Child-focused Arguments Against Procreation

Andrew Pearson

Central European University

Submitted in partial fulfilment of the requirements for a Master's Degree in Philosophy

Supervisor: Simon Rippon

Acknowledgements

I wish to thank the following people for questions and conversations relating to antinatalism and previous drafts of this essay: Hanoch Ben-Yami, Emma Bullock, Sam Dumitriu, Daniel Hanson Dzah, Philip Goff, Mike Griffin, Marek Iwaniak, Yuk-Pui Lam, Brian McLean, Andres Moles, Laura Nicoara, Jasmine Opie, Edi Pavlovic, José Ricon, Simon Rippon, Andrew Sabisky, Ben Southwood, and David Weberman. Furthermore, I wish to thank José Ricon, Liam Shields, and Simon Rippon for comments on previous drafts and sections of this essay.

Given the subject matter and my pro-natalist conclusions, it would perhaps also be appropriate to thank my parents, Tim and Karen Pearson, for bringing me into existence. Beyond that, I thank them for their support - emotional and otherwise - during my time studying for this degree.

Table of Contents

Introduction 1
The Asymmetry Argument 3
Terminology and Conceptual Clarification 3
Optimistic Asymmetries Rejected 8
Incomparability Dismissed 9
Benatar's Argument for "Asymmetry" 9
Genethical Parity and the Dual Benchmark 11
Asymmetry and Antinatalism 13
Benatar Fails to Justify Strong Asymmetry 14
Weak Asymmetry is Also Unjustified 17
Conclusion on Asymmetry 22
The Quality of Life argument 23
A Formal Presentation of the Quality of Life Argument 23
Is It Immoral to Bring Into Existence A Person Who Will Probably Have a Bad Life? 24
Benatar's Case Against Self-Reported Happiness 25
Has Benatar Refuted Self-Reported Happiness? 26
Benatar's Argument that Lives Go Badly 30
Does Benatar Succeed in Showing that Most Lives Go Badly? 32
Conclusion on Quality of Life 37
Implications of Pro-Natalist Conclusions 38
The morality of procreation 38
Is Procreation Mandatory, Impermissible, or Something else? 39
Is There A Duty To Provide The Maximum Benefit That One Can? 40
The politics of fertility 42
Conclusion on the morality of procreation 44
Conclusion 45

Introduction

Antinatalism is the philosophical position that it is immoral to bring new people into existence. It has been defended by various philosophers on various grounds; in this dissertation, I shall critically discuss and respond to some of these arguments, in particular those made by David Benatar in his 2006 book *Better Never to Have Been: The Harm of Coming Into Existence*.

It is worthwhile to place arguments for antinatalism into two categories: *child-focused* arguments and *other-focused* arguments. Child-focused arguments are those arguments which locate the immorality of procreation in some alleged harm to the person brought into being. By contrast, other-focused hold procreation to be wrong for some reason other than harm to the new person. An example of an other-focused argument would be the claim, exemplified by Conly (2015), that increasing the human population is bound to reduce quality of life for existing humans, and that this constitutes a harm which it is immoral to inflict.

My focus is entirely upon child-focused arguments. I mean to discuss two lines of argument, both made in *Better Never to Have Been*. The first of these is what Benatar refers to as the *Asymmetry* argument, according to which coming into existence is inevitably (though not necessarily) a harm, on the grounds that the pains of existence are bad while the pleasures of existence are not good - at least, not relative to an alternative of non-existence. I show that there are several types of asymmetry which Benatar fails to distinguish between. In particular, there are "strong" and "weak" versions of Benatar's asymmetry; his arguments only support a weak asymmetry, while his conclusion requires a strong asymmetry. Moreover, strong asymmetry has several highly implausible consequences. Furthermore, I suggest that ultimately we ought additionally to reject even a weak asymmetry, in favour of the belief that coming into existence can be either a benefit or a harm depending upon the circumstances of one's life, and provide arguments in favour of this view.

The second argument is the *Quality of Life* argument, according to which most people live bad lives and so ought not to be brought into existence. I show that this argument rests upon several unacknowledged and controversial (though not implausible) claims about what constitutes the good life. Furthermore, I challenge the empirical basis upon which Benatar rests his argument, showing that it does not establish what Benatar believes it does. Firstly, the empirical claims which he relies on are more controversial than Benatar claims; secondly, there are hidden premises in his argument.

In the final section of this dissertation, I discuss what conclusions might follow from my arguments. I tentatively suggest that if a couple are both happy with how their lives are going, it will generally be morally acceptable, and indeed supererogatory, for them to procreate. I suggest that, depending upon one's other moral premises, my arguments may or may not provide an argument against the moral permissibility of abortion. Finally, I suggest that it is plausible to think that childbearing ought to be very heavily subsidized by the state.

The Asymmetry Argument

Terminology and Conceptual Clarification

Before getting deep into any arguments, I wish to set out a few distinctions and definitions. First, we need a distinction between *narrow* and *all-things-considered* benefits and harms. Suppose you are struck by a severe illness and must take an unpleasant medicine: apart from its foul taste, the medicine frequently provokes vomiting. However, the effect of the medicine is to accelerate your recovery. We may then say that the medicine was both beneficial and harmful: beneficial, because it expedited the return to health, but also harmful due to the suffering involved in taking the medicine and in throwing up. These are all examples of narrow benefits and harms.

We can, however, say that the medicine was overall beneficial or harmful - that either the narrow benefits outweighed the narrow harms, or vice versa. This is what I call all-things-considered benefit and harm. In the medicine case, so long as the enhanced recovery is of greater importance to you than the unpleasantness of taking the medicine and of vomiting, we would say that the medicine is overall beneficial.

While an action might simultaneously beneficial and harmful for an agent in the narrow sense, in the all-things-considered sense an action can only be one of these. The all-things-considered value is frequently, but by no means equal, to the sum of the values of narrow benefits and harms. Sometimes a narrow benefit or harm will act as a disabler upon another benefit or harm, meaning that it can cancel out the good or bad of that other facet of the event, but cannot outweigh it. For example, suppose someone's drink is laced by a bizarre prankster both with a mild poison and with the antidote to that poison. The inclusion of the antidote represents a genuine benefit to the drinker, but only as a disabler on the harm of the poison[1].

[1] McLean (2015) suggests that the moral reasons we have for creating people act as moral disablers on our reasons against creating them, though in personal communication he suggested an inclination towards thinking that this is a feature of morality rather than prudence and that people really can benefit all-things-considered from coming into existence.

It is usually impermissible to impose all-things-considered harms on people in a way that does not apply to narrow harms. For example, suppose that Alan is about to be hit by a moving vehicle when Betsy pulls him out of the way. Due to the force with which she yanks him, Betsy breaks Alan's wrist. In this case Betsy imposes a significant narrow harm upon Alan, but because this harm is outweighed by the narrow benefit to Alan of not being hit by the vehicle, it is entirely permissible for her to impose this harm. If Betsy had broken Alan's wrist while imposing no benefit, then she would instead have impermissibly inflicted an all-things-considered harm on him.

This is not to say there is no general rule against imposing narrow harms. If Betsy could have pulled Alan out of the vehicle's path without breaking his wrist (and without causing any greater narrow harms) then she would have been obliged to do so. But the imposition of narrow harms can, at least some of the time, be justified by also bestowing narrow benefits which outweigh the narrow harms and which could not be bestowed without the corresponding harms.

One worry to have with this line of thought is that, at present, it ignores the role of consent. While it may have been permissible for Betsy to break Alan's wrist, it would not be permissible to compel a rational person to imbibe the medicine mentioned above, even if taking it would be all-things-considered for the person. This can be taken two ways, to suggest two different (though) related challenges to the above account of narrow harms and the permissibility of their imposition.

The first suggestion is that perhaps narrow harms are not really harmful so long as we consent to them. I don't think this is a plausible view, however: consider a choice between two medicines, both of which have the same curative properties but one of which also induces vomiting. If consent cancels harms, then the medicine which causes vomiting cannot be considered harmful relative to the alternative.

A more plausible suggestion is that it is permissible to impose narrow harms only when they are consented to. This means that in general it will be permissible to impose narrow harms which are necessary for the conferral of narrow benefits which outweigh the harms, but preserves a strong role for consent. In general I would embrace this suggestion, but would note that it cannot explain all cases. In cases where it is infeasible to obtain consent, such as the case of Alan and Betsy, it would be absurd to claim that no narrow harms may be imposed. The two main ways we can take this are to suggest that in such cases, a cost-benefit analysis of narrow

benefits and narrow harms comes to determine possibility; or, more in the spirit of consent theory, to suggest that the person subject to benefits and harms "would consent" if they were able to do so.

An alternative view is provided by Seana Shiffrin (1999), who argues that it is impermissible to impose narrow harms even this is necessary to provide benefits. Shiffrin accepts (at least for the sake of argument) the claim that people can enjoy an all-things-considered benefit from being brought into existence, but contends that procreation is nevertheless rendered morally impermissible by the fact that the child will inevitably suffer narrow-scope harms. She explains cases such as Betsy's breaking Alan's arm not in terms of the benefit to Alan, but rather in the fact that Betsy averted a greater harm than the one she caused.

There does intuitively seem to be something to the idea that, as Shiffrin suggests, it is permissible to cause certain harms to an unconscious - and therefore unable either to give or deny consent - medical patient when this is necessary for averting or limiting greater harms, but not for the purposes of "supernormal benefits, such as supernormal memory, a useful store of encyclopedic memory, twenty IQ points worth of extra intellectual ability, or the ability to consume immoderate amounts of alcohol or fat without side effects." (ibid p127) However, I am inclined to reject this for three reasons: first, the original intuition seems highly unreliable. It has been shown (for example by Bostrom and Ord, 2006) that many people have an irrational bias in favour of the states of affairs which currently obtain, judging them prudentially and ethically superior to alternatives which, if viewed without knowledge of which obtains, seem clearly superior. Coupled with other known biases such as the endowment effect (Kahneman, Knetsch, & Thaler, 1990) this suggests that people irrationally underweight the benefits received in such cases, providing a debunking explanation for our intuition in the thought experiment concerning the medical patient.

Second, even if it were true in some "base-level morality" that it is impermissible to inflict even narrow harms except for the purposes of preventing or alleviating greater harms, it seems sensible for an agent to consent to a social rule according to which it is permissible to inflict narrow harms when this is necessary for the bestowal of what are agreed to be greater benefits. Suppose that there is some highly invasive and unpleasant surgery, which has the sole effect of significant boosting the strength and intelligence of the patient. This is rarely carried out due to

its painfulness, but when one is unconscious one would not care so much about this[2]. It seems then that many people would agree in advance that, were they for any reason to fall out of consciousness in a way likely to be temporary, the operation should be performed on them while they would be unable to feel the pain. If this operation involved a financial expense that was non-trivial but nonetheless near-universally agreed to represent tremendous value for money, presumably the people who would consent to this operation being performed on them would also consent to paying the cost afterwards. If we would be so willing to allow the imposition of harms and costs upon us in this case, why have a social ruling against this as a general principle?

