

Asymmetric Population Axiology: Deliberative Neutrality Delivered

Abstract. Two related asymmetries have been discussed in relation to the ethics of creating new lives: First, we seem to have strong moral reason to avoid creating lives that are not worth living, but no moral reason to create lives that are worth living. Second, we seem to have strong moral reason to improve the wellbeing of existing lives, but, again, no moral reason to create lives that are worth living. Both asymmetries have proven very difficult to account for in any coherent moral framework. I propose an impersonal population axiology to underpin the asymmetries, which sidesteps the problematic issue of whether or not people can be harmed or benefited by creation or non-creation. This axiology yields perfect asymmetry from a deliberative perspective, in terms of expected value. The axiology also yields substantial asymmetry for large and realistic populations in terms of their actual value, beyond deliberative relevance.

1. Introduction

Much recent work on the ethics of procreation is focused on what Jeff McMahan dubbed the *Asymmetry*: We have strong moral reason to avoid creating a life that is not worth living, but no moral reason to create a life that *is* worth living (McMahan 1981, 100; for a more recent reformulation, see McMahan 2009, 49). This asymmetry is intimately related to another asymmetry, which McMahan critically investigates in a recent article: We have strong moral reason to increase wellbeing by making a presently existing life better, but no moral reason to increase wellbeing by creating a new life with positive wellbeing (McMahan 2013).¹ Both of these asymmetries are discussed and endorsed by Jan Narveson in his early investigations of population ethics (1967, 1973). They are intimately related. Both asymmetries have as one of their relata additional positive lives, or what I will call "positive additions", that is lives with positive lifetime wellbeing that are added to the population. In the first asymmetry, we are comparing positive additions to what I will call "negative additions", that is lives with negative lifetime wellbeing that

¹ McMahan does not endorse these asymmetries. His treatment is more exploratory than argumentative. He does tend to group lives that will exist in the future with presently existing lives, drawing a line between these two classes on the one hand and on the other lives that may or may not exist in the future (e.g. 2013, 11-12). I will not distinguish between future lives that will and future lives that may or may not exist. One reason not to is that from a collective, deliberative perspective, there are no future lives that will exist regardless of what we do - any future life can be prevented and is in that sense contingent.

are added to the population.² In the second asymmetry, we are comparing positive additions to improvements in the wellbeing of presently existing lives. Regarding the second asymmetry, Narveson famously insists that additional positive lives have neutral value: "We are in favour of making people happy, but neutral about making happy people" (1973, 80).³ Regarding the first asymmetry, Narveson is equally decided: "If, therefore, it is our duty to prevent suffering and relieve it, it is also our duty not to bring children into the world if we know that they would suffer" (1967, 71). As far as I know, neither Narveson nor McMahan ever state that the strong moral reasons we allegedly have to avoid negative additions are equally as strong as the strong reasons we allegedly have to promote the wellbeing of lives that already exist. However, this seems a plausible claim to make, at least for anyone who endorses the two asymmetries. Taking this claim on board, we may formulate the full *asymmetry view* so: Negative wellbeing for additions, as well as changes in wellbeing for existing lives, has full moral importance; positive wellbeing for additions has no moral importance.

I aim to provide an axiological underpinning of the asymmetry view. I hope this will serve to make this view more palatable. The view includes two potentially controversial claims about value: 1) Negative additions have full moral importance, i.e. full negative value. 2) Positive additions have no moral importance, i.e. neutral value. That wellbeing effects for presently existing lives have full value I take to be tautological - "full value" here arguably *means* that value which wellbeing effects for presently existing lives has. Controversy around 1 includes the debate over whether poor existence is comparable in value with non-existence (see e.g. McMahan 2009). However, as far as pre-theoretical positions go, I take 1 to be less controversial than 2. People seem to have very different intuitions about 2. Some have *nativist* intuitions; they think with J.J.C. smart that it is only "humane and sympathetic" to value positive additions (quoted by Narveson 1967, 67).⁴ Others have *neutralist* intuitions; they think with Narveson and John Broome that our "moral instinct ... is caring for the people there are" and that we "have no natural interest in having more people about." (Broome 2004, p. 401).

It is not my purpose to argue for the intuitive appeal of neutrality. I will just briefly note that, at least for me, there are two different drivers of this intuition. The first

² Lifetime wellbeing is the only sort of wellbeing I will consider throughout and so I will drop the "lifetime" descriptor in the following. I will not discuss the nature of wellbeing, but I will assume that it is interpersonally comparable.

³ My discussion is not necessarily restricted to *people*, or to *human* lives, though most of the previous debate on these matters has been framed in terms of human wellbeing and human lives.

⁴ For an recent and unusually comprehensive defense of nativism, see Ord 2014.

is my lack of concern with the creation of happy people. I lack the nativist intuition that more happy people is a nice thing. I can see how more people might make for more diversity and variation, which may be good things, and how they may contribute to culture, science and technology, which may make already existing people happier, or increase average happiness. I do not see, however, how the mere fact of their existence, of another life lived, would be a good thing. The second driver is my concern with the wellbeing of existing people: Without neutrality, it is always better to create some number of happy lives, than to protect or promote the wellbeing of already existing people, whatever the cost or benefit to these existing people. It seems outrageous to me that it should be better to create new content people than to help existing people out of misery to contentment. It seems equally outrageous that it should be better to create new very happy people than to make already existing and merely content people very happy.

