

Comment

Thou Shalt Not Create; But If Thou Dost, Thou Shalt Kill? A Comment on Belshaw

Jeff McMahan

Faculty of Philosophy, University of Oxford, Radcliffe Humanities, Woodstock Road, Oxford OX2 6GG, UK; jeff.mcmahan@philosophy.ox.ac.uk

Submitted: 19 April 2024, accepted: 20 April 2024, published: 29 April 2024

Abstract: In his article “Anti-Natalism and the Asymmetry” in this issue, Christopher Belshaw defends the common-sense view that, while there is a reason not to cause individuals to exist whose lives would not be worth living, there is no reason to cause individuals to exist just because their lives would be worth living. But the reasons why these claims are true, he argues, also imply that it is wrong to cause individuals to exist even if their lives would be worth living. I argue that the moral asymmetry between intrinsic goods and intrinsic evils that is the basis of his view is too strong, and that a more defensible view includes a recognition that there is a moral reason to cause well-off individuals to exist, though it is less strong than the reason not to cause an equivalently badly-off individual to exist, and that this weaker asymmetry supports the view that procreation is generally permissible.

Keywords: Christopher Belshaw; the Asymmetry; anti-natalism; procreation; noncomparative benefits; noncomparative harms; Sadistic Conclusion

How to cite: McMahan, J. Thou Shalt Not Create; But If Thou Dost, Thou Shalt Kill? A Comment on Belshaw. *Journal of Controversial Ideas* 2024, 4(1), 6; doi:[10.35995/jci04010006](https://doi.org/10.35995/jci04010006).

©2024 Copyright by the authors. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY 4.0) license.



Most people, or at least most secular people in the West, accept three claims about the morality of procreation. Two of these claims together constitute a view that, in the area of moral philosophy known as population ethics, are often referred to as

The Asymmetry

There is a strong moral reason not to cause individuals to exist when they would have lives that would not be worth living; but there is no moral reason to cause individuals to exist just because they would have lives that would be worth living.

I will refer to the first of these claims as the *Affirmative Claim* and to the second as the *Negative Claim*. The third common-sense claim is that, although there is no reason to cause an individual to exist whose life would be worth living, doing so is nevertheless permissible, other considerations being equal. I will refer to this as the *Neutral Claim*.

In his article “Anti-Natalism and the Asymmetry” in this journal, Christopher Belshaw aims to defend the Asymmetry but to show that the Neutral Claim is false. His argument consists of a complicated dialectic consisting variously of challenges to the three claims, defenses against the challenges, responses to the defenses, and so on.

Belshaw’s core argument begins with a challenge to the Affirmative Claim of the Asymmetry – namely, David Heyd’s “actualist” claim that only individuals who are at some time actual matter morally, from which it is thought to follow that there cannot be a moral reason not to cause a person to exist. Belshaw responds to this claim by observing that it seems obvious that, if the individual we might create would have a life that would be so bad that there would be a strong moral reason to euthanize that individual as soon as possible after it began to exist, then there must also be a reason not to create the individual in the first place. This response might be strengthened by noting that Heyd’s rejection of the Affirmative Claim implies that there would be no reason not to cause an individual to exist whose life would contain nothing but intense suffering even if neither euthanasia nor suicide would be possible, so that the individual would be condemned to existing for decades in agony. That is impossible to believe. And the best explanation of why there *is* a reason not to cause a miserable individual to exist is that to do so would be *bad for* – though not be *worse for* – that individual.¹

Belshaw next perceptively and rightly points out that this defense of the Affirmative Claim, if correct, threatens to undermine the Negative Claim; for if the reason to end a bad life as soon as it has begun implies a reason not to create that life, then the reason to save a good life as soon as it has begun should also imply a reason to create it.