There are good reasons for why we do not ordinarily allow the imposition of harms even for greater benefit - the cases where mature agents could receive such benefits and are unable to give or deny consent are few and far-between. But we routinely impose such deals on immature agents, particularly through political membership: children receive healthcare and education (which in their lack of maturity they frequently fail to recognise even as benefits to them), and in return take on partial responsibility for the national debt. (The principle which would justify this is in fact vastly *more* permissive than that which would justify procreation: we could if we so chose pay down the national debt and pass none of it onto our children, whereas it is utterly impossible to provide children with lives which they will come to view as worthwhile and valuable without also causing them to suffer certain narrow harms). Would the many of us who enjoy our lives not consent, ex post, to the various narrow harms which have been visited upon us?

Next, it is worth digging deeper into what we mean by "asymmetry". I shall suggest that there are a variety of different types of asymmetry, and that we should be particularly interested in the differences between the positions I refer to as *strongly pessimistic prudential asymmetry, weakly pessimistic prudential asymmetry,* and *prudential symmetry*.

First: prudential versus moral asymmetry. Prudential asymmetry concerns the question of whether one can all-things-considered benefit, suffer harm, either, or neither, as a result of

[2] Shiffrin provides a similar thought experiment concerning a wealthy island-dweller who, unable to have direct contact with the inhabitants of a neighbouring island, airdrops blocks of gold bullion at their feet. I feel that the intuitive force of this case comes at least partly from a widespread unwillingness to accept trade-offs between bodily health and autonomy versus merely financial considerations (Tetlock, 2003).

coming into existence. Moral asymmetry, by contrast, concerns whether or not it is possible to morally wrong someone by bringing or failing to bring them into existence. We are primarily concerned in this section with prudential asymmetry, and any mention of asymmetry not preceded by either “prudential” or “moral” should be assumed to refer to prudential asymmetry.

Second, optimistic and pessimistic asymmetries. Optimistic asymmetries rule out the possibility of suffering an all-things-considered harm in being brought into existence: to the extent one experiences narrow harms as a result of coming into existence, they can only function as disablers. Pessimistic asymmetries are precisely the reverse of this, ruling out the possibility of experiencing an all-things-considered benefit through being brought into existence and holding that narrow benefits from coming into existence, if they even exist, can act only as disablers on narrow harms from coming into existence.

In addition to the possibilities of optimistic and pessimistic asymmetries, one might hold that both narrow benefits and narrow harms can be experienced as a result of coming into existence, and either can outweigh the other, a position we may call *symmetry*. The remaining logical space is *incomparability*, according to which existence and non-existence are just plain different things, and one can experience neither narrow benefits nor narrow harms, in the truest senses of the words, as a result of coming into existence.

Finally, weak and strong asymmetries. Strong asymmetries maintain that even in the narrow sense, one cannot experience both benefits and harms: according to strong optimistic asymmetry, one can only benefit and cannot be harmed in any way by being brought into existence, while strong pessimistic asymmetry maintains that one can only be harmed and cannot benefit in any way by being brought into existence. Weak asymmetries allow for both narrow benefits and narrow harms, but maintain that one of them can only act as a disabler on the other. Thus weak pessimistic asymmetry, for example, holds that one can experience narrow benefits by coming into existence but these cannot outweigh the narrow harms one experiences - at best, the narrow harms will be entirely cancelled out.

Thus with regard to the question of prudential asymmetry, we have six possible answers: strong optimistic asymmetry, weak optimistic asymmetry, incomparability, strong pessimistic asymmetry, weak pessimistic asymmetry, and symmetry. To offer a preview of where this is going: I shall argue that Benatar is committed to strong pessimistic asymmetry, but his arguments do not support this position: at best one can provide justification for weak pessimistic

asymmetry. I shall furthermore argue that we should reject prudential asymmetry in favour of symmetry, which is the most sensible position.

Before moving onto Benatar's view, then, I shall briefly discuss the optimistic asymmetries and incomparability, and why I think these are non-starters as positions.

Optimistic Asymmetries Rejected

Just as there are pessimistic prudential asymmetries, there are optimistic prudential asymmetries, according to which one cannot be harmed by being brought into existence. These are coherent positions, and I daresay that at least some people hold such views. A person might hold on religious grounds, for example, that every life is a precious gift bestowed by a benevolent and omniscient God and therefore cannot be a mistake. This dissertation is about ethics rather than the philosophy of religion, so I shall not be responding to such arguments.

Taking purely secular considerations into account, we have strong reasons not to accept such an asymmetry. Most straightforwardly, it seems obvious that some people *are* harmed by being brought into existence. Consider a baby born with a severe defect, so that he dies shortly after birth having known nothing but extreme pain. It seems obvious that the baby would have been better off never existing.

Second, there are at least some people who, we would generally agree, would be better off ceasing to exist. While legalisation of euthanasia remains controversial, arguments against it have tended to focus upon the danger of a slippery slope towards involuntary euthanasia, or the risk of such laws being abused[3]; if people thought that one could only be harmed, and never benefited, by dying, then we would expect this to be used as an argument against euthanasia. Given the apparently general agreement that one can benefit by ceasing to exist, then, there is a burden upon anyone defending optimistic asymmetries to explain why one can from birth be benefitted by ceasing to exist but cannot be so benefited by not being brought into existence.

In short, optimistic asymmetries regarding procreation may, I think, be safely rejected.

[3] See for example https://www.care.org.uk/our-causes/sanctity-life/arguments-for-and-against-euthanasia, a presentation of the debate by a pro-life charity.

Incomparability Dismissed

So much for optimistic asymmetries: what can we say for *incomparability*, the view that existence and non-existence simply cannot be compared, and that it is a nonsense to speak either of benefiting or of being harmed by being brought into existence?

This position does have some appeal. It is not immediately obvious what is mean by comparing existence to non-existence: can it really be so simple as asking whether an agent, given mature and careful consideration, would prefer never to have existed? I think that avoiding this issue creates more problems than it solves, however.

Incomparability suffers from the same troubles that I have suggested for optimistic prudential asymmetries, i.e. that some people are obviously harmed by existing, and that there is a tension between the view that one cannot be harmed by being brought into existence and the view that some people benefit by ceasing to exist. It also fails utterly as a guide to practical morality: one would have to be sadistic or at least uncaring to deliberately bring into existence someone who one knows will despise their life, yet incomparability struggles to prohibit such behaviour.

Perhaps a defender of incomparability would suggest that while our notions of “harm” and “benefit” are inappropriate to apply to cases of being brought into existence, it is nevertheless possible to produce true moral judgements about procreation grounded in the welfare of the person brought into being. A defender of such a position might develop a language allowing for such judgements without reference to benefits and harm; the key question would then be whether or not my arguments in this dissertation can be translated into that language.

Feinberg (1992) suggests that there ought to be sharp limits on the extent to which people who suffer harm as a result of their birth ought to be granted a legal tort against those responsible for bringing them into existence; however, I do not think Feinberg is plausibly read as a defender of incomparability.

Benatar’s Argument for “Asymmetry”

Benatar begins his discussion of the alleged asymmetry with a quartet of statements, as follows:

(1) The presence of pain is bad
(2) The presence of pleasure is good
(3) The absence of pain is good, even if that good is not enjoyed by anyone
(4) The absence of pleasure is not bad unless there is somebody for whom this absence is a deprivation (p.30)

Benatar apparently finds these fairly intuitive as is, but recognises that many others will not think this. For this reason he provides a set of four statements which he takes to be highly intuitive, in need of explanation, and best explained by the combination of claims above. These statements are:

(a) "The asymmetry between (3) and (4) is the best explanation for the view that while there is a duty to avoid bringing suffering people into existence, there is no duty to bring happy people into existence." (p.32)
(b) "Whereas it is strange (if not incoherent) to give as a reason for having a child that the child one has will thereby be benefited, it is not strange to cite a potential child's interests as a basis for avoiding bringing a child into existence." (p.34)
(c) "Bringing people into existence as well as failing to bring people into existence can be regretted. However, only bringing people into existence can be regretted *for* the sake of the person whose existence was contingent on our decision." (p.34, emphasis in original)
(d) We make "asymmetrical judgements about (a) (distant) suffering and (b) uninhabited portions of the earth or the universe. Whereas, at least when we think of them, we rightly are sad for inhabitants of a foreign land whose lives are characterised by suffering, when we hear that some island is unpopulated, we are not similarly sad for the happy people who, had they existed, would have populated this island." (p.35)

Benatar also takes the time to respond to some of the more obvious responses. For example, a common response to (a) is to suggest that it is explained by the fact that negative duties are more stringent than positive duties. Bringing someone into a bad life actively harms them, and so is generally agreed to violate the negative duty to refrain from acts which cause harm to others. By contrast, failing to bring into existence someone who would live a good life is merely denying them a benefit, and so one is at most violating a positive duty of beneficence. Benatar acknowledges that this has considerable force with regard to the usual way in which people are brought into existence, i.e. pregnancy. In general there is no duty to provide someone with a

benefit if it is at high cost to oneself - and raising a child is in many ways a difficult and expensive undertaking.

But suppose we remove these costs. Imagine instead that there is a button which one can press to create a new person. This new person will live a life that is generally happy, and their existence will not decrease the welfare of existing people. One has, in the opportunity to press the button, the ability - if coming into existence can be a benefit - to provide another person with a benefit, at no cost to oneself. Even a very modest duty of beneficence, Benatar thinks, would require one to press this button if coming into existence can be beneficial. Laying the argument out formally:

(1) We have basic duties of beneficence, including a duty to benefit other people when doing so comes at no cost to ourselves.
(2) If coming into existence can be beneficial, then pressing the button will benefit another person at no cost to oneself.
(3) There is no duty to push the button.

(C) Coming into existence cannot be beneficial.

I shall return to this argument later to suggest that it is not as sound as it may seem[4]. But let us accept it for now.

Genethical Parity and the Dual Benchmark

One perpetual worry for antinatalists is that they will end up arguing not only for the creation of no new people, but also for the extermination of existing people. It is hard to see how Benatar's quality-of-life argument (see chapter two) would avoid this implication, but he is eager to avoid this as an implication of his asymmetry argument. For this reason he argues that drawing a distinction between coming into existence and continuing to exist has genuine explanatory power; I believe that by the end of this chapter we will see there is little motivation for this view, but it is worth explaining this position and correcting one easy misapprehension of the view.

[4] Also note that some people have opted to simply deny one of the premises - e.g. Francois Tremblay denies (1), that we have any positive duties at all (despite agreeing with Benatar that antinatalism is correct!) (https://francoistremblay.wordpress.com/2017/03/17/i-like-my-life-therefore-antinatalism-is-false/)

Bayne (2010, p.32) divides the terrain into *no-fault*, *genethical parity*, and *dual benchmark* models. No-fault models are equivalent to what I have implicitly labelled *procreative moral incomparability*: it is inappropriate to bear evaluative attitudes, either positive or negative, to the event of one's having come into existence. To the arguments I made against this position earlier, Bayne adds the charge that it is arbitrary to suggest we should bear no attitude towards our coming to exist without also suggesting we should be neutral towards the event of our own death.

Genethical parity, Bayne's preferred position, claims that any life which it would be rationally preferable to continue living, it would also be morally acceptable to begin (or at least, any wrongness would be on account of harms to people other than the person brought into existence). Dual benchmark theory, by contrast, claims that there are lives which, once started, are worth continuing, but which ought not to be started. There seems something odd in this claim - why should the fact of whether a life has yet started affect whether or not it is worth living? - but it is buttressed by examples, the most prominent being that of a person born with a missing limb (Benatar 2006, p.23).

