

KRISTER BYKVIST

THE BENEFITS OF COMING INTO EXISTENCE

ABSTRACT. This paper argues that we can benefit or harm people by creating them, but only in the sense that we can create things that are good or bad for them. What we cannot do is to confer comparative benefits and harms to people by creating them or failing to create them. You are not better off (or worse off) created than you would have been had you not been created, for nothing has value for you if you do not exist, not even neutral value.

1. INTRODUCTION

One important part of morality concerns benefits and harms. Most of us agree that we have at least a *prima facie* obligation to benefit people and avoid harming them. Utilitarians would say that this is the whole truth about morality, but even non-utilitarians usually accept a *prima facie* obligation of beneficence or non-malevolence. Problems arise when we apply these obligations to cases in which the existence of people is at stake -- cases that are commonly called *different people* choices.¹ It is widely agreed that we can harm and benefit people that already exist, or will exist in the future independently of our present choice. What is controversial is whether we can benefit or harm people by creating them or by failing to create them.

This problem is relevant to many important moral decisions. For instance, choosing late abortion does not seem to be morally equivalent to choosing contraception, and this difference does not seem to be a difference in mere side-effects on prospective parents and the society as a whole. The crucial feature is rather that there is no one who would have been directly affected for worse or better were we to successfully use contraception. In contrast, in the late abortion case, there

is someone who would be directly affected. If this widely accepted assumption is false and we can affect people's well-being by creating them or failing to create them, we seem to be unable to draw a moral distinction between the two cases, at least from the perspective of the morality of beneficence.²

The problem about whether we can benefit or harm people by creating them or failing to create them has triggered a big debate and there are numerous articles on the topic. Three main views can be discerned:

- (1) There is no sense in which we can benefit or harm already existing people but not benefit or harm people by creating them (or perhaps even by failing to create them).³
- (2) We cannot in any plausible sense benefit or harm people by creating them or failing to create them.⁴
- (3) In one sense, we can harm or benefit people by creating them, but in another, we cannot. However, there is no sense in which we can benefit or harm people by failing to create them.⁵

Views (2) and (3) both make room for drawing a moral distinction between choosing contraception and choosing late abortion. If one accepts (2) one can simply say that the morality of beneficence does not apply to different people choices. This is quite an unattractive option, however, since, for example, creating a miserable person seems to involve some sort of harm. View (3), by contrast, makes room for a more sensible form of discrimination. If one accepts (3), one could argue that the harms and benefits involved in different people choices are morally relevant, but their moral importance is not the same as the moral importance of harms and benefits of already existing people. In particular, one could claim that since there is a sense in which we can harm and benefit existing people but not harm or benefit people by creating them, we have stronger reason to benefit already existing people than to benefit people by creating them.

In this paper, I am going to argue against the two extreme views (1) and (2), and thus give an indirect argument in sup-

port of the more moderate position (3). However, I should stress that I am not going to try to decide how the harms and benefits of created people should be weighed against the harms and benefits of people who will be around no matter how we choose. Indeed, I am not even going to argue that we should make use of the distinction between the different kinds of harms and benefits. The focus of the paper is the rejection of (1) and (2).

2. BENEFITS AND HARMS

To make the problem more precise we need to say something more about benefits and harms. First of all, benefits and harms are things that are good and bad, respectively. Moreover, they are not free-floating; a benefit or harm is always *someone's* benefit or harm. We usually capture these features by saying that benefits and harms are things that have *value for* individuals.⁶

It is also important to distinguish between absolute and comparative benefits and harms. Roughly put, a person enjoys (suffers) an *absolute* benefit (harm) if and only if she is in state that is good (bad) for her, whereas a person enjoys (suffers) of a *comparative* benefit (harm) if and only if she is in state that is better (worse) for her than some alternative state. The notions of enjoying a benefit and suffering a harm are not supposed to have any hedonist connotations. To enjoy a benefit or suffer a harm is simply to be the *recipient* or *subject* of this benefit or harm. I do not wish to deny that non-experiential states of affairs, such as being deceived, are harms that we can suffer.

The characterisation of comparative benefits and harms is rough and needs to be refined. For instance, it is plausible to think that in order to suffer a comparative harm one needs to suffer a loss of something that is in some sense essential or important.⁷ This qualification will not, however, affect my arguments. Later I will introduce some necessary refinements, but here it is only important to stress that the distinction is commonplace and that there are plenty of everyday examples. A person who wins on the lottery and is finally able to lead a

rich and fulfilling life is benefited both absolutely and comparatively, whereas a dying cancer patient who undergoes some treatment that lessens her suffering but does not take it away is benefited only in the comparative sense. Similarly, a person who contracts some horrible form of cancer is harmed both absolutely and comparatively, whereas a moderately rich person who loses a fair amount of money need not be harmed in an absolute sense.

‘Good for’ and ‘better for’ are elastic terms that can be stretched to cover many different things. In the context of this paper, ‘good for’ should *not* be read as ‘good according to’, ‘judged to be valuable by’, or ‘desired by’. Nor is it supposed to be read as ‘has reason to bring about’ or ‘has reason to desire’. ‘Good for’ is supposed to capture facts about a person’s *well-being*, what makes a person’s life good in itself for that person. It seems clear that what makes my life good for me in this sense does not necessarily coincide with any of the listed items.

3. A MORE PRECISE STATEMENT OF THE PROBLEM

With these distinctions at hand we can split the target question into the following more precise questions (‘a person’s existence’ and ‘a person’s non-existence’ should be read, respectively, as ‘the state of affairs that the person exists at some time’ and ‘the state of affairs that the person never exists’):

- (i) Can a person’s existence be an absolute benefit or harm?
- (ii) Can a person’s non-existence be an absolute benefit or harm?
- (iii) Can a person’s existence be a comparative benefit or harm?
- (iv) Can a person’s non-existence be a comparative benefit or harm?