Belshaw seeks to rescue the Negative Claim, while nevertheless preserving his defense of the Affirmative Claim, by arguing that there is a moral asymmetry between immediately ending a life that will not be worth living and immediately saving a life that will be worth living. There is, he argues, a reason to do the former, but no reason to do the latter. This is because “death is bad ... only when ... [the life it ends] is one that its possessor wants to continue.” In short, an individual’s desiring to continue to live, or not to die, is a necessary condition of death’s being bad for that individual. Assuming that we were once fetuses and then neonates, it follows that death would not have been bad for us at those ages, as we were then incapable of desiring to continue to live. And, if death is not bad for a fetus or neonate, there is no reason to save its life for its own sake. But, if there is no reason to save its life once it has begun to exist, there is surely no reason to cause it to exist in the first place. The threat to the Negative Claim posed by Belshaw’s defense of the Affirmative Claim is thus eliminated if it is true that death cannot be bad for an individual that has no desire to continue to live.

There is, however, another move in the dialectic that Belshaw does not consider. If death can be *bad* for an individual only if that individual desires to continue to live, then it should also be true that death can be *good* for an individual only if that individual desires *not* to continue to live. But most or all animals, and certainly all fetuses and neonates, lack a sufficient degree of self-consciousness to be capable of desiring not to continue to live. If this is right, then, just as death cannot be bad for a fetus or neonate, so it cannot be good for a fetus or neonate. And if it cannot be good for a fetus or neonate to die, there can be no reason to end its life for its own sake, even though its continued life would be intrinsically bad, or not worth living. But if there is no reason to end its life once it exists, then it seems, by the logic of Belshaw’s own argument, that there is no reason not to

¹ See Jeff McMahan, “Problems of Population Theory,” *Ethics* 92 (1981), p. 105.

cause it to exist. Belshaw's defense of the Negative Claim thus seems to undermine the Affirmative Claim, which he rightly believes to be true.

Belshaw has a reply to this, though he does not state it explicitly. If an individual's future life would be intrinsically bad, or not worth living, it will necessarily contain intrinsic evils. These, he believes, are bad for an individual whether or not the individual desires that they not occur; thus, there is a reason to prevent them from occurring even if the potential victim does not desire in advance that they not occur.

These claims need to be examined more carefully. Death, Belshaw claims, is not bad for an individual unless it frustrates some of the victim's desires – in particular, the categorical desire to continue to live. But suffering is bad independently of any desire of the victim to avoid it, and there is always a reason to prevent it, provided that the victim neither deserves nor is morally liable to experience it. These claims raise questions that Belshaw does not explicitly address. In what cases are an individual's desires necessary for some state or event to be intrinsically good or bad for the individual, and what are the connections between a state or event's being good or bad for an individual and there being a moral reason to cause or not to cause, or to prevent or not to prevent, that state or event?

Death is not an intrinsic evil. It is bad primarily because it prevents an individual from having further intrinsic goods that the individual would otherwise have had. Suffering, by contrast, *is* an intrinsic evil. It is thus a conceptual truth that it is bad for the victim when it occurs. So perhaps Belshaw is assuming that, while intrinsic evils are always bad for the victim, intrinsic goods are good for the individual who has them only if that individual has desired them. But that cannot be right; for just as it is a conceptual truth about intrinsic evils that they are bad for the individual who suffers them when they occur, so it is a conceptual or necessary truth about intrinsic goods that they are good for the individual in whose life they occur when they in fact occur. Just as it makes no difference to the badness for an individual of an intrinsic evil whether that individual desired to avoid it, so it makes no difference to the goodness for an individual of an intrinsic good whether that individual desired it before it occurred.

So Belshaw's assumption about the relevance of desire must be different. It is not, I think, an assumption about the relation between desire and what is good or bad for an individual at a particular time. It is, instead, an assumption about the relation between desires and reasons. This is suggested by the brief discussion of one intrinsic good that Belshaw discusses: pleasure. He writes that,

even if pleasure is intrinsically good – good because of how it is in itself, its immediate raw feel, rather than because of its effects – there is no reason to hold also that it is intrinsically valuable, so that the more of it the better. So there is no reason to bring into or keep in existence all the locations, all the organisms where pleasure might be found.

