

Procreation



# Preferences

Edited by  
Christoph Fehige and Ulla Wessels



Walter de Gruyter · Berlin · New York

1998

⊗ Printed on acid-free paper which falls within the guidelines of the ANSI to ensure permanence and durability.

*Library of Congress Cataloging-in-Publication Data*

Preferences / edited by Christoph Fehige and Ulla Wessels.  
p. cm. – (Perspektiven der analytischen Philosophie ; Bd. 19  
= Perspectives in analytical philosophy)  
Based on a conference held in June 1992 in Saarbrücken,  
Germany, and in Saarlouis, Germany.  
Includes bibliographical references and indexes.  
ISBN 3-11-015007-7  
1. Preferences (Philosophy) – Congresses. I. Fehige,  
Christoph. II. Wessels, Ulla. III. Series: Perspectives in analyti-  
cal philosophy ; Bd. 19.  
B105.P62P74 1998  
128'.3–dc21 97-28699  
CIP

*Die Deutsche Bibliothek – Cataloging-in-Publication Data*

**Preferences** / ed. by Christoph Fehige and Ulla Wessels. – Berlin ;  
New York : de Gruyter, 1998  
(Perspektiven der analytischen Philosophie ; Bd. 19)  
ISBN 3-11-015007-7 brosch.  
ISBN 3-11-015910-4 Geb.

© Copyright 1998 by Walter de Gruyter GmbH & Co., D-10785 Berlin  
All rights reserved, including those of translation into foreign languages. No part of this  
book may be reproduced or transmitted in any form or by any means, electronic or  
mechanical, including photocopy, recording or any information storage and retrieval system,  
without permission in writing from the publisher.

Printed in Germany

Typesetting: Thomas Fehige, Münster  
Printing: Arthur Collignon GmbH, Berlin  
Binding: Lüderitz & Bauer, Berlin  
Cover design: Rudolf Hübler, Berlin

All I want is to sit on my arse and fart and think of Dante.

Samuel Beckett

Desires and wants, however intense, are not by themselves reasons in matters of justice. The fact that we have a compelling desire does not argue for the propriety of its satisfaction any more than the strength of a conviction argues for its truth.

John Rawls

Take any demand, however slight, which any creature, however weak, may make. Ought it not, for its own sole sake, to be satisfied? If not, prove why not. The only possible kind of proof you could adduce would be the exhibition of another creature who should make a demand that ran the other way. The only possible reason there can be why any phenomenon ought to exist is that such a phenomenon actually is desired.

William James



## Preface

*Preferences* is a collection of essays on the concept and the role of preferences (desires, and the like) in practical reasoning. Ground covered includes welfare, prudence, rational decision making, and all areas of moral philosophy: ethics (applied and not so applied), metaethics, and deontic logic. A special symposium looks at *possible* preferences and their significance in matters of life and death, including the notoriously thorny question how many people there should be. All the essays are published here for the first time.

The book is not just for specialists. We have given it an introduction that, though it may move swiftly, at least starts from scratch; a selected bibliography is also provided.

Most of the authors were able to meet in advance, and to present, discuss, and then revise their contributions. But the line has to be drawn somewhere, and authors who receive a reply in this volume were not permitted to adjust their papers in the light of the final version of the reply. The initial exchange took place in Saarbrücken and Saarlouis in June 1992.

\*

Everybody has been very kind to us. Georg Meggle – selfless and cheerful as usual – co-designed the project and supported it from beginning to end. When we proposed the meeting, we were backed up by Franz von Kutschera and Wolfgang Lenzen. Barbara Schumacher helped prepare and run it.

The editors of *Perspectives in Analytical Philosophy*, Georg Meggle and Julian Nida-Rümelin, have welcomed the book in precisely the form we suggested. The authors have been co-operative and patient throughout. Christopher Abbey and Seán Matthews have given valuable advice, linguistic and otherwise, to many of us. Kornelius Bamberger was able, and kind enough, to convert most of the data that the contributors sent us. Thomas Fehige gave these data a neat, uniform lay-out. Patrick Agsten, Monika Claßen, Franziska Muschiol, Ulf Schwarz, and Valentin Wagner have assisted us, efficiently and in numerous respects; the same holds true of Karin Thom. With this list in chronological order, one important acknowledgement comes last: de Gruyter publishers. Working with Hans-Robert Cram was a pleasure; ditto, at the technical end, with Grit Müller.

The conference that gave rise to this book was made possible by the financial assistance of: Deutsche Forschungsgemeinschaft (DFG), Fritz Thyssen Stiftung, Ministerium für Wissenschaft und Kultur des Saarlandes, Universität des Saarlandes, and Vereinigung der Freunde der Universität des Saarlandes. The DFG (research project “Was zählt?”) has also funded our own work on this volume.