Bayne does much to show that this intuition is dubious, being driven less by a belief that a life without a limb is especially bad than a judgement that it is worse than some other life we think the person might or ought to have had. (Suppose everyone were born missing a limb. Would we still think it were wrong to bring someone into existence missing a limb? If not, then this suggests that our intuition is driven by something other than a belief that a life without a limb is objectively especially bad). (Bayne 2010, p.46)

I think this amounts to a slight misinterpretation of Benatar, however. Benatar does not argue that a life needs to be better to be worth starting than to be worth continuing; rather, he is more accurately seen as arguing that, since the benefits of a life do not yet exist, a life is never made better by being brought into existence. To sum up the conclusion of this section: I am sceptical of the "dual benchmark" which posits lives which are worth continuing from the beginning but not worth starting, but I also doubt that Benatar needs this precise claim to justify his asymmetry-based antinatalism.

Asymmetry and Antinatalism

Benatar does not, alas, straightforwardly lay out how he thinks his asymmetry is intended to lead to antinatalism. However, if one accepts strong pessimistic asymmetry then the argument is very easy to construct:

(1) People inevitably suffer all-things-considered harms as a result of being brought into existence.
(2) It is morally impermissible to inflict all-things-considered harms on people without their consent.
(3) People do not consent to being brought into existence.

(C) It is morally impermissible to bring people into existence.

Premise (1) follows immediately from strong pessimistic asymmetry combined with various sad facts of life that Benatar makes reference to on page 29: that everyone who ever lives will suffer - at some points in their life - pain, grief, illness, and death. It is indisputable that we suffer narrow harms, and so if one denies that there are narrow benefits to existing then one will of course conclude that existence is itself an all-things-considered harm.

Premise (2) is perhaps not true in every imaginable case - would it really be impermissible to (for example) chop off the finger of someone who did not consent, if this was known to be the only way of preventing an earthquake that would kill hundreds of people and render thousands more homeless? Outside such fanciful scenarios, however, (2) will do very nicely.

Premise (3) is so obvious as to barely be worth stating: people do not consent to being brought into existence, for the simple reason that before existence was imposed upon them there was no-one there either to give or to refuse consent.

The argument is valid, so it seems that if one accepts strong prudential asymmetry then one is bound to accept the antinatalist conclusion that procreation is morally impermissible.

Does one require strong pessimistic asymmetry though? Would the argument succeed if we were only to accept weak pessimistic asymmetry?

(1') People are put at risk of harm, with no corresponding chance of benefit, by being brought into existence.

(2') It is morally impermissible to impose risks of harm on people when they do not consent and there is no corresponding chance of benefit to them.

(3') People do not consent to being brought into existence.

(C) It is morally impermissible to bring people into existence.

Premise (1') follows almost as naturally from weak pessimistic asymmetry as (1) followed from strong pessimistic asymmetry. (3') is identical to (3).Given that the argument is valid, if the argument is substantially weaker than the difference will lie in premise (2').

Premise (2') is indeed questionable, since it pays no attention to the amount of harm risked or the likelihood of the harm coming to pass. We routinely impose trivial risks on other people, such as by driving between places by car, and there is not generally taken to be anything immoral about this.

This is not to say that an argument for antinatalism cannot make use of weak pessimistic asymmetry. If children cannot benefit from coming into existence, then we should at the very least be wary of bringing them into being, and insofar as one carries out a cost-benefit analysis (or some other such calculation) when deciding whether or not to have children, weak pessimistic asymmetry will consistently weaken the case for having children. But any such argument will end up appealing to empirical evidence concerning what proportion of lives actually do go badly and how badly lives are at risk of going. One cannot simply conclude that weak pessimistic asymmetry is true and move straight from there to antinatalism.

Benatar Fails to Justify Strong Asymmetry

Benatar's earlier considerations were intended to justify "asymmetry" - but since Benatar lacks the distinction between strong and weak asymmetry, it is unclear from this which he means. In this section I will show that the considerations he gave in favour of "asymmetry" in fact only support weak asymmetry. Strong asymmetry has counterintuitive implications (antinatalism

being one of these, but by no means the only such issue), while weak asymmetry is easier to argue for but cannot - at least by itself - justify antinatalism.

Remember, the key difference between strong and weak asymmetry is that strong asymmetry insists that people are always harmed in an all-things-considered sense by being brought into existence - no matter how overwhelmingly the good in their life might appear to outweigh the bad. According to strong asymmetry, a person who lives an utterly charmed life involving deep and abiding satisfaction from successful projects, meaningful relationships with fellow people, and a rich variety of intense pleasures is nevertheless harmed by existing if they suffer so much as a pin-prick during the course of their lifetimes. The defender of weak asymmetry is under no pressure to say that such a person is harmed by existence - merely that the person does not do better from existing than they would by not existing.

Benatar's first consideration was that there is a duty to avoid bringing into being those whose lives are characterised by suffering, but no such duty to bring into being those whose lives are overwhelmingly happy. Weak asymmetry can explain these intuitions easily: people with bad lives are harmed by existence, while people with good lives do not benefit. Strong asymmetry can also explain these intuitions - but it goes too far. Consider the statement "there is a duty to avoid bringing into being anyone whose life involves a non-zero amount of suffering, no matter how trivial the suffering may be or how much happiness they may experience by living." This is a highly counterintuitive statement, and the defender of weak asymmetry is happy to reject it - but the defender of strong asymmetry is bound to accept this statement.

The second consideration is that, if one were deliberating over whether or not to have a child, it would be strange to take as a reason in favour of procreation the fact that the child would be likely to enjoy life. It would not be in the least bit strange, however, to take as a reason against procreation the fact that the child would be unlikely to enjoy life. Again, weak asymmetry can explain this: the child who enjoys life is not genuinely benefited, whereas the child who hates life is genuinely harmed. Again, strong asymmetry overshoots the mark, implying that one ought to take as a reason against procreation the fact that the child would experience some suffering in life - no matter how much good there might be in the child's life to balance it out.

Both weak and strong asymmetries can explain the third consideration, that one might regret bringing someone into existence for that person's sake but would not regret failing to bring someone into existence for that non-existent person's sake. Yet again, strong asymmetry goes

far beyond our intuitions by implying that we ought to regret the existence of anyone whose life contains even the shortest moment of the most trivial suffering.

Finally, weak asymmetry can explain our judgements of sadness for distant people who are suffering and of indifference towards parts of the universe that are uninhabited, on the grounds that the first involves actual people suffering while the latter only involves the potential for people to not even benefit from existing. Strong asymmetry also gives these explanations, but implies that we ought to be sad for those people who exist happily.

The pattern is clear: weak asymmetry provides an adequate account of our judgements. Strong asymmetry can in principle deliver the same judgements, but also gives a variety of highly counterintuitive claims that we ought to regret the existence of people even when they are living extraordinarily happy lives.

There are further problems with strong asymmetry, such as its inability to judge precisely *how* harmful coming into existence is. Suppose there are two people, Catherine and Dafydd; Catherine undergoes pleasure worth 100 and suffering worth 10 in her life, while Dafydd only experiences 5 worth of pleasure but also experiences 10 worth of suffering. Even if one accepts the claim that both of them - and not just Dafydd - are harmed by their existences, it seems obvious that Dafydd is harmed significantly more by being brought into existence than Catherine is. Yet according to strong pessimistic asymmetry, pleasure simply has no value and so both suffer harm equal to the sum of bad things in their life. That is to say, strong pessimistic asymmetry judges Catherine's and Dafydd's existences to be equally miserable. This is patently absurd.

We can change the numbers further and get even more ridiculous results. Ernie experiences 1 billion worth of pleasure compared to only 5 pain, while Freya experiences no pleasure at any point in her life but suffers 4 pain. (Perhaps Freya is a baby born with a severe congenital disease which kills her, very painfully, within a few hours of being born). According to strong pessimistic asymmetry, Ernie suffers the greater harm from being brought into existence.

In sum, we have no reason to accept strong pessimistic asymmetry over weak pessimistic asymmetry, and multiple reasons to reject it. The asymmetry argument for antinatalism will therefore have to rest upon weak pessimistic asymmetry, and so requires significant empirical backing to suggest that the risk of living a bad life is severe.

Weak Asymmetry is Also Unjustified

Up until this point, I have taken Benatar's "considerations supporting asymmetry" at face value. In this section, I shall suggest that each of these intuitions is either fully explainable within a framework of procreative symmetry, or is otherwise unreliable.

Following this, I present an argument that - so long as one accepts that the relation of "better-than" is transitive - weak pessimistic asymmetry entails a contradiction. On these grounds, I argue that we should reject all prudential asymmetries in favour of prudential symmetry.

As a reminder, Benatar's first consideration in favour of asymmetry was that there is a duty to avoid creating people who live unhappy lives, but there cannot be any such duty to create happy people, even if there is no cost to ourselves. Benatar argues persuasively that this cannot be explained in terms of the fact that there is normally a significant cost to bringing people into existence, or by the greater stringency of negative duties relative to positive duties.

The answer to this, I believe, is that our intuitive sense of duties, at least from a modern western liberal perspective, is that duties must be owed to someone (and arguably to someone other than the duty-bearer). We are deeply suspicious of the notion that there are duties which are not explained by a need to uphold the welfare or dignity of another moral patient. (While this conception of duty is very particular to the modern liberal world, so also is the intuition that there is and can be no duty to procreate. Indeed it is surprising that Benatar finds the idea that there can be no duty to procreate so obvious as to serve as a premise of his argument, when he has stated earlier in the book that many people *do* believe in such a duty (p.9)!)

This does not by itself mean that we cannot bear duties to agents whose existence is contingent upon the commission of our duties. It does, however, entail that we cannot wrong someone by failing to uphold such a duty: if we uphold it then we have not committed a wrong, and if we do not uphold it then there is no-one to have been wronged. By contrast, bringing someone into a bad existence involves an identifiable and tangible victim.

Have I admitted the existence of a kind of asymmetry here? Perhaps, although it can only be of a very weak kind. Firstly, I am able to maintain that one *can* benefit from coming into existence, and suggest merely that it may never be wrong to deny such a benefit. Second, our intuitive morality may well be subject to revision - indeed, the non-identity problem (Parfit 1982, p.351) is

usually taken to show that the intuition that all duties must be borne towards some agent is false, and there are some duties which cannot be explained in these terms. In either case, the intuition provides no evidence for pessimistic asymmetry over symmetry.

The second reason for believing in asymmetry was it is (in Benatar's view) strange to cite as a reason for having children the fact that they would enjoy life, whereas it is not at all strange to avoid having children due to a belief that they would have bad lives. The second half of this is surely true; the first is less clear. I do not share this intuition; nor does Bryan Caplan, who goes so far as to suggest that parents ought to be willing to make themselves less happy by having more children than they really want if this will allow their children to exist and be happy (Caplan 2011, p.2) (though the main focus of Caplan's book is arguing that parents can and in most cases will be more happy by having more children, and suggesting ways to derive even more happiness from one's procreation).

