

This is a pre-print of paper in *Bioethics*. Please cite the published version only.

Future Minds are Not a Challenge to Anti-Natalism: A Reply to Gould

Abstract: Deke Caiñas Gould (2021) argues that the possibility of future non-human-like minds who aren't harmed by coming into existence poses a challenge to David Benatar's well-known Asymmetry Argument for anti-natalism. Since the good of these future minds has the potential to outweigh the current harms of human existence, they can be appealed to in order to justify procreation. I argue that Gould's argument rests on a fundamental misunderstanding of Benatar's argument. According to the Asymmetry Argument, if a person experiences *any* harm at all, then bringing them into existence is unjustified. It does not depend upon on-balance judgments about the relative harms and benefits of existence. It therefore remains impermissible to procreate *right now in our world*, regardless of the prospect of future humans contributing to the successful development of beings who are not harmed by existence. I conclude by addressing two alternate readings of Gould, which for the sake of argument permit such on-balance judgments, and show why they fail to rescue his case. Benatar's Asymmetry Argument might be unsound, but not for any reason identified by Gould.

Keywords: Anti-Natalism; The Asymmetry Argument; Benatar; Futurism; Gould; On-Balance Judgments

1. Introduction

In his recent article, “Future minds and a new challenge to anti-natalism,” Deke Caiñas Gould (2021) argues that the possibility of non-human-like minds who aren’t harmed by coming into existence poses a challenge to David Benatar’s well-known Asymmetry Argument for anti-natalism. Since the good of these future minds has the potential to outweigh the current harms of human existence, they can be appealed to in order to justify procreation. In Section two, I explicate what I understand to be the reasoning behind Gould’s assertion. In Section three, I claim that Gould’s argument rests on a fundamental misunderstanding of Benatar. According to the Asymmetry Argument, if a person experiences *any* harm at all, then bringing them into existence is unjustified. It does not depend upon on-balance judgments about the relative harms and benefits of existence. It therefore remains impermissible to procreate *right now*, even with the prospect of contributing to the successful development of beings who are not harmed by existence. Gould’s argument mistakenly seems to ask whether procreation is permissible in some other possible world, but Benatar’s concern is whether it’s permissible in the actual world.

In Section four, I address two possible alternative readings of Gould, assuming for the sake of argument that on-balance judgments are indeed permitted by the Asymmetry Argument. The first of these says that it is not the positive experiences of future minds that furnish us with the reasons right now to build them, and hence continue procreating, but it is instead the prospect that those minds will make our current lives go better. I argue that Gould’s use of this amounts to an illicit means-ends justification for procreation. The second says that it is not the positive experiences of future minds that are the justification for building them and hence for continuing the human species. Instead, the justification might be due to non-welfare kinds of value such as

the identification of knowledge production as the most intrinsically valuable thing and its subsequent achievement. I submit that Gould appeals to ignorance as he fails to offer concrete evidence that such non-welfare value will in fact be achieved. It is also doubtful that even allowing this sort of comparative judgment would justify procreating because it's difficult to see how an individual's current suffering could be justified for some impersonal good in the very distant future. Thus, even if the type of on-balance judgments to which Gould appeals are appropriate (something I do not admit), his appeal to non-human future minds does not defeat Benatar's anti-natalism.

2. Gould's Challenge to Benatar's Asymmetry Argument

Gould's article concerns future minds that are unlike human minds.¹ While human-like future minds are superintelligent, they also possess the "capacity to endure chronic pains, mental anguish, depression, crippling anxiety" (Gould 2021, 794). But *non-human* future minds are "intelligent, conscious, capable of having fulfilling projects, but are *incapable* of experiencing any sort of suffering, depression, anxiety, etc., that afflicts any typical human person now" (Gould 2021, 794). It is the possibility of such minds that Gould alleges creates a challenge for Benatar's Asymmetry Argument for anti-natalism. Instead of spending time outlining Benatar's argument, I will move right to Gould's main argument (I refer the reader to Benatar 2006, 28-49 for the relevant details of the Asymmetry Argument).

It's striking that Gould acknowledges early on that "if the basic asymmetry [i.e., the one Benatar defends] is correct, then it is *always* a harm for a human being to come into existence"

¹ See Danaher 2018 and Shiller 2017 for two articles that influence Gould's thinking.

