Being and Betterness

JENS JOHANSSON

Stockholm University

In this article I discuss the question of whether a person's existence can be better (or worse) for him than his non-existence. Recently, Nils Holtug and Melinda A. Roberts have defended an affirmative answer. These defenses, I shall argue, do not succeed. In different ways, Holtug and Roberts have got the metaphysics and axiology wrong. However, I also argue that a person's existence can after all be better (or worse) for him than his non-existence, though for reasons other than those provided by Holtug and Roberts.

I. INTRODUCTION

If your life is on the whole good, it may be natural to think that your existence is better for you than your non-existence. That is, of the two alternatives, coming into existence and never coming into existence, the former is better for you than the latter. Likewise, if your life is on the whole bad, it might be natural to think that your existence is worse for you than your non-existence.

But think again. The view faces a serious difficulty. If it is better (or worse) for you to exist than never to exist, it seems to follow that if you had never existed, then never existing would be worse (better) for you. But nothing can be better or worse for someone who never exists.[1]

In spite of this argument, some philosophers – most notably Nils Holtug and Melinda A. Roberts – have recently provided extensive defenses of the view that existence can be better or worse for a person than non-existence.[2] These defenses, I shall argue, do not succeed. In different ways, Holtug and Roberts have got the metaphysics and axiology wrong. Nonetheless, the thesis that they fail to defend is defensible after all. Thus, I shall argue that your existence can indeed be better or worse for you than your non-existence, and that the above argument to the contrary can be met.

Some clarifications and background assumptions are in order. First, the thesis under consideration here is: 'A person S's existence can

---

[1] See e.g. J. Broome, 'Goodness is Reducible to Betterness: The Evil of Death is the Value of Life', *The Good and the Economical*, ed. P. Koslowski and Y. Shionoya (Berlin, 1993), p. 77; K. Bykvist, 'The Benefits of Coming into Existence', *Philosophical Studies* 135 (2007), pp. 335–62; D. Parfit, *Reasons and Persons* (Oxford, 1984), p. 489.

[2] N. Holtug, 'On the Value of Coming into Existence', *The Journal of Ethics* 5 (2001), pp. 361–84; M. Roberts, 'Can it Ever Be Better Never to Have Existed At All? Person-Based Consequentialism and a New Repugnant Conclusion', *Journal of Applied Philosophy* 20 (2003), pp. 159–85. Roberts concentrates on the view that existence can be *worse* for a person than non-existence.

be better or worse for S than S's non-existence'. I assume that the statement that S's existence is (e.g.) better for S than S's non-existence is to be understood as saying that (i) S's existence, (ii) S and (iii) S's non-existence are related, in that order, by (iv) the *x is better for y than z* relation.

Second, I follow many other participants in the debate in taking 'S's existence' to denote the state of affairs that S exists at all (viz. at some point), and 'S's non-existence' to denote the state of affairs that S never exists.[3] States of affairs are fine-grained, propositional entities that can exist without obtaining.

Third, like nearly all parties to the debate, I endorse 'actualism', which in this context means that an object has to actually exist (in a possible world w) in order to exemplify a property or relation (in w).[4]

Fourth, as some of the above phrases suggest, I am taking the liberty of writing as if there really are such things as relations and properties. More specifically, I am writing as if predicates like, 'x is good for y', 'x has zero value for y', 'x is bad for y', 'x is better for y than z', 'x is equally good for y as z', 'x is worse for y than z', etc. refer to genuine things – presumably universals of a certain kind. This is not a crucial assumption: I am speaking like this merely because I find it convenient. What is important in the above three points is that, in order for S's existence to be better for S than S's non-existence, the relevant states of affairs and S have to exist and jointly satisfy the predicate, 'x is better for y than z'. So do not feel excluded if you are a nominalist.

Fifth, it will only muddle the issues to suppose that merely possible people *exist*, only not as people. That view seems to be endorsed by Josh Parsons, who says that a 'merely possible person is a representation of a person as having a certain level of welfare'.[5] He goes on to argue that a merely possible person cannot be benefited or harmed, since a representation cannot be benefited or harmed. But this view is incompatible with the natural idea that if something is a merely possible person, then it is possibly a person. Surely no representation of a person could have been a person (just as no representation of a tree could have been a tree).

---

[3] Note that I take 'S's existence' to denote S's *mere* existence, rather than, e.g., the state of affairs that S exists and leads a life with a particular content. See further Section IV.

[4] Palle Yourgrau is an exception (by which I mean that he denies actualism, not that he succeeds in exemplifying properties or relations without being actual). See his 'The Dead', *The Journal of Philosophy* 86 (1987), pp. 84–101.

[5] J. Parsons, 'Axiological Actualism', *Australasian Journal of Philosophy* 80 (2002), pp. 137–47, at p. 139. Cf. Bykvist, 'Violations of Normative Invariance: Some Thoughts on Shifty Oughts', *Theoria* 73 (2007), pp. 98–120, at p. 101n.: 'One could . . . identify merely possible people with certain complexes of actually existing properties.' If such complexes exist, this suggestion leads to the view that merely possible people exist.