Further evidence for Benatar's intuition not being commonly shared comes from one of the main arguments against complete vegetarianism. This argument, presented in Crisp (1988), is that ceasing to eat meat causes people to cease rearing animals. If these animals experience lives that are on balance worth living, as seems plausible in at least some cases (Ricon, 2016), then ceasing to eat meat will therefore cause worthwhile lives not to be shared. If Benatar's intuition was common then it would provide an obvious response to this argument: what's wrong with not bringing these lives into existence? The fact that this counter-argument is not made - even by Benatar himself, who has written in defence of vegetarianism! (Benatar, 2001) - suggests that the intuition is not so widely held.

Benatar's third justification for asymmetry lies in the fact that we may regret bringing into existence those who live bad lives for the sake of those brought into existence, whereas our regret at failing to bring into existence those who would have lived happy lives "is remorse for ourselves - sorrow about having missed childbearing and child-rearing experiences." Curiously, Benatar states this several times but makes no attempt to defend it. Would it really have been so difficult to look over or conduct some interviews with people suffering from infertility, to examine whether their remorse really was entirely for themselves rather than for the children they were unable to have?

At least some of the difference in our strength of feeling can be explained by the fact that living people are verifiably unhappy - we can see them suffering. By contrast, we are never entitled to

certainty that a potential person would have had a good life - one may have all the advantages imaginable, and yet be unhappy or unsuccessful. One might suffer an unexpected accident, or fall prey to some rare disease. Moreover, the suffering of a real person is clearly visible and therefore visceral, whereas the potential success of a potential person is always vague and undefined. Instead of considering the fate of some imaginary and ephemeral person, let us instead consider a real person who is living a happy and successful life. Hopefully, the reader will know of a wealth of examples. With this person firmly entrenched in our minds, as rich in detail as any suffering person one knows - would we not think it regrettable, for this person's own sake, if they had never come into existence? I believe we would.

Finally, Benatar appeals to the differing judgements we make about distant suffering people and distant uninhabited places. When we hear of people living in poorer countries, living in tiny shacks with insufficient food and unreliable access to clean water, our hearts are moved to pity. When we hear of distant places that could sustain life, however, we do not typically think it a great shame that there is no-one in these places living good lives.

As with the third case, I believe that this intuition is driven partly by the greater weight we intuitively give to suffering relative to happiness, and partly by status quo bias. If there were an option to create two people - one happy and one unhappy, so that in terms of their contribution to the sum of human welfare the whole thing was a wash - most people, I believe, would be opposed to creating the people. Many people who do not believe economic inequality to be inherently unjust still think that the world would nevertheless in some sense be better if it were more equal, which is to say that the suffering of the poorly off is not entirely compensated for by the happiness of the well-off. This would explain why we care more about the distant suffering than the non-existent, but it still suggests that we would care to some extent about those people who could be happy were they instantiated.

I believe this lack of caring is explained by a bias towards the status quo (Kahneman et al, 1991). We do not think it a shame that non-existent potential happy people are not instantiated, but we think it would be a great shame if actually-existing potential happy people were not instantiated. The German city of Bielefeld is home to some 333,000 souls; it were to turn out that these people had never existed, and that Bielefeld's very existence were a myth, would we not be saddened to discover that the world was smaller than we had thought?

Having suggested that each of Benatar's claimed grounds for belief in asymmetry either fail to support it or are themselves explained by biases in our judgement, I shall now present two arguments in favour of prudential symmetry with regard to procreation. The first is a simple but I think widespread intuition; the second is a longer argument that rejecting symmetry leads to absurdity.

First, and quite simply: I believe that I am better off for existing. Moreover, this seems to be a widespread belief. Indeed I doubt Benatar would deny it, but would be more likely to argue that such a belief is mistaken. In-depth discussion of whether or not we are mistaken will take place in the chapter of this essay on the Quality of Life argument; until then, one would think that introspection and intuition regarding personal cases ought to be worthy of serious consideration, and that there is a strong burden of evidence on the person who believes people claiming to benefit by existing should be ignored.

Second, I present the *transitivity* argument. Consider the three following (non-exhaustive) ways in which someone's existence might have gone. (Let us refer to this person as Hannah).

(a) Hannah exists, and has an absolutely wonderful life.
(b) Hannah never comes into existence.
(c) Hannah exists, and has a life which is good but is not as good as the life in (a).

Furthermore, let us make the following assumptions:
(1) It is possible to compare existence to non-existence: that is to say, any particular existence is better than, worse than, or equally preferable to non-existence. (This does not rule out the quite likely possibility that there is considerable vagueness over when a life becomes worth or not worth living).

(2) The relation "x is better than y" is transitive: that is to say, if x is better than y, and y is better than z, then x must be better than z. (A classic example of a transitive relation is taller-than: If Angus is taller than Billy, and Billy is taller than Carmen, then it follows that Angus must be taller than Carmen).

To reject (1) is to embrace incomparability, as discussed above. In my discussion I noted that we have various reason to avoid this view; moreover, a rejection of the comparability of existence and non-existence is a rejection of pessimistic asymmetry every bit as much as it is a rejection of symmetry, and so there is no particular reason why it should lead to antinatalist conclusions.

(2), the transitivity of the better-than relation, has been argued against by several philosophers - most notably Larry Temkin (1987) and Stuart Rachels (1998). This is not the place to engage in a defence of the transitivity of that relation; I will simply concede that if one accepts Temkin and Rachels' thesis, one is entitled to reject my argument (or perhaps indeed to consider my argument an example of their thesis!)

Providing one accepts that better-than is transitive, one must of necessity also admit that worse-than-or-equal-to is transitive. (Note that this is different from the not-better relation, which is probably intransitive due to there being some things which seem to be genuinely incomparable: it seems reasonable to simultaneously believe that true love is not better than £500,000 and that £1,000,000 is not better than true love, yet £1,000,000 is clearly better than £500,000).

This stated, I give the following deductively valid argument:

(1) If weak pessimistic asymmetry is true, then (a) is worse than or equal to (b).
(2) If weak pessimistic asymmetry is true, then (b) is worse than or equal to (c).
(3) If weak pessimistic asymmetry is true, then (a) is worse than or equal to (c).
(4) It is not the case that (a) is worse than or equal to (c).

---

(5) Weak pessimistic asymmetry is false.

(1) is true on any pessimistic asymmetry, indeed it is precisely what we mean by pessimistic asymmetry: a life cannot be better than non-existence, it can only be worse than non-existence or at best no better than non-existence.

(2) is true on weak pessimistic asymmetry, which allows for the good things in a life to cancel out (but not outweigh) the bad: a good life has a value of 0, rendering it morally equivalent to non-existence (from the perspective of the person living the life). If we specified that lives (a) and (c) contained no suffering at all, being differentiated only by (a) containing a greater quantity or quality of happiness, then (2) would also be true on strong pessimistic asymmetry; however, this would make the case less realistic, and so create a possible escape route for the defender of pessimistic asymmetry.

(3) follows simply from (1), (2), and our assumption that the worse-than-or-equal-to relation is transitive.

(4) is true simply in virtue of the way we have set up the cases.

(5) follows from (3), (4), and the rule of *modus tollens*.

I believe this to be a sound argument against weak pessimistic asymmetry. Since we have previously demonstrated that there is no reason to believe in strong asymmetry, and moreover that it has highly counterintuitive consequences, this leaves symmetry as the last position standing. It is conceptually possible both to benefit, and to be harmed, by being brought into existence.

Conclusion on Asymmetry

There are a variety of different possible asymmetries regarding the potential beneficiality and harmfulness of coming into existence. We should differentiate between “strong pessimistic asymmetry”, according to which existence is inevitably harmful, and “weak pessimistic asymmetry”, according to which existence is never beneficial but need not be harmful. Benatar’s antinatalist conclusions are only directly justified by strong pessimistic asymmetry, a position which we have no reason at all to support. Weak pessimistic asymmetry is more plausible, but for it to support an argument for antinatalism would require significant empirical work alongside it. Furthermore, I suggest we have good reasons to reject pessimistic asymmetry, and instead to accept “symmetry”: the view that people can either benefit or be harmed by being brought into existence.

The Quality of Life argument

A Formal Presentation of the Quality of Life Argument

Having discussed and criticised Benatar's arguments for an asymmetry between the pleasures and pains of existence, I now turn to his second line of argument: the "quality-of-life" argument (Benatar & Wasserman, 2015). As with the asymmetry argument, Benatar does not lay the argument out premise by premise, but it can easily be reconstructed as follows:

1. It is immoral to bring into existence a person who is likely to have a bad life.
2. All people's lives go badly.
3. If all people's lives go badly, a person brought into existence is likely to have a bad life.

---

4. It is immoral to bring people into existence.

I shall discuss premise 1 briefly, and suggest that it is an oversimplification. However, it can replaced with a modified premise 1' that it is immoral to bring into existence a person who is incapable of living a good life. This suggests a recasting of the argument, which I believe Benatar would be happy with and which his arguments can equally well be used to support.

The problem Benatar by his own acknowledgement faces when defending premise 2 is that most people intuitively feel reasonably happy with their lives. Before he can directly defend the premise, then, he needs to provide an argument for why we should trust what he says over our lived experiences. This argument he makes on the basis of various conclusions from psychological research which he believes show that our self-assessments of happiness and satisfaction are unreliable and unduly biased towards positive assessments. I shall argue that this attempt is insufficient, on multiple grounds. First, Benatar makes various hidden assumptions about what makes - or rather what does not make - for a good life, and so fails to show that the psychological phenomena to which he appeals are best seen as biases or as features of our assessments of personal well-being. Second, even if one accepts the required claims about well-being, Benatar fails to provide a compelling empirical case against knowledge

of one's own well-being, due to his utter failure to discuss the size of the biases in question. All of this suggests that we should be wary of accepting premise 2.

Even if Benatar were correct that we have little or no reliable awareness of how well our lives are going, it would remain an open question as to whether our lives actually go well. Benatar's strategy for his direct defence of premise 2 is to consider three leading theories of the good life and to argue that, regardless of which one we accept, we ought to conclude that most lives go badly - or at least, much worse that we typically think. I shall question whether the theories possess the internal resources to answer question of whether lives go well relative to non-existence, and suggest that two of them - the hedonic theory and the objective list theory - give us no grounds for such a comparison. The desire-satisfaction theory does allow for such a judgement, and when properly interpreted it strongly suggests that our lives do, for the most part, go well.

Premise 3 I shall accept for present purposes.

In sum, I conclude that as a result of the failure of premise 2, the quality-of-life argument against procreation fails as a whole: the majority of available evidence suggests not only that most people are glad to exist, but that they are entirely reasonable to feel this way.

Is It Immoral to Bring Into Existence A Person Who Will Probably Have a Bad Life?

Sometimes a person is brought into such an existence that it is impossible for them to live a good life. No matter how they behave, no matter how brave a face they try to put upon those things which happen to them, the possibility of a good life is closed to them right from the start. An unquestionable example of this would be the baby born with a severe congenital disease, which kills her after an exceptionally short life full of pain and misery. Perhaps we might think that many people born into countries afflicted by war and famine are similarly prevented from even a chance at an acceptably good life.

This is not the position of all people, however. A person *E* with no particular disabilities, and equipped with a solid education and upbringing, ought to be able to live a good life. If *E* fails to do so, then the blame for their poor life lies less with their parents than *E* herself. This raises tough questions: if a person *F* may be expected on average to live an unsatisfying life, but *F* will nevertheless have a sufficiently high chance at a good life that *F* will be largely to blame for his

own failure to live a good life, do *F*'s parents wrong him by bringing him into existence? Would it not be deeply paternalistic (in the morally troubling sense) to deny someone a chance at a good life for their own good?