I shall argue that a person’s existence can have absolute value for her and that it is in this special sense we can talk

about benefiting or harming a person by creating her. However, I shall deny that non-existence can have any absolute value for a person, not even neutral value. Finally, I shall argue that neither existence nor non-existence can be a comparative benefit or harm.⁸

4. PRESUPPOSITIONS

Before I move on, I need to introduce some crucial presuppositions.

(a) *Actualism*. If an individual exemplifies a property or stands in a relation in a world, it must exist (i.e., be actual) in this world.

An individual exists in a world iff had the world been actual, then the individual would have existed. Actualism should not be confused with the trivial principle that a property-exemplification requires an exemplifier. Even possibilists could accept this. They would add that a non-actual merely possible individual can exemplify a property. Actualism is of course not an uncontroversial assumption. It is an attractive assumption to make if you share my robust sense of reality and think that mere possibilities should be rejected. But it is also accepted by many who want to defend the view that we can benefit people by creating them.⁹

(b) I take the value-for relations *good for*, *bad for* and *neutral for*, at face value, that is, as genuine relations holding between state of affairs and (actual) persons – *mutates mutandi* for the comparative counterparts *better for*, *worse for*, and *equally as good as for*. This gives us a neat explanation of why the following inferences are valid:

Exercise is good for me.

So, there is someone for whom exercise is good.

Everything that is good for me is bad for you.

Exercise is good for me.

So, exercise is bad for you.

Exercise is better for me than binge drinking.

So, there is something that is better for me than binge drinking.

(c) The things that are good for us are abstract states of affairs, some of which constitute whole possible lives. They are abstract in the sense that they can exist without obtaining. If we assumed that the value-bearers were obtaining states of affairs, or concrete events, then given (a) and (b) it would follow directly that non-existence cannot have value for a person.

5. THE VALUE OF EXISTENCE

Let us start with (i), the question about whether existence can be seen as an absolute benefit or harm. Surely creating a person can be *instrumentally* good or bad for the person in the sense that if the person were created, she would lead a life that is intrinsically good or bad for her, that is, good and bad, *in itself*, for her. Causing someone to exist who will lead a happy life is instrumentally good for that person.¹⁰ Causing someone to exist who will lead a wretched life is instrumentally bad for that person. In a similar way, coming into existence can be instrumentally neutral. Just imagine causing someone to exist who will have a life that contains only neutral states of affairs or who will have their bad patches exactly balanced by good patches. It is more controversial whether a person's existence can be intrinsically good or bad. Vitalists would say that existence would be intrinsically good, at least if we assume that a welfare-subject cannot exist without being alive. But the standard view is that it is intrinsically neutral. So, it seems clear then that by creating a person we can harm or benefit a person in an absolute sense. (To avoid cluttering the exposition, I will, in the following, adopt the convention that when a value term occurs unqualified in the text, as in 'x is good', it should be read as implicitly referring to both the intrinsic and the instrumental forms of value, as in 'x is intrinsically or instrumentally good'.)

Some would object here that even if it is true that existence can be good or bad it is not clear that this shows that creating someone can be a genuine case of benefiting or harming because these harms and benefits would lack some features shared by all other harms and benefits. One might think that the problem with benefiting or harming people by creating them is that there is no person to benefit or harm *prior* to the act of creation.¹¹ How can we benefit or harm a person if the person does not yet exist? But this is not a problem. Consider Parfit's example about the future child who is wounded by some broken glass that he left in the woods.¹² It seems clear that his act of leaving this broken glass in the woods harmed the child even though the child did not yet exist.

A much more controversial feature of benefiting (harming) someone by creating her is that these actions do not make her better (worse) off than she would otherwise have been.¹³ But note that this feature is not even present in all cases where we harm or benefit already existing people. For instance, if I have to cause you severe pain and I choose to cause you as little pain as possible, surely I am still harming you in an absolute sense. This is what makes the situation so tragic; I cannot avoid harming you (in an absolute sense). But you would not have been better off if I had done otherwise. Similarly, if I cause you some intense pleasure when my only alternative is to cause you even more intense pleasure, then I am still benefiting you, but I am not making you better off than you would otherwise have been.

6. THE VALUE OF NON-EXISTENCE

What about a person's non-existence? Can this state of affairs be seen as an absolute benefit or harm? I think it is obvious that non-existence cannot be *intrinsically* good or bad for a person. If non-existence could be good (bad) for me, then it would be a benefit (harm) that it is impossible that I enjoy (suffer), for I could not be around to enjoy (suffer) it. But, surely, a benefit (harm) must be possible to enjoy (suffer). No

doubt the state of affairs of my non-existence will still *exist* when I am around, but for you to enjoy a benefit it is not enough that there exists a state of affairs that is good for you. The state of affairs must *obtain* for you to enjoy it. Note also that these conditions on enjoyability hold for *any* kind of benefit, not only welfare benefits. For instance, to be benefited morally, financially, or aesthetically by a state of affairs you must exist and the state of affairs obtain. This shows that one cannot object to my argument by saying ‘Of course, welfare benefits must be enjoyable since well-being has to do with *being* – the value of a *life*. But the benefits of non-existence have nothing to do with well-being’.¹⁴ For no matter what kind of benefits one has in mind, they at least need to be enjoyable.

A related trouble with benefits and harms that cannot be enjoyed is that they could not make a life better or worse. But this is absurd. To say that something is good or bad for a person is to say that it could add or detract this value to his life.¹⁵

This leaves us with the possibility that non-existence is neutral for a person. Note that if we can show that non-existence has neutral value for an individual, then we can also show that a person’s existence can be better or worse than her non-existence. For if non-existence has neutral value and existence can have positive or negative value, then it follows that non-existence can be better or worse than existence. Since existence is a state of affairs that one can be in, indeed has to be in if actualism is true, we could then give an affirmative answer to question (iii) and say that existence is a comparative benefit when it is better than non-existence, and a comparative harm when it is worse than non-existence.