Sixth, in analyzing modal statements, I will make use of the notion of possible worlds. As I see it, a possible world, w, is a maximal state of affairs that could have obtained. It is maximal in the sense that, for every state of affairs A, either w entails A or w entails not-A.[6] To say that something, e.g. S, exists 'in' or 'at' w is just to say that w entails that S exists. I like this account of possible worlds best, but similar accounts – e.g. ones that identify possible worlds with some other abstract entities than states of affairs – would work fine as well.

Seventh, it is important in what follows that we distinguish between intrinsic and extrinsic value for a person. Neither I nor anyone else whose views I will discuss here (as I understand them) claims that existence – i.e. mere existence – can be *intrinsically* better or worse for a person than his non-existence. That would presuppose the dubious idea that someone's existence or non-existence can have non-zero intrinsic value for him. Rather, what is at issue is whether existence can be *extrinsically* better or worse than non-existence for a person.

It may be useful if I state at this early point what I think makes a state of affairs A extrinsically better (worse) for a person S than not-A. As I see it, this is that the intrinsic value for S of the closest possible A-world (viz. the closest possible world where A obtains) is higher (lower) than the intrinsic value for S of the closest possible not-A-world. Details will be discussed later.

The article is organized as follows. In Sections II–III, I present and criticize Roberts's account. In Section IV I turn to Holtug's proposal. Then, in Sections V–VI, I put forward an account I prefer and defend it against some objections. In the final section, VII, I consider some related normative issues.

II. ROBERTS'S APPROACH

Roberts considers a case where a person, Nora, is born into an unequivocally anguished existence. Roberts advocates:[7]

> BNB   Any alternative at which Nora never exists at all would have been better for Nora than Nora's anguished existence in fact is.

By 'alternative' Roberts means a whole possible world, not a simple state of affairs like Nora's non-existence. Thus, strictly speaking, BNB does not by itself answer the question we are interested in regarding Nora, namely, whether Nora's existence is worse for her than her non-existence. However, it is easy to see how BNB might be taken

---

[6] 'Not-A' refers, of course, to the state of affairs that A does not obtain.
[7] Roberts, 'Can it Ever Be Better Never to Have Existed At All?', p. 168.

to support an affirmative answer. Supposing that 'better' in BNB means intrinsically better, and that Roberts intends 'Nora's anguished existence' to refer to the closest possible world where Nora exists (i.e. the actual world), the combination of BNB and the view of extrinsic value I just outlined apparently yields the thesis that Nora's existence is extrinsically worse for her than her non-existence.

What then about the 'serious difficulty' noted in the beginning – the objection appealing to the claim that nothing can be better or worse for someone who never exists? Roberts interprets this charge as an instance of the general actualist idea that non-existent objects cannot exemplify properties or relations. Her main strategy is to argue that her position only requires Nora to exemplify the relevant properties and relations in the actual world, where Nora does exist. Hence, Roberts thinks her view respects actualism.

Roberts asks us to consider the statement that Sadie has more money on deposit at the bank in the actual world w1 (where she exists) than in world w2 (where she doesn't). Roberts suggests that while the truth of this statement requires Sadie to have certain properties in w1, such as *having more money on deposit at the bank in w1 than in w2*, it does not require her to have any properties in w2. 'We are rather comparing two *amounts* – the amount of money Sadie in fact has on deposit at the bank and the amount of money Sadie has on deposit at the bank at [w2].'[8] The appropriate amount, $0, is what needs to exist at w2, and, Roberts insists, it does.

Similarly, Roberts suggests, with the statement that Nora has less well-being in the actual world w1 (where she exists) than in world w2 (where she doesn't). While the truth of this statement requires Nora to have certain properties in w1, such as *having less well-being in w1 than in w2*, it does not require her to have any properties in w2. We are rather comparing two *amounts* – the amount of well-being Nora in fact has (which is negative) and the amount of well-being Nora has at w2. The appropriate amount, zero well-being, is what needs to exist at w2, and, Roberts insists, it does.

But does not the claim that Nora has zero well-being in w2 imply that Nora exemplifies the property of *having zero well-being* in w2? No, says Roberts. To begin with, not every predicate expresses a property: take, for instance, 'x such that x is not a member of x' (where x is any set). Moreover, the claim that Nora has zero well-being in w2 merely requires that it be false in w2 that Nora has the properties needed for positive or negative well-being.

---

[8] Roberts, 'Can it Ever Be Better Never to Have Existed At All?', p. 177.

These two reasons strike me as unconvincing. And I accept neither BNB nor the appeal to the money analogy. Let me explain.

III. PROBLEMS WITH ROBERTS'S APPROACH

Consider BNB first. As we saw, Roberts wants to respect actualism. However, BNB violates actualism. According to BNB, if Nora had not existed, then that world, w2, would be better for her than the actual world w1, where she exists. Alternatively put, if Nora had not existed, then the following would be the case: w2 *is* better for her. But actualism precludes anything from being better for someone in w2 who does not exist in w2. Thus, Roberts is wrong to think that BNB merely requires Nora to exemplify properties and relations in the actual world, w1.