The radical asymmetry that Belshaw proposes thus seems to be that, whereas intrinsic evils are always *disvaluable*, in that there is always a reason to prevent or eliminate them, intrinsic goods are not correspondingly *valuable*, as there is no reason to produce or sustain them, *unless they are desired*.

This view has one immediate implication that many of us will find quite implausible. Suppose that a couple want to have a child but the woman has a condition that will cause any child she conceives to have almost no sense of smell, so that this individual would be unable to enjoy food as much as people with a normal sense of smell can. If, however, she takes a medication for six weeks before conceiving a child, her condition will be cured.

But the couple are eager to have a child as soon as possible and so conceive a child now with a diminished sense of smell. Most of us think that she ought to have taken the medication, and then conceived a child with a normal sense of smell. But this would have been a different child and the only predictable difference between its life and that of the couple's actual child is that the alternative child would have had more intrinsic good in its life. And on Belshaw's view, the couple had no reason to conceive a different child who would have had more intrinsic good in its life rather than their actual child, who will have less; for the child who could have had more good in its life did not exist and in fact never will exist.

According to Belshaw, intrinsic goods have neither a reason-giving function nor an offsetting function *unless they are desired*. This is the basis of his claim that there is a moral reason to kill not only all animals that are incapable of having categorical desires, but also all human beings until they become persons and can have categorical desires for intrinsic goods. Given the radical asymmetry he assumes there to be between intrinsic evils and intrinsic goods, there are reasons to prevent the evils that individuals of these types will suffer. But, he writes, "as a counter to these reasons," the "pleasures to come are ineffective." This is because any intrinsic goods that these individuals will experience when they are not persons cannot offset the evils, as the goods are normatively inert unless they are desired. This is true, he says, even if the goods greatly outweigh the evils.

These are highly counterintuitive conclusions. Indeed, expressed in his preferred terms of prohibitions and requirements, Belshaw's claim is that allowing individuals of these types to continue to live is prohibited. Killing them is required. He writes, for example, that "We shouldn't start the lives of ... babies, and if started then – so long as we are thinking just of them – we should end them." We learn at the end of the essay, however, that killing infants is not really required, as we are never thinking just of them. There he writes that "other-regarding reasons are often important reasons." If those of us who are persons want to have animals around or want to have children, these reasons of self-interest are, it seems, sufficient to override the moral reasons to kill animals, fetuses, and infants. So, at least in their implications for matters of practice, his conclusions are, he suggests, not necessarily counterintuitive after all.

This may, however, be overly complacent. Belshaw never explains how reasons of self-interest can override the reason not to inflict repeated, uncompensated suffering on an innocent infant. I therefore think his claims about killing are sufficiently counterintuitive that we should reject them unless there is a highly convincing argument for accepting them.

But Belshaw does not present an argument or justification for the claim that there is always a reason to prevent, or not to cause, an intrinsic evil, though no reason to produce, or not to prevent, an intrinsic good, unless that good is desired. And while the claim that there is no *moral* reason to produce undesired intrinsic goods for people in the future seems dubious, the claim that there is no *prudential* reason to produce undesired intrinsic goods for oneself in the future seems clearly false.

In the latter case, one may not desire some intrinsic good because one mistakenly believes that it would not be good for one. One nevertheless has what Bernard Williams calls an "external reason" to bring it about.² Indeed, one has an external reason to *desire* it. That the intrinsic good would be *good for* one does not depend on whether one desires it in advance. When Belshaw acknowledges that one might desire to live to 90 even when

² Bernard Williams, "Internal and External Reasons," in *Moral Luck: Philosophical Papers 1973–1980* (Cambridge: Cambridge University Press, 1981), pp. 101–113.

that would not be good for one, he is conceding that what makes something good for an individual is not that the individual has desired it or desires it at present.

