

Gustaf Arrhenius and Wlodek Rabinowicz The value of existence

Book section

Original citation:

Originally published in Arrhenius, Gustaf and Rabinowicz, Wlodek (2015) *The value of existence*. In: Hirose, Iwao and Olson, Jonas, (eds.) *The Oxford Handbook of Value Theory*. Oxford handbooks. Oxford University Press, Oxford, UK, pp. 424-444. ISBN 9780199959303

© 2015 [Oxford University Press](http://www.oxforduniversitypress.com)

This version available at: <http://eprints.lse.ac.uk/66820/>

Available in LSE Research Online: June 2016

LSE has developed LSE Research Online so that users may access research output of the School. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LSE Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain. You may freely distribute the URL (<http://eprints.lse.ac.uk>) of the LSE Research Online website.

This document is the author's submitted version of the book section. There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

The Value of Existence

(Published in I. Hirose and J. Olson *Oxford Handbook of Value Theory*, Oxford UP 2015,)

Gustaf Arrhenius & Wlodek Rabinowicz

Department of Philosophy, Stockholm University

Department of Philosophy, Lund University

Gustaf.Arrhenius@philosophy.su.se

Wlodek.Rabinowicz@fil.lu.se

I. Introduction

Can it be better or worse for a person to be than not to be, that is, can it be better or worse for her to exist than not to exist at all? This old and challenging philosophical question, which we can call *the existential question*, has been raised anew in contemporary moral philosophy. There are roughly two reasons for this renewed interest. First, traditional “impersonal” ethical theories, such as utilitarianism, have paradoxical and very counterintuitive implications for procreation and our moral duties to future, not yet existing people. Second, it has seemed evident to many that an outcome can be better than another only if it is better for someone, and that only moral theories that are in this sense “person affecting” can be correct. The implications of this so-called *Person Affecting Restriction* will differ radically, however, depending on which answer one gives to the existential question. Hence, many of the problems regarding our moral duties to future generations turn around the issue of whether existence can be better or worse for a person than non-existence.

Some think so, others adamantly deny it. Thus, for example, Nils Holtug (1996, 2001), Melinda Roberts (1998, 2003) and Matthew Adler (2009) have defended an affirmative answer to the existential question. Contrariwise, Derek Parfit (1984), John Broome (1999), Krister Bykvist (2007) and others have worried that if we take a person’s life to be better for her than non-existence, then we would have to conclude that it would have been worse for her if she did not exist, which is absurd: Nothing would have been worse or better for a person if she had not existed.

We shall start by explaining in more detail why the existential question has in recent years moved to the forefront of moral philosophy. We shall then discuss some of the proposed

answers in the literature and our own suggestion. On our view, one can plausibly claim that it is better or worse for a person to exist than not to exist without thereby implying any absurdities. Finally, we shall consider and rebut some objections to our position.

II. The Person Affecting Restriction and the Existential Question

The Person Affecting Restriction, put as a slogan, states that one outcome can be better than another only if it is better for someone. The restriction has a strong intuitive appeal and it has been suggested that it is presupposed in many arguments in moral philosophy, political theory, and welfare economics.¹ Moreover, several theorists have argued that the counterintuitive implications in population ethics of “impersonal” welfarist theories arise because such theories violate this restriction. This applies in particular to the well-known Repugnant Conclusion, which – as has been pointed out by Parfit – is entailed by classical utilitarianism.²

It is not easy to discern what exactly the distinction between “impersonal” and “person affecting” theories amounts to in the literature, partly because different authors have had a different take on the distinction and partly because other ideas have been conflated or combined with the Person Affecting Restriction. One can interpret the restriction in a weak way that makes it perfectly compatible with impersonal welfarist theories such as classical utilitarianism.³ For example, it could be understood as the idea that moral goodness exclusively supervenes on facts concerning individual wellbeing. We are, however, interested in a stronger reading of the restriction which stresses the individualist aspect of value even more by claiming that axiology is essentially *person comparative*:

The Person Affecting Restriction: If an outcome A is better (worse) than B, then A is better (worse) than B for at least one individual in A or in B.

¹ See Temkin (1993a, b). The label “Person Affecting Restriction” was introduced by Glover (1977), p. 66, but see also Narveson (1967).

² See Parfit (1984), p. 388. For an overview of these counterintuitive implications, see Arrhenius et al. (2010) and Arrhenius (2000, 2014). The Repugnant Conclusion is the claim that for any world inhabited by people with very high welfare, there is a *better* possible world in which everyone has a life that is barely worth living, other things being equal. Imposing the Person Affecting Restriction might block the derivation of the Repugnant Conclusion if it is conjoined with a negative answer to the existential question. Then it is arguable that a world in which everyone has a life barely worth living cannot be better than a world in which all individuals have a very high quality of life, since the former is not better for anyone, not even for the people who exist in the former but not in the latter world. Since we are going to argue that the existential question should be answered in the affirmative, however, we are sceptical about this manoeuvre.

³ See Arrhenius (2003, 2009, 2014).

In cases involving only the same people in the compared outcomes, this restriction is, we surmise, widely accepted by theorists with welfarist inclinations.⁴ In comparisons between outcomes involving different people, for example in cases involving people whose existence is contingent on our choices, the restriction however becomes ambiguous. An outcome A is better than B for Peter if Peter has a higher welfare in A as compared to B. We can assume that much. But what if Peter exists in outcome A but not in outcome B? Is A then better than B *for Peter*? More generally, can existence be better or worse for a person than non-existence?⁵ Depending on the answer to the existential question, the Person Affecting Restriction has very different implications regarding how to morally evaluate different possible futures.

III. Neither Better nor Worse to Be than Not to Be

A popular answer is to claim that existence cannot be better or worse for a person than non-existence, nor equally as good for that matter, since existence and non-existence are, in some sense, incomparable in value for a person. Thus, David Heyd argues that the view according to which existence could be worse than non-existence “is inconsistent with a person-affecting theory as it presupposes the comparability of non-existence with life of a certain quality”.⁶

In his early pioneering work in population ethics, Narveson seems to share Heyd’s concern regarding comparability, although he formulates it in terms of happiness comparisons rather than comparisons in value:

If you ask, “whose happiness has been increased as a result of his being born?”, the answer is that nobody’s has. - - - Remember that the question we must ask about *him* is not whether he is happy but whether he is happier as a result of being born. And if put this way, we see that again we have a piece of nonsense on our hands if

⁴ Three qualifications: (i) The label “Person Affecting” might be misleading, since many theorists would, sensibly we think, weaken the restriction so as to also cover other sentient beings. Cf. Holtug (1996). (ii) Since the Person Affecting Restriction is formulated without a *ceteris paribus* clause, value pluralists are not likely to accept it since it leaves little room for other values apart from welfarist ones. Clearly, one might embrace non-welfarist values such as virtue, reward in accordance to desert (cf., for example, Feldman 1995a, b, 1997), beauty (cf. Moore 1903, section 50), variety of natural species, or what have you (for a general discussion of value pluralism, see Rabinowicz & Rønnow-Rasmussen 2004; for a discussion of this issue in connection with the Person Affecting Restriction, see Arrhenius 2003, 2009, 2014). We shall, however, only discuss implications of the restriction in cases where one can assume that non-welfarist values are not at stake. Hence, the arguments below also apply to the *ceteris paribus* version of the restriction, that is, the version that is of interest also to the value pluralists. (iii) Certain welfarist theories are ruled out by the restriction already in the same people cases, such as some extreme versions of welfarist egalitarianism (Arrhenius 2009, 2014).