But didn't Roberts's money analogy show that that is all BNB requires? In fact, if anything, the analogy shows the opposite of what Roberts intends. Again, Roberts grants that if Sadie has more money on deposit at the bank in the actual world w1 than in w2, then Sadie exemplifies the property of *having more money on deposit at the bank in w1 than in w2*. Given this, it would be unnatural to deny that Sadie, w1, and w2 enter into the relation of *x has more money on deposit at the bank in y than in z*. Now, surely if Sadie has more money on bank deposit in w1 than in w2, then in w2 she has less money on bank deposit than in w1. Of course, the relation of *x has less money on deposit at the bank in y than in z*, no less than *x has more money on deposit at the bank in y than in z*, holds only if its exemplifiers exist. They do, so there is no violation of actualism so far. But, since Sadie does not exist in w2, the relation does not hold *there*. So it is false (even in w1) that, if Sadie had not existed, then Sadie *would* have had less money on bank deposit than in w1. But this false claim is what is analogous to BNB. Accordingly, if the money analogy is a good analogy, it does not establish BNB, but its negation.

It may be easy to be misled here by the reasonableness of the claim that, e.g., even though Sadie has never been, and never will be in Scotland, she has more money on bank deposit in the US (where she has quite a lot) than in Scotland (where she has nothing) – and less in Scotland than in the US. This case is different from the case of w1 and w2, however. For, in the relevant sense, Sadie does exist in Scotland: someone in Scotland can express the proposition that Sadie exists and thereby utter a truth. But no one can in w2 express the proposition that Sadie exists and thereby utter a truth.

It may be illuminating to compare the statement that Sadie has more money than her non-existent little brother. That statement is not true: actualism precludes the non-existent from entering into relations. And

then surely we must say the same thing in the case of Sadie's non-existent 'w2 self'.

Let us return, finally, to the charge that, if Nora has zero well-being in w2, then Nora must, in w2, exemplify the property of *having zero well-being*. Roberts's first point, that not every predicate expresses a property, is plausible. Indeed, her example – 'x such that x is not a member of x' – establishes a stronger claim: that not every predicate that is true of certain objects expresses a property. But, as noted in §I, the important thing is not whether there are properties, but that Nora has to exist in order for the predicate 'has zero well-being' to apply to her. She clearly has to, because of the logical equivalence of 'Fx' and '∃y(y = x & Fy)'. What about Roberts's second point: i.e. that the claim that Nora has zero well-being in w2 merely requires it to be false in w2 that Nora has the properties needed for positive or negative well-being? Well, it is false that the number 2 has the properties needed for positive or negative well-being, but it does not thereby have zero well-being; it is not worse off than I am.[9] Moreover, the proposal does not sit well with Roberts's concession that, in the money analogy, the amount $0 has to exist in w2. (Why would it have to, if 'Sadie has $0' could be true simply by virtue of the non-obtaining of Sadie's having any other amount?) In any case, whatever we say about this matter won't affect the fact that BNB violates actualism.

IV. HOLTUG'S APPROACH

Holtug says that if a person, Jeremy, is in fact well off, then Jeremy's existence is *intrinsically* better for him than his non-existence. This is not as radical as it may sound, however, because by 'Jeremy's existence' Holtug does not mean Jeremy's mere existence, but Jeremy's *life with all its contents*.[10] (By 'Jeremy's non-existence', on the other hand, he means the same as I do, namely, the state of affairs that Jeremy does not exist.) In contrast to Roberts, Holtug suggests that if Jeremy had *not* existed, his life would not be better for him than his non-existence. (Likewise – since 'x is better for S than y' implies 'y is worse for S than x' – Jeremy's non-existence is worse for him than his life, but would not be if he had not existed.) This idea apparently respects actualism, because it only claims that the *x is better for y than z* relation holds between the state of affairs of Jeremy's life, Jeremy, and the state

---

[9] This is probably because it does not have any capacity for mental states (nor does, of course, Nora in w2). Cf. Section VI below.

[10] To avoid confusion, I shall use the term 'Jeremy's life' instead of using 'Jeremy's existence' in Holtug's sense. I will come back to how Holtug's view is related to the question of whether Jeremy's mere existence is better for him than his non-existence.

of affairs of Jeremy's non-existence in worlds where all three relata exist.[11]

To support the thesis that Jeremy's life is better for him than his non-existence, Holtug considers an '"object" account of preferences, according to which it is the object of an intrinsic preference that has intrinsic value'.[12] Suppose Jeremy intrinsically prefers his life to never existing. Given the object account, Jeremy's life is then intrinsically better for him than his non-existence. Fortunately, however, Holtug does not want to rely on this account. I say 'fortunately', for suppose that Jeremy intrinsically prefers his *mere* existence to his non-existence (or vice versa). Then the account seems to imply that the former is intrinsically better than the latter (or vice versa) for Jeremy.[13] As I said in Section I, this is a dubious idea; moreover, it is an idea that Holtug says he does not endorse.

The other accounts Holtug considers are a 'satisfaction' account of preferences (according to which it is the state of affairs that a person's preference is satisfied, rather than the object of the preference, that has intrinsic value for him), a mental-state view (e.g. hedonism), and an objective list theory. According to Holtug, the thesis that a person's life can be intrinsically better for him than his non-existence is plausible on each of these. First of all, we are supposing that 'Jeremy's life contains a net surplus of positive values (preference-satisfactions, positive mental states or items on an objective list)'.[14] Further, Holtug contends, Jeremy's non-existence has no intrinsic value for him. And 'it seems to be better for a person to have a surplus of positive value than to have no value accrue to him'.[15] Hence, Jeremy's life is intrinsically better for him than his non-existence.[16]

Holtug's approach seems to me to have two main deficiencies. First, in order for Jeremy's non-existence to be intrinsically worse for Jeremy than something else, it must have an intrinsic value for Jeremy – at least zero. (If a person does not have any amount of dollars, including $0, we cannot compare 'it' with the amount of dollars I possess.) This

---

[11] Holtug says that he got this way of respecting actualism from Wlodek Rabinowicz (in personal communication): Holtug, 'On the Value of Coming into Existence', p. 374n. In fact, when I discussed my ideas with Rabinowicz in May 2007, it emerged that what he had had in mind was closer to my own view (introduced in Section V below) than to Holtug's.