I will also note that neutrality seems embedded in widespread everyday intuitions and conversations, as well as in various institutions in society. As Broome observes, when we consider priorities in health services or in road safety, we typically do not consider population effects (2004, 144-5). If we were nativists, we should find it more important to save fertile people than infertile, since these people are likely to create positive additions. For example, it would make sense to investigate which accidents are most frequent for women below the age of 45 and to make the prevention of those accidents a priority. In more personal circumstances, if we were nativists we should be concerned with our friends' childlessness partly because it prevents the creation of positive additions. In fact, if we were nativists and some fellow nativists learned of our friend's situation and decided to have one or two extra children to make up for the lost positive additions, this should be quite a relief for us. However, very few of us would feel any such relief. These brief observations will hopefully serve to explain the intuitive appeal that neutrality has for some of us.

Among philosophers, there has been much more engagement with the first asymmetry (positive additions vs. negative additions) than with the second (positive additions vs. existing lives). Many attempts at defending the first asymmetry are normative as opposed to axiological: They are focused on the moral duty not to create lives that are not worth living.⁵ Axiological attempts at defending the first asymmetry typically assume some version of the so-called person-affecting view, such that wellbeing

⁵ Two significant recent normative defenses of the first asymmetry, which also contain ample references to earlier attempts, are Roberts 2011 and Algander 2013.

outcomes are only good or bad to the extent that they are so for particular persons.⁶ These normative and axiological approaches are united in their focus on whether or not creation can be harmful and/or beneficial to the created and in what way.

Similarly, the second asymmetry has also become associated with the person-affecting view. Broome assumes that neutrality "is part of the broader way of thinking known as the 'person-affecting view'", by which he means the view that "[a] change is [ethically] neutral unless it makes someone either better or worse off than she would otherwise have been" (2004, 145).⁷ However, the second asymmetry only claims neutrality for positive additions. It does not deny that other things than changes in wellbeing can have value. Neutrality is consistent with the inherent value of such welfare "pattern goods" as equality or population longevity.⁸ It is also consistent with the inherent value of non-human goods such as preserved wilderness. It therefore seems unnecessarily roundabout and controversial to defend neutrality via a defense of this person-affecting view. Furthermore, neutrality only follows from this view together with the assumption that positive additions are not made better off by their creation, which is also quite controversial (e.g. Bradley 2013). Moreover, if positive additions are neutral because they are not made better off by their creation, coherence seems to require that negative additions are neutral too, because they are not made worse off by their creation. If so, we get neutrality from the person-affecting view only at the price of a negation of the first asymmetry.

A purely impersonal axiological defense of the asymmetry view has the strong advantage that it removes the need to solve the puzzles around harm and benefit to future and potential people. Furthermore, some impersonal axiology is anyway unavoidable: If it is not impersonally good to create a new life with positive wellbeing, the fact that creating this life would yield a benefit to the created seems beside the point (as we can hardly owe anyone this benefit). Conversely, if positive additions *are* impersonally good, it seems difficult to escape the conclusion that we have some moral reason to create them, whether or not they are thereby benefitted, otherwise harmed, etc. Classical population axiologies, which are in conflict with neutrality, are impersonal and so independent of any claims regarding harm and benefit: According to the *total view*, positive additions are

⁶ Recent significant examples include Hare 2007 and Arrhenius 2009 (the latter is more an investigation than an outright defense).

⁷ In a recent book, Broome even identifies neutrality with a version of the person-affecting restriction: "The [neutrality] intuition is that, when something changes in the world, we can evaluate the change on the basis of how good it is for the people who exist". (2012, 171)

⁸ Broome recognizes that these are possible goods early in *Weighing Lives* (2004, 43-45).

always good. According to the *average view*, positive additions are good if they increase average wellbeing. It would seem an obvious strategy for a defender of neutrality to offer a competing impersonal axiology. This is the strategy that I will pursue.

That relatively little effort has been spent on similar projects may be explained by widespread pessimism regarding their feasibility. One reason for pessimism is Broome's argument against axiological neutrality. Broome convincingly shows that, given some very plausible assumptions, axiological neutrality, which he calls *the principle of equal existence*, conflicts with another appealing principle: *The principle of personal good*. Applied to additions, the principle of personal good states that if a new life is to be created, and if the wellbeing of other lives is constant, it is better if this new life has higher wellbeing. Broome's argument for the incompatibility of the two principles is essentially this: According to axiological neutrality, the status quo is equally as good as the addition of a life with wellbeing w . According to the same principle, the status quo is equally as good as the addition of the same life with wellbeing $w-1$. It follows, by the transitivity of 'equally as good as', that the addition of the life with wellbeing w is equally as good as the addition of the life with wellbeing $w-1$. This contradicts the principle of personal good. (Broome 2004, 146-148, cf. Broome 1999, 222-224).⁹ More generally, axiological neutrality conflicts with various pareto-like principles thought to apply to differences in wellbeing.¹⁰

In light of Broome's argument, I can only see two possible strategies for achieving impersonal, axiological neutrality. One is to accept pareto-like principles for populations in general but to introduce an exception for positive additions. Various such exceptions have been proposed. Additions, it has been argued, do not belong to the class of presently existing lives, or do not belong to the class of actual lives, whether present or

⁹ Narveson appears not to have realized how problematic it is to hold both that positive additions are neutral and, as he did, that a more positive addition is better than a less positive one. He claims that "if you are going to produce people, then you should produce the happiest ones consistently with the equal consideration of all others concerned", but fails to notice that this claim is in strong tension with his other views (1978, 53). As he goes on to say, "the utility of those to be produced need to be consulted only if they actually are produced" (ibid.). But then you cannot, as instructed in the first quote, consult the relative happiness of various groups of potential people in order to determine who are the happiest.

In his most recent refutation of axiological neutrality, Broome has modified his argument, perhaps in order to account for the possible value of equal distribution of wellbeing (2012, 176-7). Since I am not concerned with such equality, and since I find the modified argument less clear and less convincing, I stay with the earlier version.