⁵ To forestall possible misunderstandings, the question we are interested in is not whether *mere* existence, or life *as such*, can be better or worse for a person than non-existence. Instead, the question is whether having a life of a certain quality can be better or worse for a person than not having a life at all.

⁶ Heyd (1988), p. 161. See also Heyd (1992), pp. 124-5. Heyd states that his view is “grounded in an ‘anthropocentric’ conception of value according to which value is necessarily related to human interests, welfare, expectations, desires and wishes – that is to say to human volitions” (1988, p. 164).

we suppose the answer is either “yes” or “no”. For if it is, then with whom, or with what, are we comparing his new state of bliss? Is the child, perhaps, happier than he used to be before he was born? Or happier, perhaps, than his alter ego? Obviously, there can be no sensible answer here.⁷

Similarly, John Broome states that “...it cannot ever be *true* that it is better for a person that she lives than that she should never have lived at all”.⁸ Cf. also Dasgupta (1995), p. 383:

Recall our definition of the zero level of well-being. This isn't a standard arrived at through a comparison with ‘non-existence’. Such comparisons can't be made. The ‘unborn’ aren't a class of people. It makes no sense to attribute a degree of wellbeing, low or high or nil, to the ‘state of not being born’.

Likewise, Alan Buchanan et al. (2000), p. 234, claim that “when the alternative is nonexistence, there is no individual who is made worse off by being conceived and born. Nonexistence is not a condition that is better for an individual only in rare cases like having Lesh-Nyhan or Tay Sachs disease; it is no condition at all, and so it is not better or worse than any other condition.”

The negative answer to the existential question in combination with the Person Affecting Restriction has such counterintuitive conclusions that it is hard to believe that anyone would seriously endorse the conjunction of these two views. Consider the Future Bliss or Hell Case:

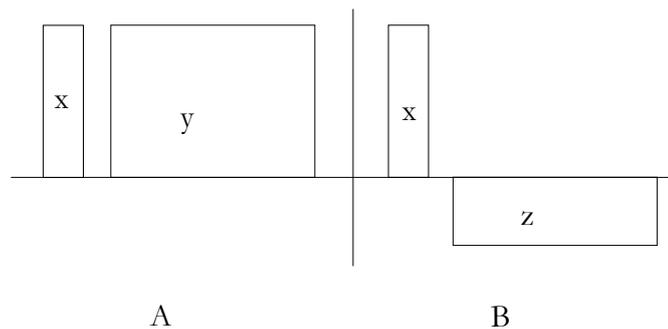


Diagram 1

⁷ Narveson (1967), p. 67 (emphasis in original).

⁸ Broome (1999), ch. 10, p. 168 (emphasis in original). See also Parfit (1984), pp. 395, 489. However, these three authors, unlike Heyd and Narveson, do not endorse the restriction.

The blocks in the above diagram represent populations. The width of each block represents the number of people, whereas the height represents their welfare. Assume that we can either see to it that all the people in the future have excellent lives (the y-people in outcome A) or that they have hellish lives (the z-people in outcome B). Assume further that these two possible future populations are of the same size but consist of different people, and that these two outcomes are equally good for us, the currently existing x-people.

We take it that outcome A is clearly superior to outcome B. However, since the y- and z-people don't exist in both outcomes, the negative answer to the existential question implies that outcome A is neither better nor worse for the y- and z-people as compared to B. Moreover, the two outcomes are equally good for the x-people. Hence, according to the Person Affecting Restriction, A cannot be better than B since it is not better for any individual. Nor is of course B better than A. In other words, if combined with the negative answer to the existential question, the Person Affecting Restriction ranks these outcomes as either equally good or as incomparable in value. But that is clearly the wrong diagnosis of the Future Bliss or Hell Case.

This and other counterintuitive implications of the Person Affecting Restriction in combination with the negative answer to the existential question have led philosophers to abandon the restriction (the majority) or to accept not only that existence can be better or worse for a person than non-existence but also that (i) a non-existent person has a certain welfare level (namely, zero welfare, *pace* Dasgupta) and that, consequently, (ii) non-existence can be better or worse for that non-existent person than a life at some specified level of welfare. As we shall show, we neither need to give up the restriction nor make moves (i) and (ii).

IV. The Argument from Absurdity

One worry that seems to motivate the negative answer to the existential question is the following: If the question were given an affirmative answer, that is, if we took a person's life to be better or worse for her than non-existence, then we would have to conclude that *it would have been worse or better for her if she did not exist* (henceforth, the *Absurd Conclusion*). Clearly, this is unacceptable: Nothing would have been worse or better for a person if she had not existed. Parfit puts this worry as follows: "in being caused to exist, someone can be benefited. ... [W]e need not claim that this outcome is *better* for this person than the alternative. This would imply the implausible claim that, if this person had never existed, this would have been worse for this person."⁹ The absurdity of the conclusion is well brought out by Broome:

⁹ Parfit (1984), p. 395, emphasis in the original; cf. also p. 489. He asserts, however, that causing someone to exist still can be *good* for the person in question. Good, but not better. It seems to us that Parfit needn't have been so cautious. His discussion of the matter contains all that is needed for the bolder betterness claim (see below).

...[I]t cannot ever be *true* that it is better for a person that she lives than that she should never have lived at all. If it were better for a person that she lives than that she should never have lived at all, then if she had never lived at all, that would have been worse for her than if she had lived. But if she had never lived at all, there would have been no her for it to be worse for, so it could not have been worse for her.¹⁰

However, the Absurd Conclusion does not follow. A triadic relation consisting in one state (having a certain life) being better for a person p than another state (non-existence) cannot hold unless its three relata exist. Now, the states in question are abstract entities and thus can be assumed to exist even if they do not actually obtain. Consequently, the triadic relation in question can indeed hold as long as also the third relatum, person p , exists. However, if persons are concrete objects, which is the received view (and, we surmise, the correct one), a person exists only insofar as she is alive. Therefore, this relation could not hold if p weren't alive, since the third relatum, p , would then be missing.^{11, 12} Consequently, even if it is better for p to exist than not to exist, assuming she has a life worth living, it doesn't follow that it *would have been worse* for p if she did not exist, since one of the relata, p , would then have been absent.¹³ What

¹⁰ Broome (1999), ch. 10, p. 168 (emphasis in original). Note that this argument, if correct, would also work against the idea that existence could be worse for someone than non-existence: If it were worse for a person that she exists than that she should never have existed, then it would have been better for her if she had never existed. But if she had never existed, then there would have been no her, so it could not have been better for her. Thus, it cannot be true that it could be worse for a person to exist than not to exist.