[12] Holtug, 'On the Value of Coming into Existence', p. 364.

[13] Cf. Bykvist, 'The Benefits of Coming into Existence', p. 346.

[14] Holtug, 'On the Value of Coming into Existence', p. 366.

[15] Holtug, 'On the Value of Coming into Existence', p. 366.

[16] Someone might object here that though Jeremy's life *contains* a surplus of positive intrinsic value, this does not imply that Jeremy's life itself *has* a positive intrinsic value (which it must have in order to be intrinsically better – have a higher intrinsic value – for Jeremy than his non-existence). But I won't press this point.

is incompatible with Holtug's contention that it does not have any intrinsic value for Jeremy.[17] Even if the state of affairs *that no value accrues to Jeremy* is intrinsically worse for him than Jeremy's life (as Holtug suggests), this does not show that a state of affairs that has no value for him can be intrinsically worse for him than something else. Arguably, *that no value accrues to Jeremy* has zero intrinsic value for Jeremy.

However, Holtug also suggests that, if necessary, we may assign zero intrinsic value to Jeremy's non-existence after all. He points out that an existing person's life has zero intrinsic value for him if no positive or negative values accrue to him. 'Likewise, no positive or negative values befall a person who does not exist. For the same reason, then, we may assign zero [intrinsic] value to her non-existence.'[18] But this move sacrifices Holtug's main advantage over Roberts. For if something could have zero intrinsic value for a non-existent person, then surely w2, where Nora does not exist, has zero intrinsic value for Nora in w2. Thus, we would have to accept Roberts's BNB, the problems of which we have seen (Section III):

> BNB   Any alternative at which Nora never exists at all would have been better for Nora than Nora's anguished existence in fact is.

For something that has zero value for Nora has of course higher value for her – i.e. is better for her – than something with negative value for her. In any case, surely it would be preferable to be able to say that Jeremy's existence is better for him than his non-existence without having to claim that anything has zero value for a non-existent person. The account I shall put forward in the next section relies on no such claim.

Holtug's approach has a second drawback, which is independent of these actualist qualms. It is strange to focus on whether Jeremy's life, with all its contents, is intrinsically better for him than his non-existence. In other cases, this is not the kind of comparison we are interested in. Consider for example the issue of whether it is worse for Jeremy to die than not. One relevant question to ask here is: Is the state of affairs *that Jeremy dies* intrinsically worse for him than the state of affairs *that Jeremy does not die*? Arguably, the answer is No: these states of affairs seem to have no positive or negative *intrinsic* value for Jeremy. However, another question is whether *that Jeremy*

---

[17] Someone might think that Holtug uses the phrase 'no value' to mean zero value. On the contrary, however, he emphasizes that he does not: 'On the Value of Coming into Existence', p. 381.

[18] Holtug, 'On the Value of Coming into Existence', p. 381.

dies is *extrinsically* worse for him than *that Jeremy does not die*. Here, we should consider the values for Jeremy of the entire scenarios – with pleasures, preference-satisfactions, etc. – in which these states of affairs obtain. But one question that would be rather pointless to consider is whether the entire scenario where Jeremy dies is intrinsically worse for him than the bare state of affairs *that Jeremy does not die*. That would be an unfortunate mixture of the above considerations. The same point, I submit, applies to Holtug's account. He wisely resists the idea that Jeremy's mere existence is intrinsically better for him than his non-existence. He also wisely appeals to the fact that, in the actual world where Jeremy exists, there is a surplus of positive values for Jeremy. But he invokes this fact in the wrong way. The way to invoke it is to compare the intrinsic value for Jeremy of the actual world, not with the intrinsic value for him of the state of affairs of Jeremy's non-existence, but with the intrinsic value for Jeremy of the *entire* scenario that would have obtained if Jeremy had not existed. The intrinsic value of Jeremy's non-existence, zero or not, is the wrong focus.[19]

V. ANOTHER APPROACH

I have already hinted at the view I favor. Now I will make it more precise. I accept:

> EV1   A state of affairs A is extrinsically better (worse) for a person S than the state of affairs not-A iff the intrinsic value for S of the closest possible A-world is higher (lower) than the intrinsic value for S of the closest possible not-A-world.[20]

In contrast to Holtug's suggestion, EV1 does not tell us to calculate the extrinsic value for S of S's non-existence by focusing on *its* intrinsic value for S. Instead, EV1 dictates that we focus on the intrinsic value for S of whole possible worlds. What does it mean, then, that a possible world has a certain intrinsic value for S? This idea, I think, is not mysterious. To ask for the intrinsic value for S of a certain possible world is to ask: 'Consider this total way that things might have been. How valuable is it, in itself, for S?' For example, given a certain familiar

---

[19] Bykvist, too, seems to overestimate the importance of whether S's non-existence has zero value for S. He devotes several pages to this question: see 'The Benefits of Coming into Existence', pp. 342–7 (as well as elsewhere in the article). True, in contrast to the relevant passages in Holtug, Bykvist does not appear to be concerned exclusively with zero *intrinsic* value for S on those pages. But that hardly justifies the preoccupation with whether S's non-existence has zero value for S, for anyone who takes S's existence to be extrinsically better for S than S's non-existence is very likely to hold that S's non-existence has *negative* – not zero – extrinsic value for S.

