, the sick person is not better off than the healthy person; the healthy person doesn’t need the ability she’s lacking, so it is neither a benefit nor a harm to her to have a quick illness-recovery speed. Similarly, Benatar claims, a person who never exists doesn’t need the goods that can be gained through existence because she already possesses the preferable state of being—nonexistence (43). Some might be unsatisfied with the claim that non-existent good is better than existent good, however, and this

---

<sup>9</sup> This assumes nothing about the goodness/badness contents of the existent person’s life.

<sup>10</sup> Whether the good in a person’s life must be good for the individual herself or can be good created by that individual for others is a further consideration that goes beyond the scope of this paper (as is the discussion of how to determine specific values for good and bad acts). For the purposes of this argument, I will assume that the good of a person’s life can counterbalance the harm of being brought into existence regardless of whether it is good for her, or good for others.

is where we return to the earlier discussion of Pollyannaism and internal bias. Benatar uses these concepts to support his claim that existence is exceptionally bad, worse than we're capable of perceiving from within its limits, but this will again prove to be a weak point for his argument.

Based on what we've accepted about the pain-pleasure asymmetry prior to this point, we've agreed that individuals' concerns about their happiness are irrelevant when considering whether good-from-existing or good-from-not-existing is better. If we did not ignore individual concerns, the good from existence and non-existence would be the same type of good, and Benatar has stated these are different kinds of goods. The trouble comes from Benatar's use of an external reference point for the good: he takes human perception out of the morality equation for his own purposes—exactly what he criticized the utilitarians for before. True, humans are no longer “vessels of happiness” under his view, but they have become “vessels of good” instead, and a non-existent person is a better vessel of good than an existent one, so non-existent persons are preferable. Furthermore, the idea that the good of a non-existent person is better than any goods of an existent person suggests that a lack of perceptions, thoughts, and opinions is preferable to the presence of these mental processes, which further suggests that any persons and potential persons are only useful insofar as they contribute good to the universe.

So, premise 7 brings Benatar back to the problem of vessels—if the good being created is not good for anyone (and it will not be good for anyone, because the potential persons it is allegedly good for will never exist), then humans are reduced to vessels of the good, rather than the good being something that benefits humans. Benatar may respond that the good is different from happiness, and human thought and perception can only influence the latter, not the former; however, this will cause his objection to utilitarianism to fail, as—according to utilitarianism—

happiness *is* the good. It is clear, then, that Benatar needs to revisit his criticism of utilitarian duties, and possibly re-evaluate the implications of his claims.

Both of Benatar's controversial premises encounter problems with examining morality from an internal versus an external perspective. His second premise relies on four asymmetries that emerge from internal bias, and he does not adequately address the concern that these asymmetries may be illusory—a result of our inability to escape internal perception. His seventh premise suggests that the good of a non-existent human is preferable to the good of an existent human, raising Benatar's objection to utilitarianism back against him, by implying that human thoughts and emotions about the good are not important, and should be eliminated in favor of achieving maximal good in the universe. While Benatar's argument appears strong, it is clear there are some holes in his claims that should be patched up before we agree to accept them. In the next section, I will address these issues and suggest an alternative way we can utilize Benatar's observations.

Up to this point, Benatar has expressed conflicting perspectives on whether good and bad are determined by what is good or bad for individuals. He claims that individual opinion matters when determining how to achieve maximal happiness, but also that a universe without individual opinion is ideal, because there is more good in a universe that lacks individuals (and, therefore, pain) entirely. In the following section, I will argue that Benatar's view entails the conclusion that a universe without moral agents is ideal, and demonstrate why this cannot be the case because of the role moral agents play in determining morality. I will then conclude by adjusting Benatar's argument to a milder form that does not result in a goal of extinction but still limits the number of persons harmed by coming into existence.

To begin our discussion, we must familiarize ourselves with moral agents. Moral agents are those individuals who possess an awareness of right and wrong, and the ability to act using this awareness as a reference, to determine the goodness or badness of their actions. We commonly understand moral agents to include most humans (with the exception of those whose minds are not fully developed or possess some other cognitive deficiency<sup>11</sup>), but there is continued controversy about what other beings—if any—should also be considered moral agents. That debate is beyond the scope of this paper, however, so I will only assume that there are some individuals in existence who can be considered moral agents: these individuals are the relevant subjects in our discussion. With an understanding of right and wrong and the ability to act freely, moral agents can sway the balance of good and bad on the moral spectrum. It is commonly thought that an ideal world would have the highest good ranking possible on the moral spectrum, meaning moral agents would perform right actions more often than the wrong actions, if not always.