(2021, 795; emphasis mine). The reason it is striking is because as I will show in Section three, Gould's argument doesn't fully grasp the implication of this idea. Still, Gould reasonably states that his:

New objection that Benatar's asymmetry applies only to minds as they contingently are, not to minds as they are necessarily. Given the freedom that genetic enhancements, brain- computer interface implants, or even artificial neural networks might allow, it is conceivable that future minds will enjoy the benefits of pleasure without pain. Such non- human future minds wouldn't yield the same asymmetry results that actual human minds do. So, an antinatalist conclusion that it would be wrong to bring them into existence wouldn't apply, at least not for reasons having to do with the asymmetry (2021, 796).

This is true if Gould just means that Benatar's Asymmetry Argument doesn't necessarily apply to creatures who are differently constituted from humans. If a being cannot experience pain (or other forms of harm), then it is true that Benatar's argument doesn't apply to them. Benatar's argument ultimately rests on contingent facts about our actual world. So far, so good.

Gould asks us to consider two different options prior to coming into existence (in addition to rightly asking the reader to set aside the paradoxes involved in doing so):

Option A: you will live a normal mortal existence, in which you are capable of pursuing typical productive pursuits experiencing fulfillment, etc., with typical human responses to suffering, illness, fear of death, etc. Let's also assume that Option A also comes with a guarantee that you will live a relatively comfortable life by today's standards, enjoying the typical pleasures available in today's modern society: health care, relative economic comfort, and so on, but with the usual cognitive limitations,

and eventual death from cancer or some similar end. Option B, however, is a comparatively atypical existence in which you are capable of experiencing those same productive pursuits, fulfillment and more, but you also enter an existence with an exceptional disposition to endure what would normally be considered painful experiences with no suffering, no tendency toward depression, debilitating fear, etc. In other words, a mind that is closer on the spectrum to a non-human future mind. Considering these two options, from the vantage point of seeing out from the future-entity's point of view, gives us that future-entity-centered perspective we should adopt, rather than an anthropocentric one. And this thought experiment suggests a new argument, one that will lead us to the new challenge to anti-natalism:

1. If you were presented with this choice, Option B would be preferable to Option A.
2. If that's so, then future generations of enhanced future minds are preferable to future generations of typical human beings.
3. If future generations of enhanced future minds are preferable, then we have reason to pursue the development of enhanced future minds instead of typical human procreation.
4. Therefore, we should prefer the development of enhanced future minds over the usual reproduction of human offspring (Gould 2021, 797).

Gould's basic point here is that enhanced future minds are better than the usual future human offspring. But this fact on its own doesn't defeat anti-natalism. Gould explains that:

The asymmetry argument urges us against procreation as the harms of existence outweigh the goods, as compared on balance against non-existence. Yet for whatever intrinsic goods you identify— whether in pleasure, knowledge-acquisition,

eudaimonia, etc.— the prospect of nonhuman future minds promises an existence that isn't subjected to our contingent embodiment of the asymmetry (2021, 798).

From these considerations, he formulates his specific challenge to anti-natalism:

5. Axiological asymmetry holds for all of the persons we know (viz., human beings), but only contingently so.
6. If the asymmetry holds for persons only contingently, and there are means to ensure that persons are not afflicted with this asymmetry, then we ought to pursue the development of such persons and not the ordinary sort affected by the asymmetry.
7. The possibility of developing non- human future minds would present a means to ensure the existence of persons who are not asymmetry- afflicted.
8. Therefore, we ought to pursue the development of future minds that are not relevantly human- like, but we ought not procreate, or develop relevantly human- like future minds (Gould 2021, 798; premise numbers mine).

Since Gould believes that the development of non-human like minds is an open possibility, it is permissible to continue to procreate humans in order to achieve that end. This doesn't entirely defeat Benatar's Asymmetry Argument because if we were to discover that creating non-human-like minds is impossible, then we would not be justified in continuing to procreate.

3. Gould's Gaffe: Harms and On Balance Judgments

Gould should be commended for bringing discussion about human enhancement and futurism to bear on anti-natalism. Critics of Benatar's Asymmetry Argument often attempt to identify issues with the asymmetry itself (or simply express anger at its conclusion) and Gould's article shifts the focus in an interesting way. Unfortunately, Gould either misinterprets and/or misunderstands

Benatar such that once these confusions are cleared up, it is apparent that no challenge to the Asymmetry Argument exists.