These are heavy issues, but Benatar does not need to take up any position on them. He would instead argue that *E*, regardless of who he or she may be, is incapable of living a sufficiently good life: on Benatar's view, *all* lives go badly and not just those which we ordinarily accept to be of low quality. We may therefore present a revised form of the quality-of-life argument, with the emphasis shifted from wellbeing itself to the possibility of an acceptably good life:

1'. It is immoral to bring into existence a person who has an unacceptably low chance of living a good life.

2. All lives go badly.

3'. If all lives go badly, a person brought into existence has an unacceptably low chance of living a good life.

---

4. It is immoral to bring people into existence.

This argument, like that presented at the beginning of this chapter, is logically valid; the bulk of the work is done in premise 2, and it is that to which we now turn.

Benatar's Case Against Self-Reported Happiness

Most people, as Benatar readily admits, are happy to exist and firmly believe that they are better off for existing. He therefore accepts the challenge of showing that we lack knowledge of our own well-being, and in fact are biased to think our lives go much better than we are. There are a variety of plausible reasons why this might be - for example, that evolution would select for people inclined towards optimism and disinclined towards suicide, or that children will tend to adopt the views of their parents - who presumably thought that their children's lives would be worth living (p.100). However, Benatar focuses upon three particular psychological phenomena which he sees as responsible for throwing off our judgement of how well our lives are going: the "Pollyanna principle", hedonic adaptation, and a tendency to comparison with other people. The first of these causes our judgements to be more optimistic than is merited by the objective

quality of our lives; hedonic adaptation and comparison do not by themselves push in any direction, but they add random variation to our judgements and can amplify the effect of the Pollyanna principle in accentuating the positive bias in our judgements.

Has Benatar Refuted Self-Reported Happiness?

In this section I shall advance two ways - one normative, one empirical - in which Benatar's attempt to discredit our intuitive beliefs about how well our lives are going fails. Following that, I suggest some other ways he could have attempted to make this argument, but conclude ultimately that our judgements - while biased and imperfect - are more useful than Benatar would have us believe.

Firstly, we must recognise that any argument that people are unable to accurately assess their own well-being depends not just on empirical claims, but also upon normative claims. Such an argument must take the form of an argument like this:

(1) Actual human well-being is identical with a property or bundle of properties X. *(Normative premise)*
(2) Individual self-assessments of well-being do not give us knowledge of X. *(Empirical premise)*

---

(3) Individual self-assessments of well-being do not give us knowledge of actual human well-being.

Benatar devotes some energy to the second premise of this argument, but - perhaps because he never sets out the explicit structure of his argument - apparently takes the truth of the first for granted. In his defence, the considerations he raises in defence of premise (2) apply to a very wide range of plausible metrics with which X might be identified. The Pollyanna Principle, is a tendency for humans to emphasise the positive in their thought in a wide range of contexts - Matlin and Stang (year, pp.1-2) list no fewer than 23 different ways in which this pattern displays itself at the outset of their book on the principle, and describe numerous other ways during the course of the book.

There is, however, at least one thing with which we might identify X that is not subject to these concerns: the self-assessment itself. That is, one might argue that the level of well-being of any

human can be identified simply with what that person judges their level of well-being to be. This is not an especially popular theory, but nor (I shall argue) is it clearly false. To demonstrate bias in our judgements of how well our lives are going, Benatar needs to demonstrate that our levels of well-being and our assessments thereof are separate phenomena.

Some very crude forms of what we might call the *pure self-assessment theory* of well-being can clearly be rejected. Suppose a teenager has an argument with his parents and angrily proclaims "I hate my life! I wish I'd never been born!" We would not take this judgement, made in such circumstances, to provide an authoritative view of how the teenager's life is doing. But this does not overturn the pre self-assessment theory; rather, it demonstrates that it should be based not on any old judgement but specifically upon sober and considered judgements. If the teenager, given a few days to calm down and think things over, remained convinced that their life was not, on the whole, worth living - would we not be inclined to take this judgement rather more seriously?

One might think that the pure self-assessment theory says too little about objective standards of living. After all, it seems likely on this theory that an optimist living in extreme poverty would be judged to live a better life than a much richer, healthier, and more accomplished person who is psychologically incapable of being satisfied. Is this an utterly ridiculous implication, though? It seems on the surface very similar to certain ascetic views, such as Buddhism and Stoicism, which suggest that the key to living the Good Life lies not in achievement but in our attitude towards them.

My intention is not to argue that this view is true. Rather, I wish to suggest that it is at least plausible, and that Benatar's argument against self-knowledge with regard to well-being requires the rejection of this position - a case which Benatar does not make. It is not the case that Benatar fails to recognise this as a position - in a footnote on pages 67-8 he makes reference to Richard Easterlin assuming such a position. But rather than argue against it, Benatar simply dismisses it: "But [Easterlin's position] presupposes that there is no difference between one's perceived (subjective) and actual (objective) level of well-being." So Benatar - quite reasonably - rejects the pure self-assessment theory of well-being, but despite it being apparently plausible enough that a fellow academic took it for granted, does not feel the need to provide any actual arguments against it!

A second, and in my opinion more damning, problem with Benatar's attack is his complete failure to consider the effect size of the phenomena he describes. Consider an analogy: there are doubtless certain biases to each individual's perception of colours, due to a range of factors: varying numbers of cone cells in our eyes, for example, or a type of acclimatisation so that we judge the colour of objects relative to other objects around it. Yet the cases in which our perception of colour is unreliable are notable as exceptions to the general rule that we can trust our senses: colour-blind people are a small minority, for example (and even they can tell apart most colours, typically struggling only to distinguish between red and green), or the famous picture of a black-and-blue dress which many people thought to be white and gold[5]. Our perceptions of colour are biased, but they are nonetheless generally reliable.

There are other cases where biases render our intuitive judgements to the point of uselessness. For example, Tversky and Kahneman (1983) famously found that the "representativeness heuristic" causes people to make judgements about the relative likelihood of different circumstances which violate the laws of probability (the famous "Linda problem"). 85% of respondents to a survey judged that "Linda", a woman who had in past years been active in left-wing student politics, was more likely to be a feminist and bank teller than a bank teller, feminist or no. Knowledge in these cases cannot be drawn from intuition or direct perception, but relies upon effortful thought, theorising and measurement.

In between these two extremes are examples where we are subject to biases which throw off our intuitive judgements, but which we are nonetheless able to correct for. An example of this would be a spear-fisher, who consistently perceives fish as lower in the water than they are due to the different refraction of light through air and water. If the fisher stabs at where he perceives the fish to be, he will miss; a skilled fisherman can nonetheless reliably catch fish by stabbing higher than where his eyes tell him. Our perceptions may be consistently biased, but nonetheless useful for ascertaining truth.

Is the bias in our perception of personal well-being like the bias in colour perception, the bias in spear-fishing, or the bias in judging the relative likelihood of aspects of Linda's life and personality? On the information Benatar brings up, there is simply no way to tell. This is problematic, since Benatar clearly believes it is disanalogous to colour and does not seem to

[5] https://en.wikipedia.org/wiki/The_dress

consider that we might be able to correct for Pollyannaism and so continue making use of our intuitions.

Some evidence that these biases do not completely throw off our ability to make useful judgements about well-being comes from the fact that there are significant differences in the average levels of reported happiness and life-satisfaction between citizens of different countries. If these biases were all that strong, we would expect to see different societies having approximately the same reported levels of happiness: there is no particular reason to expect pollyannaism to affect some cultures more strongly than others (particularly if, as Benatar suggests, the existence of Pollyannaism is an evolutionary adaptation which reduces one's risk of suicide - 2006 p.69). If we all return to set points of happiness, then while it would be unsurprising for this set point to vary between individual humans there is no obvious reason why it should be systematically higher in some cultures than others. And if we judge our well-being by comparison to others, then we will judge relative to other people within our community - meaning that once again, there would be no reason to expect significant differences in happiness between societies.

A couple of decades ago, the evidence would have looked very good for this: Richard Easterlin (1974) found that while self-reported happiness was positively correlated with income within societies, there was no clear correlation between living in a wealthier society and being happier. Thus for some years the received wisdom was indeed that we judge our well-being relative to the other people around us. However, more powerful analyses in recent years have consistently found significant differences between the average happiness of citizens of various countries and cultures: Stevenson and Wolfers (2008) found a positive correlation between national average income and national average happiness, a finding that was backed up by the World Happiness Report 2017 (which also found positive associations across societies between happiness and life expectancy). These differences are substantial: on the 2017 World Happiness the average respondent claimed a happiness level of 5.31 out of 10 (Chapter 2, p.14), but the averages for countries and regions varied significantly: North America, Australia and New Zealand had an average happiness of 7.046, Southeast Asia an average of 5.369, and Sub-Saharan Africa an average of only 4.292. The *World Values Study Group* survey of 1994 (quoted in Diener & Suh 1999, p.436) looked at a more restricted range of countries but found similarly large differences: Bulgarians, the least happy nation surveyed, has an average life satisfaction score of 5.03, while at the other end Swiss had an average life satisfaction score of 8.39. The existence of

these significant differences in happiness between members of different cultures provides evidence against the idea that psychological biases prevent us from accurately assessing our own well-being.

Benatar's Argument that Lives Go Badly

Let us now suppose that Benatar is entirely correct about the epistemology of well-being: that individual self-assessments are hopelessly biased, and that the only way we can determine whether lives go well is by agreeing upon a theory of the good life and then applying it to actual lives. There are a variety of theories of the good life which are prominent in the literature, and Benatar considers three particular theories: the hedonic theory, the desire-satisfaction theory, and the objective list theory. I shall briefly explain Benatar's arguments for why, on each of these theories, almost all human lives are bad.

First, there are hedonic theories. These take the measure of well-being to be a function of the subject's mental states. Typically they involve some kind of summation or averaging of various positive mental experiences - joy, satisfaction, interest... - grouped under the heading of "pleasure" and various negative mental experiences - fear, shame, grief… - grouped under the heading of pain. The most traditional hedonic theory, following in the Benthamite tradition, consists in attaching numbers to the pleasure minus pain experienced in each individual moment, summing the values achieved across all moments of a live, and taking this to be the entire value of a life. A hedonic theory need not be of this precise form, however - many have thought that it matters when in one's life utility occurs, or that the pleasures not be too infrequent and the pains not be too intense.

Benatar suggests that much of our waking life is in fact characterised by unpleasant experiences - hunger and thirst, tiredness, itches, and the need to relieve oneself. These are facts of life, requirements which accumulate over time but which we satisfy in an instant. This means that one will generally be hungry for a while before one eats, will suffer bladder distension for a while before one uses the toilet, and tiredness for a while before one is able to sleep (Benatar 2006, pp.71-2). Precisely because these are facts of life we take them for granted, and do not factor them into our assessments of how well we are doing - but on hedonic theories of wellbeing, they ought to constitute significant pains to which we are all subject.