[20] Cf. F. Feldman, 'Some Puzzles about the Evil of Death', *The Philosophical Review* 100 (1991), pp. 205–27, at p. 216.

hedonist approach, if S's total amount of pleasure minus pain in the world in question is +100, the answer is, '+100'. (Other approaches, of course, are more complex.)

Suppose that the actual world (where S exists) is intrinsically good for S. Given EV1, S's existence is then extrinsically better for S. For the closest world where S does not exist has an intrinsic value of zero for S. (I can hear you protesting here. The protest will be addressed shortly.) Of course, S's existence wouldn't be extrinsically better for S if S had not existed, for S would not then stand in the relevant value relations.

But how can a merely possible world – a possible world that does not obtain – have any intrinsic value for S (even zero)? You might think this violates actualism. However, I am assuming that each possible world is a state of affairs, which exists even if it does not obtain (see Section I). So actualism does not preclude any possible world, actual or non-actual, from standing in value relations to S.

But S, you may insist, may be precluded from entering into such relations. Surely, you might point out, if S had not existed, the world that would obtain then would not have any intrinsic value for S, not even zero. Indeed, wasn't that my main objection against Roberts's BNB? However, this charge rests on a misunderstanding. Unlike Roberts, I am not contending that that world, or anything else, *would* have any value for S if S had not existed. Rather, I am saying that that world *in fact has* a certain intrinsic value for S, namely zero. S does not stand in any relations in worlds where S does not exist; but S does (in fact) stand in relations *to* worlds where S does not exist. The difference between Roberts and me can be brought out if we consider two possible readings of EV1:

> EV2   A state of affairs A is extrinsically better (worse) for a person S than the state of affairs not-A iff the intrinsic value that the closest possible A-world *would have* for S if it were actual is higher (lower) than the intrinsic value that the closest possible not-A-world *would have* for S if it were actual.

Roberts's approach rests on something like EV2. That's why it violates actualism.[21]

> EV3   A state of affairs A is extrinsically better (worse) for a person S than the state of affairs not-A iff the intrinsic value that the closest possible A-world *in fact has* for S is higher (lower) than the intrinsic value that the closest possible not-A-world *in fact has* for S.

---

[21] Of course, EV2 itself does not violate actualism.

The present approach rests only on this second interpretation of EV1. That's why it does *not* violate actualism.

There seems to be no good reason to deny that possible worlds where S does not exist have an intrinsic value of zero for S (provided that S actually exists). In order to calculate the intrinsic value of a possible world for S – i.e. the intrinsic value that it in fact has for S – we should look at the basic intrinsic values for S of the states of affairs that obtain in that world (i.e. the states of affairs that would obtain if that world were actual).[22] It is hard to see why we should not say that a world has zero intrinsic value for S if all the states of affairs that obtain in that world have zero basic intrinsic value for S. Moreover, it is hard to see any reason to deny that every state of affairs that does not have positive or negative basic intrinsic value for S – e.g. *S's non-existence* or *Prince Charles's existence* – has zero basic intrinsic value for S (assuming, again, that S actually exists).

You might find it problematic to claim that something can have an intrinsic value (basic or not) for S without ascribing any property to S. It may be held that in order for a state of affairs to have any intrinsic value for S, the state of affairs has to entail, for some property, that S has that property. This requirement seems plausible as regards intrinsic goodness and badness for S. It is indeed difficult to see how a state of affairs that in itself has nothing in particular to do with S – a state of affairs that, taken all by itself, leaves S 'intact' – can be intrinsically good or bad for S. But this reason does not apply to zero intrinsic value for S: on the contrary, to leave S 'intact' seems to be a straightforward way of having zero intrinsic value for S. Furthermore, notice that I do not deny that states of affairs of the type *A having zero intrinsic value for S* ascribe properties to S. What I deny is that, in order for such a state of affairs to obtain, A has to ascribe a property to S. If you find the latter denial objectionable, it may be because you conflate it with the former denial, which I do not make. Moreover, if we once again assume hedonism, it seems natural to hold that, where S and I are numerically distinct, *S experiencing pleasure* is intrinsically better for S than *my experiencing pleasure*. Also, *S experiencing pleasure* seems to

---

[22] To say that something has basic intrinsic value is to say that it has intrinsic value in the most fundamental way. For instance, given a simple form of hedonism, states of affairs with positive basic intrinsic value have the form *S experiencing pleasure to degree n at time t*. For more on basic intrinsic value, see G. Harman, 'Toward a Theory of Intrinsic Value', *The Journal of Philosophy* 64 (1967), pp. 792–804, and F. Feldman, 'Basic Intrinsic Value', *Philosophical Studies* 99 (2000), pp. 319–46. They are concerned with intrinsic value *simpliciter* rather than intrinsic value for a person. But regardless of whether we want to determine a world's intrinsic value *simpliciter* or its intrinsic value for a person, we need to appeal to the basic, as opposed to non-basic, intrinsic values (either *simpliciter* or for a person) of the states of affairs that obtain in that world. Otherwise there will be double-counting and other problems.

be intrinsically better for S than it is for me. These judgments appear to presuppose that a state of affairs can have (zero) intrinsic value for a person without attributing any property to him. (Similar remarks apply to non-hedonist views.)