Now, let us consider how morality functions according to Benatar's argument. Benatar claims that the ideal world possesses no pain, implying (uncontroversially) that pain is something bad. Because all humans who exist are made subject to pain by being brought into existence, it is a harm to bring a person into existence. The harm comes from allowing bad things to happen to that person, when the same bad things could not happen to a non-existent person. The moral wrongness, then, according to Benatar, comes from moral agents (in this case, humans) choosing to procreate and bring other humans into existence, because this act subjects the newly existing individuals to pain.

---

<sup>11</sup> It is worth noting that sociopaths hold an interesting position on this scale. For my purposes, I will assert that they are not moral agents, as they are incapable of telling right from wrong, however, this view is not universally accepted.

I am sympathetic to Benatar's view that bringing persons into existence is always a harm, for reasons we have touched on before. It is true that, if a person never exists, that person cannot experience pain; in that sense, it is surely a harm to invite pain onto another being that they wouldn't have experienced otherwise. However, despite this, I do not share his belief that we should cease procreation altogether; rather, I think this provides us with a compelling reason to believe procreation is something that *should* be participated in, albeit sparingly. I limit Benatar's claim in this way because I believe his argument, which suggests we pursue the ultimate end of humanity, leads to a conclusion that is incompatible with morality as we currently understand it.

Let us assume, for the sake of argument, that Benatar is correct in his hypothesis and the best world is the world in which humans do not exist. This world is ideal because it dramatically reduces the level of pain in the world by eliminating the presence of suffering humans, and removing entirely the person-to-person harm of bring new people into existence. Still, one might wonder, isn't it true that other animals would still suffer pain, even if humanity no longer existed? It is unclear whether Benatar would concede to this point or not, and it is a matter of heated debate which animals are capable of feeling pain to begin with; however, it is clear that Benatar's argument relies on the existence of pain as something bad, and his argument should be extended to any beings that can be considered moral agents, since these beings should be able to recognize the harm of bringing others into existence (even non-humans). I will therefore revise Benatar's claim to state that an ideal world is one without moral agents, because it is a harm to bring any moral agent into existence. Thus, to prevent this harm from occurring, moral agents should be obligated not to procreate.

Now we find ourselves faced with an interesting thought experiment. What would morality look like in a world without moral agents? It must be true that morality either persists or

vanishes without the presence of moral agents, and Benatar suggests that morality persists, as he claims the lack of harm occurring to never-existing persons is a positive on the moral scale. This suggests that morality exists on some scale that extends beyond the scope of persons who can influence it (moral agents). True, Benatar claims that the good he speaks of is only good insofar as it is “good for” those persons who may have existed, but this conclusion seems impossible without any moral agents to evaluate it. If there is no existing agent to determine what is good or bad, who is there to claim that something is good or bad for anything? The only alternative is some sort of objective moral judge—an ideal observer<sup>12</sup>—existing outside of experience itself, who can tally the universal values of good and bad even without the existence of any individual who can change these scores. The notion of an ideal observer is somewhat questionable, but we will accept this idea so we may posit that morality can persist beyond the existence of moral agents (who judge morality by their own standards).

Imagine morality as a scorecard, with points on opposing sides for good and bad. The only way to alter those point values is to take actions (or inactions), which will lead to good or bad results, that are added to the scorecard. If this is how morality functions, then the scorekeeper is only necessary if there is a way to change the score. However, if we follow Benatar’s view, we come to a problem: When we have achieved agent extinction there will be no moral agents to perform good or bad actions. Yet, if no one is performing good or bad actions there is no way for the score to change, unless there are good and bad things in the universe that can be created or destroyed without the influence of moral agents, and are not good or bad for anyone at all.

---

<sup>12</sup> An ideal observer theory is one way to view impartiality in relation to morality that provides a useful tool for discussing morality without agents. As Benatar’s claims have had utilitarian leanings thus far, I will assume our observer follows a utilitarian moral scale as well.