Second, Benatar turns to desire-fulfillment theories, which take the goodness of a life to consist in the fulfilment of certain desires that a person holds or would hold under certain idealised conditions. For unclear reasons Benatar thinks that Pollyannaism will be even stronger on this theory of the good than under hedonism (p.74) - is there any reason why we are more liable to falsely believe that our desires have been satisfied than we are to falsely believe that they have not been satisfied? In any case, Benatar argues that "rather little of our lives is characterised by satisfied desires and rather a lot is marked by unsatisfied desires" (p.74). In the best case, Benatar claims, there is a delay between one's forming a desire and it being satisfied, and so there is some period during which the desire is frustrated. The satisfaction of a desire will often be a long time coming - if indeed it comes at all. Moreover, there is a whole class of desires which will inevitably be frustrated, namely those desires to keep that which we already have (and in particular, life and health).

In addition to all of this, Benatar suggests that we face a "treadmill of desires". If we get the things we want, then they become the new normal: instead of desiring to be rich, one desires to keep one's wealth and furthermore to be respected by one's peers; achieving a desire does not reduce one's worries, but merely frees the space for a new desire to creep into one's mind. We are stuck in a permanent state of striving, with no genuine end to be found.

Finally, there are objective list theories, which define the good life in terms of various things which are considered to be good for us as humans. Examples of these things can include happiness, knowledge, appreciation of beauty, health, and numerous other things. Benatar does not deny that we have access to these things, nor does he suggest that we are especially deluded as the extent we enjoy their presence in our lives. However, he challenges that the scale we use to judge any life good on an objective list theory is incredibly parochial and unambitious. Man may have his fourscore years and ten, but why think that this represents a good life rather than, say, a life which could have been like that of Methuselah but which was tragically short? Or indeed, why not compare our lives to a scale lasting potentially thousands of years. Why think any of our lives are good, rather than to be pitied for their ephemerality?

Second, Benatar raises the question of meaning in life. Most people seem to care that their life have some kind of meaning, and that this meaning be given by someone or something outside of them - community, God, or any other source of ultimate value. "A meaningless life would be lacking an important good, even if it had other goods," Benatar notes.(p.82). But from an impersonal and universalistic perspective, there is nothing special about any of our lives. The

only meaning our lives have are those which we give them ourselves - arguably a pale and weak substitute for full-blooded objective meaning in life.

Does Benatar Succeed in Showing that Most Lives Go Badly?

In this section, I shall argue that the hedonic and objective list theories of the good life lack the resources to compare existence to non-existence. I shall furthermore argue that this need not be the end of these theories, as they can still perform the most important tasks that we require of theories of the good life: providing advice for individuals on how to live, and advice for policy-makers on what they should aim to achieve for citizens of political communities.

The desire-satisfaction theory, by contrast, *does* give us resources to make a comparison between existence and non-existence. However, Benatar's arguments rely upon an esoteric interpretation of this theory, and not one that we have any particular reason to accept. Upon a more plausible understanding of the theory, we end up with no particular reason to think anything more than a minority of lives go badly.

Griffin (1986, p.93) provides a useful typology of the different scales by which things may be ranked. Most basic is the ordinal scale, which allows us to give an ordering to values but has no further meaning; an example of this would be positions in a race, where knowing that Jay came 3rd and Kay came 6th tells us that Jay was faster than Kay, but tells us nothing at all about how much faster Jay was. Beyond this is the interval scale, where each step along the scale represents an equal difference in the quantity being measured: the Celsius scale of temperature measurement is like this, in that the difference in heat between 10 degrees and 30 degrees is the same as that between 25 degrees and 45 degrees, but 30 degrees is not three times as hot as 10 degrees. Going even further we reach the ratio scale, in which not only are the intervals between steps constant but the starting point of our counting matters: the measurement of lengths in centimetres is such a system.

A theory of well-being can be of some use on any of these scales. If one is choosing between two readily achievable alternatives, all one needs to know is which is better - which requires only that they be placeable on an ordinal scale. More complicated decisions will require more complex theories of well-being; for example, choosing between alternatives whose success is

subject to chance. (For example: suppose one is choosing between the fish and the lasagna at a restaurant. One need only know which of the two is better, without any concern for how much better. But now suppose that the chef has a tendency to mess up one's usually preferred option, so that it has a 30% chance of being inedible. Assuming one is indifferent to risk - and a similar calculation must be made if one is not - then one needs now to know whether the difference between the values of the better and worse option is 30% or less of the difference between the better option and going hungry. Hence one needs an interval scale of measurement).

For ordinary decisions, it is usually taken to be the case that an interval scale is fine - indeed, the Independence axiom of expected utility theory makes this fully explicit[6]. However, for decisions concerning existence and non-existence this simply will not do. From the perspective of the individual non-existence has neither value nor disvalue and so may be taken to have a value equivalent to 0, the value of a bad life is objectively negative, and the value of a good life (I argued in the previous chapter) is objectively positive. To assess whether any particular life is better, equal to, or worse than non-existence, we need a theory of well-being which gives lives value on a ratio scale. Are the theories that Benatar discusses amenable to such valuations? I shall suggest that the hedonic and objective list theories are not.

It is in principle quite possible to assess the quantity of pleasure and the quantity of pain in a life, in a fairly directly mathematical fashion. People are quite able to assess momentary happiness on dimensions of happiness and unhappiness, although their later recollection of this happiness can be very fallible (Kahneman 1999, p.5). Although goodness and badness are separate dimensions of momentary assessment (p.12), people are nevertheless usually able to assess whether the good or bad outweighs the other in any particular moment (p.7) and despite the varying qualities of pleasures and pains, they are nevertheless possible to assess along a common measure of intensity (p.10). Despite our changing perception over time of what counts as "zero" for happiness, it is nonetheless meaningful to speak of neither happiness nor unhappiness outweighing the other in the moment. In short, there is little difficulty in constructing a ratio scale of happiness and unhappiness regarding individual moments.

[6] The Independence axiom states that agents presented with a choice of two gambles will not change their rank preference between the two when a third independent gamble is added to each of the original two. This assumption is not without criticism, most famously the Allais paradox.

One can make use of the hedonic theory of well-being without assuming that there is some "exchange-rate" between pleasure and pain, such that a certain amount of pleasure is sufficient to cancel out a certain amount of pleasure and vice versa. However, this use is very limited indeed: one is limited only to saying that a particular life *A* is better than an alternate life *B* only if *A* contains more pleasure and no more pain than *B*, or if *A* contains less pain and no less pleasure than *B*. Given that non-existence contains no pleasure or pain, one is unable to definitively say that any life is better than non-existence if it contains the slightest sliver of suffering, but one is equally unable to definitively say that any life is worse than non-existence if it contains the tiniest morsel of pleasure. In short, to make useful hedonic comparisons between existence and non-existence, we require pleasure and pain to be comparable on the same scale.

In individual moments, we are indeed able to make such comparisons: pleasure and pain are experienced as separate feelings. The possibility of such comparison over extended time periods, however, is much less clear. Indeed, it is hard to say what would make any particular exchange rate "correct" and others wrong. Perhaps we defer to individual judgement - in which case we seem to arrive back at the pure self-assessment theory of wellbeing. Perhaps we compare the absolute length of time in which an individual is happy overall with the absolute length of time in which the individual is unhappy overall - but this ignores the vast differences between moments in our strength of feeling happy or unhappy. Trying to weight this by strength of feeling presupposes that one already has an exchange rate - or else what allows us to say that a moment with happiness *x* cancels out a moment with unhappiness *y*?

Furthermore, it seems entirely likely that as with so many other psychological traits, the exchange-rate that one individual would deem correct will be different from that which another considers to be appropriate. Informally, I am aware both of people who proclaim unwillingness to accept a year of torture even in exchange for a thousand years of constant happiness, and people who would accept less than a year of happiness for a year of torture. We should not place too much stock in such intuitions, which are made for the most part on relatively little consideration and with no experience either of torture or of such sustained happiness; nevertheless, I think the point that we should expect variation in the pleasure-pain trade-offs that people are willing to accept stands.

All this suggests to me that insofar as pleasure and pain can be traded off against each other, the rate at which they do so is probably to a considerable degree subjective. This, in turn,

means that there is no way for us say, even if we accept Benatar's argument that most individual moments - or at least many more than we realise - are on balance painful.

Next, we should consider the objective list theory. I take Benatar to have largely made my argument for me here: objective list theories allow us to compare lives within a group or species, but offer us no way of saying whether lives are "good" in some absolute sense. As Benatar writes, "[objective list theories] are not taken to be objective in the sense of judging what a good life is *sub specie aeternitatis*." (p.82).

It is worth briefly dissecting his attempts to suggest that human lives are nevertheless bad on objective list theories. Would human lives be vastly better for being much longer? I think they would, but it strikes me that to double the length of a life while preserving its average quality does not add a fixed amount to the value of that life, but rather doubles it. Similarly, intensely shortening a life while maintaining the same average quality may greatly reduce the good of the life, but does not change the plus sign in front of its value to a minus (or vice versa). Short human lives are often pitiable because they are shortened by horrendous diseases and sufferings, not because of the shortness in itself. Given this, I do not think the suggestion that our lives are worse for being so short by universal standards holds much water. If we could have had near-infinite lives of happiness, a reduction to our current lifespan would represent an immense loss of value for our lives - but would not, and indeed could not, change our lives from good to bad.

Meaning is a more plausible candidate for something whose absence could render our lives bad. Even then, Benatar's claims seem at best under-defended: he claims "it would surely be much better if our lives had meaning independently of our own human perspective" but does not make any serious attempt to defend it, apparently taking it to be fully intuitive. But leaving all this aside we must wonder: why would human lives be made worse for not possessing something that cannot exist? I take it that either the absence is bad in and of itself - a case which Benatar needs to make instead of relying on intuitions that "life would be better if it were more meaningful" - or it is bad because of its negative effects on other things that matter in themselves - most obviously, our happiness.

Perhaps many people genuinely do find their lives much worse for the absence of an overwhelming narrative provided by a non-human narrative to life. This is already a vastly

weaker claim, however, since it subordinates any claim that our lives are devoid of meaning and therefore of reduced worth to personal judgement. In any case, such an absence is of little use for comparing life to non-existence, as we have seen in discussion of the hedonic theory. If on the other hand the absence of meaning is taken by itself to render lives not worth living, then an argument is required for why meaning is of such ultimate and intrinsic importance to our lives.

I think Benatar and I are largely in agreement, however, that the objective list theory does not possess the internal resources to compare existence with non-existence. We can say that some lives are better than others, which is in many cases sufficient for individuals and for policy makers, but we do not have the kind of ratio scale of well-being that would be necessary for objective comparison with non-existence.

Finally, there is the desire-satisfaction theory. It is true, as Benatar notes, that many of our desires go unfulfilled. It is also true that many are fulfilled. One might adopt *frustrationism*, the view that a life cannot be good and is bad to the extent that it involves unsatisfied desires; however, this is a highly counterintuitive view. In general we take it that developing the capacity to experience certain kinds of fulfillment makes one's life go better: if two lives contain the same amount of bad, but one person develops an appreciation of fine art and is able to fulfil their desire to see it, this life goes better than the other one. The question, then, is how do we know whether the desires which are fulfilled or those which go unfulfilled are stronger?

The clearest way is to appeal to what the individual thinks. And most people have a fairly strong desire which is fulfilled and seems to outweigh for them all the desires which go frustrated: the desire to exist! This is not true of all people, to be sure: some people commit suicide, or are held back from it by the fear of what it would mean for their loved ones. But such people are a small minority: in most people, the fact that we keep on living is strong evidence that we desire to be alive.