It is popular to claim that non-existence has neutral value, but the claim is not often argued for. Indeed, many writers seem to think that this can be established by mere stipulation.¹⁶ Holtug is an exception and he provides the following argument:

Let us say that existence (or non-existence) has zero value for a person, if and only if no positive and negative values befall her or the positive and negative values cancel each other out. Now suppose that a person exists but that no positive or negative values befall her. Since no positive or

negative values befall her, her life has zero value. Likewise, no positive or negative values befall a person who does not exist. For the same reason, then, we may assign zero value to her non-existence.¹⁷

To this Broome has objected that having no value is not the same thing as having zero value. For instance, logic has no temperature, but that does not mean that it has zero temperature and thus is colder than the ocean. Holtug's answered that even if the presence of certain properties is crucial for zero temperature, it is the absence of certain properties that is crucial for zero value.¹⁸

I side with Broome on this issue but I think the real problem lies elsewhere. The problem is not just whether the absence of certain properties shows that something has neutral value for a person. The more important problem is that in the case of non-existence there is no person for whom things can have neutral value. Holtug moves too quickly from:

- (1) *If a person exists* and no positive or negative values befall her, then her existence is neutral for her.

to

- (2) *If a person does not exist* and thus no positive or negative values befall her, then her non-existence is neutral for her.

If *having value for* is a relation and actualism is assumed then only (1) can be accepted. If a person does not exist, then she is not standing in any relations. So, nothing can be neutral for her. Holtug's formulation masks this difference because he talks about no values befalling a person 'who does not exist', which suggests, misleadingly, that there is person there who lacks something, it is just that he does not exist. But this is of course incoherent given my and Holtug's actualist assumptions.¹⁹

One could reply that 'neutral for' differs from 'good for' and 'bad for' in that it does not express a relation between a person and the presence (or absence of) certain states of affairs. 'neutral for' only expresses the mere lack or the absence of the instantiations of the relations of being good for and

being bad for. More precisely, ' p is neutral for S ' means (or is at least implied by) something like 'the relations being good for and being bad for do not hold between p and S '. Since the relations being good for and being bad for do not have me as one of the relata in a world in which I do not exist, it is then perfectly fine to say that non-existence is neutral for me in this world. In fact, *everything* will be neutral for me in a world in which I do not exist, including states of affairs such as *my suffering horrendously*.

I think it is clear that this is not what 'neutral value for' means.²⁰ First, by giving up the idea that 'neutral for' expresses a relation we would have to treat the following intuitively plausible inference schema as invalid:

p is neutral for S .

So, there is someone for whom p is neutral.

Second, we would have to accept very strange evaluations. Anything that *cannot* be good or bad for me would automatically be neutral for me. For instance, the number 6, the empty set, and the proposition that 2 plus 2 equals 5 would be neutral for me, since these entities cannot be good or bad for me. More importantly, we would have to accept some very peculiar comparative value judgements. Consider a world in which you exist but I do not. Take any state of affairs that is bad for you in this world. Since I do not exist in this world this state of affairs is neutral for me. But then we get the conclusion that it is better for me than it is for you. This means that if you deny that 'neutral for' expresses a relation, you also have to deny that 'being better for A than for B ' expresses a relation.

Finally, this view on neutral value would have very strange implications for egalitarian theories that 'mind the gap' and give weight to differences in well-being. The only way to achieve perfect equality is to make sure that everyone is leading a life that is neutral for them. Why? Well, in any possible world everything will be neutral for the persons who do not exist. So, in order to equalize well-being it is not enough to make sure that all existing people have the same positive

well-being. We will have to make sure that they all hit zero. This is an amazing levelling-down objection that is too good to be true. It cannot be that easy to refute egalitarianism.

Of course, one could reply by saying that vacuous neutral values should not have any moral importance. Only neutral value for existing people matters. But then we cannot use the notion of vacuous neutral value to support the claim that in a morally relevant sense we can make it better for people by creating them.

7. SHIFTY NEUTRALITY

What we have been discussing so far is the claim that a person's non-existence would be neutral for the person even if she did not exist. But perhaps it is enough to claim that non-existence *is* neutral for a person and deny that it *would* still be neutral for a person if she did not exist. Since state of affairs can exist without obtaining, an actualist can happily accept that the non-obtaining state of affairs of someone's non-existence stands in a certain relation to this person, viz. the relation of being neutral in value for her. According to this view, one would still be able to say that existence can be better or worse for persons than non-existence. It is just that these comparative evaluations only hold when the persons exist.

7.1. *An Argument from Desire Theory*

Why would anyone accept this view? If one thinks desires should play a crucial role in defining a person's well-being, one could claim that a person's non-existence is neutral for him if he takes a neutral attitude towards his non-existence, for one could claim that, in general:

p is neutral for *S* iff *S* takes a neutral attitude towards *p*.

Non-existence will then be better or worse for a person if he takes a positive or negative attitude towards his existence, for one could claim that in general:

p is good for S iff S takes a positive attitude towards p .

p is bad for S iff S takes a negative attitude towards p .

Of course, the comparative judgements could be reached more directly by considering a person who prefers or disprefers her existence to her non-existence and by simply assuming that

p is better for S than q iff S prefers p to q .²¹

Now, these arguments are only as plausible as the underlying desire-based theory, and I doubt that many would find this theory attractive. Even if you think that desires play a crucial role in defining well-being, there is a compelling reason to reject this particular version of the desire-based theory: it succumbs to a serious scope problem. If we put no constraints on the objects of the desires, then *any* state of affairs that is logically possible to desire can have absolute or comparative value for a person. But this is too permissive. To be part of a person's well-being a state of affairs must in some clear sense involve the person, and it is not enough that the state of affairs is desired by the person, since his desires might be *other-regarding*. An unconstrained desire theory is therefore not a well-being theory.