VI. ACCESS DENIED

It may be argued that I have not established that S's existence can be better for S than S's non-existence in a sense that shows that S's existence *benefits* S. And therefore, it may be held, the thesis is not very interesting. To be a (comparative) benefit to S seems to require more than merely being extrinsically better for S than something else. For instance, it sounds strange to call a state of affairs a benefit to S if it does not obtain. Of course, S's existence does, so that is not a problem. But it may be held that more is required. Krister Bykvist writes:

> 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 benefit constitutes a gain in value and a comparative harm a loss in value.[23]

This requirement, he points out, is captured by the following principle:[24]

Accessibility   If a state of affairs A is better (worse) for S than a state of affairs B, then A would be better (worse) for S than B even if A obtained.

If the actual world is intrinsically good for S, my account implies that S's non-existence is worse for S than S's existence. Given Accessibility, S's non-existence *would* then also be worse for S than S's existence if S had not existed. But this claim contradicts actualism, and I have tried to avoid it like the plague.

However, Accessibility is questionable. An initial problem emerges if we recall that my claim is that S's existence can be *extrinsically* better or worse for S than S's non-existence. (Also, let us focus on the comparison between A and not-A rather than between A and any state

---

[23] Bykvist, 'The Benefits of Coming into Existence', p. 348 (emphasis his). I have corrected one typo in this passage.

[24] Bykvist, 'The Benefits of Coming into Existence', p. 348. More exactly, he points out that the requirement is captured by Accessibility and the principle that A is better for S than B iff B is worse for S than A. (This principle seems obviously true.) But we can disregard this, for my approach violates Accessibility directly, since EV3 is a principle of both betterness and worseness. I have changed some of the notation in Accessibility to make it cohere with the one I use in this article.

of affairs B.) Therefore, the relevant version of the principle appears to be this:

Accessibility (E1)   If a state of affairs A is extrinsically better (worse) for S than the state of affairs not-A, then A would be extrinsically better (worse) for S than not-A even if A obtained.

Consider the following case. S's weight is 120 pounds. S would be more attractive, and lead a better life in general, if S's weight were at least 160 pounds. Let A be the state of affairs of *S weighing at least 160 pounds*. The intrinsic value for S of the closest possible A-world, w2, is higher than the intrinsic value for S of the closest possible not-A-world, w1 (viz. the actual world). So – given either of EV2 and EV3 (Section V) – A is extrinsically better for S than not-A. But it may very well be that if A had obtained – i.e. if S had weighed at least 160 pounds – then A would be extrinsically worse for S than not-A. For, arguably, in that world, w2, the following is true: If S *hadn't* weighed at least 160 pounds, then S's weight would be just slightly less – 159 pounds, say. And suppose that that weight would have made S even more attractive and lead an even better life. Then it is true in w2 that the not-A-world closest to w2, i.e. w3 (where S weighs 159 pounds), is intrinsically better for S than w2. Hence, if A had obtained, A would not be extrinsically better for S than not-A. In this case, therefore, the antecedent in Accessibility (E1) is true and the consequent false. So Accessibility does not work for extrinsic betterness, for reasons independent of the issue of existence and non-existence.[25]

But the version of Accessibility we should focus on might instead be this:

Accessibility (E2)   If a state of affairs A is extrinsically better (worse) for S than the state of affairs not-A, then even if A obtained, the intrinsic value for S of the world that is *in fact* the closest possible A-world would be higher (lower) than the intrinsic value for S of the world that is *in fact* the closest possible not-A-world.

---

[25] Bykvist has told me (in personal communication) that he intended Accessibility to cover only cases where 'symmetry' holds, so that, for example, 'w1 is the closest possible not-A-world and w2 is the closest possible A-world' holds in w1 iff it holds in w2. However, it is still worthwhile to see why the unrestricted version of the principle fails. Moreover, notice that it is far from clear that symmetry holds in a case of a person's existence and non-existence. Suppose S exists in the actual world w1 and that w2 is the closest world where S does not exist. It may well be that, relative to w2, w3 (where w1 ≠ w3) is the closest world where S exists. And it may well be that the intrinsic value for S of w3 is very different from the intrinsic value for S of w1.

The counterexample to the previous principle does not threaten this one. That example needed three worlds; Accessibility (E2) allows us only to consider two. Moreover, Accessibility (E2), too, contradicts my approach. This is easily seen, if we let A be S's non-existence.

Still, why accept Accessibility (E2)? I am inclined to deny it simply because of actualism. There is a straightforward explanation of why the consequent is false in the case of non-existence: S does not stand in any relations in worlds where S does not exist.[26] However, this may strike you as too facile. Recall Bykvist's argument above. How can S's existence be extrinsically better for S than S's non-existence in a sense that allows it to be a comparative benefit to S, although S would not have had a lower well-being level if S had not existed? But if this sounds strange, this seems to me to be because one conflates it with a related claim, which I do not make. What would be strange to say is that a state of affairs is a comparative extrinsic benefit to S even though S would have had at least as high a well-being level if it hadn't obtained. However, if S's existence hadn't obtained, it would not be the case that S has at least as high a well-being level as in the actual world.