One objection to this would be that before one exists, one has no desire to come into existence for the simple reason that one has no desires at all. But this is an overly restrictive conception of the theory: what matters is not that one desires something beforehand, but that if a desire is fulfilled one will appreciate it (Griffin 1986, p.27). Thus one may have no desire to read some classic novel, but having read it be glad to have done so: on a sensible desire-fulfillment theory, this will be considered to have improved one's life. Desires need not be consciously held before

they are satisfied, but may instead be “discovered”. This, incidentally, points to the problem with Benatar’s observation that “no desires for that which we lack are ever satisfied immediately”: this is true, but our desires go far beyond that which we lack, into many things that we take for granted.

A second objection would be that we ought not to count this kind of desire, since it is not desirable in itself but rather a precondition for other things which are desired. Even if one accepts this, I do not think it changes the substantive conclusion that people are typically better off existing: many cognitive capacities are also preconditions for possession and fulfillment of desires, and we would happily acknowledge that if one exists, one is made better off by having these capacities.

Conclusion on Quality of Life

Benatar apparently fails to realise the normative claim implicit in the idea that we do not accurately apprehend our own well-being, and so fails to defend it. This undermines his claim that the Pollyanna principle, comparison, and acclimatization remove our ability to assess our own well-being. Furthermore, he does not look at the sizes of the effects of these on our subjective assessments of well-being: empirically, there are well-established correlations between income, health, and well-being, suggesting that the effects of these phenomena are not so great as to allow our assessment to be thrown out.

The hedonic and objective list theories lack the internal resources for comparison between normal existences (in which there is both bad and good) and non-existence. The hedonic theory does not allow us a complete ordering between pairs of lives where one does not strictly dominate the other by containing more pleasure and no more pain, or less pain and no less happiness. The objective list theory is determined relative to the kind of beings that we are, and so is only suitable for comparison within our particular mode of existence.

The desire theory, however, does allow comparison with non-existence. Once one allows that we “discover” certain desires by having them fulfilled, we realise that many - though by no means all - people have a strong and abiding desire to exist.

Implications of Pro-Natalist Conclusions

In this section I shall discuss what my conclusions in this essay mean for our everyday morality and our politics. Many of the inferences I draw here are highly subject to uncertainty based upon our features of the moral landscape: for example, the debate between deontology, utilitarianism and their rivals is hardly likely to be resolved any time soon, yet it has a strong impact upon whether procreation is generally obligatory or merely supererogatory.

Furthermore, I shall suggest that my conclusions have implications for how we ought to think about procreative beneficence and the alleged duty to bring into being the children capable of living the best lives (Savulescu, 2001). Such a duty may well exist, but so long as children do benefit all-things-considered from coming into existence, this duty is probably not especially strong; moreover, in such cases we may have strong reasons against legally enforcing such a duty.

The morality of procreation

The position I have defended thus far is a fairly conventional one, a defence largely of common sense morality. I defend the fairly intuitive position that bringing people into existence is typically permissible, but impermissible in cases where the child brought into existence would be likely to have a bad life. The conclusion that people are benefited, rather than merely harmed, is not obviously part of common sense, but nor (I would suggest) is its denial. This means that my position is unlikely to have vast implications for how we should change our moral practices and attitudes - if my arguments showed something to be immoral, we would probably already believe it to be immoral.

That said, my arguments are not utterly without implications for how we ought to change our behaviour. I shall suggest that having children is a morally good though not mandatory act, and that we should to be more sanguine about having children who are in various ways impaired or disadvantaged. Adoption, typically taken to be a virtuous act, may in fact be much less morally good than we typically think. Depending upon one's theory of government one might also think that child-rearing ought to be encouraged; however, most popular political philosophies do not have especially strong pro-natalist conclusions, even accepting my argument that most people benefit by existing.

Is Procreation Mandatory, Impermissible, or Something else?

In general, actions are taken to be *prima facie* permissible. Various things might override the permissibility of any individual action - most obviously, harms to other parties caused by the action. The main thrust of this dissertation has been to argue that bringing someone into existence is not generally harmful to the person brought into existence; this leaves open the possibility that it might be harmful to someone other than the person brought into existence.

Might it be mandatory to bring happy people into existence? If it is, this would be a positive duty, and so probably not especially stringent. Most plausibly, I think procreation should be regarded as supererogatory so long as one expects one's potential child to have a life that is on balance worth living - that is, neither forbidden nor mandatory, but better to perform than to not perform.

Sometimes one is obliged to perform actions which improve the well-being of others, but these are generally either actions necessary to avoid harming others - that is, to avoid violating a negative duty - or actions which relieve need or deprivation in another person. A person who does not exist cannot suffer deprivation, indeed cannot suffer at all. Needs can be either instrumental ("One *needs* education in order to acquire intellectually satisfying work") or basic ("It is human nature to *need* food, warmth and community") (Griffin 1986, p.42) but neither of these notions is appropriately applied to the non-existent: one cannot have goals to which other things might be necessary means unless one already exists, or at least have existed in the past. The notion of a basic need is typically explained in terms of those goods without which a typical person cannot achieve an acceptably good life. One reading of this would exclude someone who does not live, suggesting that non-existence constitutes a violation of one's basic needs; I think, however, that this is an example of our usual definitions being less precise than our intuitions. Our intuitive notion of a basic need is that without it one will be suffering deprivation: Griffin notes that the failure to obtain a basic need is connected with "malfunction, harm, or ailment" (p.43) - and just as a non-existent person cannot suffer deprivation, they also cannot lack a basic need on the ordinary understanding of the term.

All this suggests that bringing a happy person into existence should be considered in the same terms as providing a benefit to a person who is already living a quite acceptable life: good, and to be encouraged, but not in general a moral requirement.

Is There A Duty To Provide The Maximum Benefit That One Can?

If one is providing a benefit to another person, and it would be morally option to provide the entire benefit, it is a natural - but, I shall argue, mistaken - move to make to assume that therefore it must be permissible to provide less than the entire benefit. A consequence of this would be that there is no strong moral reason against having children who are some way disadvantaged, so long as their lives are still preferable to non-existence. That is to say, it would generally be morally permissible for a couple to choose to have a child who is missing a limb, or who is blind, even if they could just as easily have a child who will grow up in good health with full use of their body. After explaining how this argument will run, I shall outline a route towards resisting it, and then suggest that nevertheless we might have reasons to respect parental choices in certain cases of this kind.

A thought-experiment (originally due to Matthew Clayton, unpublished) that has appeared in various places concerns a man who is able to risk a drowning woman, but at sufficient risk to himself that the rescue is merely supererogatory rather than mandatory. He is also unwilling to risk damaging the suit which he happens to be wearing, a risk which would be entailed by pulling the woman out of the water by hand. However, his boat has a claw hammer, and it would be possible for him to thrust this into her eye socket and pull her from the water this way. It would be permissible for the man to do nothing to rescue the woman, or to provide the full possible benefit by pulling the woman out of the water by hand. Yet intuitively, it is impermissible for the man to pull the woman from the water using the claw hammer, even though the woman is better off in this case than if she drowns.

Without wishing to commit to any particular explanation of this situation, I would suggest that it provides an analogy for parents who would prefer to have children who are significantly disadvantaged in some way. Creating a child missing a limb, or blind, constitutes a severe harm or a severe lack of respect towards the child which cannot be outweighed by the parents' interest in having a preferred child. That said, while these harms to children are impermissible, an argument from Victor Tadros shows that we may nevertheless have reasons to allow these harms to be caused.

Tadros presents a slight modification of the original case, in which one is on the boat belonging to the man, but is unable to rescue the woman oneself - perhaps being confined to a wheelchair. You know the man's motives, and that he will not rescue the woman by hand; he asks you to pass him the claw hammer so that he may pull the woman from the water. Tadros

observes that although the man's action is impermissible, since the woman benefits one ought nevertheless to aid him in his wrongdoing by passing the claw hammer (2011, p.161). In the same way, so long as the children brought into existence would be expected to have lives worth living, we should tolerate and perhaps even abet the bringing into existence of new disadvantaged people. There is probably some margin along which we should be willing to accept fewer people being brought into existence in order to ensure higher welfare for those people brought into existence, so I would not suggest a complete absence of regulation of which children may be brought into existence. But we may have good reasons for being broadly tolerant of children being brought into disadvantaged existences.

One objection to this would be that the drowning woman in the *claw hammer* case would be actively harmed by being drowned, whereas children not brought into existence would merely fail to benefit. This threatens the analogy between the cases. Perhaps the thought experiment can be changed so that the woman would merely be "failing to benefit" (maybe she attempted suicide by jumping off a bridge, then changed her mind part-way through?) so as to restore the analogy. In any case, I would suggest, as I did in response to Shiffrin's argument in the earlier chapter discussing asymmetry, that this attributes too much significance to the difference between conferring benefits and averting harm. There is an important moral difference between *inflicting harm* and *failing to benefit* someone, but this difference does not entail the former one. To the extent that we have greater reasons to avert harms to others than to provide benefits, it is (I would suggest) due to epistemic concerns - one can generally be confident that one is helping someone by averting a harm to them, whereas there is far less agreement concerning what is beneficial for someone. If I protect someone's property from damage or theft (or gift them replacements for goods which they lost to one of these causes), I can be quite confident that I am protecting something they actually wish to possess. Gifting them the same goods, when they did not formerly possess such goods, I have far lower confidence that I am *genuinely* providing a benefit. When the alleged benefit is agreed by all parties to be genuinely beneficial to the recipient, there is no intrinsic difference from averting a harm to a patient. If this is right, then the analogy between allowing the boat-owner to use a claw hammer to rescue the woman and allowing parents to bring into existence a disadvantaged child holds.

The politics of fertility

The focus of this dissertation has been *child-focused* arguments against procreation, but as the discussion turns to actual practice it would be remiss of me to not mention once again that there are other arguments. Any conclusions reached on the basis of the well-being of people brought into existence are therefore *ceteris paribus*: other considerations might push us to be more pro-natalist (for example, the need to maintain societal institutions such as pensions - Folbre 1994) or more anti-natalist (for example, concerns about environmental degradation and global warming - Conly 2015).

I shall briefly argue that, unless one is a total utilitarian, my conclusions in this dissertation have surprisingly little relevance to politics. When practicing total utilitarianism, it is not difficult to see that we should wish to bring into existence anyone whose life is worth living. Average utilitarianism, on the other hand, suggests that we should bring into existence only those people necessary to achieving the highest average well-being in life; this might be a very large number, due to a larger population allowing greater specialisation and greater gains from trade, or might alternatively be very low due to a higher population meaning smaller individual shares of various scarce resources. Either way, considerations about the extent to which people benefit by coming into existence are likely to be dwarfed by considerations concerning how the total number of people affects average welfare (though facts about who will be happiest to exist may affect *which* people we bring into existence, insofar as it is at all possible to predict someone's happiness in life before they exist).

Rawlsianism, the dominant view in political philosophy, explicitly excludes questions of this sort. In Rawls' formulation of the Original Position, people know that they will be instantiated in society (1971, p.118) even if they nothing about how or when they will come to exist[7]. As such, they would not be concerned with measures to increase (or reduce) fertility, except for the kinds of reasons that an average utilitarian would be concerned about.