A constrained desire theory that still counted non-existence as a proper object of desire would not do better. For if non-existence can have neutral value in virtue of being the object of a neutral attitude, then, surely, non-existence can have positive or negative value in virtue of being favoured or disfavoured. But we have already seen that it does not make sense to say that a person's non-existence has positive or negative value for him.

To avoid this implication by accepting only the link between preferences and betterness and claiming that things can only be better or worse for people, never good or bad, would give us a seriously incomplete theory of well-being. Any plausible well-being theory must have something to say about what is good and bad for people. It is also unclear whether this is a possible position to take given that compar-

ative welfare judgements and absolute welfare judgements seem to be conceptually linked. For instance, if a state of affairs is better for a person than something that is indifferent for the person, then it is good for the person. The absolute sense of indifference seems to be linked with a comparative sense of indifference: something is indifferent for a person just in case its presence is equal in value for the person to its absence. If these links hold, then any desire-theorist who accepts the link between preferences and betterness will also have to accept that there are good or bad things. We only need to find a person who prefers or disprefers something to a state of affairs he is indifferent towards.

7.2. *Inaccessible Comparative Benefits*

The argument from desire-theory does not succeed in showing that existence can be a comparative benefit. There is also a direct argument against the idea that existence can be a comparative benefit. To show this, I have to introduce the promised refinements of my rough characterization of comparative benefits and harms. However, before I do this, I will show that *non-existence* cannot be seen as a comparative benefit or harm.

I said before that a person enjoys a comparative benefit if and only if she is in a state that is better for that person than some alternative state. The left-to-right direction of this biconditional will of course rule out that non-existence can be a comparative benefit or harm, since, given actualism, non-existence is not a state anyone can be in. If you deny this, you would have to accept that there can be comparative benefits (harms) that, if realized, could not make the person better off (worse off).²² This is absurd. Surely, a benefit or harm, no matter whether it is absolute or comparative, must be something that can be enjoyed.²³

As in the case of absolute benefits, there is no point in replying that a person's non-existence is still something that *exists* when the person is around. If absolute benefits need to obtain for a person to enjoy them, then, surely, the same must hold for comparative benefits. Since this must hold for any

kind of comparative benefit, there is also no point in saying that the comparative benefits of non-existent beings are not strictly about well-being, or, to be more precise, better-being.

It is also futile to reply that even if it is true that, strictly speaking, non-existence is not a state you can be in, the sentence 'It is better for *S* to be in a state of non-existence' could still be true, for it only requires that 'Non-existence is neutral for *S*' is true in a world in which *S* does not exist. This possibility is excluded, if my arguments in the previous section are sound.

Although the definition of comparative benefits and harms rule out that non-existence can be a comparative benefit or harm, it does not rule out the same possibility for *existence*. A person's existence *obtains* when he exists, and is thus a state he can be in. So, if we can show that existence is better or worse for the person than non-existence, then my definition of comparative benefits and harms implies that existence can be a comparative benefit or harm. It is here we need to introduce some refinements of the definition. To be benefited (harmed) comparatively is not just to be in a state that is better (worse) for you than some alternative state; it must also be true that things *would* have been worse (better) for you in the alternative state. It is in this sense a comparative benefits constitute a gain in value and a comparative harm a loss in value. Since nothing has value for you if you do not exist, not even neutral value, you are not better off (or worse off) created than you would have been had you not been created.²⁴

This refinement can be captured by the following conditions on the notion of betterness (worseness) involved in comparative benefits (harms):²⁵

Accessibility: If *p* is better (worse) for *S* than *q*, then *p* would be better (worse) for *S* than *q* even if *p* obtained.

Converse: *p* is better for *S* than *q* iff *q* is worse for *S* than *p*.

It is easy to see that these conditions together entail that existence cannot be a comparative harm. If existence is better (worse) than non-existence, then by *Converse* non-existence is worse (better) than existence. According to *Accessibility*, this

means that non-existence would be worse (better) than existence even if it obtained. But this is impossible.

I am not claiming that these conditions hold for all notions of value for. For instance, *Accessibility* does not hold if we read 'better for' simply as 'judged better by' or 'preferred by'. But remember that we are interested in the interpretation of 'better for' that is conceptually linked to well-being or, more broadly, to benefits and harms.

7.3. *The Relation Between Betterness and Worseness*

My conclusion that existence cannot be a comparative benefit rests on these extra refinements. Ryberg accepts *Accessibility* but denies *Converse*.²⁶ This means that he can accept that creating a happy person would be to make it better for him without accepting that not creating the person would be to make it worse for him. For in this case, whereas the happy person's existence is better for him than his non-existence, his non-existence is not worse for him than his existence.

This is a desperate move to make, as desperate as denying that 'taller than' and 'shorter than' are conceptual mirror-images. If there are any value principles that merit a conceptual status *Converse* is clearly one of them. Of course, there are still some *pragmatic* differences between 'better' and 'worse'. For instance, we tend to use 'better' when the higher-ranked alternative is good, and 'worse' when the lower-ranked alternative is bad. But this does not affect the conceptual point that 'worse' is the converse of 'better'.

Note also that if we follow Ryberg here, we cannot even say that neutral things are always worse than good things. For if non-existence is neutral for a person and his life is good for him, we are not allowed to conclude that his non-existence is worse for him than his happy life. Nor can we any longer say that neutral things are always better than bad things, for if non-existence is neutral for a person and his life is bad for him, we cannot conclude that his non-existence is better for him than his unhappy life.