It also seems clear that there are moral reasons to produce intrinsic goods for others in the future even if those others do not at present desire them. Suppose there is an infant with no desires for its future life. This infant may or may not live to the age of 30. One can now, at no cost to oneself or anyone else, ensure that this infant, if it reaches the age of 30, will then have substantially more intrinsic goods in its life than it otherwise would. And one's action would not affect this individual's life in any way before it would produce the additional goods. There is, I believe, a moral reason, even if not a duty, to bring about the additional goods, even though the infant is incapable of desiring them. This suggests that there is also a reason not to kill the infant, and a reason to save its life. For, if there is a reason to enable it to have *more* intrinsic goods in the future rather than *fewer*, there should also be a reason to enable it to have *some* rather than *none*. And if there is a reason to produce undesired future goods for an infant, then it seems that future goods should be able to offset intrinsic evils in the infant's present life as well.

Matters are, however, more complicated than this. According to a view I have defended, to which Belshaw refers on occasion, whether and to what extent an individual has a present interest in having some intrinsic good in the future depends on whether and to what extent various psychological relations would hold between the individual now and that same individual when that the good would occur.³ On this view, an intrinsic good that would occur when an individual is 30 could not *compensate* that same individual for an evil experienced in infancy; for the virtual absence of psychological relations between the infant and its later self at 30 means that the infant has, if any interest at all, an interest of only negligible strength in having the later good. Even though the infant and the 30-year-old would be the same individual, the 30-year-old would be so psychologically remote from the infant as to be relevantly like an entirely different individual.

These considerations do not, however, support Belshaw's view. His view is that an intrinsic good in the life of an infant that would occur no more than an hour later than a lesser intrinsic evil cannot *offset* that evil, thereby making it the case that on balance there would be a reason to prevent both the evil and the good by painlessly killing the infant rather than to allow both to occur. The good is "ineffective" "as a counter to" the evil. He recognizes, however, that infants are "stuck in the present," without future-directed desires or relevant psychological relations to themselves even in the very near future. The reason to prevent an infant from suffering tomorrow cannot, therefore, be a reason to act *for its own sake now*. Whereas pleasure is not, Belshaw claims, intrinsically *valuable* – by which he means "the more of it the better" – suffering is intrinsically *disvaluable*, in that the less there is of it, the better, so that there is always a reason to prevent it – *even* if the only way to prevent it is to prevent the existence of the individual who would experience it. The reason to prevent suffering, and other intrinsic evils, must, therefore, be impersonal. But, if this is so, then it should be possible, in principle, for the impersonal badness of suffering to be offset by the impersonal goodness of intrinsic goods. Of course, Belshaw denies that there are impersonal reasons to produce intrinsic goods. This is, again, the radical asymmetry between intrinsic evils and intrinsic goods that is the basis of his arguments for the anti-natalist claim that there is no reason to cause individuals to exist who would

³ Jeff McMahan, *The Ethics of Killing: Problems at the Margins of Life* (New York: Oxford University Press, 2002), pp. 69–86 and 165–174.

be well off and for the “pro-mortalist” claim that there is a reason to kill animals, fetuses, and infants.

But, as I noted, Belshaw does not provide an argument for this radical asymmetry, which seems too strong. If there is an impersonal reason to prevent intrinsic evils even when the individual who would suffer them does not exist and thus cannot desire that they not occur, why is there *no* impersonal reason, not even a much weaker one, to produce intrinsic goods of equivalent magnitude unless the potential beneficiary exists as a person and desires them? Belshaw concedes that we all rationally accept certain intrinsic evils, such as a “debilitating operation,” in order to have greater intrinsic goods by living longer. If there is no radical *intrapersonal* asymmetry between intrinsic evils and benefits, why should we suppose that there is a radical impersonal asymmetry?

I noted earlier that there is a view about the nature of interests that I have defended that implies that a good that an individual experiences at the age of 30 cannot compensate that same individual for an evil suffered in infancy. I originally developed this view as an account of the misfortune of death. It offers, I believe, a better understanding of the matters that Belshaw discusses than his desire-based account of the misfortune of death.