Here is the basic problem for Gould. He himself acknowledges the Asymmetry Argument implies that it is *always* harm to come into existence, but his argument doesn't account for this point. This is because in his defense of Option B over Option A he writes of making an 'on balance' assessment. But this is confused. The Asymmetry Argument does *not* rest on the claim that when we weigh the harms against the goods of existence, we shouldn't procreate. It doesn't rest on an 'on balance' assessment. Every person brought into existence in our world experiences some degree of harm, even people who live comparatively good lives. This is all Benatar needs because his argument implies that if a person who will be brought into existence experiences *any* amount of harm, no matter how quantitatively or qualitatively insignificant, then they ought not to be brought into existence (see Benatar 2006, Ch.2).

Even if it's true that non-human minds are better than human minds, this fact does not defeat Benatar's claim that we shouldn't procreate *right now in our world*. Benatar's argument is about procreation in the actual world. If humans were constituted differently such that we didn't experience pain or if the world was perfect such that it contained no evil or pain, then bringing someone into existence would *not* cause them harm. Benatar need not deny this is so because while procreation in such circumstances would be permissible, it is entirely irrelevant to the soundness of the Asymmetry Argument which is about the permissibility of creating *right now in the actual world*. If a person experiences *any* harm at all for just one *moment*, it is impermissible to bring them into existence if Benatar's Asymmetry Argument is sound. On balance judgments about the relative quantity and quality of pain/evil to happiness/pleasure a person will experience if brought into existence are irrelevant to Benatar's asymmetry thesis.

Some of what Gould seems to be saying is that if there is a possible world ‘Perfect’ where non-human creatures aren’t harmed by existence, then we have an obligation to bring ‘Perfect’ into existence (or at least it’s permissible to bring it into existence). The difficulty is that Gould fails to recognize that in order to bring about ‘Perfect’, we have to continue to procreate in our actual world ‘Alpha’. In other words, we need ‘Alpha’ in order to get to ‘Perfect’. But just because it’s permissible to procreate in ‘Perfect’, it doesn’t follow that it’s permissible to procreate in ‘Alpha’. This is so even if continuing to procreate in ‘Alpha’ will eventually get us to ‘Perfect’.

Another way to see the problem is to reconsider Gould’s premise (6) (which is (2) of the second argument in his text). (6) is incomplete because it doesn’t include the clause that people created who are inflicted by the asymmetry are *created right now in the actual world*. But once this clause is included, then (6) is false or at least doesn’t obtain in the actual world. There is no means by which we can ensure that persons we create right now in the actual world do not experience any harm, and so aren’t ‘afflicted’ by Benatar’s asymmetry.² If Benatar’s Asymmetry Argument is sound, then the goal of non-human-like minds does not justify creating human-like minds who are guaranteed to experience harm once brought into existence.

Now, Gould anticipates some of my response when he writes:

[S]o long as the overall benefits of a world with non-human future minds outweighs the costs of the suffering of previous or even intervening generations, it might be acceptable to reproduce just enough people to arrive at this techno-utopian generation. Call this the ‘just enough humans’ answer. The latter answer might identify knowledge-acquisition as the most important intrinsic good, for example, and provided

² I recognise the utility of theorizing about the best possible worlds in order to help realize them (as is sometimes done in ideal political theory). This type of theorizing is not, in itself, what is at issue here. The problem is trying to achieve these better or best worlds involves bringing minds into existence in the actual world that do in fact experience harm.

that we have good reason to believe the non-human future minds generation will achieve returns we can barely conceive, it might be worth it (2021, 798)

But this response again makes clear that Gould does not recognize the basic implications of the Asymmetry Argument. There are no ‘outweighing’ value judgments permitted. Benatar never denies that there are good things in life (though he does claim people overestimate how well their lives are going). This is why he consistently highlights the distinction between a life worth living and a life worth creating (Benatar 2006, 22-27). The vast majority of people have an interest in continuing to live which is why most of us do not commit suicide. Indeed, working on technology that would lead to non-human-like minds in the future may very well contribute to the meaning of a person’s life. But it is still a fact that any person brought into existence *right now in our world* will be harmed. Gould fails to defeat Benatar’s Asymmetry Argument.

4. What if ‘On-Balance’ Judgments are Permitted?

I believe that the reading of Gould I offer in Section two is the most accurate one on offer and that as such my response in Section three is sufficient to show that his rejection of Benatar’s Asymmetry Argument fails. However, supposing for the sake of argument that on-balance judgments are actually permitted by the Asymmetry Argument, there are two possible ways Gould could try to avoid my objections. In this section, I briefly examine these possibilities and show that there are plausible replies to be made on behalf of anti-natalism.³

4.1 The Prospect of Future Minds Improving Our Lives

³ The second and third interpretations of Gould, in addition to the replies offered, are owed to an anonymous referee. Some of my descriptions of the possible readings are taken almost verbatim from the referee report.