¹¹ On the other hand, if persons were viewed as abstract objects – as collections of properties which can exist without being instantiated, then there would be nothing absurd in claiming that if a person were not instantiated, this state could have been worse for her than her actual state, since all three relata would then exist as abstract entities. This interpretation of persons as abstract objects is suggested in a forthcoming paper by Fleurbaey and Voorhoeve (2014). (Though the authors do not clearly distinguish their proposal from a view according to which persons can exist even in worlds in which they are merely possible persons; cf. below, fn. 13.) In our view, the interpretation of persons as abstracta is not a tenable position. Abstract objects cannot have the kind of properties we ascribe to persons: they have no temporal existence, they have no bodies, they cannot be happy or sad, nor can they make any decisions.

¹² It should be noted that in this paper we simply presuppose that relations require the existence of relata. This presupposition could, of course, be questioned. (We are indebted to Staffan Angere for pressing this point.) Thus, for example, it could be argued that intentional attitudes are best interpreted as relations between a subject and an intentional object, which would mean that a relation can obtain after all even if one of the relata (the intentional object) happens not to exist. On our view, such a relational account of intentional attitudes is *prima facie* attractive, but in the end unsatisfactory. An ontology that allows obtaining of the relations without existing relata leads to a whole host of problems, but the discussion of these issues would take us too far afield.

¹³ Another approach to this issue has been suggested by Tim Williamson (2013). On his “necessitist” view, everything that exists does so necessarily, in all possible worlds. Consequently, a person p exists even in a possible world in which she never has lived at all and will never live in the future. She still does exist, as a merely possible person. As such, she is according to Williamson neither a concrete nor an abstract object. This would mean that even if p is a merely possible person, all three relata of the “better for”-relation exist. However, one might still question whether anything could be better or worse for a merely possible person and whether such an entity could be assigned any level of wellbeing. Williamson himself leaves this possibility open pending further philosophical investigation (cf. *ibid.* p. 29). We are sceptical on this point. Even if merely possible persons do exist (which is a huge “if”, of course), they can hardly be assigned more than a merely possible level of wellbeing (or rather a whole range of such possible levels). And it can hardly be claimed that being merely possible is worse (or better) for that

does follow is only that non-existence *is* worse for her than existence (since ‘worse’ is just the converse of ‘better’), but not that it *would have been* worse if she didn’t exist. Hence, Broome’s argument is a non-sequitur and the Absurd Conclusion doesn’t follow from the idea that existence can be better or worse for a person than non-existence.¹⁴

It might be that Broome assumes that the following general principle is true:

Subjunctive Connection 1 (SC1): An outcome A is better (worse, equally as good) for *p* than (as) another outcome B only if outcome B would be worse (better, equally as good) for *p* than (as) A if B came about.¹⁵

Krister Bykvist has suggested a similar principle which he calls *Accessibility*: “If A is better (worse) for S than B, then A would be better (worse) for S than B even if A obtained”.¹⁶ However, as we pointed out above, it doesn’t follow logically from “it is better for *p* to exist than not to exist” that “it would have been worse for *p* if she did not exist” since in the latter case one of the relata, *p*, would be absent.¹⁷ So it is not clear to us why one should go for SC1 rather than for the following connection between “better for” and “would be worse for”:

merely possible person than having a life. A philosopher who actually was prepared to make this seemingly preposterous claim was Richard Hare, but his proposal, when it was made, fell for dead ears (cf. Hare (1975), (1988a), (1988b), (1998)).

Yet another approach, which we disregard in this paper, is David Lewis’s ‘modal realism’, according to which other possible worlds are as real as the actual world. The objects in other possible worlds are not mere possibilities, as on Williamson’s view; they are as concrete as the objects in the actual world. On modal realism, there seems to be no problem in postulating the obtaining of ‘trans-world’ relations: a relation can obtain despite the fact that its relata exist in different worlds. (It is somewhat unclear though *where* such a relation is supposed to obtain; it does not seem to be located *in* a world.) From the point of view of modal realism, the argument from absurdity does not get off the ground. If it is better for *p* to have a life of a certain quality than not to exist, then this would still be better for *p* even if *p* did not actually exist: Even then *p* would have a real existence, as a resident of some possible world, and therefore the betterness relation in question would still obtain. For a hilarious investigation of the implications of modal realism for transworld relationships see Sinhababu (2008).

¹⁴ Rabinowicz suggested this argument already back in 2000 in personal conversation with Arrhenius, Broome, Bykvist, and Erik Carlson at a workshop in Leipzig; and he has briefly presented it in Rabinowicz (2003), fn. 29, and in more detail in Rabinowicz (2009a), fn. 2. For a similar argument, see Arrhenius (1999), p. 158, who suggests that an affirmative answer to the existential question “only involves a claim that if a person exists, then she can compare the value of her life to her non-existence. A person that will never exist cannot, of course, compare ‘her’ non-existence with her existence. Consequently, one can claim that it is better ... for a person to exist ... than ... not to exist without implying any absurdities.” Cf. also Holtug (2001), p. 374f. In fact, even though he accepted the negative answer to the existential question (and instead went for the view that it can be good but not better for a person to exist than not to exist), Parfit (1984) came very close to making the same point as we are making when he observed that there is nothing problematic in the claim that one can benefit a person by causing her to exist: “In judging that some person’s life is worth living, or better than nothing, we need *not* be implying that it would have been worse for this person if he had never existed. --- Since this person *does* exist, we can refer to this person when describing the alternative [i.e. the world in which she wouldn’t have existed]. We know who it is who, in this possible alternative, would never have existed” (pp. 487-8, emphasis in original; cf. fn. 9 above). See also Holtug (2001), Bykvist (2007) and Johansson (2010).

¹⁵ As we have seen, this principle is assumed by Parfit (1984), pp. 395, 489.

¹⁶ Bykvist (2007), p. 348. Symbols have been changed in this quotation, for the sake of consistency.

¹⁷ Moreover, it seems clear that SC1 is false, *mutatis mutandis*, for related concepts, such as “considered better by” / “would be considered worse by” and “preferred by” / “would be dispreferred by”. This is acknowledged by Bykvist (2007), p. 349. Bykvist (2014) considers the possibility that subjunctive connection obtains specifically for

Subjunctive Connection 2 (SC2):

- (i) If a person p exists in both outcomes A and B, then A is better (worse, equally as good) for p than (as) B only if B would be worse (better, equally as good) for p than (as) A, if B obtained.
- (ii) If a person p exists in A but not in B, then A can be better (worse, equally as good) for p than (as) B although B would not be worse (better, equally as good) for p than (as) A, if B obtained.