¹⁰ Principles of this sort invoked in population axiology include Blackorby, Bossert & Donaldson's *strong Pareto axiom* (2005, 69) and Gustaf Arrhenius' *Pareto Dominance* (2013, 213).

future, or do not belong to the class of lives that are necessary relative to some decision (for an overview and discussion, see Arrhenius 2013, chapter 10). The main problem with all of these views, I believe, is that they imply a sort of relativism such that the value of wellbeing either changes with the circumstances or depends on what we do, even if the wellbeing itself remains constant: For *presentism*, the value of wellbeing changes with the passing of time, for *actualism*, this value changes depending on who we cause to exist, and for *necessitarianism*, it changes with the choice we face and the alternatives it includes. A relativistic population axiology implies that outcomes differ in value depending on perspective. Such relativism is in strong tension with the supposed impersonal value of wellbeing, as well as with the strong link between value and morality: When the value of outcomes changes with the passing of time, or the creation of lives, of the decisions we face, so does outcome-determined moral duties. This means that what is morally required at one time may be morally prohibited at another, which seems to me unacceptable. (For a fuller critique of this sort of relativism, see Broome 2004, 71-76).

The second, remaining strategy for achieving impersonal, axiological neutrality is to deny that the impersonal value of wellbeing depends only on individual levels of wellbeing. This is the strategy that I will pursue. If some independent factor bears on the value of wellbeing, then pareto-like principles that operate directly and exclusively on aggregate individual wellbeing are simply not relevant. If, for example, the value of wellbeing is partly dependent on the equality of its distribution, then an addition with positive wellbeing w that increases inequality may not be better than an addition with positive wellbeing $w-1$ that does not increase inequality. To say that this form of strict or non-instrumental *egalitarianism* violates pareto principles operating directly on wellbeing is true, but not a good or interesting objection to this view. The egalitarian view must be evaluated by considering arguments for and against thinking that the value of wellbeing partly depends on the equality of its distribution (I take this to be one main point made by Larry Temkin 2003). That aggregate individual wellbeing + equality of wellbeing does not track aggregate individual wellbeing is not an argument against the significance of the former, amalgam value, unless we assume from the outset that the latter is the only significant value. Another view that severs the direct link between aggregate individual wellbeing and the value of wellbeing is Fred Feldman's *justicism* (1995). Feldman proposes that the value of a person's wellbeing depends in part on that person's desert

level.¹¹ So, for example, an addition with positive wellbeing w and a high desert level may have higher impersonal value than an addition with positive wellbeing $w-1$ and a low desert level. Just as for egalitarianism, to say that justicism violates pareto principles operating strictly on aggregate individual wellbeing is true, but hardly a good or interesting objection to this view. Instead, the view must be evaluated by considering the reasons for and against thinking that the value of wellbeing depends on desert.

I do not endorse either egalitarianism or justicism. Instead, I will present a view according to which the value of wellbeing is relative to expected wellbeing. The point of and main argument for this view is that it yields an asymmetric population axiology for all the choices we can face, i.e. all alternatives that are relevant from a deliberative perspective. I will describe the view in the following two sections. I will then address some objections or possible concerns in three different sections, before I conclude.

2. Axiological asymmetry delivered

I call the view I am about to propose *expectism*, in recognition of the central role it awards expected wellbeing. I will discuss expected wellbeing and the probabilities that it involves further in the next section. Expectism is a *critical level* view. On critical level views, the value of a life is the difference between its wellbeing and some set level of wellbeing - the critical level. Critical level views are relatively popular in contemporary population ethics, defended by for example Broome (2004) and by Blackorby, Bossert & Donaldson (2005). On standard critical level views, the critical level is the same for all lives. Feldman's (1995) justicism, however, at least in Gustaf Arrhenius interpretation (2013), can be understood as a critical level view with individual critical levels, which depend on individual desert. Expectism is another individual critical level view.¹²

*Expectism assigns as the critical level for any life the expected wellbeing of that life when it starts, with one restriction: Critical levels cannot be negative, so the critical level for a life with negative expected wellbeing is zero.*¹³ The strong link to expected

¹¹ As noted by Ingmar Persson (1997), Feldman's presentation oscillates between different ideas on how to factor in desert. However, Arrhenius (2013, chapter 8) has reconstructed Feldman's view such that it is unambiguous and consistent.

¹² It will be imperative to understanding my views to remember the distinction between value and wellbeing, as well as the distinction between either of these being either expected or actual.

¹³ I take no stand on the controversial issue of when exactly a life starts. In other words, I follow Narveson in "neglecting the question about the point at which a person comes into existence" (1967, 63).

One could introduce an upper limit for critical levels in addition to the lower limit. However, I find this restriction less intuitive than the lower limit. Also, there is no obvious level of

wellbeing has two important consequences, given an account of probability that is robust enough: First, we cannot choose to create lives with positive value. When we consider the value of a possible future life, we know that it may be positive or negative, depending on how its actual wellbeing deviates from its expected wellbeing. We may hope for positive deviation, perhaps we can even aim for it, even if our aims, to the extent that they manifest in behavior, may influence the expected wellbeing of the life under consideration, and so undermine their own fulfillment. What we cannot do, however, is to choose to create a life with positive value, or even positive expected value. We know that the expected value of the wellbeing of any future life is zero or negative. Thinking of alternatives in terms of their possible outcomes and their associated probability, we know that any such bundle of outcomes and probabilities will, as far as it concerns future lives, have zero or negative expected value. Expectism therefore delivers straightforward, impersonal axiological asymmetry for all deliberatively relevant alternatives. This is its main virtue.

The second consequence is that the actual value of the wellbeing of large groups of lives will, in the aggregate and under realistic circumstances, tend towards zero. Expectism therefore provides substantial axiological asymmetry for large and realistic populations, even beyond deliberative relevance. Such populations are the only populations that are relevant in political and other pragmatic and large-scale contexts. This tendency towards actual neutrality may not seem a very strong advantage of expectism. Once we leave the deliberative perspective, the view may seem to have very counter-intuitive implications: A life with very low wellbeing can be better than a life with very high wellbeing, if the expected wellbeing for the first life is low enough and/or the expected wellbeing for the second life high enough. In section 5, I will argue that such consequences of expectism are not as problematic as they may seem.