If 'on-balance' judgments are permitted, then one possible reading of Gould says that it is not the positive experiences of future minds that furnish us with the reasons right now to build them (and hence to continue procreating) but it is the prospect that those minds will make our current lives go better.

Benatar's analysis of cloning or procreating in order to save a child who already exists is instructive here. There are cases where parents have a child in the hopes of having a match for a bone marrow transplant for their child who already exists and suffers from Leukaemia. In some cases, the parents have a baby and hope that it will be a match. But in other cases, embryos are screened in advance and only chosen if they are in fact a match (Benatar 2006, 128-129). Benatar explains that many object to such cases of procreation or cloning because it creates a child merely in order to achieve some end (i.e., in this case, to provide bone marrow).

Here is part of Benatar's analysis of this case, along with the typical objection to it which worries about the means-end justification it uses:

What those who raise the Kantian-like objection routinely ignore is that in so far as it applies to cloning and cases of having a child in order to save a child, it applies at least as much to ordinary cases of having children. *This is true irrespective of whether one accepts that coming into existence is always a serious harm.* Clones and those children who are produced to save the life of a sibling are not brought into existence for their own sakes. This, however, is no different from any children. Children are brought into existence not in acts of great altruism, designed to bring the benefit of life to some pitiful non-being suspended in the metaphysical void and thereby denied the joys of life. In so far as children are ever brought into existence for anybody's sake it is never for their own sake (2006, 129-130; emphasis mine).

Notice the force of this response when applied to Gould. Even if one rejected the idea that coming into existence is always a serious harm (in addition to ‘on-balance’ judgments being permitted), this objection still applies to Gould. If it is wrong to procreate in order to achieve some end, it doesn’t matter if the end in question is good. Indeed, Benatar observes that the typical reasons for ordinary reproducing are “(a) to satisfy their procreative or parenting interests; (b) to provide siblings to existing children; (c) to propagate the species, nation, tribe, or family; or (d) for no reason at all. These are all clearly weaker reasons for producing a child than is the goal of saving the life of an existing person” (2006, 130).

When applied to Gould’s argument, it becomes clear that anti-natalists don’t have to deny that the very prospect of non-human future minds might improve our current lives. However, continuing to procreate in order to improve persons’ lives who already exist is impermissible to anyone who accepts that it is wrong to justify procreation through means-end reasoning. I suspect that this kind of means-end justification would be a bitter pill for Gould to swallow. And even if this is a route he is willing to take, it would serve to show his argument rests on highly controversial moral intuitions.

4.2 Future Minds and Non-Welfare Value

The other possible reading of Gould is ‘on-balance’ judgments are permitted suggests that it is not, strictly speaking, the positive experiences of future minds that is the justification for building them and hence for continuing the human species. Instead, the justification might be due to non-welfare kinds of value. For example, recall that Gould refers to the value of an important type of knowledge production “[we] might identify knowledge-acquisition as the most important intrinsic good, for example, and provided that we have good reason to believe the non-human future minds generation will achieve returns we can barely conceive, it [i.e., procreation] might be worth it” (2021, 798).

Yet from the anti-natalist's perspective, there are still a number of problems with this example. Perhaps the most conspicuous is that it's difficult to see why Gould hasn't committed the fallacy of appealing to ignorance. Even supposing it is true that there is no evidence to think future minds won't (a) be able to discover what constitutes the most intrinsic good and (b) be able to realize that good, it does not follow that there are reasons to suppose they will, in fact, be able to do so. What is Gould's evidence that this mere logical possibility is guaranteed or even made likely by the existence of future minds?⁴

Even if the anti-natalist were willing to entertain comparative value judgments in assessing the permissibility of procreation, something not required by their position, I submit that this strategy from Gould still fails. More carefully, anti-natalists should reject the idea that the harms current and intermediate generations experience could be justified by the creation of value in the distant future. Consider prospective parents in a concentration camp in WWII. Suppose they reasoned along the following lines: Our bringing a child into existence in these terrible conditions is justified because in the distant future we will discover that knowledge is the most intrinsically valuable good and be able to realise that knowledge. This reasoning is dubious inasmuch as it offers little comfort to the child born into such appalling circumstances. Not only is it entirely disconnected from their current life, it also trades a very distant *impersonal* good for their individual suffering. Finally, we might fairly wonder what evidence the parents possess to think the world will be so much better in the distant future.