Of course, one might find SC1 more attractive than SC2, perhaps because one finds it more in line with our common way of thinking and reasoning: If we consider one outcome as better for someone than another outcome, then we usually are prepared to conclude that the other outcome, if it obtained, would be worse for that person (and not just that it *is* worse). This, however, might simply have to do with the fact that we are accustomed to compare outcomes in both of which the affected person exists. If this habitual presupposition is given up, the old ways of thinking have to be adjusted accordingly. Note that we do not simply reject SC1 in order to avoid the Absurd Conclusion; we explain what's wrong with this condition in the context where people's existence is at stake.

V. The Argument from Welfare Level Comparisons

To save the Person Affecting Restriction from cases like the Future Bliss or Hell case, Melinda Roberts has suggested that we should accept not only that existence can be better or worse for a person than non-existence, but also the apparently absurd conclusion that in cases like this it would have been better or worse for a person not to exist than to exist. The reason is that according to Roberts a non-existing person has a certain welfare level, namely, zero welfare:¹⁸

...Nora does not have any properties at all at any alternative at which she does not exist and ..., where Nora has no properties at all, all the properties that she does have — that empty set — add up to a zero level of wellbeing. --- It would have been better for Nora not to have any wellbeing at all — to have zero wellbeing —

values even though it doesn't hold for preferences and the like. In particular he considers what he calls the principle of Axiological Invariance: "If [A] has absolute or comparative value for a person, it would have this value, no matter whether [A] were to obtain or not." He points out, however, that this principle is untenable: A person's existence might be good (or bad) for that person but it couldn't be good (bad) for her if she did not exist.

¹⁸ Adler (2009), p. 1506, tentatively embraces a similar position: "Existence can be better or worse for an individual than nonexistence. Nonexistence can be better or worse for an individual than existence. Where an outcome set contains potential nonexistents, their interests should be taken into account by assigning them a utility level of zero in the outcomes where they do not exist".

than to have the negative level of wellbeing that she in fact has. It would have been [better] for Nora ... never to have existed at all than it is for Nora to exist.¹⁹

However, in our view it is quite nonsensical to ascribe any wellbeing level at all to a person in a state in which she does not exist. Wellbeing presupposes being. Having a zero degree of wellbeing is arguably the kind of property the instantiation of which requires the existence of property bearers.²⁰ Indeed, Roberts' view seems contradictory: While she ascribes to the non-existent Nora the property of having zero degree of wellbeing, she states that the non-existent Nora "has no properties at all".²¹

As we have shown above, one can endorse an affirmative answer to the existential question without affirming that non-existence could have been better or worse for a person and without assigning any welfare levels to persons who don't exist.

Yet, one might insist that the suggestion we make still doesn't make sense: that we cannot make sense of one state, A, being better for p than another state, B, if we cannot compare the wellbeing levels of p in the two states in question. This might be what Heyd and Buchanan et al. had in mind in the quotes above. Likewise, when Bykvist claims that SC1 (his Accessibility principle) is true about any "interpretation of 'better for' that is conceptually linked to well-being...", it seems that his idea is that "better for"-claims are analyzable in terms of comparisons between well-being levels possessed by a given individual in different outcomes.²² This would entail SC1 given that no individual has any level of wellbeing in an outcome in which she does not exist.²³ The idea is that there is a necessary connection between "better for" and "has a higher welfare than":

Welfare Level Connection (WLC): An outcome A is better for a person p than another outcome B if and only if p has higher welfare in A than in B.

¹⁹ Roberts (2003), pp. 168-9. Moreover, Roberts (1998), p. 64, writes: "I am thus supposing that it is at least possible that s has more well-being in a world in which s does not exist than s actually has. Suppose s 's existence in X is unavoidably *less* than one worth living ... and that s has, in any world in which s does not exist, a zero level of well-being. Under these conditions, s 's level of well-being at zero is actually *greater* than s 's well-being in X..." (emphasis in original). See also Roberts (this collection).

²⁰ Of course, it isn't evident that all properties are of this kind, for example, properties of fictional objects or tropes (that is, concrete property-like particulars whose existence is supposed to be ontologically independent of the existence of property bearers).

²¹ We are indebted to Jonas Olson for this last point.

²² See Bykvist (2007), p. 348. Adler (2009), p. 1503, considers a similar conceptual connection, which he presents as the connection between "worse for" and "worse off than". However, unlike Bykvist, Adler's discussion leads him to reject, at least tentatively, this supposed conceptual link (cf. *ibid.*, p. 1505).

²³ This seems to hold even if we were to construct persons as abstract objects that can obtain or not obtain. A specific welfare level is something an abstract person can possess only in a world in which she obtains. Unless, of course, we bring in that level in the abstract specification of the person in question. But then this need not be the zero level; it can be any level.

However, as we shall argue in the following section, “better for” comparisons can be made without comparisons of welfare levels. Consequently, one should reject the suggested tight connection between “better for” and comparisons of welfare levels as expressed by WLC.

VI. Guardian Angels and Fitting Attitudes

Instead of relying on WLC, one might explicate “better for” in terms of what a benevolent impartial observer would prefer for a person when she is only considering what is in the interests of the person in question, or – better – in terms of what that person’s guardian angel would prefer.²⁴ According to this view, an outcome A is better for a person than another outcome B if and only if this is what her guardian angel would prefer for her sake. If a person exists in the two compared outcomes, then trivially the guardian angel will prefer the state in which her charge has the highest welfare level. However, if the guardian angel compares a state in which her charge has a life with negative welfare with a state in which that person does not exist at all, she prefers the latter. Moreover, if the guardian angel compares a state in which her charge has a life with a positive welfare with the state in which her charge does not exist, she prefers the former.

We can think of this idea of a guardian angel as just a *criterion* for the “better for”-relation. On this criterial interpretation, we can try to find out what is better for a person by putting ourselves, in imagination, in her guardian angel’s shoes and then try to determine what our preferences would be in that hypothetical position. Additionally, on a view that is philosophically more far-reaching, the idea of a guardian angel can also be seen as a metaphor for a certain *analytical* proposal. More precisely, on this reading, we should take it as an application to “better for” of the so-called *fitting-attitudes analysis of value*. Along the lines of this format of analysis, we could say that

²⁴ This passage on guardian angels draws on Rabinowicz & Arrhenius (2010) but has now been re-formulated. The original formulation, which was in terms of the pairwise choices a guardian angel would make, invited an unfortunate misinterpretation of our view, as we have become aware after reading a forthcoming paper by Fleurbaey and Voorhoeve (2014).