That lives have individual critical levels does in no way affect how presently existing people should be treated. The value of a life depends both on its actual wellbeing and on its critical level, but once it has started, only its actual wellbeing can be affected. Therefore, expectism behaves exactly like the total view, and standard critical level

wellbeing at which to fix it. Therefore, and to simplify my presentation, I will ignore this possibility in the following. It should be kept in mind, however, especially by those who are somewhat inclined towards the total view, since expectism with an upper limit on critical levels approaches the total view as this upper limit approaches zero wellbeing. In fact, expectism with a relatively low upper limit behaves much like a standard critical level view, while it avoids, at least from the deliberative perspective, the counter-intuitive implication that lives with positive wellbeing can have negative value.

views, in regard to existing lives: Any change in the wellbeing of any existing life has the same value, and this value is independent of what other lives there are and the wellbeing of those lives.¹⁴

Expectism has implausible consequences for many population comparisons, realistic and unrealistic. All these comparisons, however, involve alternatives that we cannot choose to bring about. In that sense, they are irrelevant from the deliberative perspective.¹⁵ It would be satisfying to have a non-relativistic axiology that underpinned the asymmetry view for all logically or at least all physically possible populations. Alas, there are none on offer as far as I know. Fellow neutralists who share my pragmatic orientation may be somewhat satisfied with population axiological asymmetry for deliberatively relevant alternatives. Fellow neutralists who are not so easily satisfied may perhaps consider expectism a partial success in the quest for axiological asymmetry. Perhaps expectism is a piece in a larger puzzle, perhaps it must ultimately be discarded but may help indicate some superior axiology.

As shown by Arrhenius, population axiology is notoriously haunted by the impossibility of simultaneously satisfying various seemingly plausible desiderata (e.g. Arrhenius 2013, chapter 11). Arrhenius emphasizes that population ethics should ideally be applicable to any population comparison, but he recognizes that the development of an axiology for normatively relevant cases "is an important task, or, as some people might argue, *the* important task in population axiology." (2013, 274) Outcomes that we cannot choose but that may result from our choices are in one sense normatively relevant. However, what is most obviously normatively relevant are the alternatives that we can choose. I am myself prepared to incur the theoretical cost of limited applicability in order to achieve systematic and plausible verdicts to guide choice.

I can now spell out how expectism avoids Broome's argument against the neutral value of positive additions. Broome's argument depends crucially on the principle of

¹⁴ In other words, expectism is consistent with what Broome calls *separability of lives*. Broome derives separability of lives from the principle of personal good, but it has independent appeal. Notably, the average view is inconsistent with separability of lives, since it implies that the value of any change in the wellbeing of an existing life is relative to the size of the population (individual changes are more important the smaller the population). It seems to me that the intuition that the value of a life should not depend on population size is a close cousin to neutrality: If a population is not improved by positive additions, then neither should these additions affect the value of changes in the wellbeing of existing people.

¹⁵ If one adopts, as does Broome, an expected value view of the normative import of consequences (whether or not one is a consequentialist or only considers consequences one morally relevant factor), then this deliberative perspective is also the only normatively relevant perspective.

personal good and its implication that an additional life with (positive) wellbeing w must be better than the same additional life with (positive) wellbeing $w-1$. This principle is simply not relevant to axiologies in which the value of a life depends on other things than its wellbeing. To insist on the validity of the principle is merely to insist on the irrelevance of anything other than absolute, individual wellbeing. According to expectism, the value of a life depends in part on its expected wellbeing when starting. If the expected wellbeing of the life when it has wellbeing w is w and its expected wellbeing when it has wellbeing $w-1$ is $w-1$, then the alternatives are equally good. This is a clear violation of the principle of personal good. It is, however, a sought-for violation. It is the very purpose of expectism to undermine the principle of personal good by implying that all positive additions are equal in value from the deliberative perspective. In the face of this challenge, the defender of the principle of personal good should present some independent reason to deny that expected wellbeing affects the value of a life. Mere insistence on the principle is question-begging.

We might of course add a *ceteris paribus* clause to the principle of personal good and assume that expected wellbeing is constant over the alternatives. On that assumption, adding the life with wellbeing w is indeed better than adding the life with wellbeing $w-1$. However, this means that at least one of the alternatives cannot be chosen, and so is irrelevant from the deliberative perspective. As I have declared, my ambitions are limited to providing a plausible asymmetrical axiology for all deliberatively relevant populations. Broome's argument does show that expectism has consequences that contradict *objective* axiological asymmetry, though these consequences will tend to even out in the larger perspective for any real population.

3. Expected wellbeing, probability and neutrality

I take expected wellbeing for a life to be calculated as expected wellbeing or expected utility is typically calculated: Each possible way the life may turn out given its starting conditions is an outcome. The wellbeing in each possible outcome is multiplied with the probability of that outcome being the actual outcome. The expected wellbeing of the life is the sum of all these products. This familiar account of expected wellbeing is consistent with any account of probabilities, as long as they obey standard probability calculus.