My view, to which I gave the label the “Time-Relative Interest Account” of the misfortune of death, is that the extent to which death is bad for an individual is a function of two variables: the goodness of the life that death prevents an individual from having and the degree to which various psychological relations, to which I refer as the “prudential unity relations,” would have held between the individual at the time of death and that same individual during the later life that death prevents. On this view, even though death may prevent a fetus or infant from having a very long and good life, the fetus or infant would be unconnected to itself in the future by any of the central prudential unity relations, such as memory and continuities of character, desires, beliefs, values, and so on.

A fetus or infant would, however, be connected to itself both in the immediate future and in the distant future by what I refer to as “continuity of consciousness,” by which I mean that all its experiences would occur in one and the same consciousness, or the same mind. Continuity of consciousness, I have argued, is sufficient for interests of at least minimal strength in events in at least the relatively near future.⁴ This is supported by Bernard Williams’s thought experiment in which one knows that all the elements of one’s distinctive psychology (such as one’s beliefs, desires, character traits, and memories) will be expunged over the next few hours while one remains conscious, after which one’s body will be subjected to horrible tortures.⁵ It seems prudentially rational to fear the torture even though the only psychological relation between oneself now and the conscious individual a few hours from now will be continuity of consciousness. So just as, on the view that continuity of consciousness is sufficient for minimal egoistic concern about the near future, an infant has an interest – albeit a comparatively weak one – in not experiencing suffering tomorrow, so it also has a comparatively weak interest in having intrinsic goods, such as pleasure, tomorrow. On this view, therefore, the intrinsic evils suffered by fetuses and infants are normally offset by greater intrinsic goods. Fetuses and infants, in other words, normally *are* compensated for the evils they suffer by also having greater intrinsic goods that they experience either shortly before or shortly after they suffer the evils. It is just that the compensating goods must occur in temporal proximity to the evils rather than in

⁴ *The Ethics of Killing*, p. 75.

⁵ Bernard Williams, “The Self and the Future,” in *Problems of the Self* (Cambridge: Cambridge University Press, 1973), pp. 46–63.

the later future. According to Belshaw's view, by contrast, a fetus or infant that suffers a pain is not compensated to any extent by a greater pleasure either an hour earlier or an hour later.

The Time-Relative Interest Account does, though, have implications for the badness of death for fetuses and infants that are similar to, though I think more plausible than, those of the desire-based account that informs Belshaw's arguments. Because the psychological relations that connect a conscious fetus or infant to itself in the future are few, weak, and relatively unimportant, these individuals' interest in continuing to live is very weak; hence death cannot be a substantial misfortune for them. It is possible, of course, that they nevertheless have significant moral status or a right not to be killed that makes killing them seriously wrong even though being killed would be only minimally bad for them. I think there are no good grounds for attributing such status to them but I cannot argue for that here.

There are several reasons why I think the Time-Relative Interest Account is more plausible than the view that death is bad only if it frustrates categorical desires. Not only can it explain why there is seldom a moral reason to end the life of a fetus or infant, but it also offers a plausible justification for the common-sense view that death becomes a greater misfortune for a fetus or infant the more psychologically developed it becomes. It therefore has more plausible implications about, for example, the death of an 18-month-old child. It is doubtful that such a child can have any conception of death. If not, it cannot desire to continue to live. It does, presumably, have some desires for the immediate or near future, and the fact that death would frustrate those desires might mean that, on Belshaw's view, death would be minimally bad for it. But even if Belshaw's view does allow for death to be bad because it frustrates desires other than the desire for continued life, it still implies that death for an 18-month-old child can be no worse than the frustration of a few comparatively trivial desires. And this seems implausible.