7.4. *Normative Variance*

More problems are generated when this inaccessible betterness is used to determine what is morally right or wrong. These problems occur if one accepts the following reasonable principles of non-malevolence:²⁷

- (NM1) If the outcome of action *A* is worse for someone and not better for anyone than that of any alternative action, then, other things being equal, *A* is wrong.²⁸
- (NM2) If the outcome of *A* is not worse for anyone than that of any alternative action, then, other things being equal, *A* is not wrong.

If better-for is inaccessible in the sense that a person's existence is worse than his non-existence but could not be worse if he did not exist, then we will have a violation of normative invariance:

Normative Invariance: An action's normative status does not depend on whether or not it is performed.²⁹

To see this, consider the following example in which the agent faces a choice between creating me and thereby realizing world 1 or not creating me and thereby realizing world 2

World 1	World 2
I am created and have an unhappy life	I am not created
My existence is worse for me than my non-existence.	<i>It is false that</i> my existence is worse for me than my non-existence.
Creating me is wrong (by NM1)	<i>It is false that</i> creating me is wrong (by NM2) ³⁰

(no other persons will be affected by the choice):

So, whether it is right to create me depends on whether I am going to be created.

This normative variance is troubling since it is difficult to use a moral theory as a guide to action if whether you ought to do something depends on whether you will do it. When you use a theory as a guide to action, you use the theory in

your deliberations about what to do. On the basis of this deliberation you then make up your mind and decide what to do. But if an action's rightness depends on whether it is performed, then in order to decide whether an action is right you first have to know whether or not you are going to perform it. But there is no point in deliberating about whether to perform an action if either you believe that you will perform it, or you believe that you will not perform it. If you believe that you will perform the action, the issue is settled for you, and there is no point in deliberating about it further. If you believe that you will not perform the action, the action is no longer a serious possibility, that is, something that is compatible with what you believe (even if it might be something you can do); so again there is no point in deliberating about whether to perform it.³¹

To forestall a possible misunderstanding, I should stress that my view is *not* that once you have formed the belief (or disbelief) that you are going to do *A*, then you are no longer able to deliberate about whether to do *A*. If you give up the belief or the disbelief, you may of course start deliberating again. The claim is rather that *while* you are in the grip of the belief or disbelief that you will do *A*, it is not possible for you to deliberate about whether to do *A*.

A second and related objection is that the shifty better-for generates moral dilemmas: some situations involve unavoidable wrong-doing in the sense that whatever you were to do, you would do something that would be wrong. The following

World 1	World 2
I am created and have an unhappy life, you are not created	You are created and have an unhappy life, I am not created
My existence is worse for me than my non-existence.	<i>It is false that</i> my existence is worse for me than my non-existence.
<i>It is false that</i> your existence is worse for you than your non-existence.	Your existence is worse for you than your non-existence.
Creating me is wrong (by NM1)	Creating you is wrong (by NM1)

example is an illustration. (Again, the agent has only two options, but this time he can choose between creating me and creating you.)

What we have here is a kind of dilemma, but one that differs from the usual kind. Normally, a dilemma is seen as a situation in which all available actions *are* wrong. This is not the situation here. No matter how you act, there is an available act that *is* right. If I were to be created, then it would be right to create you. If you were created it would be right to create me. But this is not much comfort, for you cannot act in such a way that you comply with the theory: there is no action such that if you were to perform that action you would act rightly. No matter how you *were* to act, you *would* be damned.

Why is it important that a normative theory is possible to comply with? Again, the answer has to do with the practical role of a normative theory. A normative theory is supposed to guide action and thus be usable at least by agents who are ideal in the sense that they lack all cognitive and motivational defects. But if we cannot comply with a theory in a situation, we cannot be guided by it in that situation, since to be guided by a theory is to use that theory when deliberating about what to do and on the basis of this deliberation act in compliance with the theory.

The problem of normative variance is not insurmountable, however. A defender of inaccessible betterness could avoid it by revising the common-sense principles of non-malevolence in some suitable way. Indeed, this is what both Holtug and Ryberg seem to do.³² They do not take a definite stand on how the normative status of an action should be determined, but they suggest a theory that is extensionally equivalent to a consequentialist theory, according to which the value of an outcome is determined by the good and bad states included in it, and the right action is the one that brings about the best outcome. Note that this consequentialist theory does not have to accept that people can be benefited or harmed comparatively by being created. It focuses exclusively on what is absolutely good or bad for people. When the theory goes on

to claim that we ought to bring about the best outcome, we will not have any problems with normative variance, since facts about what absolute benefits and harms an outcome contains do not depend on which action is performed. But given this straightforward option, one could wonder what the point is with accepting inaccessible betterness. If, in the end, what we get is a moral theory that tells us something that we could have said without accepting inaccessible betterness, why bother taking an axiological detour that forces us to accept that either there are comparative benefits that cannot be enjoyed or that 'worse for' is not the converse of 'better for'? Surely, it would be better to accept that we cannot make it better or worse for people by creating them and agree that the simple principles of non-malevolence stated above are true only when applied to same people choices. It is, of course, another matter whether we should accept a consequentialist theory that focuses exclusively on absolute harms and benefits. This theory is not attractive if you think there is a morally relevant difference between choosing contraception and having a late abortion. If you want to draw a distinction between these two cases, then you will have to complicate your moral theory by assigning different weights to the absolute benefits or harms of created people and the comparative benefits or harms of already existing people.

7.5. *The Guardian Angel*

Rabinowicz has suggested a possible defence of inaccessible betterness.³³ It seems that someone who exclusively cares about me – my guardian angel – could truthfully say that he prefers, for my sake, that I had not come to existence if my life is bad for me. But that my guardian prefers my non-existence to my existence, for my sake, seems to show that my non-existence is better for me than my existence. We do not need to say that this link provides a definition of well-being. To avoid any charges of circularity, we could say that the link expresses a condition of adequacy rather than a definition.