The point might perhaps be strengthened by the following consideration. In order to have a well-being level in a world, I take it, one has to have, at least at some point of one's career in that world, a capacity for mental states. Consider some animal, e.g. a wasp, W, that lacks such a capacity in the actual world w1. This does not seem to be an essential feature of W: it seems that W *could* have had the capacity.[27] And if this much is granted, we should also grant that there is a world, w2, where W has a positive or negative well-being level. Now, suppose that W has a positive well-being level in w2. It seems that, in w2, it *benefits* W that it has the capacity for mental states. And not only in some purely non-comparative way: i.e. by causing something that is intrinsically good for W or preventing something that is intrinsically bad for W. Rather, in w2, having the capacity is a benefit to W because it is *better* for W to have it than to lack it; w2 is, in w2, intrinsically better for W than any world where W lacks it. Crucially, moreover, this cannot be because W, in w2, has a higher well-being level than W would have had if W had lacked the capacity. For W would not have had any well-being level at all (not even zero) without the capacity. Thus, an

---

[26] Cf. Holtug, 'On the Value of Coming into Existence', p. 375.
[27] You might object that if there is a world where W has mental states (which, apparently, there must be if W has the capacity in one world), this shows that W does after all have the capacity for mental states in all worlds where W exists. But this completely distorts the notion of a capacity. I do not have the capacity of flying, even if there is a possible world where I fly.

extrinsic comparative benefit need not make the subject have a higher well-being level than the subject would otherwise have had.[28]

VII. NORMATIVE ISSUES

It may seem attractive to hold that a scenario would be morally wrong to bring about only if it is worse for someone than some alternative scenario that the agent could have brought about instead. The axiological approach I suggested above allows an advocate of that normative view to hold that, in Roberts's case (Section II), Nora's mother acted morally wrongly in causing Nora to exist, even though the alternative would be that Nora had never existed at all. For the approach implies that the actual world, where Nora exists, is intrinsically worse for her than any world where she does not.

However, this is not the line I take. Notice that if Nora is *not* brought into existence, then the world where Nora exists is not worse for Nora than the actual world. (For if Nora does not exist, nothing is anything for her.) Here, then, a proponent of the view at hand has to deny the wrongness of bringing Nora into existence. In short, the wrongness of creating Nora depends on whether Nora is in fact created. But the idea that an action's normative status (e.g. its wrongness) depends in this way on whether it is performed is troublesome: for example, it has been argued that it makes deliberation impossible.[29]

Someone might hold that this consequence is avoided if our normative view appeals, not to what is better and worse for people, but to what is (non-comparatively) good and bad for people. For instance, you might suggest that Nora's mother acts wrongly in creating Nora, not because the world she brings about is worse for Nora, but because it is *bad* for her (and everything else is relevantly similar to the alternative world where Nora is not created). However, observe that the world where Nora is created *is* bad for her only if she is in fact created; as before, it is not anything for her if she never exists. So if we are to avoid the idea that an action's normative status depends on whether it is performed, there is little point in moving to what is non-comparatively good and bad for people.

Instead, the normative view might appeal to what has value *simpliciter*, and make no appeal to what has value *for* people. The

---

[28] Someone might object that, necessarily, all beings with psychological features have psychological identity conditions, so that nothing that has psychological features in w2 can be identical to W. But it would be a great cost for my opponents to have to rely on this principle, which is rejected by most philosophers.

[29] For a much more thorough discussion of this and related problems, see Bykvist, 'The Benefits of Coming into Existence', pp. 350–3, 'Prudence for Changing Selves', *Utilitas* 18 (2006), pp. 264–83, at pp. 273–4, and 'Violations of Normative Invariance'.

values *simpliciter* of the world where Nora exists and of the world
where she does not need not depend on whether she does in fact
exist. However, like the previous suggestion (i.e. the appeal to what is
good and bad for people), this 'impersonal' approach apparently makes
morally irrelevant the approach to betterness for people that I have
proposed. In order to be of worth, shouldn't that approach be used in
our theory of right and wrong?

This is not obvious. One could still use the approach to justify certain
emotions: for instance, many people are happy that they have come into
existence. Nevertheless, the approach is of course of much *more* use if
it can be used to determine what actions are morally right and wrong;
therefore, let me sketch one such normative outlook. It should be clear,
however, that it goes beyond (i.e. is not entailed by) my defense of the
thesis that a person's existence can be better or worse for her than her
non-existence. Moreover, other normative views than this one could be
based on that defense.