The Time-Relative Interest Account, by contrast, can recognize that such a child would be connected to itself later as an older child and even as an adult by continuities of memory (the individual at two years has memories from 18 months; at five years, memories from two years; and so on) and continuities of character traits, desires, beliefs, values, etc. Assuming that the 18-month-old child's future life would be long and well worth living, these forms of psychological continuity are sufficient to give the child a relatively weak but not negligible interest in continuing to live. Because death would frustrate this interest, death would be bad for the child. This interest in continuing to live thus grounds a reason to preserve the infant's life that is substantially stronger than the reason to ensure that whatever desires it might have for objects in the near future are not frustrated.

I have argued that the moral asymmetry between goods and evils in Belshaw's view – namely, that there is no reason to produce intrinsic goods by causing individuals to exist, though there is a reason *not* to cause intrinsic evils by causing individuals to exist – is too strong. There is, however, a moral asymmetry, though it is difficult to identify its precise nature. In the remainder of this essay, I will gesture towards a more modest and, I believe, more plausible asymmetry between intrinsic goods and intrinsic evils. (All the general claims I will make should be assumed to be qualified by an implicit *ceteris paribus* clause.)

I refer to intrinsic goods and evils produced by causing individuals to exist as “noncomparative benefits” and “noncomparative harms” respectively. They are “noncomparative” because, although they are good or bad for the individual in whose life they occur, they are not better or worse for that individual. This is because the alternative to their occurring is for the victim or beneficiary never to exist. (Here I ignore cases in

which there are three options: causing an individual to exist with a particular good or evil in that individual's life, causing the same individual to exist without that good or evil, and not causing the individual to exist.) Being caused to exist cannot be better or worse than never existing; for if it were, that would imply that never existing would be worse or better; but there cannot be anyone for whom never existing could be worse or better. By contrast, I refer to goods and evils caused in the lives of individuals whose existence is independent of the act that causes those goods and evils as "comparative benefits" and "comparative harms" – for these are better or worse for those in whose lives they occur.

I believe, though I cannot argue for it here, that there is a moral reason to confer a noncomparative benefit by causing an individual to exist.⁶ That is, I reject the Negative Claim. Because this reason is stronger the more net good the individual's life would contain, there is a reason – contrary to what Belshaw's view implies – to cause a better-off individual to exist rather than a different, less well-off individual, when the only difference between the two lives is that the better-off individual's life would contain more intrinsic goods. But the reason to confer a noncomparative benefit (an intrinsic good) by causing an individual to exist is weaker than the reason to confer an equivalent comparative benefit on an existing individual or an individual in the future whose existence is independent of the conferral of the benefit. This is because the reason to confer a benefit, or produce an intrinsic good, is stronger when conferring it would be better for the beneficiary, so that not to confer it would be worse for that individual.

By contrast, the reason not to inflict a noncomparative harm by causing an individual to exist is not weaker, or not significantly weaker, than the reason not to inflict an equivalent comparative harm on an individual whose existence is independent of the infliction of the harm. So one important asymmetry between intrinsic goods and evils is that, whereas it makes a moral difference whether the absence of an intrinsic good would be worse for someone, it makes no difference, or comparatively little difference, whether the occurrence of an intrinsic evil is worse for someone.

Another asymmetry between goods and evils that is widely recognized is that the reason to cause a comparative benefit or intrinsic good in the life of an existing individual is weaker than the reason not to inflict an equivalent comparative harm or intrinsic evil on an existing person. (A good and an evil are of equivalent magnitude if it would be prudentially rational for a person to be indifferent between having both the benefit and the harm and having neither.) Because the reason not to inflict a noncomparative harm is as strong, or almost as strong, as the reason not to inflict an equivalent comparative harm, the asymmetry in strength between the reason to confer comparative benefits and the reason not to inflict comparative harms is weaker than that between the reason to confer a noncomparative benefit and the reason not to inflict a noncomparative harm.

These claims yield the following hierarchy of reasons, ordered from the strongest to the weakest type of reason:

- (1) the reason not to inflict a comparative harm (which would be worse for the victim)
- (2) the reason not to inflict an equivalent noncomparative harm (which would not be worse for the victim)
- (3) the reason to confer an equivalent comparative benefit (which would be better for the beneficiary)
- (4) the reason to confer an equivalent noncomparative benefit (which would not be better for the beneficiary)

⁶ I have argued for this claim in "Human Extinction and the Morality of Procreation," (unpublished).