Rabinowicz suggested the guardian angel approach in 2000 (see fn. 14) and Arrhenius (2003) proposes the benevolent impartial observer approach. See also Bykvist (2007). Broome (2004), p. 63, credits Rabinowicz with a suggestion that is simpler but less plausible: A history (or a world) X is better for *p* than a history Y if and only if *p* prefers X to Y. As Broome points out: “A person may prefer one history to another even if she does not exist in both of them” (ibid.). Obviously, however, this simple proposal is not satisfactory as it stands. The advantage of appealing to the preferences of the guardian angel rather than to those of the benevolent impartial observer is that the latter are supposed to track what is impersonally good (good, period) rather than what’s good for the person under consideration. A benevolent impartial observer tracks impersonal goodness even when she only focuses on the interests of that person and of no one else. There is a conceptual distinction between how good a prospect that only concerns the interests of a given person is for the person in question and how good that prospect is from an impersonal perspective. And, at least for prioritarians, that distinction makes a difference: It affects the ordering of prospects. (Cf. Rabinowicz 2001, 2002.) The task of a guardian angel is different in that respect: His responsibility is to track goodness for the person in his care.

A is better for p than B if and only if one ought to prefer A to B for p 's sake.²⁵

This analytic proposal could be made to work provided we can make some sense of locutions such as “preferring A to B for p 's sake”.²⁶ Again, it seems reasonable to say that in the choice between bringing p into existence with negative welfare or not bringing her into existence at all, one ought to prefer the latter for p 's sake. Likewise, in the choice between bringing p into existence with positive welfare or not bringing her into existence at all, one ought to prefer the former for p 's sake.

On both these interpretations, the criterial and the analytic one, if a person p has higher welfare in an outcome A as compared to another outcome B, then A is better for p than B, but the reverse doesn't always hold. Hence, there is a connection between “better for” and “has higher welfare than” but this connection isn't as tight as WLC would have it.

VII. The Issue of Value Bearers

The affirmative answer to the existential question faces a serious problem, however. We have assumed that the “better for”-relation obtains between abstract states and a person. One might object, however, that what is better for a person is not an abstract state itself but the *obtaining* of that state.²⁷ To assume that states and not just their obtainings are better for a person seems counterintuitive, since the existence of such abstract states doesn't make things better for a

²⁵ Cf. Darwall (2002) for this proposal. As Darwall puts it: “[W]hat it is for something to be good for someone *just is* for it to be something one should desire for him for his sake, that is, insofar as one cares for him” (p. 8). See also Toni Rønnow-Rasmussen (2007, 2011), where this fitting-attitudes account of value-for is elaborated and defended. That this account can be used to clarify comparisons between existence and non-existence has been suggested in Rabinowicz (2009a), fn. 2.

²⁶ The challenge here is whether the “for p 's sake”-locution can be independently understood, without presupposing the notion of “better for” as already given. If preferring something for p 's sake just means “preferring it insofar as one only cares for what is better for p ”, then the analysis becomes circular. On the other hand, if “for p 's sake” is given an independent interpretation, then it is not obvious that all that one ought to prefer for p 's sake *is* better for p . In particular, perfectionist considerations complicate matters at this point. Thus, for example, it might be argued that one is a better person if one experiences sorrow at the thought of others' suffering and that one ought to desire this for p 's own sake. Yet, it is not obvious that such a state of mind would be good for p . (We are indebted to Johan Brännmark for pressing this point.) Maybe, therefore, a circular analysis of “sake” would, after all, be preferable. It should be noted that even circular analyses can be instructive to some extent: They can be used to exhibit structural connections between concepts appearing in the analysans and the analysandum. Thereby, they can provide relevant information to those who already possess the concepts involved but are not clear about their mutual relationships. Thus, to take an example, David Wiggins adheres to the sentimentalist version of the fitting-attitudes account even though he explicitly recognizes the charge of circularity. Still, as he argues, the account is informative in its “detour through sentiments” (See Wiggins (1987), p. 189. Cf. Rabinowicz & Rønnow-Rasmussen (2004, 2006). The circularity Wiggins has in mind is different from the one mentioned here, though. He thinks that it might be essential to the fitting sentiments with regard to objects that these attitudes themselves already involve evaluations.

²⁷ We are grateful to Erik Carlson for pressing this point. A somewhat similar point has subsequently been made by Bykvist (2014). Following Sven Danielsson, Bykvist writes: “to say that a state of affairs [A] is good for you is to say that you are such that it is possible that the world ... would be good for you (at least to some extent), if it exemplified [A]. --- The slogan is: states of affairs that are good for you could rate the universe a plus for you [if they were exemplified by the universe, that is, if they obtained].”

person in any way. It is only the obtaining of such states that can make things better for a person. Now, it is clear that we cannot claim that the obtaining of the state of a person p 's existence is better for p than the obtaining of p 's non-existence since these two obtainings and the person under consideration do not co-exist. Indeed, they cannot co-exist and hence they cannot stand in any relation to each other. As we have assumed throughout, a relation cannot obtain without the existence of all the relata.

Actually, this is a general problem in axiology. It concerns not only "better for" but also "better (and "good", "good for" etc.). We quite often want to say that a state, S, is better than another state, T, when S and T are incompatible with each other. Yet, we cannot reduce this relation between abstract states to a relation between the obtainings of those states. We cannot say that if S were to obtain, this would be better than if T were to obtain, if this is meant to state a relation between the obtainings of S and T, respectively. For if S and T are mutually incompatible, then if one of them obtains, the other does not. So the two obtainings cannot co-exist and consequently cannot stand in any relation to each other.

As this is a general problem in axiology, which isn't specific to the comparisons of existence and non-existence, we might ignore it in the present context. As an objection to the affirmative answer to the existential question, it has too wide a reach to be convincing. Most of our thinking in axiology, both in philosophy and everyday life, presupposes that we can meaningfully say that a certain state is better than another state even when the obtainings of these states are mutually incompatible.²⁸

Let us, however, say a bit about how we believe this issue could be dealt with. One might simply take the view that not only obtainings of states of affairs can be good or better, but also those very states themselves. On this reading, S is literally better than T and this claim is not reducible to a claim about the relationship between the obtainings of these states.

How can one then deal with the objection that the world is not made any better by the existence of such states? Well, a simple and straightforward answer is to say that abstract states can be valuable but they don't contribute any value to the world. As Michael Zimmerman

²⁸A caveat: There is a way of dealing with this general problem. We could argue that statements of the form "The obtaining of S would be better (for p) than the obtaining of T" should be read as "The obtaining of S would have a higher degree of goodness or lower degree of badness (for p) than the obtaining of T". Unlike obtainings, degrees of goodness and badness are abstract objects and there is therefore nothing that hinders their co-existence.

While this proposal would solve the general problem of value comparisons between obtainings of states of affairs, it would lead to the negative answer to the existential question: The obtaining of the state consisting in p 's non-existence could not have any level of goodness or badness *for* p , since p would not exist if that state had obtained. Consequently, the obtaining of that state could not be compared in its degree of value for p with the obtaining of p 's existence.