Benatar, in claiming that a lack of pain is good for potential persons (31-32), demonstrates his belief that good and bad must be viewed in terms of some subject, therefore we can eliminate the existence of these potential “other influences” that would not be good or bad for anyone. Thus, we are left with a moral scorecard that cannot be altered meaningfully absent moral agents. These values will remain the same indefinitely, so our universal moral judge has no further purpose—the extinction of moral agents cements the moral judgment of the universe.<sup>13</sup> One might argue this is how it should be: The extinction of moral agents is inevitable, after all, as the universe’s end will entail an ending of all that morality as well. When this happens, if there is a universal moral judge, she will take her final tally, determine whether the universe had more total good or bad value, and have completed her task. This hardly seems impossible, and may even seem likely to some readers.

So now we have a way for morality to exist even absent moral agents, and it seems that Benatar’s view can still hold. He is suggesting that we escalate the process of eliminating moral agents to achieve a better overall moral score, absent the harm values brought on by further procreation by moral agents. Since each new person brought into existence is a negative mark on the moral scorecard, stopping procreation is an excellent way to make fewer moral negatives to mar our universe. Slowly the insurmountable negatives would disappear, then the moral score would come to a stop with the end of moral agents, ideally higher than it would have been if agents had continued to procreate despite the harm it caused. Yet there is something undeniably

---

<sup>13</sup> Some might argue that there are still other things that impact morality. Some non-agent animals may still exist and they can hurt each other or live idyllic lives and contribute to the moral count. My response is that this does, in fact, change the moral scale; however, no non-agent will be judged for its actions because it cannot understand right and wrong. Agents can be judged, and thereby influence the scorecard in a meaningful way. Non-agents, on the other hand, may change the values, but they have zero effect on the ultimate moral ruling because these rights and wrongs are coincidence, not deliberate actions to promote good or ill.

strange about the way this addition works, which we can see if we examine the world at its beginnings.

Prior to the existence of moral agents, the universe was filled with objects and beings without moral awareness, incapable of committing harms by procreating because they were unable to comprehend good and bad. Returning to this state would be an ideal universe for Benatar, but note that, by his own calculations, our universe can likely never return to as good of a state as it was in at that time, since its moral value has steadily depreciated as moral agents have procreated again and again through the years. Interestingly, the first moral agent to come into existence was possibly also the only moral agent who came into existence without any wrong being done to her, since she would have been born to non-agent parents (assuming that morality is an ability gained by humanity over time), but from that point on procreative acts began to acquire more negative value, as more and more agents ignored the harm caused and participated in acts of procreation. In this sense, we can begin to see the aforementioned oddity—if the harm of being brought into existence can never be overcome or even balanced out by the good of existence, then the very emergence of moral agents was a bad development from the start.

Now we have learned that each moral agent is a negative mark on the scorecard of morality. Yet, moral agents are the only beings who can change the moral score, as mentioned before. This makes morality a game we always have to lose: Prior to the existence of moral agents, the score is zero. After the existence of moral agents has ceased, it is necessarily negative. Here is where I can no longer accept the implications of Benatar's view: I find his assertion that we can only possess negative moral worth, individually and collectively utterly unreasonable.

Morality is easily understood in terms of balance, as illustrated through the scorecard analogy. It is a game of numbers between the good and the bad, and ultimately one will win out over the other when no other moral or immoral actions can occur. Benatar's view suggests that this will *always* be the bad, because the positives in the lives of moral agents can never outweigh the negatives, and there is no good or bad except that which is good or bad for someone (some moral agent). It seems to me that, if there is a reason for us to desire something about morality, we should desire to achieve a morally good universe. We want to act morally—to do good—because that will influence the moral scorecard and we will move closer to the fully good universe we desire. However, if we accept Benatar's view, we are fighting a losing battle against negative moral values, and no matter how much good we bring to the world we'll never even achieve neutrality, let alone achieve a good universe.