There is, however, a further complication that arises if the Time-Relative Interest Account is correct. Suppose that there is a fetus that has only recently become conscious. It will be psychologically connected to its future self as an adult only by continuity of consciousness. There will be no other prudential unity relations between the fetus now and itself as an adult. Suppose that one can act now in a way that will enable the adult into whom the fetus will develop to have a very great intrinsic good that this adult would not otherwise have. Although there is a moral reason to produce this good, the Time-Relative Interest Account implies that this reason is *much* weaker than it would be if the act that would produce it could be done only many years later, when the beneficiary would be an adult and the good would occur shortly after the act was done. This is because the fetus has at most an interest of negligible strength in having the later good, whereas the same individual would, as an adult, have a very strong interest in having it. When the individual is only a fetus and has either no present interest or only a negligible present interest in having the good, the reason to produce that good is only slightly stronger than a reason to produce an equivalent good by causing an individual to exist. In short, the reason to produce a comparative benefit when the potential beneficiary has virtually no interest in having it is only slightly stronger than the reason to produce an equivalent noncomparative benefit.

Yet a further asymmetry between goods and evils follows from three claims I have made earlier:

- (1) The reason not to inflict a noncomparative harm is as strong, or almost as strong, as the reason not to inflict an equivalent comparative harm. (In other words, it makes little or no moral difference whether the infliction of an intrinsic evil is worse for the victim.)
- (2) A fetus with the capacity for consciousness will be psychologically connected with itself as an adult only by continuity of consciousness and therefore, according to the Time-Relative Interest Account, can have a present interest of only negligible strength in having goods or avoiding evils when it becomes an adult.
- (3) Affecting a fetus now in a way that will cause it to have an intrinsic good or suffer an intrinsic evil only after it becomes an adult is therefore morally little or no different from doing what will both cause an individual to exist and cause that individual to have the same good (a noncomparative benefit) or suffer the same evil (a noncomparative harm) as an adult.

Assuming that the reason not to inflict a significant intrinsic evil on an innocent adult is very strong, then, given claim 1, the reason not to inflict an equivalent evil by causing someone to exist is also very strong. Given claim 3, it follows that the reason not to do to a fetus what will cause it later to suffer an equivalent evil is also very strong, despite the negligible strength of the prudential unity relations and despite the fact that the fetus thus has only a negligible present interest in avoiding that future evil. We can therefore infer that, although the relative strength of the prudential unity relations significantly affects the strength of the reason to produce intrinsic *goods* in the lives of existing individuals, the strength of these relations does *not* significantly affect the strength of the reason not to cause intrinsic *evils* in the lives of existing individuals.

In short, the reason to produce an intrinsic good or benefit is substantially stronger if the benefit would be comparative rather than noncomparative. And if the good or benefit would be comparative, the strength of the reason to produce it varies with the strength of the prudential unity relations between the beneficiary at the time of action and the beneficiary at the time when the good would occur. The stronger the relations, the stronger the reason to produce the intrinsic good. But the reason not to cause an intrinsic evil or

harm is no stronger, or not much stronger, if the harm would be comparative than it would be if the harm would be noncomparative. And if the evil or harm would be comparative, the strength of the reason not to cause it varies only slightly, if at all, with the strength of the prudential unity relations between the victim at the time of action and the victim at the time when the evil would occur. In other words, the strength of the reason not to inflict an intrinsic harm on an existing individual varies only slightly, if at all, with the strength of that individual's interest in avoiding that harm.

I suspect that the reason not to inflict an intrinsic harm is indeed somewhat stronger if the victim has a present interest in not suffering it, and that the strength of the reason varies at least to some extent with the strength of that interest. If this is right, then it does make a moral difference whether an intrinsic harm would be worse for the victim and, therefore, whether the harm would be comparative or noncomparative.⁷ But these considerations matter rather little in comparison to the extent to which they matter in the case of acts that produce intrinsic goods.