I will not take a stand on the plausibility of different accounts of probability. Nor will I pick or develop a particular account suitable for expectism in particular. However, I will propose that the account should be objective or, if subjective, strongly idealized, for

three independent reasons.¹⁶ First, expected wellbeing should tend to coincide with actual wellbeing as far as possible, since this makes for more or stricter neutrality. This is mostly just to say that probability should coincide with actual frequencies when there is a large data set, which is a general adequacy criterion for any account of objective probability. Second, it is hard to see whose subjective probabilities should be used on a more concretely subjective account. A person obviously cannot herself make predictions about her life when it starts.¹⁷ Subjective probabilities could be sampled from other people, such as those most directly casually responsible for the life (typically the parents), but this seems rather arbitrary. Third, expectism implies that a population would be better if we could somehow lower expected wellbeing without lowering actual future wellbeing, thus creating positive deviations from critical levels. I take it that objective probability or strongly idealized subjective probability will exclude this possibility.¹⁸

Calculating the expected wellbeing for a life when starting requires the identification of all possible lifetime outcomes and the attribution of a probability to each outcome. We cannot actually make such detailed calculations. At best, we may have a general idea of the expected wellbeing for a life given some salient features of its starting conditions. This epistemic situation, however, is not a problem for expectism. What we do know is that whatever the expected level of wellbeing for a future life, as long as it is positive, this is also its critical level, and hence we know that its expected value is zero. In contrast, competing axiologies, which assign value to lives only in relation to their actual wellbeing, are at a great epistemic disadvantage when it comes to predicting the

¹⁶ Note that I am here referring to objective *probability*, while I previously referred to an objective perspective in the sense of all-knowing, retrospective or actualist. The objective probability of throwing a 6 on a standard dice may be 1/6, but the actual outcome either is 6 or is not 6.

¹⁷ Nor of course could non-human beings, unless perhaps some advanced alien or technological intelligences that come into being highly developed.

¹⁸ Idealized subjective probabilities could for example be such as would be attributed by a fully rational team of diverse scientists with unlimited time and access to all information collectible with the best current (or future) scientific methods as well as the best methods for analyzing that information.

It is no threat to expectism that subjective probabilities are idealized to the point that they are perfect predictions, such as could perhaps be made by an omniscient being. In fact, this would just help underpin neutrality. However, critical levels must not be defined directly as actual outcomes, since this would imply that increased wellbeing has no value.

Note that with objective probabilities, the deliberative perspective is not the subjective perspective. A person that has not accepted expectism may reasonably but falsely believe that the expected value of a life she creates is positive and so be subjectively justified in creating it. Any person that accepts and understands expectism, on the other hand, will know that the expected value of any future life is zero or negative, purely on theoretical grounds.

value of future lives.

It may seem that an expectist should not only be concerned with the expected value of a future life, but also with the finer details that form the basis for the calculation of this value - the possible outcomes and their associated probabilities. In particular, it may seem we should be concerned with possible outcomes in which wellbeing is negative. I believe this is true in one sense, and not true in another.

To create a life is a gamble in impersonal value, according to expectism as well as according to competing axiologies. For two lives with the same expected value, the variation in possible value may be much larger for one than for another. If we are risk averse, we should have some reason to prefer, *ex ante*, the safer bet. However, I fail to see any reason to be risk averse in this context, where the risk we take concern impersonal value. We might reasonably be averse to particularly poor outcomes, such as lives with negative wellbeing. This, however, would be a reason to attribute lower value to such lives, in line with axiological prioritarianism (e.g. Holtug 2010, chapter 8). It would not be a reason, once any such priorities are incorporated into the value of outcomes, to prefer some amount of value for sure, or with greater certainty, over a risky bet with the same expected value.¹⁹ Note further that according to expectism, the potential variation in the value of a life has little to do with the absolute wellbeing of this life. A riskier bet in terms of value may be a much safer bet in terms of absolute wellbeing (e.g. a life with equiprobable outcomes evenly distributed from wellbeing level 100 to 1000 may be more impersonally risky than a life with a similar distribution from level 0 to 10, yet the former does not entail a risk of low wellbeing).

The sense in which we might reasonably be risk averse when it comes to future lives is that we may not be able to accurately determine whether or not a future life will have negative *expected* wellbeing. There is a risk, in creating a life, that it will have negative expected wellbeing, and such lives are likely to have negative value, and more so the more negative. In practice, therefore, I believe we should be concerned, not with the variation in the bundles of outcomes and probabilities, but with the resulting epistemic uncertainty around expected wellbeing. This aspect of expectism may be a virtue, as it explains why we should be concerned with keeping expected wellbeing well above zero, as opposed to being satisfied with any positive level.

¹⁹ Prioritarianism cannot be incorporated into expectism as straightforwardly as for totalism or averaging, but it can be added as a separate element, such that the value of a life is the divergence of its actual wellbeing from its critical level, plus its priority value. To maintain neutrality, the priority value must be based only on negative and not on positive wellbeing (the simplest version would be to let it be identical to negative wellbeing).

When lives are considered in retrospect, the epistemic advantage of expectism turns into a disadvantage. Like competing axiologies, expectism operates on actual wellbeing in order to determine the value of a life *ex post*. However, unlike competing axiologies, it also operates on the expected wellbeing for a life when starting. As noted above, this is very difficult to determine at the time, and it is not much easier in retrospect. Retrospective evaluation, however, is not deliberatively relevant. What is deliberatively relevant is the value of future lives, and the value of future changes in the wellbeing of existing lives. It may seem that retrospective evaluation of lives would be relevant for decisions on how to relate to these past lives. For example, whether or not lives should be celebrated, statues erected, people canonized. However, it is not clear that the impersonal value of a life is anything to celebrate. The lives we should celebrate are arguably good lives - good either in the sense that they have high wellbeing, or in the sense that they contribute instrumentally to impersonal values. For such evaluations, expectism behaves just like standard views.

4. Ad hoc?

Expectism is designed to underpin the asymmetry view. Attributing to each life an individual critical level that depends on the expected wellbeing of that life when starting is a technical maneuver that has no other warrant than that it happens to deliver deliberative axiological asymmetry. The axiology may therefore appear to be *ad hoc*. I believe this appearance would be deceptive.