Rabinowicz's argument does have some force. It seems to be true that in some sense my guardian angel can, for my sake, desire that I had not been born if my existence is horrible. But I think we can make sense of this attitude without assuming that existence can be a comparative harm. My guardian angel cares not just about my comparative benefits and harms. He also cares about the absolute harms and benefits that befall me. I have already acknowledged that by creating people we benefit or harm people in an absolute sense if we make sure that they will lead lives that are good or bad for them. Surely, my guardian angel will be sensitive to these facts about my absolute harms and benefits. He will favour my existence if it is good for me, and disfavour it if it bad for me. I have, therefore, no problems accepting that something is *good* (bad) for me just in case my guardian angel *favours* (disfavours) it.

What about the comparative preference? Can I make sense of my guardian angel's preference for my non-existence over my existence? I think I can. My guardian angel will not just disfavour things that are bad for me; he will also want that things that are bad for me are prevented from occurring. In some situation, the only way bad things can be prevented is to make sure that the person is never born. My guardian angel's preference for my non-existence can then be seen as a preference that is based on his interest in bad things being prevented from happening to me. His preferences is not based on an assumption that I would be better off if I did not exist, or that I would be worse off if I existed. So, there is no sense in which I am saved from a terrible fate by not being created. Note that it is exactly this fact that makes room for drawing a moral distinction between harming or benefiting people by creating them and harming or benefiting people who already exist.

This leads to me to think that Rabinowicz and I agree more than we disagree. We both agree that nothing can be good, bad, or neutral for people who do not exist. We also agree that we cannot make people better off or worse off by creating them. The sticking points are whether non-existence can be neutral for people who exist, and whether we should allow that

non-existence can be better or worse for people than existence. The first thing to note is that the guardian angel criterion does not show that we have to assign intrinsic neutrality to non-existence. Nor does it show that non-existence has instrumental value of any kind, that is, value in virtue of having intrinsically valuable consequences. The only thing the criterion commits us to is a notion of *preventive* value.³⁴ Some things have value not in virtue of itself or its consequences, but in virtue of what it prevents or precludes. This value concept is not mysterious. It is, for instance, employed in cases where we judge a vaccine to be good. A vaccine is not usually good in virtue of itself or its consequences. It is good in virtue of what it prevents, that is, an illness of some sort. In a similar vein, we could accept that non-existence is sometimes the only effective vaccine there is against suffering certain bad things. We could then say that non-existence can be *preventively* good for a person, to some degree, in the sense that things would not have been bad, to the same degree, for the person, if he had not existed. This also shows that it can make sense to wish that one had not been born. If one's life is hell and there is no way out of it, one's non-existence is preventively good.

We could also move on to say that if non-existence is preventively good, then it is preventively better than existence. For to say that *p* is preventively better in some respect than *q* is just to say that (a) there is something that is bad and that is prevented by *p* but not prevented by *q*, or (b) there is something that is good and that is prevented by *q* but not by *p*. It is only in this limited sense we can say that non-existence is better for a person than existence.

It is important to stress, however, that this sense is indeed limited, for the preventive betterness of non-existence can never be enjoyed. Remember that to enjoy an absolute or comparative benefit you will have to be in a state that is good or better for you than some alternative state. Since no one can be in a state of non-existence, no one can enjoy the preventive goodness or betterness of non-existence. This means that if you believe that preventive goodness of non-existence is a kind of benefit, you would also have to believe the absurd view that

there are benefits that no one can enjoy. So, even if Rabinowicz has succeeded in showing that there is a limited sense in which non-existence has value, he has not shown that this value can be seen as a genuine benefit.

8. CONCLUDING REMARKS

The findings of this paper are the following:

- (1) Existence can be good, bad or neutral for a person, at least if we assume that the values are instrumental. So, there is a clear sense in which we can benefit or harm a person by creating her. We make sure that a person enjoys an absolute benefit if we create her and her life is good for her. We make sure that a person suffers an absolute harm if we create her and her life is bad for her.
- (2) Non-existence cannot be an absolute benefit or harm, for all benefits and harms must be enjoyable. The only sense in which we can say that non-existence has value is that it prevents certain intrinsic values. But this preventive value cannot be seen as a benefit or harm.
- (3) Existence cannot be a comparative benefit or harm. Although existence is a state a person can be in, it cannot be seen as comparative benefit or harm because to receive a comparative benefit or harm is to gain or lose some value in relation to an alternative situation. So, it has to be true that things *would* have been worse or better for the person in the alternative state. But things would not have been better or worse for the person if he had never existed. They would have lacked all value for the person, including neutral value. Existence can, at most, be preventively better or worse for a person. But this preventive value cannot be seen as a comparative benefit or harm.
- (4) Non-existence cannot be a comparative benefit or harm, since all benefits and harms must be enjoyable. Non-existence, can, at most, be preventively better or worse for a person. But this preventive value cannot be seen as a comparative benefit or harm.

One important reason why these conclusions might be resisted is that it is easy to be misled by the slippery terms ‘better for’ and ‘worse for’. As I acknowledged in the introductory section, these terms can stand for many different things. What I have argued is that if they are seen as capturing facts about well-being, or, more generally, facts about benefits and harms, it makes no sense to say that non-existence can be better or worse for a person than existence. If, in contrast, they are seen as conveying descriptive facts about what the person judges to be better, or what the person prefers, then the comparisons between non-existence and existence do make sense. But to take this latter interpretation as a reason for claiming that non-existence can be a comparative benefit or harm would be a serious case of equivocation.³⁵

NOTES

¹ Parfit (1992, pp. 355–356).

² No revision is of course called for if one thinks that late abortion does not affect any existing person or individual. But if one thinks that a person starts to exist later, some time after she is born, then other revisions will be called for. In particular, infanticide and use of contraception will be on a par.