I wish I could claim that the view I have just outlined has implications for procreation that are not only more plausible than those of Belshaw's view but also entirely consistent with common-sense moral intuitions. But, without further refinements or qualifications, this view, too, has implications that many will find counterintuitive. First, and obviously, the view I have sketched asserts that there is a reason to cause well-off individuals to exist just because they would be well off. Some people find this plausible (though most of these recoil at the idea that this reason, if unopposed by conflicting reasons, rises to the level of a duty). But most people, it seems, reject this claim.

Second, the view implies not only that there is a reason to cause a well-off individual to exist rather than not to cause anyone to exist but also that there is a reason to cause a better-off individual to exist rather than a different, less well-off individual. Yet the second of these reasons seems to be far more widely accepted than the first. And even among those who accept that there are both of these reasons, most believe that the second is stronger than the first. Yet it is hard to see how that could be true. The second is a reason to produce a greater good rather than a lesser good, whereas the first is a reason to produce a good – perhaps the same greater good – rather than not produce any good at all. When the reasons are described in this way, it seems that the second reason should be *stronger*. Surely it could not be *worse* to produce the lesser good (the less well-off individual) than not to produce any good at all.

Third – though probably not finally, as there are bound to be further implausible implications of which I am unaware – the view I have sketched may imply what Gustaf Arrhenius calls a “Sadistic Conclusion.”⁸ The view asserts that the reason not to cause noncomparative harms is substantially stronger than the reason to produce equivalent noncomparative benefits. This suggests that for it to be permissible to cause the inevitable noncomparative harms in causing someone to exist, the noncomparative benefits one would also cause must be substantially greater. But this implies that it would be wrong to cause a person to exist if the benefits would not exceed the harms to a sufficient degree – even if the person's life would be worth living. But if there is a moral reason not to cause one such person to exist, the reason not to cause two, or 100, to exist should be even

⁷ Others disagree. Johann Frick, for example, writes that “Our moral reasons against creating new lives that are miserable seem just as weighty as our reasons against making existing lives miserable.” “Conditional Reasons and the Procreation Asymmetry,” *Philosophical Perspectives* 34 (2020), p. 5.

⁸ Gustaf Arrhenius, “An Impossibility Theorem for Welfarist Axiology,” *Economics and Philosophy*, 16 (2000), pp. 247–266.

stronger. Thus, for some number of such people who would have lives worth living, the reason not to cause them to exist must be stronger than the reason not to cause a single person to exist whose life would be intrinsically bad, or not worth living. This is a “Sadistic Conclusion.” While it can be argued that the view I have sketched does not actually entail such a conclusion, I cannot develop those arguments here.⁹

Despite these serious problems, the view I have sketched still seems to me the most plausible view of the morality of procreation. Like Belshaw, I have mainly outlined my view and indicated what some of its implications may be rather than arguing for it at length. That the view has some implications that conflict with common-sense intuitions should not be surprising. If decades of work in population ethics have revealed one certainty, it is that common intuitions about the morality of causing people to exist are deeply confused. Our various common beliefs are probably irreconcilable and therefore collectively incoherent. Some will have to be abandoned, or at least heavily revised. But Belshaw’s view, while perhaps coherent, abandons more, and more fundamental, moral beliefs – for example, that it is wrong to kill babies and animals who would have lives worth living – than I think is likely to be necessary to achieve coherence.

Funding: This research received no external funding.

Acknowledgments: I am grateful to Fiona Clarke for discussion and helpful comments.

Conflicts of Interest: The author declares no conflict of interest.

⁹ I have attempted to do so in “Replies to the Commentaries,” in Jeff McMahan, *Wrongful Life*, ed. Hon-Lam Li (Cham, Switzerland: Palgrave Macmillan, forthcoming).