However, a serious weakness of the proposal under consideration lies in its reduction of comparative value notions (such as 'better' and 'worse') to the non-comparative ones ('good', 'bad'). On the view taken by many philosophers, not least those who have been influenced by measurement theory, the order of conceptual priority is precisely the opposite one: We arrive at such notions as levels of goodness and badness only by making value comparisons; thus, betterness is prior to goodness and not the other way round. See e.g., Broome (1999), ch. 10, p. 164. Cf. also Chisholm and Sosa (1966).

puts it: “One might ... contend that the way to avoid having to declare that our world is better simply for the *existence* in it of the state of affairs of everyone being happy is to say that what counts in the evaluation of a world is not what valuable states of affairs *exist* in it but what valuable states of affairs *obtain* in it.”²⁹

Zimmerman himself rejects the view we have just suggested as inadequate: “Perhaps it is reasonable to say that a world’s parts consist of those states of affairs that obtain in it, rather than of those that exist in it, so that it’s no surprise that *its* final value derives only from the former. But this still leaves us with the problem of saying that states of affairs (including worlds) themselves have final value, even when *they* don’t obtain”³⁰ He doesn’t explain however, why the latter problem he mentions is a problem to begin with. It seems to us that the view in question is tenable provided we allow for the possibility of valuable entities that do not contribute in value to the actual world. More precisely, a state of affairs *S* *can* contribute in value to a world understood as a possible way things might be, that is, to a world understood as a maximal consistent abstract state of affairs, if that world contains *S*. However, when we talk about “the world” or “the actual world”, we normally mean the totality of what obtains and an abstract state of affairs that *doesn’t* obtain cannot, of course, contribute in value to that totality. To assume otherwise would be to confuse distinct ontological categories.³¹

That abstract states can be valuable is supported by the fitting-attitudes analysis of value. This point was already noted by Lemos (1994), p. 24: “The weightiest reason for an affirmative answer [to the question whether states of affairs are intrinsically valuable] arises from thinking of intrinsic value in terms of correct or fitting emotional attitudes. ... Whether a state of affairs is worthy of love or a pro-attitude does not depend on whether it obtains or is a fact.”³² In fact, it seems to us that both abstract states and their obtainings can be fitting objects of pro-attitudes. For states it might be fitting to, say, contemplate them with pleasure or desire them to obtain, whereas for obtainings it might be fitting to be pleased with them or to welcome their existence.

VII. The Person Affecting Restriction Revisited

As for the connection between “better” and “better for”, the Person Affecting Restriction remains an attractive option. It does seem plausible to claim that, to the extent we only focus

²⁹ Zimmerman (2001), p. 48 (emphasis in original).

³⁰ Ibid (emphasis in original).

³¹ We are indebted to Jonas Olson for pressing us on this point.

³² Nevertheless, Lemos rejects the suggestion that non-obtaining states of affairs possess intrinsic value and draws the conclusion that there might be something wrong with the fitting-attitudes account of value: “if some nonobtaining states of affairs are worthy of love, then we should say that being intrinsically good implies being worthy of love, but being worthy of love does not imply being intrinsically good.” (ibid.)

on welfare, an outcome cannot be better than another outcome without being better for someone. While this restriction would lead to counterintuitive implications if combined with the negative answer to the existential question (see the case of Future Bliss or Hell above), we have argued that the existential question can be given an affirmative answer.

It should be noted, however, that even coupled with the affirmative answer to the existential question, the Person Affecting Restriction, as we have stated it above, leads to counterintuitive implications, unless it is appropriately weakened. The reason is that the betterness relation between outcomes does not require the actual existence of the affected persons. Persons enter as relata in the triadic ‘better for’-relation and therefore must exist for that relation to obtain, but they are not relata in the dyadic betterness relation that obtains between outcomes. This contrast between the triadic and the dyadic relations of betterness explains why the Person Affecting Restriction cannot be correct as it stands.

Thus, to give an example, consider a variant of the Future Bliss or Hell Case above in which only the x-people exist in outcome A (i.e., in this variant, outcome A does not contain any future y-people) while outcome B still in addition contains z-people that lead hellish lives, and suppose that outcome A is the one that actually obtains. The Person Affecting Restriction implies, counterintuitively, that A is not better than B, since – as things actually are – there exists no one for whom A is better than B: The added z-people in the hypothetical outcome B, for whom A would have been better, do not actually exist. Intuitively, however, if A would have been better than B had B obtained (and B is not better than A if A obtains, as in our case), then A *is* better than B irrespective of whether A or B obtains.³³ To solve problems like this, Holtug (2004) has argued that we should replace the restriction with a weaker version, which in our formulation runs as follows:

The Wide Person Affecting Restriction: If an outcome A is better than B, then A would be better than B for someone that would exist if either A *or* B were to obtain.

In the example above, it is the second disjunct of this weaker restriction that is applicable. Clearly, it is only this wide, disjunctive version of the restriction that deserves serious consideration.³⁴

³³ This counterfactual invariance of the dyadic betterness relation is possible only because its relata (outcomes) can be assumed to exist even if they do not obtain. By contrast, the triadic relation of ‘better for’ can only satisfy a weaker condition of counterfactual invariance: If A would have been better for *p* than B if B obtained, then A is better for *p* than B even if B does not obtain, *provided that p exists*.

³⁴ In fact, even this disjunctive version of the restriction might be too strong. In principle, it is conceivable that A is better than B because it would be better for some individual who only exists in the actual outcome but not in A or B. For example, suppose that C is the actual outcome and *p* exists in C but not in A or B. Suppose further that *p* in C devotes her whole life to a certain goal *x*. For *p*, this goal is categorical: she wishes it to be realized even if it weren’t her goal in the first place. The realization of *p*’s categorical goals is of one of the things that determine how

Another problem is that it might seem that the restriction doesn't have much bite given the affirmative answer to the existential question. If one compares two outcomes A and B, then the restriction won't exclude any rankings of A and B as soon as A contains some persons with positive wellbeing that don't exist in B and B contains some persons with positive wellbeing that don't exist in A.

However, the appearances are misleading. The wide restriction does have some force, irrespectively of whether it is coupled with a positive or negative answer to the existential question. For example, it rules out all welfarist theories which imply that a mere addition of lives with positive welfare can make a population worse. Prominent examples of such theories are Average and Critical-Level Utilitarianism.³⁵ To see this, consider the following case:

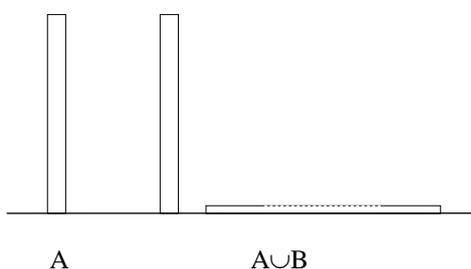


Diagram 2

According to Average Utilitarianism, $A \cup B$ is worse than A since the average welfare is lower in $A \cup B$ as compared to A. According to Critical-Level Utilitarianism, the contributive value of a person's life is her welfare minus a positive critical level and the value of a population is calculated by summing these differences for all individuals in the population. Assuming that the B-people are below the critical level, Critical-Level Utilitarianism reaches the same verdict as Average Utilitarianism. The addition of B-people has a negative contributive value given that their welfare is below the critical level.