³ See, for instance, Ryberg (1995), Holtug (2001), Persson (1995) and Hare (1975). Among these authors, only Holtug and Hare explicitly claim that we can harm people, at least extrinsically, by not creating them.

⁴ See, for instance, Heyd (1992, pp. 30–31, 122), Williams (1973, pp. 82–100) and Bigelow and Pargetter (1988). Narveson (1978) defends the claim that we can harm but not benefit people by creating them.

⁵ Parfit (1992, pp. 487–490), Nagel (1979, p. 7), McMahan (1981) and Hurley (2003, pp. 118–119).

⁶ In the following, I am going to use ‘individuals’, ‘persons’, and ‘people’ interchangeably. These terms are supposed to be understood loosely as referring to whatever we consider being proper welfare-subjects. I am not going to take a stand on what qualifies someone to be a welfare-subject.

⁷ Gustaf Arrhenius reminded me about this.

⁸ I am, of course, not denying that a person can be comparatively benefited or harmed by the *way* we choose to create him. A woman who chooses to take heavy drugs during sexual intercourse might affect the person’s well-being in the sense that the person would have been better off if the woman had stayed away from the drugs while having intercourse.

What is at issue in the present discussion is whether the act of creation itself, not the way it was performed, can bring about a comparative benefit.

⁹ It is not clear that possibilism will help. According to possibilism, an individual does not have to be an actual thing in order to have properties. But that does not show that a possible person qualifies as a welfare-subject. In the same way that a merely possible fat man is not a fat man a merely possible person is not a person. A merely possible person is something that is only *possibly* a person. So, a merely possible person is at best a merely possible welfare-subject.

¹⁰ As an anonymous referee pointed out, 'instrumental value' is sometimes used in a narrower sense, referring to the value something has in virtue of standing in a strict *causal* relation to something that is intrinsically valuable. In this narrower sense, it might not be true to say that creating someone has instrumental value, since it seems a bit odd to say that creating someone causes her to have a happy life.

¹¹ This objection was raised by an anonymous referee.

¹² Parfit (1992, p. 356).

¹³ See, for instance, Parfit (1992, p. 489) and Österberg (1996, p. 99).

¹⁴ Wlodek Rabinowicz made me aware of this objection.

¹⁵ I do not want to say that it *would always* add or detract this value to the person's life, since there might be cases of organic unities in which adding a good thing can make the whole life worse due to some relational features between the added state and the rest of the life.

¹⁶ See, for instance, Feldman (1991, p. 311).

¹⁷ Holtug (2001, p. 381).

¹⁸ *Ibid.*, pp. 381–382. Broome's objection is stated in footnote 38 on p. 381.

¹⁹ The same misleading formulations occur when Holtug summarizes a more direct argument for the claim that existence can be better for a person than non-existence. In defence of this claim he writes: '...it seems to be better to have a surplus of value than to have no value. Contrariwise, it seems to be worse to have no value than it is to have a surplus of value.' Holtug (2001), p. 373. This suggests that a person who did exist would have something, namely 'no value'. But, according to actualism, if a person did not exist, he would not have anything, not even 'no value'.

²⁰ Holtug (in conversation) claims that he is only talking about the *truth-makers* of the sentence '*p* has neutral value for *S*', not its meaning. However, my objections apply equally to a truth-maker account.

²¹ This is what Ryberg (1995) does on pp. 106–110. Holtug (2001) also mentions this possibility on pp. 365–366, but he does not take a stand on whether this desire-theory is true. He claims only that it is a *possible* well-being theory.

²² That there is a connection between ‘better for’ and ‘better off’ is often stressed in the literature. See, for instance, Broome (2004, p. 63) and Carlson and Olsson (2001, p. 89).

²³ Holtug explicitly denies that this applies to harm. He claims that a person can be (comparatively) harmed by his non-existence. This is a remarkable claim. But what Holtug means is only that ‘[the person] *S* does not exist and had he existed, his existence would be intrinsically better for him than his non-existence’, Holtug (2001, p. 376). Clearly, this is a flawed analysis of comparative harms. Whether existence has absolute or comparative *intrinsic* value should be left open. More importantly, to be harmed comparatively is to actually be in a worse state. For instance, we cannot truthfully say that the jolly Santa Claus *is* harmed by not being created. What we can say is, at most, that he *would* (or *might*) have been benefited (in an absolute sense) if he had existed. Note also that by defining harm in this way Holtug is, in effect, conceding that we cannot harm non-existing people in the same sense as we harm already existing people.

²⁴ I would say the same thing about temporal comparative benefits. If you do not persist from one time to another you cannot undergo a temporal change. You cannot, therefore, be better off now than you were at a time before you were born.

²⁵ A qualification needs to be added about comparative harms. Even if *p* is worse than *q*, we may want to say that *p* is a possible harm only if the better state *q* involves something essential and important that is missing in *p*. The millionaire who does not win an extra thousand pounds on the lottery is perhaps not harmed even though he is worse off. This qualification does not matter to my argument, however, since when we say that non-existence is instrumentally worse than existence we are comparing non-existence with a *whole life*, which can include all the relevant essential and important factors.

²⁶ Ryberg (1995, p. 109). This idea was also endorsed in Holtug (1996, p. 77), but he later changed his mind and opted for an inaccessible betterness relation. See Holtug (2001, p. 374).

²⁷ Of course, it is reasonable to apply them to different people choices only if we assume, mistakenly, that there is no way we can affect people without making things better or worse for them. Since I am arguing that we can affect people by making things *good* or *bad* for them without making things better or worse for them, I think the principles should be applied only to same people cases.

²⁸ By ‘other things’ I mean to include all moral considerations that fall outside the scope of the morality of beneficence (or non-malevolence). Examples of such considerations are equality, rights, and deontological constraints and options.