If we allow people to be concerned with coming into existence, then the next question of relevance is what we take the distribuenda of justice to be. If it is well-being, then there is a case to be made that so long as there is some class in society who are permanently condemned to

[7] Carl Shulman (2012) provides an interesting discussion of this in "Rawls' original position, potential people, and Pascal's mugging", arguing that potential people would realise that their chances of existing are incredibly low and so would become less concerned with getting to live in society than with aesthetic features of the society.

living lives that are not worth living - perhaps people suffering from certain mental illnesses - Rawlsianism might push in favour of the view that humans ought to extinguish themselves. However, this relies on the claims both that such people form a distinct class within society, rather than merely being members of other classes, and that the main good which society distributes is well-being. Rawls' original formulation held that the distribuenda of justice are "social primary goods" (1971, p.79): goods which are useful to one regardless of one's ends, such as money and political power. On these terms, Michael Bayles (1980, p.117) has argued that the worst off are those who never exist, since they never possess any social primary goods. Benatar (2006, p.181) objects to this on grounds of asymmetry, arguing that a non-existent person can never be deprived and so the non-existent cannot be the worst-off within society; however, I take his arguments for prudential asymmetry to have been defeated in the first chapter of this dissertation.

Might my conclusions here have any relevance to democratic theory? There is an argument to be made that in principle democracies ought to give some weight to the desires and interests of future generations - perhaps immense weight. Abizadeh (2008) argues in a celebrated article that "the *demos* of democratic theory is in principle unbounded", and that political institutions must be justified to all who are subject to them. Abizadeh's article was meant as an attack on democracies which ignore the interests of foreigners, but perhaps a similar move could be made with regard to future people. Future people are not "subject to" our institutions in entirely the same sense, since they are not being directly coerced on a day to day basis. But their lives are greatly affected by our political decisions, both in terms of whether they come to exist and what levels of well-being they will have if and when they come to exist. It is difficult to see how these concerns might be implemented in a real liberal democracy, but they are at least suggestive that, given my arguments that people benefit from coming into existence, democracies ought to be more pro-natalist and more in favour of measures that will increase future well-being (such as preserving a human-friendly environment and encouraging economic growth).

However, this argument is probably unsuccessful. Domestic democracy can actively harm foreigners, whereas by being less pro-natalist than it might otherwise have been it merely fails to benefit future people. If the possibility of being benefitted were sufficient to grant a right to some form of participation in a political community, then the fact that countries can simply make gifts to each other would be sufficient to establish Abizadeh's thesis. The significance of border control for democratic theory comes from the fact that it represents - we may believe - a

limitation on what is available to people beyond what they may reasonably expect to have access to. People do not have any “reasonable expectation” of coming into existence, and so the possibility of their existence being effected or prevented by acts of a democracy does not suffice to give them any right to participation in that democracy.

Conclusion on the morality of procreation

I have argued that, so long as one has good reasons to expect one’s child to have a life that is on balance good, procreation is supererogatory: it is a good thing and to be commended, but one is under no obligation to perform this action. One ought where possible to bring into existence a child without any significant disadvantages (such as missing limbs or congenital diseases), but we have reasons to be somewhat tolerant of parents who nevertheless perform this action so long as their children are still better off existing than not.

The conclusions of this dissertation may have some relevance to the policies of a political community, but this is dependent upon the adoption of particular moral frameworks such as total utilitarianism. Other frameworks, such as standard Rawlsianism, will find no relevance for the claims of my earlier chapters. Moreover, even in the cases where my arguments have relevance, there are multiple other important considerations such as the effects of procreation on the quality of life of already-existing people.

Conclusion

I have argued strongly against Benatar's asymmetry argument and in favour of the proposition that it is conceptually possible to benefit from coming into existence. I rejected Benatar's argument by exposing a logical gap between "weak pessimistic asymmetry", according to which one cannot benefit from coming into existence, and "strong pessimistic asymmetry", according to which one is always harmed by being brought into existence. Benatar's antinatalist conclusions require strong asymmetry, but his arguments only support weak asymmetry. Furthermore, strong asymmetry has several highly counterintuitive implications - for example, that the amount of good in a life is irrelevant to how bad it may be to come into existence.

Weak pessimistic asymmetry is more plausible than strong pessimistic asymmetry, but I nevertheless argued that if the better-than relation is transitive, then we ought also to reject weak asymmetry. Given the obvious falsity of optimistic asymmetries which claim that one can only benefit from coming into existence, and the fact that "incomparability" (the position according to which one can neither benefit nor suffer harm from coming into existence) is useless for guiding our actions, I suggested that we should adopt symmetry: the position that one can either benefit or suffer harm as a result of coming into existence.

After this, I turned to Benatar's argument that most lives are actually bad. Benatar argues that we lack reliable knowledge of how well our lives are going; without denying the reality of the psychological phenomena which he believes disturb our ability to assess our own well-being, I suggested that they may be features of, rather than problems with, our psychology; furthermore, even if they are biases, Benatar failed to show that they are of great significance.

Even if one accepts that we lack individual-level knowledge of how well our lives are going, I then argued that Benatar fails to successfully show that our lives go badly. Hedonic and objective list theories of the good life lack the internal resources to compare life that contain both good and bad with non-existence; desire-satisfaction theories of the good life do, and suggest that most people actively desire to exist.

Finally, I discussed the morality of procreation based on my other conclusions within the essay. Any conclusions here are tentative and reliant upon one's beliefs about ordinary normative ethics; with that caveat, I suggested that bringing predictably happy people into existence is supererogatory. My conclusions here have surprisingly little impact upon the kinds of policies which our institutions ought to pursue.

Bibliography

Abizadeh, A. (2008) “Democratic Theory and Border Coercion”, *Political Theory* vol.36 issue 1

Bayles, M (1980) *Morality and Population Policy*, Tuscaloosa: Alabama University Press

Bayne, T. (2010) “In Defence of Genethical Parity”, pp.31-56 in Benatar, D. & Archard, D. (editors) *Procreation and Parenthood: The Ethics of Bearing and Reading Children*, Oxford: Clarendon Press

Benatar, D. (2001) “Why the Naive Argument against Moral Vegetarianism Really is Naive”

(2006) *Better Never to Have Been: The Harm of Coming Into Existence*, Oxford: Oxford University Press

Benatar, D. and Wasserman, D. (2015) *Debating Procreation: Is It Wrong to Reproduce?*, Oxford: Oxford University Press

Bostrom, N. & Ord, T. (2006) “The Reversal Test: Eliminating Status Quo Bias in Applied Ethics”

Caplan, B. (2012) *Selfish Reasons to Have More Kids*, New York, New York: Basic Books ltd

Clayton, M. (unpublished) “Claw Hammer” thought experiment, discussed in Tadros (2011)

Conly, S. (2016) *One Child: Do We Have A Right To More?*, Oxford: Oxford University Press

Crisp, R. (1988) “Utilitarianism and Vegetarianism” *International Journal of Applied Philosophy*, Vol.4 No.1, pp.41-49

(1997) *Routledge Philosophy Guidebook to Mill on Utilitarianism*, New York, New York: Routledge

Diener, E. & Suh, E. M. (1999) “National Differences in Subjective Well-being”, pp.434-450 in Kahneman, D., Diener, E., and Schwarz, N. (editors) *Well-being: Foundations of Hedonic Psychology*, New York: Russell Sage Foundation

Easterlin, R. (1974) “Does Economic Growth Improve the Human Lot? Some Empirical Evidence”, *Nations and Households in Economic Growth*, Vol.89, pp.89-125

Feinberg, J. (1992) “Wrongful Life and the Counterfactual Element in Harming”, pp.3-36 in *Freedom and Fulfillment*, Princeton: Princeton University Press

Grey, W. (1996) “Possible Persons and the Problems of Posterity”, *Environmental Values*, Vol.5 No.2, pp.161-179

Griffin, J. (1986) *Well-being: Its Meaning, Measurement and Moral Importance*, Oxford: Clarendon Press

Kahneman, D. (1999) “Objective Happiness”, pp.3-25 in Kahneman, D., Diener, E., and Schwarz, N. (editors) *Well-being: Foundations of Hedonic Psychology*, New York: Russell Sage Foundation

Kahneman, Daniel; Knetsch, Jack L.; Thaler, Richard H. (1990). "Experimental Tests of the Endowment Effect and the Coase Theorem". *Journal of Political Economy*. Vol. 98 No. 6, pp.1325–1348

Kahneman, D., Knetsch, J. L., & Thaler, R. (1991) “Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias”, *Journal of Economic Perspectives*, Vol. 5 No. 1, pp.193-206

Matlin, M.E. & Stang, D.J. (1978) *The Pollyanna Principle: Selectivity in Language, Memory and Thought*, Cambridge, Massachusetts: Schenkman Publishing Company

McLean, B. (2015) “What’s So Good About Non-Existence?”, *Journal of Value Enquiry* Vol. 49, pp.81-94

Metz, T. (2011) “Are Lives Worth Creating?”, *Philosophical Papers* Vol.40 No.2, pp.233-255

Parfit, D. (1984) *Reasons and Persons*, Oxford: Oxford University Press

Perry, S. (2014) *Every Cradle is a Grave*, Charleston, West Virginia: Nine Banded Books

Rachels, S. (1998) “Counterexamples to the transitivity of better than”, *Australasian Journal of Philosophy*, Vol. 76 No. 1, pp.71-83

Redelmeier, D. & Kahneman, D.I. (1996). "Patients' memories of painful medical treatments: real-time and retrospective evaluations of two minimally invasive procedures". *Pain*. Vol. 66 No. 1, pp.3–8

Ricon, J. L. (2016) “On the living standards of animals in the United Kingdom”, https://nintil.com/2016/12/17/on-the-living-standards-of-animals-in-the-united-kingdom/, last accessed Sunday 7th May 2017

Savulescu, J. (2001) “Procreative Beneficence: Why We Should Select the Best Children”, *Bioethics* Vol.5 Issue 5-6, pp.413-426

Shiffrin, S. (1999) “Wrongful Life, Procreative Responsibility, and the Significance of Harm”, *Legal Theory* vol.5, pp.117-148

Stevenson, B. and Wolfers, J. (2008) “Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox”, *Brookings Papers on Economic Activity*, Spring 2008 pp.1-87

Tadros (2011) *The Ends of Harm: The Moral Foundations of Criminal Law*, Oxford: Oxford University Publishing

Temkin, L. (1987) “Intransitivity and the Mere Addition Paradox”, *Philosophy and Public Affairs*, Vol.16 No.2, pp.138-187

Tetlock, P. (2003) “Thinking the unthinkable: sacred values and taboo cognitions”, *Trends in Cognitive Sciences* Vol.7 No.7, pp.320-324

Tremblay, F. “I like my life, therefore antinatalism is false!” https://francoistremblay.wordpress.com/2017/03/17/i-like-my-life-therefore-antinatalism-is-false/

Tversky, A. & Kahneman, D.l. (October 1983). "Extension versus intuitive reasoning: The conjunction fallacy in probability judgment". *Psychological Review*. **90** (4): 293–315

https://en.wikipedia.org/wiki/Allais_paradox

https://en.wikipedia.org/wiki/The_dress

https://www.care.org.uk/our-causes/sanctity-life/arguments-for-and-against-euthanasia