An axiology is a system for assigning value to things, or in other words to order them according to their value. The quality or plausibility of such a system, I propose, can only be evaluated by looking to its implications. This is how authors have discussed and compared the total view and the average view, for the most part. The average view is generally considered to imply that whether or not a life is good depends on the wellbeing of all other lives in the population. This speaks against the view. The total view is generally considered to imply the repugnant conclusion (introduced by Parfit 1984) - that a large population consisting only of rather poor lives can be better than a much smaller population consisting only of very good lives. This speaks against this view (I will consider below how expectism fares in relation to the repugnant conclusion). That the total view uses summation and the average view uses averaging does not speak either for or against either view. It is not relevant whether summing is more or less *ad hoc* than averaging, or whether they are equally *ad hoc*. Both summing and averaging are, I

believe, technical maneuvers that have no other warrant than that they happen to deliver particular value orderings.

I believe that the total view and the average view gain a sort of superficial plausibility from the fact that they are simple and straightforward ways of ranking so called same-people cases, i.e. cases where distributions of wellbeing differ between alternatives but the people that this wellbeing is distributed over does not (Parfit 1984, 356). For such cases, expectism, the total view and the average view all entail the same ranking of populations (given that people are individuated by their starting conditions). Because they do, an expectist might as well simplify her ranking of such populations by disregarding expected wellbeing and looking only at total or average wellbeing. This appealing convergence of the different theories disappears, however, as soon as we consider population outcomes with different people or with different numbers of people. It is in light of these more difficult cases that we have reason to consider expectism a serious contender. Theoretical simplicity is good when the subject matter is simple; it is no virtue when the subject matter requires more complex theoretical solutions.

A more specific worry is that the restriction that critical levels cannot be negative is *ad hoc*. This restriction draws a line at zero wellbeing and so helps deliver an axiology that is in perfect agreement with the asymmetry view. For the reasons just stated, I cannot see that any restriction or modification of an axiology can be *ad hoc*. Restrictions or modifications can be surprising and they may make axiologies complex and difficult to apply, but that is another matter. However, let me anyway note that the difference between a life worth living and one not worth living is quite real and very important. It therefore seems to me particularly non-*ad hoc* to treat positive and negative wellbeing differently from an axiological perspective.²⁰ It seems clearly more *ad hoc*, or at least arbitrary, to define, as do standard critical level views, a general critical level at some positive level of wellbeing: What should define this level? Should it be at two units of wellbeing or at twenty? In contrast to the zero level, nothing sets any positive level apart from any other. Similarly, on so-called number-sensitive or number-dampened views, the rate at which the value of additional wellbeing declines with population size seems particularly arbitrary. Having mentioned these views, I note in passing that they share with expectism a certain complexity in that they require complex operations on wellbeing

²⁰ Compare Arrhenius argument that the distinction between lives which are good to live and lives which are not is substantial and "has nothing to do with numbers". We can represent the distinction between these two types of lives by the value zero on a scale, but this does not mean that the distinction is merely an arbitrary number on a scale (Arrhenius 2013, 38-39).

distributions in order to determine population value (see Blackorby, Bossert & Donaldson 2005, pp. 140-147).²¹

While another matter entirely, I will close this section by emphasizing that not only is expectism not *ad hoc*, but neither is it *relativistic*. In my introductory section, I dismissed presentism, actualism and necessitarianism as relativistic. I believe this is a fundamental flaw. On these views, what value a life has depends in one way or other on perspective. It would be easy to formulate a relativistic asymmetric axiology: Existing lives and future negative lives have value according to their wellbeing, while future positive lives have no value.²² With this axiology, however, what value a life has depends on whether or not it has started. This is not a property of that life but rather a matter of perspective. According to this relativistic axiology, a life with wellbeing 10 that starts in 2018 has zero value in 2017 but a value of 10 in 2019. In contrast, like the total view and the average view, expectism implies that the value of a life is independent of perspective. Assume that the expected wellbeing of the life with wellbeing 10 is 7. This means that the value of this life is 3. This value is not relative to times or to perspective. The value of the life is always 3. This cannot be known, of course, before the life is lived, if then. We should therefore, in deliberating about choices that might affect the existence of this life, consider only its expected value *ex ante*, which is zero.

5. Positive lives with negative value

Consider a life with high expected wellbeing and slightly lower actual wellbeing. According to expectism, this life has negative value. This means that, *ceteris paribus*, the population, or the world, would be better without this life. From the deliberative perspective, of course, what matters is that the expected value of any life is either neutral or, if the expected wellbeing is negative, negative. Still, we know, even from the deliberative perspective, that most lives are quite likely to have negative value, even when their expected wellbeing is high. This may seem very counter-intuitive.

We must remember, however, that when I now refer to the value of a life, I am disregarding any net contribution to the wellbeing of other lives or to any impersonal values, such as cultural or other diversity. We must also remember that the fact that a life has negative value in no way means that it is a poor life. Nor does it mean that we have any less reason to regret any missed opportunity to improve its wellbeing. In fact, the

²¹ These views are, like the average view and unlike expectism, inconsistent with separability of lives. Cf. footnote 14.

²² Arrhenius (2013, 278) calls this view "asymmetric presentism".

absolute impersonal value of a life is a very abstract thing once it has started. This value cannot guide any practical decisions, since we have the same reason to cause a relative improvement in the value of a life irrespective of its absolute value. However, the fact remains that when we consider the creation of a life with positive expected wellbeing, it is as likely to have negative as it is to have positive non-instrumental value. This is the sort of neutrality that expectism delivers.