²⁹ For a thorough discussion of this principle, see Carlson (1995, pp. 100, 114–115, 162). The principle and its name originate with Wlodek Rabinowicz.

³⁰ This might look like I am committed to saying that a non-actual person has a relational property, namely the property of being such that creating her would be wrong, and this would be to betray my actualist conviction. But the state of affairs of its not being the case that creating me is wrong does not ascribe a property to me, it ascribes a property to the proposition or states of affairs *that I am created*.

³¹ This argument for Normative Invariance is spelled out in Carlson (1995, pp. 101–102). Bricker (1980, p. 395), touches on a similar argument. The general idea that the prediction of one's actions crowds out deliberation has wide-spread support. See, for instance, Goldman (1970, p. 194) and Taylor (1966, p. 174). Cf. also Ginet (1962), Levi (1997), Schick (1988) and Spohn (1977).

³² Ryberg (1995, pp. 114–116) and Holtug (1999, esp. p. 36).

³³ In conversation.

³⁴ For a more thorough treatment of preventive value see Zimmerman (2001, pp. 256–257) and Bradley (1998).

³⁵ Earlier versions of this paper were presented at Stockholm University, Lund University, and Uppsala University. I am grateful for comments from the audiences on these occasions. I would like to thank, in particular, Gustaf Arrhenius, John Broome, Johan Brännmark, Wlodek Rabinowicz, Folke Tersman, Torbjörn Tännsjö, and Michael Zimmerman. For written comments on earlier versions, I would like to thank Gustaf Arrhenius, Nils Holtug, Petter Karlsson, and Wlodek Rabinowicz.

REFERENCES

- Bigelow, J. and Pargetter, R. (1988): 'Morality, Potential Persons and Abortion', *American Philosophical Quarterly* 25, 173–181.
- Bradley, B. (1998): 'Extrinsic Value', *Philosophical Studies* 91, 109–126.
- Bricker, P. (1980): 'Prudence', *Journal of Philosophy* LXXVII(7), 381–401.
- Broome, J. (2004): *Weighing Lives*, Oxford: Oxford University Press.
- Carlson, E. (1995): *Consequentialism Reconsidered, Theory and Decision Library*, Dordrecht: Kluwer.
- Carlson, E. and Olsson, E. (2001): 'Existence, Beneficence, and Design', in E. Carlson and R. Sliwinski (eds.), *Omnium-gatherum, Uppsala Philosophical Studies* 50, Uppsala: Department of Philosophy, Uppsala University.
- Feldman, F. (1993), 'Puzzles About the Evil of Death' in *The Metaphysics of Death*, Y.M. Fisher (ed.), Stanford: Stanford University Press.
- Ginet, C. (1962): 'Can the Will Be Caused?', *The Philosophical Review* 71, 49–55.

- Goldman, A.I. (1970): *A Theory of Human Action*, Princeton, NJ: Princeton University Press.
- Hare, R. M. (1975): 'Abortion and the Golden Rule', *Philosophy and Public Affairs* 4, 147–167.
- Heyd, D. (1992): *Moral Issues in the Creation of People*, Berkeley, CA: University of California Press.
- Holtug, N. (2001): 'On the value of coming into existence', *The Journal of Ethics* 5, 361–384.
- Holtug N. (1996): 'In Defence of the Slogan', in W. Rabinowicz (ed.), *Preference and Value: Preferentialism in Ethics*, Studies in Philosophy, Department of Philosophy, Lund.
- Holtug, N. (1999): 'Utility, Priority and Possible People', *Utilitas* 11(1), 75–84.
- Hurley, S. L. (2003): *Justice, Luck, and Knowledge*, Cambridge, MA: Harvard University Press.
- Levi, I. (1997): 'Consequentialism and Sequential Choice', in his *Covenant of Reason* (pp. 701–101), Cambridge University Press.
- McMahan, J. (1981): 'Problems of Population Theory', *Ethics* 92, 96–127.
- Nagel, T. (1979), *Mortal Questions*, Cambridge: Cambridge University Press.
- Narveson, J. (1978): 'Future People and Us', in R.I. Sikora and B. Barry (eds.), *Obligations to Future Generations*, Philadelphia: Temple University Press pp. 38–60.
- Österberg J. (1996): 'Value and Existence: The Problem of Future Generations', in S. Lindström et al. (eds.), *Odds and Ends*, Uppsala Philosophical Studies 45, The Department of Philosophy, Uppsala University, Uppsala.
- Parfit, D. (1992): *Reasons and Persons*, Oxford: Clarendon Press.
- Persson, I. (1995): 'Genetic Therapy, Identity and Person-Regarding Reasons', *Bioethics* 9(1), 17–25.
- Ryberg, J. (1995): 'Do Possible People have Moral Standing?', *Danish Yearbook of Philosophy* 30, 96–118.
- Schick, F., (1988): 'Self-knowledge, Uncertainty, and Choice', in P. Gärdenfors and N-E. Sahlin (eds.), *Decision, Probability, and Utility*, Selected Readings (pp. 270–286), Cambridge, UK: Cambridge University Press.
- Spohn, W. (1977): 'Where Luce and Krantz Do Really Generalize Savage's Decision Model', *Erkenntnis* 11, 113–134.
- Taylor, R. (1966): *Action and Purpose*, Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Williams, B. (1973): *The Makropulos case: reflections on the tedium of immortality* in his *Problems of the Self* (pp. 82–100), Cambridge, UK: Cambridge University Press.
- Zimmerman, M. (2001): *The Nature of Intrinsic Value*, Maryland: Rowman and Littlefield.

Jesus College, Philosophy Faculty

Oxford University

Turl Street OX1 3DW

Oxford, Oxfordshire

UK

E-mail: kriste.bykvist@jesus.ox.ac.uk