By analogy, appealing to the very distant good of knowledge does little to assuage the current suffering of individuals brought into existence right now. It is of little comfort to claim that

⁴ However, even permitting 'on balance' judgments, the chance of obtaining such knowledge probably has to be non-trivially significant. The judgment will also depend on the qualitative and quantitative suffering that (maybe many generations of) minds must endure in order to achieve the knowledge in question. If he wanted to pursue this strategy, the burden of proof is on Gould to explain how such 'on balance' judgments are supposed to work.

in the face of the harm of existence, there is the very distant future possibility of realizing some great intrinsic good, a good that the suffered won't experience. Asking what evidence there is that this intrinsic good will be identified and subsequently realized also brings us back to the worry that Gould has appealed to ignorance to justify his position.⁵

In sum, the type of 'on-balance' or comparative judgments to which Gould appeals to defeat Benatar's Asymmetry Argument is rejected by Benatar. However, in this section, I have demonstrated that even if such judgments were permitted by the argument, the kind offered by Gould does not defeat the argument. For they either make illicit appeals to ignorance or to counterintuitive means-end reasoning.

5. Conclusion

As stated, I commend Gould for taking a different angle on Benatar's Asymmetry Argument. It may well be possible that we can create non-human minds in the future who won't be harmed by being brought into existence. Unfortunately, unless they can be created by humans *who already exist*, they cannot be successfully appealed to in order to defeat Benatar's Asymmetry Argument. Since any human currently brought into existence in our world will be harmed by their existence, it is impermissible to bring them into existence. This is so even if we have reason to think we could eventually create the non-human-like minds Gould has in view. In other words, if the Asymmetry Argument is sound, then the ends Gould identifies cannot justify the means to continue to procreate

⁵ Gould might try to rally by claiming a good need not be experienced by an individual in order for it to be good. Perhaps the knowledge-acquisition in question here isn't really impermissible because a mind in the future does in fact experience it. However, though it's true that such knowledge isn't an impersonal good for those future minds, it remains the case that this knowledge is not possessed by the minds that come before them. It remains impermissible to procreate those minds because the harm that *they* experience is not outweighed by the knowledge (they never achieve). None of this involves denying that arriving at such knowledge is in fact good.

humans. Benatar's argument does not rely upon 'on balance' judgments about the relative harms and pains of existence. Any harm, no matter how slight, is all Benatar needs to secure his conclusion on the assumption that his asymmetry thesis is correct. It therefore remains the case that a successful rebuttal of Benatar's Asymmetry Argument needs to face his asymmetry thesis head-on. Gould doesn't attempt to do this which is why I think these problems arise. It's true that there are possible worlds where even if Benatar's Asymmetry Argument is sound, it is permissible to procreate because beings aren't harmed by existence in those worlds. This changes nothing of what the asymmetry thesis tells us about procreation right now in the actual world.

But even supposing 'on-balance' judgments are permitted by the Asymmetry Argument (something I do not admit), does not help Gould's case. The first of this type of strategy says that it is not the positive experiences of future minds that furnish us with the reasons right now to build them, and hence continue procreating, but it is the prospect that those minds will make our current lives go better. But this makes use of means-end justification that many people will want to reject in relevantly similar cases. The second says that it is not the positive experiences of future minds that are the justification for building them and hence for continuing the human species. Instead, the justification might be due to non-welfare kinds of value such as the identification of knowledge production as intrinsically invaluable and its subsequent realisation. Here Gould appeals to ignorance as he fails to offer concrete evidence that such non-welfare value will in fact be achieved. It is also difficult to see how an individual's current suffering could be justified for some impersonal good in the very distant future.

Though I deny that ‘on-balance’ judgments are permitted by the Asymmetry Argument, it ultimately does not matter. No matter which strategy Gould employs, Benatar’s anti-natalism emerges unscathed by Gould’s appeal to future non-human-like minds.⁶

⁶ Acknowledgements.

Bibliography

Benatar, D. (2006). *Better Never to Have Been: The Harm of Coming into Existence*. Oxford University Press.

Danaher, J. (2018). “Why we should create artificial offspring: Meaning and the collective afterlife.” *Science and Engineering Ethics* 24: 1097– 1118.

Gould, D.C. (2021). “Future minds and a new challenge to anti-natalism.” *Bioethics* 35: 793-800.

Shiller, D. (2017). “In defense of artificial replacement.” *Bioethics*, 31(5), 393– 399.