To the extent that this consequence of expectism is found problematic even after the above considerations are appreciated, it should be noted that standard critical level views are partners in guilt in this regard. Any critical level view with a positive general critical level of course implies that lives with a level of positive wellbeing lower than the critical level have negative value. On standard critical level views, this implication is not restricted to the non-deliberative perspective, and so we may have deliberative reason to make sacrifices in order to avoid the existence of lives with positive wellbeing. I do not find this completely unacceptable, but those who do should remember that expectism does not have this implication.

It might be thought that there is technical fix available for expectism: Introduce the restriction that lives with positive wellbeing cannot have negative value. This would mean that lives with negative wellbeing always have negative value, and that lives with positive wellbeing have positive value to the extent that they outperform expectations, but never negative value. However, if we should introduce this restriction, expectism would no longer behave like the total view for existing lives: There would be no reason based on the value of wellbeing for improving the wellbeing of an existing life from a low level to a high level if both levels are positive and lower than the expected level. This seems to me unacceptable.

It may be an advantage of the total view relative to both the average view and critical level views that it entails that all positive wellbeing is positively valuable. However, this rather directly implies non-neutrality, which is a strong disadvantage.

6. Repugnant conclusions

Broome describes a dilemma for standard critical level theories: If the critical level is low, as for the total view, then the view is vulnerable to the repugnant conclusion. If, on the other hand, the critical level is high, then the view is vulnerable to what Broome calls the negative repugnant conclusion. There seems to be no level that is neither high nor low in this sense. (Broome 2004, pp. 210-14)

As noted above, the repugnant conclusion is that a large population consisting

only of rather poor lives can be better than a (much) smaller population consisting only of very good lives. The negative repugnant conclusion is that a small population consisting of very bad lives can be better than a (much) larger population consisting of lives with positive wellbeing.²³ Both conclusions are generally considered difficult to accept, though some find one and some the other more acceptable.

A great advantage of expectism is that, from the deliberative perspective, it is vulnerable to neither repugnant conclusion. This follows directly from the fact that all lives with positive wellbeing have neutral expected value. While it is always good to improve the wellbeing of existing lives, to create a new life is a gamble with zero or negative expected value. Axiologically speaking, therefore, we have no reason to sacrifice average wellbeing for the sake of a higher total made up of relatively poor lives, and we have positive reason to avoid creating lives with negative expected wellbeing.

From an objective or actualist perspective, expectism is vulnerable to both repugnant conclusions, just like any critical level view with positive critical levels. The way the repugnant conclusions are formulated presupposes a general critical level, at zero wellbeing or otherwise. According to expectism, of course, critical levels are individual. However, there can, in theory, be outcomes that are repugnant in the same way. For example, all lives might just happen to have the same critical level. Even if they do not, a large population of rather poor lives can be better than a small population of excellent lives, as long as the expected wellbeing of the poor lives is even lower than actual outcomes. Similarly, a small population of horrible lives can be better than a large population consisting of lives with positive wellbeing, as long as the expected wellbeing of the positive lives is higher than actual outcomes. Because of expectism's dependence on expected wellbeing, it is actually quite roundabout to invoke the repugnant conclusions to identify implausible consequences from the objective perspective. It is enough to note that expectism can rank a population where everyone suffers terribly over a population where everyone has a good life, given only that the expected wellbeing in the latter population is high enough. By varying expected wellbeing without restriction, it is quite trivial to identify very implausible consequences. Such variation of expected wellbeing is not, of course, realistic.

Broome notes, concerning the repugnant conclusions, that lack of neutrality on the level of value does not preclude a sort of normative neutrality in terms of

²³ The negative repugnant conclusion closely resembles what Gustaf Arrhenius (2000) has called “the sadistic conclusion” – that adding lives with negative wellbeing to a population can be better than adding lives with positive wellbeing.

responsibility. We may, he says, “have no moral responsibility towards population” (2005, p. 413). If we do not, we have no moral responsibility to move towards the repugnant outcomes even if they are better than our starting points. As I noted above, however, if positive additions have value, it seems difficult to escape the conclusion that we have some reason, if not responsibility, to bring them into existence. Difficult, that is, unless there is a good explanation for this lack of connection between value and responsibility or between value and reason for action. By implying that we cannot choose to bring about valuable additions, expectism provides such an explanation and so very effectively severs the connection between axiology and normativity: We can hardly be responsible for doing something that we cannot choose to do.

7. Conclusion

I have presented expectism and explained that it is an asymmetric population axiology for populations we can choose to bring about. I have argued that for the same populations, it is not vulnerable to either the repugnant conclusion, the negative repugnant conclusion, or the implication that lives with positive wellbeing have negative value. For anyone prepared to accept, perhaps as a lesser evil, the restriction to deliberately relevant populations, this should be an axiology worth considering. Even for those who are not so prepared, expectism may indicate a partial solution to the axiological underpinning of the asymmetry view.

The key to axiological asymmetry is to define individual critical levels in terms of expected wellbeing for a life when it starts. Since expected wellbeing is all we should take into consideration when we consider the creation of future lives, positive additions will be axiologically neutral from the deliberative, *ex-ante* perspective. Since critical levels have a lower limit at zero wellbeing, lives that are expected to have negative wellbeing will, from the same *ex-ante* perspective, have negative value.

Individual critical levels rather directly imply violations of common axiological principles that operate on actual wellbeing. These principles, on the other hand, rather directly imply non-neutrality. Population axiology is notoriously beset by conflicting principles and intuitions. This clash between neutrality and pareto optimality for actual wellbeing is one such conflict. I propose that the concern with actual wellbeing that forms the basis of pareto principles is quite proper when applied to existing lives, but that it cannot straightforwardly be extrapolated to future lives, whose existence is, at least from a deliberative perspective, still uncertain. I also believe, on a more pragmatic note, that, since the field is so conflicted, old dogmas must be questioned and new paths tried.