However, $A \cup B$ would not be worse than A for anyone, irrespectively of whether A or $A \cup B$ were to obtain. Consequently, the Wide Person Affecting Restriction rules out theories such as Average and Critical-Level Utilitarianism.

good an outcome is for p . Although p wouldn't exist in A or B, let us suppose that x would be better realized in the former outcome than in the latter. It then might well be the case that A is better for p than B, even though – as it happens – there would be no one in either A or B for whom the former outcome would be better than the latter. However, note that this would make better-for comparisons outcome relative, which one might find problematic. For example, if in outcome D p would have not- x as her categorical goal, B could be better than A relative to D whereas A is better than B relative to C.

³⁵ For the latter theory, see Blackorby, Bossert, and Donaldson (1997, 1995, 2005) and Blackorby and Donaldson (1984). For a discussion of both theories, see Arrhenius (2014), ch. 3 and 5.

Nevertheless, even if we deny that $A \cup B$ is worse than A , we are not yet forced to say that $A \cup B$ is better than or equally as good as A . The wide restriction is compatible with theories that declare these outcomes incommensurable. Consider, however, the following condition:

Subjunctive Weak Pareto: If A would be better than B for everyone who would exist if A were to obtain, and for everyone who would exist if B were to obtain, then A is better than B .

This condition is a weaker version of the Pareto condition adjusted for the context in which it can be better for someone to exist than not to exist. It might seem to be an irresistible condition in the present setting, as long as we disregard other values apart from welfare. Consider now the following outcomes:

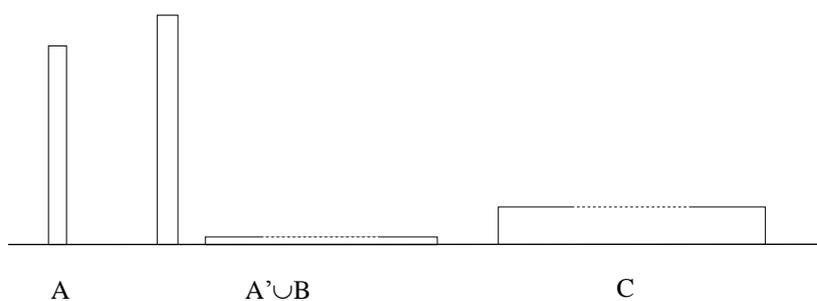


Diagram 3

Assume that A and A' consist of the same people, namely the α -people. Assume further that A is the case. Then $A' \cup B$ is better than A for all the people that exist since the α -people enjoy higher welfare in A' as compared to A . Assume now that $A' \cup B$ is the case. Since the B -people have positive welfare, $A' \cup B$ is better for them than A given the affirmative answer to the existential question. As before, $A' \cup B$ is better than A for α -people. Thus, $A' \cup B$ is better than A for everybody irrespective of whether A or $A' \cup B$ were to obtain. It follows from Subjunctive Weak Pareto that $A' \cup B$ is better than A . Hence, the affirmative answer to the existential question in conjunction with Subjunctive Weak Pareto rules out theories which imply that $A' \cup B$ and A might be incommensurable.

What is worse, it follows from any reasonable condition of inequality aversion that one can always find a B -population such that C , which consists of the same people as $A' \cup B$, is at least as good as $A' \cup B$. It follows, given transitivity, that C is better than A . Hence, the

affirmative answer to the existential question together with a couple of weak additional conditions yield the Repugnant Conclusion.

Some will find this implication unsettling and as another reason to resist the affirmative answer to the existential question. However, it can also be taken as a new argument in favour of accepting the Repugnant Conclusion *or* as a reason to be cautious with seemingly irresistible conditions such as Subjunctive Weak Pareto.

Before we finish, we should say more about ordinary language formulations of value comparisons between outcomes. The reader might have got an impression that, on our view, counterfactual claims such as “not to exist would have been better/worse for p than to exist” are absurd. But are such claims really so unpalatable? Think of Melinda Roberts’ statement about Nora: “It would have been better for Nora never to have existed at all than it is for Nora to exist.” This doesn’t sound absurd at all. But, if not, then perhaps the Argument from Absurdity doesn’t even get started?

However, how should a counterfactual statement like the one about Nora be understood? Here is what we’d like to suggest. When we use ordinary language formulations of the form “A would have been better/worse for p than B” what we do is something along these lines: (i) we state that the relevant triadic relation obtains between A, B, and p ; *and* (ii) we imply that A does not obtain. If we instead of “would have been” use “would be”, (ii) we only imply that it is not settled that A obtains.

On this analysis of “A would have been better/worse for p than B,” or “A would be better/worse for p than B,” when we use such formulations, we don’t take a stand on what relation would obtain between A, B and p under counterfactual or potential circumstances, i.e. if A did obtain. It is this feature that explains the absence of absurdity in Nora-type statements. However, we do take such a stand when we expressly state that “A would have been better/worse for p than B, *if C had been the case.*” Thereby we do state that the betterness/worseness relation would have obtained between A, B and p under the counterfactual circumstance C. Consequently, the following *is* absurd: “Not to exist would have been better/worse for p than to exist, *if p had not existed.*” It is this kind of statements that the Argument from Absurdity focuses on. Therefore, if that argument is to be invalidated, it has to be met head on, as we have done, rather than rejected on the grounds that there is no absurdity in what is being claimed to begin with.³⁶

³⁶ We would like to thank Matthew Adler, Staffan Angere, John Broome, Johan Brännmark, Krister Bykvist, Erik Carlson, Marc Fleurbaey, Iwao Hirose, Ingvar Johansson, Jonas Olson, Christian Piller, Melinda Roberts, and Toni Rønnow-Rasmussen for their helpful comments. Thanks also to the Collège d’études mondiales for being such a generous host during some of the time when this paper was written. Financial support from Riksbankens Jubileumsfond and Fondation Maison des sciences de l’homme through the *Franco-Swedish Program in Economics and Philosophy* is gratefully acknowledged.