My primary reason for finding this view of morality not simply bleak, but entirely unreasonable, is that it seems to nullify the entire purpose of adhering to morality to begin with. Those who believe in morality believe we are obligated to follow the rules of morality, and obligations suggest a motivating factor—usually some form debt repayment—involved with performing or failing to perform certain actions. Assuming this is true, and that morality exists regardless of moral agents' opinions about morality, then we have an obligation to create a morally good universe, yet we can never repay this obligation<sup>14</sup> because it is impossible for us to elevate our universe's morality to a positive value. And if moral agents are unable to ever repay

---

<sup>14</sup> Although I have framed this argument in terms of debt repayment, it applies across the board to whatever can be filled in for an agent's reason to act morally. Since an agent will always have negative moral value, simply by existing, she has no reason to try to act morally, since she can never contribute any positive moral value to the universal score, which is her objective.

the debt they owe, then it is difficult to accept that we should keep trying, since failure is guaranteed.<sup>15</sup>

A word regarding some potential concerns about this claim. First, some may worry that I have failed to adapt the moral system properly—rather than viewing moral values as existing solely on the negative side of the moral spectrum, I should adjust the scale according to the actual abilities of moral agents. With these alterations, small negative moral values would be equivalent to positive moral values on the revised scale, and large negative moral values would be equivalent to negative moral values on the revised scale. I respond to this proposition with a concern of my own: To adjust the scale like this would alter the good/bad divide in a strange way. When the scale is positioned normally, there is a definite divide between morally positive and morally negative values—a zero point that could be considered moral neutrality, where the good and bad in a life are equivalent. If we shift the scale to the negative side of the moral spectrum, however, there is no such zero point to divide between the good and the bad. Instead, we would have to decide arbitrarily on a point where good shifts to bad, and it wouldn't be neutral morality but a middling negative value. As deciding these values would be nearly impossible, shifting the scale is not an adequate solution to this problem.

A second concern is the thought that morality may not be objective at all, but is affected by moral agents' opinions. There has been extensive debate on the nature of morality as subjective or objective, and it is beyond the scope of this paper; however, I would like to acknowledge this possibility and remind readers that I am attempting to stay within Benatar's parameters as I formulate my argument. Given the nature of Benatar's argument, and the fact

---

<sup>15</sup> There is extensive debate about why we should be moral, even in situations where it seems meaningless, but that is beyond the scope of this paper.

that an agent's opinion does not determine what is good for her (since not existing is good for an agent), Benatar must be speaking of an objective system of good and bad, so this is the system of morality I'm assuming for my argument as well.

Finally, I may receive some pushback from those who would remind me that sometimes the most moral thing to do is minimize the harm. While this is applicable to situation-specific scenarios (such as when one is faced with a dilemma of saving one or five people)—this is a discussion of morality on its broadest scale. It seems wrong that we could never achieve an overall positive moral value if morality started at neutral and moral agents began to take positive and negative actions from that point on. I will grant that it's possible moral agents have committed so many negative actions by this point that they are impossible to surmount, but, initially, I would argue that the moral score must have been zero, so moral agents would have a chance to attain and maintain a positive moral status. There is much more to be said on this topic, and an extensive debate about the nature of morality and its origins, but to speak on this further would be to stray from this paper's topic, so now I will return to Benatar's view.

Thus far, I have asserted that an objective, universal morality that can never achieve positive value is not a functional system of morality. From this claim, I further conclude that Benatar's view that the harm of being brought into existence is insurmountable is false, since this is the only way for moral agents to have a chance at tipping the moral scales toward a positive value. Instead, I would argue Benatar has demonstrated that there is a great harm committed by bringing moral agents into existence, but also that this harm can be offset by the amount of good brought into the universe by an individual's life. Benatar's argument demonstrating the net positive value of a non-existent person shows us how bringing a person into existence only invites harm upon her, which she would have avoided had she never existed. However, Benatar

does not provide a compelling argument that the goods for existent and non-existence persons are fundamentally different, which leaves room for a variation on his thesis. The variation I propose asserts that procreation among moral agents should occur sparingly, only in cases where it appears that the offspring will be able to experience or cause enough pleasure and thereby create enough good to be valuable to moral calculations.<sup>16</sup>

I'm sure many readers will have noticed I am falling into the same trap as Benatar before me, suggesting that moral agents are nothing more than vessels for morality, existing only to push the number closer to a positive value. I admit this is true: I believe an objective morality *is* that demanding of its moral agents. An objective morality must exist mind-independently, without the need for any agents at all; however, agents are the only beings that can affect morality, and they are obligated to take positive moral actions (as previously established). This obligation makes them beholden to morality, required to increase its value as much as possible. Any utilitarian or moral "scoring" theory will arrive at this same result—we cannot avoid it while discussing the asymmetry of pain and pleasure, because this utilizes greater and lesser values of goodness and badness.