I suggest that Nora's mother acts wrongly in creating Nora because
the world where she creates Nora *would* be worse for Nora if it were
to obtain than the world where she does not create Nora. This modal
fact holds both if Nora is in fact created, and if she is not. Hence, the
wrongness of the action does not depend on whether it is performed.
Granted, if Nora is not created, then the world where she is created
is not worse for her; but it is still the case – that is, it is still the case
in the world where Nora is not created – that the world where she is
created *would* be worse for her if she were created. And that does seem
to be the relevant consideration. In assessing the normative status of
an action, we should focus both on what *would* be the case *if* it were
performed and on, for each alternative action, what would be the case
if *it* were performed. So a normative theory that appeals to betterness
and worseness for people should focus on to what extent things would
be better or worse for people in the worlds that the agent can bring
about: that is, for each possible world that the agent can bring about in
the situation, the view should focus on to what extent things *would* be
better or worse for people *if* that world were to obtain.[30]

---

[30] In 'Person-Affecting Moralities', *The Repugnant Conclusion*, ed. J. Ryberg and T.
Tännsjö (Dordrecht, 2004), pp. 136–7, it appears that Holtug intends the following 'Wide
Person-affecting Principle' to allow that an outcome, C, where everyone is badly off is
worse (*simpliciter*) than an outcome, D, where there are no individuals, regardless of
whether C or D obtains: 'An outcome, O_1, cannot be better (worse) than another outcome,
O_2, if there is no one for whom, were O_1 to obtain, O_1 would be better (worse) than O_2,
and no one for whom, were O_2 to obtain, O_2 would be worse (better) than O_1.' However,
if D obtains, there *is* no one – and hence no one for whom, were C to obtain, C would be
worse than D, and no one for whom, were D to obtain, D would be better than C. So, the
principle entails that, if D obtains, C cannot be worse than D.

To illustrate further, let us look at another example. Imagine that I have just two options: I can either create S, whose life will then be good for S, or create S*, whose life will then be good for S* to an even larger extent.[31] No one else than S and S* will be affected. It seems intuitively attractive to say that I ought to create S*, not S. Notice that if I create S, it is not the case that the alternative world, w2, where S* is created, is better for S* than the actual world, w1. But the proposed normative view is interested in more than that. It also cares about the fact that even if S* is not created, it is still true that w2 *would* be better for S* *if* it were to obtain. It should also care about the fact that w2 would be better for S* than w1, if w2 were to obtain, to a larger extent than w1 would be better for S than w2, if w1 were to obtain (which, incidentally, it does). For this reason, the view is able to explain why I ought to create S*, not S.

Some philosophers think that the fact that a person's life would be miserable constitutes a moral reason not to create him, whereas the fact that a person's life would be happy does not constitute a moral reason to create him.[32] The present normative approach suggests that the former claim is correct but the latter incorrect. Personally, I do not find this troubling, especially in light of the fact that the moral reason to create a happy person can be outweighed by many other factors. A proponent of the normative view suggested here is also able to say, if he likes, that the reason not to create a miserable person is stronger than the reason to create a happy person. For the view is compatible with the idea that comparative harms have more moral weight than comparative benefits. If we create an unhappy person, we harm him comparatively; but if we do not create a happy person, it is not the case that we harm him comparatively.

Someone might worry that there is a tension between the normative view adumbrated here and the approach to extrinsic betterness proposed earlier. My preferred principle of extrinsic betterness for S, EV3 (see Section V), appeals to the intrinsic values that the relevant worlds *actually have* for S; by contrast, the normative view just outlined appeals to the intrinsic value that each relevant world *would* have for S (and others) if it were to obtain. However, this difference should not be surprising. Whereas the account of extrinsic betterness concerns states of affairs, which exist even if they do not obtain, the normative approach concerns actions. Since actions are concrete objects, unperformed (i.e.

---

[31] It is particularly cases of this kind that are usually referred to as instances of the 'non-identity problem'. See Parfit, *Reasons and Persons*, part IV.

[32] Indeed, some philosophers are drawn to certain versions of 'person-regarding morality' precisely because they want to defend this asymmetry. Clearly, however, there are other reasons to find 'person-regarding morality' attractive: see e.g. Holtug, 'Person-Affecting Moralities', pp. 137–8.

never performed) actions do not exist.[33] Thus, when we wonder what we should do, we wonder what *would* be right if we were to do it. Likewise, when we talk about (e.g.) the wrongness of an unperformed action, we talk about the normative feature that it would have if it were performed.[34] Therefore, it is not strange that the answer to what we should do depends crucially on what outcomes *would* be better and worse for people if they were realized.

There are different ways to develop the general normative approach suggested here. (I do not feel obliged to choose among these ways in this context.) Naturally, any version of it might deliver the same verdicts in all specific cases as do some 'impersonal' views; but, arguably, *any* normative view that appeals to what is better and worse for people will do that. There is still a distinction between views whose fundamental explanations of moral rightness and wrongness refer to what is better and worse for people and views that do not.[35]

jens.johansson@philosophy.su.se

---

[33] Krister Bykvist has objected (in personal communication) that in order for me to deliberate between different alternative actions, these alternative actions have to exist. Therefore, he suggests, it is plausible to regard them as states of affairs. But it seems enough for deliberation that there are states of affairs such as *my performing action a1* and *my performing action a2*. Actions themselves – e.g. a1 or a2 – are still concrete (and exist only if performed).

[34] This does not violate the relevant version of the idea that e.g. an action's wrongness should not depend on whether it is performed. What should not depend on whether the action is performed is the truth value of the statement that the action would be wrong if it were performed. The normative view I discussed in the opening two paragraphs of the present section violates this principle. So does the view in the third paragraph.

[35] Many thanks to several anonymous referees, Matthew Liao, Rebecca Roache, Julian Savulescu, Nicholas Shackel, and, especially, to Krister Bykvist and Wlodek Rabinowicz.