References

- Adler, M. D., 2009, "Future Generations: A Prioritarian View", *The George Washington Law Review* 77.
- Arrhenius, G., 1999, *Population Axiology*, PhD -Diss., Dept. of Philosophy, University of Toronto.
- 2000, *Future Generations: A Challenge for Moral Theory*, F.D.-Diss., Uppsala: University Printers.
- 2003, "The Person Affecting Restriction, Comparativism, and the Moral Status of Potential People", *Ethical Perspectives*, no. 3-4.
- 2009, "Can the Person Affecting Restriction Solve the Problems in Population Ethics?" in M. Roberts & D. Wasserman (eds.) *Harming Future Persons*, Ashgate.
- 2014, *Population Ethics*, Oxford UP, forthcoming.
- Arrhenius, G., Ryberg, J., and Tännsjö, T., 2010, "The Repugnant Conclusion", *Stanford Encyclopaedia of Philosophy*, <http://plato.stanford.edu/>
- Blackorby, C., and Donaldson, D., 1984, "Social Criteria for Evaluating Population Change", *Journal of Public Economics* 25, pp. 13-33.
- Blackorby, C., Bossert, W., and Donaldson, D., 1995, "Intertemporal Population Ethics: Critical-Level Utilitarian Principles", *Econometrica* 65, pp. 1303-1320.
- 1997, "Critical-Level Utilitarianism and the Population-Ethics Dilemma", *Economics and Philosophy*, 13, pp. 197-230.
- 2005, *Population Issues in Social Choice Theory, Welfare Economics, and Ethics*, Cambridge UP.
- Broome, J., 1999, *Ethics out of Economics*, Cambridge: Cambridge University Press.
- 2004, *Weighing Lives*, Oxford: Oxford University Press.
- Buchanan, A., Brock D. W., Daniels N., Wikler D., 2000, *From Chance to Choice: Genetics and Justice*, Cambridge: Cambridge University Press.
- Bykvist, K., 2007, "The Benefits of Coming into Existence", *Philosophical Studies* 135.
- 2014, "Being and Well-Being", in I. Hirose and A. Reisner (eds.) *Weighing and Reasoning*, Oxford University Press, forthcoming.
- Chisholm, R. M. and Sosa, E., 1966, "On the Logic of "Intrinsically Better"", *American Philosophical Quarterly* 3, pp. 244-249.
- Darwall, S., 2002, *Welfare and Rational Care*, Princeton, NJ: Princeton University Press.
- Dasgupta, P., 1995, *An Inquiry into Well-Being and Destitution*, Oxford: Oxford UP.
- Feldman, F., 1995a, "Adjusting Utility for Justice: A Consequentialist Reply to the Objection from Justice", *Philosophy and Phenomenological Research* 55, pp. 567-85. Reprinted in Feldman (1997).
- 1995b, "Justice, Desert, and the Repugnant Conclusion", *Utilitas* 7, pp. 189-206. Reprinted in Feldman (1997).
- 1997, *Utilitarianism, Hedonism, and Desert: Essays in Moral Philosophy*, Cambridge: Cambridge University Press.
- Fleurbaey, M., and Voorhoeve, A., 2014, "On the social and personal value of existence" in I. Hirose and A. Reisner (eds.) *Weighing and Reasoning*, Oxford University Press, forthcoming.
- Glover, J., 1977, *Causing Death and Saving Lives*, London: Penguin Books.
- Hare, R., 1993 [1975], "Abortion and the Golden Rule", reprinted in R. Hare, *Essays on Bioethics*, New York: Oxford University Press, pp. 147-67.
- 1988a, "Possible People", *Bioethics* 2, pp. 279-293.
- 1993 [1988b], "When Does Potentiality Count?", reprinted in *Essays on Bioethics*, New York: Oxford University Press, pp. 84-97.
- 1998, "Preferences of Possible People", in C. Fehige and U. Wessels (eds.), *Preferences*, Berlin: de Gruyter, pp. 399-405.
- Heyd, D., 1988, "Procreation and Value: Can Ethics Deal With Futurity Problems?", *Philosophia* (Israel 18, pp. 151-170.

- *Genethics: Moral Issues in the Creation of People*, Berkeley: University of California Press, 1992.
- Holtug, N., 1996, “In Defence of the Slogan”, in W. Rabinowicz (ed.), *Preference and Value: Preferentialism in Ethics*, Studies in Philosophy, Dept. of Philosophy, Lund University pp. 64-89.
- 2001, “On the Value of Coming into Existence”, *The Journal of Ethics* 5.
- 2004, “Person-affecting Moralities”, in J. Ryberg and T. Tännsjö (eds.), *The Repugnant Conclusion. Essays on Population Ethics*, Dordrecht: Kluwer, pp. 129–61.
- Johansson, J., 2010, “Being and Betterness”, *Utilitas* 22, pp. 285-302.
- Lemos, N. M., 1994, *Intrinsic Value: Concept and Warrant*, Cambridge: Cambridge University Press.
- Narveson, J., 1967, “Utilitarianism and New Generations”, *Mind* 76, pp. 62-72.
- Moore, G. E., 1966 [1903], *Principia Ethica*, Cambridge: Cambridge University Press.
- Parfit, D., 1991 [1984], *Reasons and Persons*, Oxford: Clarendon Press.
- Rabinowicz, W., 2001, “Prioritarianism and Uncertainty: On the Interpersonal Addition Theorem and the Priority View”, in Egonsson, Josefsson, Petersson and Rønnow-Rasmussen (eds.), *Exploring Practical Philosophy: From Action to Values*, Aldershot: Ashgate, pp. 139-165.
- 2002, “Prioritarianism for Prospects”, *Utilitas* 14, pp. 2-21.
- 2003, “The Size of Inequality and Its Badness: Some Reflections around Temkin’s Inequality”, *Theoria*, 69, pp. 60 – 84.
- 2009a, “Broome and the Intuition of Neutrality”, *Philosophical Issues* 19, pp. 389 – 411.
- 2009b, “Values Compared”, *Polish Journal of Philosophy* 3, pp. 73 – 96.
- Rabinowicz, W., and Rønnow-Rasmussen, T., 2004, “The Strike of the Demon: On Fitting Pro-attitudes and Value” *Ethics* 114, pp. 391- 423.
- 2006, “Buck-Passing and the Right Kind of Reasons”, *Philosophical Quarterly* 56, pp. 114 – 120.
- Roberts, M. A., 1998, *Child versus Childmaker: Future Persons and Present Duties in Ethics and the Law*, Rowman and Littlefield.
- 2003, “Can it Ever Be Better Never to Have Existed At All? Person-Based Consequentialism and a New Repugnant Conclusion”, *Journal of Applied Philosophy* 20, pp. 159-185.
- Rønnow-Rasmussen, T., 2007, “Analyzing Personal Value” *The Journal of Ethics* 11, pp. 405-436.
- 2011, *Personal Value*, Oxford: Oxford University Press.
- Sinhababu, N., 2008, “Possible Girls”, *Pacific Philosophical Quarterly* 89, pp. 254 – 260.
- Temkin, L. S., 1993a, *Inequality*, Oxford: Oxford University Press.
- 1993b, “Harmful Goods, Harmless Bads” in R. G. Frey and C. W. Morris (eds.), *Value, Welfare, and Morality*, Cambridge: Cambridge University Press, , pp. 291- 324.
- Wiggins, D., 1987, *Needs, Values, Truth: Essays in the philosophy of value*, Oxford: Blackwell.
- Williamson, T., 2013. *Modal Logic as Metaphysics*, Oxford: Oxford University Press.
- Zimmerman, M. J., 2001, *The Nature of Intrinsic Value*, Lanham, Md.: Rowman & Littlefield.