Expectism does not yield perfect asymmetry for the many possible populations that we cannot choose but that can come about as a result of our choices and actions. For these non-chosen but realistic populations, however, the view yields substantial neutrality, since actual outcomes will on the whole and in the long run tend to coincide with expectations. While single additions will typically have positive or negative value, large and realistic additions will tend to have very little value, that is value close to neutral.

It may be thought that expectism achieves asymmetry only at the price of an unacceptable disregard for future positive wellbeing. To some extent, such disregard is simply the other side of neutrality. On the other hand, I tend to agree that there must be something positive about what we typically think of as a bright future for posterity. Two considerations help to mitigate the apparent disregard for the future:

First, positive wellbeing for future lives may promote or even be necessary for other values, such as cultural diversity and appreciation of nature, as well as many other values of human culture and civilization. If so, we have some independent reason to promote future wellbeing. The fact that there are these other values, if it is a fact, also mitigates the perhaps disturbing consequence of expectism that as far as their wellbeing is concerned, populations on the whole tend to be axiologically neutral or negative.

Second, it is difficult to determine the expected wellbeing for a future life and in particular whether or not it may be negative, and so very likely have negative value. The best way to avoid as far as possible that there are any such lives may be to strive for very high expected wellbeing for all lives.

The normative implication of expectism is that we can only have deliberative moral reason to care about existing lives and to avoid creating lives with negative expected wellbeing. This is the asymmetry view, which combines what McMahan called the Asymmetry with a second asymmetry between the value of the wellbeing of presently existing lives and the addition of new, future lives. From a political perspective, these implications are welcome because they allow us to evaluate our response to such global challenges as climate change, without considering the non-instrumental value of their effects on population size. Broome claims “we have no right to assume global warming will be bad just because it will kill so many millions of people. Perhaps its effects on population will be so beneficial as to cancel out this badness.” (2005, p. 411) According to expectism, this is highly unlikely, and irrelevant for deliberation.

Acknowledgments: The core aspects of the view developed in this paper were first presented at a workshop on Climate change policy after Copenhagen, in Uppsala in February, 2010. Versions have since been presented at the philosophy departments in Uppsala and Umeå, as well as at the ISUS conference in Yokohama 2014. I am grateful for the feedback received. Lars Lindblom and Niklas Möller provided helpful criticism of an early draft. Per Algander, Erik Carlson and Stephen Wilkinson each did the same for two separate drafts. My wife Camilla Grill asked helpful, probing questions based on my many animated explanations. Last but certainly not least, a reviewer for this journal provided very helpful comments.

References

- Algander, Per. 2013. *Harm, Benefit, and Non-Identity*. Dissertation, Uppsala University.
- Arrhenius, Gustaf. 2000. "An Impossibility Theorem for Welfarist Axiologies". *Economics and Philosophy* 16: 247-266.
- Arrhenius, Gustaf. 2009. "Can the Person Affecting Restriction Solve the Problems in Population Ethics?" In *Harming Future Persons*, edited by Melinda A. Roberts and David T. Wasserman, 289–314. International Library of Ethics, Law, and the New Medicine 35. Springer Netherlands.
- Arrhenius, Gustaf. 2013. *Population Ethics: The Challenge of Future Generations*. Manuscript.
- Blackorby, Charles; Bossert, Walter & Donaldson, David. 2005. *Population Issues in Social Choice Theory, Welfare Economics, and Ethics*. Cambridge: Cambridge University Press.
- Bradley, Ben. 2013. "Asymmetries in Benefiting, Harming and Creating." *The Journal of Ethics* 17: 37-49.
- Broome, John. 1999. *Ethics out of Economics*. Cambridge: Cambridge University Press.
- Broome, John. 2004. *Weighing Lives*. Oxford: Oxford University Press.
- Broome, John. 2005. "Should We Value Population?". *Journal of Political Philosophy* 13: 399-413.
- Broome, John. 2012. *Climate Matters: Ethics in a Warming World*. New York: W. W. Norton & Company.
- Feldman, Fred. 1995. "Justice, Desert, and the Repugnant Conclusion". *Utilitas* 7(2): 189-206.
- Hare, Caspar. 2007. "Voices from Another World: Must We Respect the Interests of People Who Do Not, and Will Never, Exist?" *Ethics* 117: 498–523.
- Holtug, Nils. 2010. *Persons, Interests, and Justice*. Oxford: Oxford University Press.
- McMahan, Jeff. 1981. "Problems of Population Theory." *Ethics* 92: 96–127.

- McMahan, Jeff. 2009. "Asymmetries in the Morality of Causing People to Exist." In *Harming Future Persons*, edited by Melinda A. Roberts and David T. Wasserman, 49–68. International Library of Ethics, Law, and the New Medicine 35. Springer Netherlands.
- McMahan, Jeff. 2013. "Causing People to Exist and Saving People's Lives." *The Journal of Ethics* 17: 5–35.
- Narveson, Jan. 1967. "Utilitarianism and New Generations." *Mind*, New Series, 76 (301): 62–72.
- Narveson, Jan. 1973. "Moral Problems of Population". *The Monist* 57: 62–6.
- Narveson, Jan. 1978. "Future People and Us". In R.I. Sikora and Brian Barry (eds.) *Obligations to Future Generations*. Philadelphia: Temple University Press.
- Ord, Toby. 2014. "Overpopulation or Underpopulation?". In Ian Goldin (ed.) *Is the Planet Full?*. Oxford: Oxford University Press.
- Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Oxford University Press.
- Persson, Ingmar. 1997. "Ambiguities in Feldman's Desert-adjusted Values", *Utilitas*, Vol. 9, No. 3., pp. 319-28.
- Roberts, Melinda A. 2011. "The Asymmetry: A Solution." *Theoria* 77: 333–67.
- Temkin, Larry S. 2003. "Equality, Priority or What?" *Economics and Philosophy* 19(1): 61–87.