Before closing, there are two final concerns with my view I wish to address: The first is that "sparingly" is a vague term for a restriction, and something so strict as morality should have more clearly defined restrictions; the second is that we cannot tell how much good an individual will experience or create in her life, and are thus incapable of determining when procreation is appropriate. These two concerns are closely related, and of notable consequence to this view.

---

<sup>16</sup> This implication stems from the acceptance of the harm that comes from bringing a person into an existence: it harms the individual because bad things will happen to her in existence but they could not have if the individual never existed.

To answer the first concern, it is impossible to give a specific value for procreating “sparingly” because personal situations are so different. One well-off, competent person could have many children, or we could discover that single-child families are ideal for middle-class persons because there is less strain on the adults and the child is not competing with any siblings.<sup>17</sup> More research must undoubtedly be performed to give us a better idea of the answer to this question, but we can say a few things about it. First, this tells us there is no obligation to have children; second, we know that “because you want to” is not sufficient grounds for procreation; and third, bringing someone into existence always harms them, so the being’s creators have some obligation to compensate their creation for the harm done (at the very least, seeing that she is looked after until she becomes a moral agent in her own right).

The second concern is similarly tricky, and evades an easy explanation, so I will respond by saying we must rely on conditions we have observed in the past to predict future patterns. Trapped as we are in our own place in time, we cannot foresee whether a being created by a moral agent will have enough good in her life to overcome the harm of her existence; however, predictions can be made based on individual family situations, as I mentioned previously, which can help guide us to whether procreation is a viable option for a couple or not. This would make the procreation process highly regimented, which would not only be difficult but also costly and complicated, but if it would truly improve the lives of all existing individuals—which it should,

---

<sup>17</sup> An extreme version of this view could suggest that it might lead us to a socio-economic cleansing of sorts. I don’t intend to suggest such a vicious goal, but rather to emphasize that all manner of environmental, social, economic, and other factors must be accounted for when assessing who is qualified to procreate. A potential solution to this issue is to develop a course for those who are interested in becoming parents. Anyone interested in having children could take this course to prepare for the task, thus ensuring they were in the best possible place to start raising a child, and the course would be available to the public.

as it would eliminate many cases of individuals who experience more pain than pleasure in their lives—then it would be a worthwhile process to adopt.

In this section I have discussed the nature of morality in a world without moral agents, demonstrated how Benatar's view forces him to the undesirable conclusion that it is impossible to create a world with positive moral value, and offered my own view—that the harm of coming into existence can be overcome if the individual in question lives a life with enough positive experiences to counterbalance the negatives—using Benatar's ideas as a basis. Benatar makes several important points throughout his book, encouraging readers to reevaluate their biased perception of existence as persons participating in that perception already. Though his arguments have their flaws, they prompt an important, necessary discussion about the future of humanity and procreation, which I hope to have contributed to through my arguments above.

## Bibliography

- Benatar, D. (2006). *Better never to have been: The harm of coming into existence*. Oxford: Oxford University Press.
- Bradley, B. (2012). Asymmetries in benefitting, harming, and creating. *The Journal of Ethics*, 17(1), 37-49. <https://doi.org/10.1007/s10892-012-9134-6>
- Gardner, M. (2015). Beneficence and procreation. *Philosophical Studies*, 173(2), 321-336. <https://doi.org/10.1007/s11098-015-0494-1>
- Jollimore, T. & Zalta E. (ed.). (2017). Impartiality. *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/impartiality/>
- Mbugua, K. (2013). Is non-procreation the solution to human suffering? *Eubios: Journal of Asian and International Bioethics*. 23(3), 77-80.
- McLean, B. (2014). What's so good about non-existence? An alternative explanation of four asymmetrical value judgements. *Journal of Value Inquiry*, 49(1-2), 81-94. <https://doi.org/10.1007/s10790-014-9455-8>
- Parfit, D. (1984). *Reasons and persons*. New York: Oxford University Press.
- Singh, A. (2012). Furthering the case for anti-natalism: Seana Shiffrin and the limits of permissible harm. *South African Journal of Philosophy*. 31(1), 104-116.
- Smuts, A. (2013). To be or never to have been: Anti-natalism and a life worth living. *Ethical Theory and Moral Practice: An International Forum*. 17(4), 711-729. <https://doi.org/10.1007/s10677-013-9461-0>