We thank all these persons and institutions for their support.

\*

We share the belief, now regarded in some quarters as both unsound and old-fashioned, that, in essence, morality is all about welfare, and welfare all about preferences. Some of the contributors to this volume would agree, some would not. With luck, this collection will help advance matters a little.

Christoph Fehige and Ulla Wessels  
Leipzig, January 1998



## Contents

Abstracts . . . . .	xiii
---------------------	------

### Introduction

CHRISTOPH FEHIGE AND ULLA WESSELS	
Preferences – an Introduction . . . . .	xx
Preferences – a Short Bibliography . . . . .	xliv

### Part I: Preference and Decision

ELIJAH MILLGRAM	
Deciding to Desire . . . . .	3
SYDNEY SHOEMAKER	
Desiring at Will (and at Pill): A Reply to Millgram . . . . .	26
CHRISTOPH LUMER	
Which Preferences Shall Be the Basis of Rational Decision? . . . . .	33
ANTONELLA CORRADINI	
Intrinsic Desirabilities: A Reply to Lumer . . . . .	57
RICHARD B. BRANDT	
The Rational Criticism of Preferences . . . . .	63
ANNA KUSSER	
Rational by Shock: A Reply to Brandt . . . . .	78
J. DAVID VELLEMAN	
Is Motivation Internal to Value? . . . . .	88
GEORG MEGGLE	
Motivation and Value: A Reply to Velleman . . . . .	103
SVEN DANIELSSON	
Numerical Representations of Value-Orderings: Some Basic Problems	114
ULRICH NORTMANN	
Interval Orders Defended: A Reply to Danielsson . . . . .	123

## Part II: Preference and Metaethics

LENNART ÅQVIST	
Prima Facie Obligations in Deontic Logic: A Chisholmian Analysis Based on Normative Preference Structures . . . . .	135
UWE BOMBOSCH	
The Meaning of “Ought, Prima Facie” and Decision Situations: A Reply to Åqvist . . . . .	156
FRANZ VON KUTSCHERA	
Values and Duties . . . . .	163
WILFRIED HINSCH	
Beyond Duty: A Reply to von Kutschera . . . . .	172
MARCUS WILLASCHEK	
Agency, Autonomy and Moral Obligation . . . . .	176
HILARY BOK	
Autonomy and Morality: A Reply to Willaschek . . . . .	204
DAVID WIGGINS	
In a Subjectivist Framework, Categorical Requirements and Real Practical Reasons . . . . .	212
DAVID GAUTHIER	
Subjective Obligation: A Reply to Wiggins . . . . .	233
ALLAN GIBBARD	
Preferences and Preferability . . . . .	239
JULIAN NIDA-RÜMELIN	
Goodness and Rational Preferability: A Reply to Gibbard . . . . .	260

## Part III: Preference and Ethics

JOHN BROOME	
Extended Preferences . . . . .	271
RUDOLF SCHÜSSLER	
Wish You Were Me: A Reply to Broome and a Comment on Harsanyi’s Extended Preference Theory . . . . .	288

RAINER HEGSELMANN	
Experimental Ethics: A Computer Simulation of Classes, Cliques, and Solidarity . . . . .	298
ULRICH KRAUSE	
Solidarity among Rational Egoists: A Reply to Hegselmann . . . . .	321
RAINER WERNER TRAPP	
The Potentialities and Limits of a Rational Justification of Ethical Norms, or: What Precisely is Minimal Morality? . . . . .	327
ANTHONY SIMON LADEN	
A Hobbesian Choice: A Reply to Trapp . . . . .	361
CHRISTOPH FEHIGE, RICHARD M. HARE, WOLFGANG LENZEN, JEFF MCMAHAN, PETER SINGER, THOMAS SPITZLEY, AND ULLA WESSELS	
Symposium on Possible Preferences . . . . .	367
CHRISTOPH FEHIGE AND ULLA WESSELS	
Introduction to Possible Preferences . . . . .	367
PETER SINGER	
Possible Preferences . . . . .	383
RICHARD M. HARE	
Preferences of Possible People . . . . .	399
WOLFGANG LENZEN	
Who Counts? . . . . .	406
ULLA WESSELS	
Procreation . . . . .	429
JEFF MCMAHAN	
Preferences, Death, and the Ethics of Killing . . . . .	471
THOMAS SPITZLEY	
McMahan on Psychological Continuity and the Value of Future Goods . . . . .	503
CHRISTOPH FEHIGE	
A Pareto Principle for Possible People . . . . .	508

Notes on Contributors . . . . .	544
Name Index . . . . .	550
Subject Index . . . . .	555



ULLA WESSELS  
Procreation\*

1. Introduction
  2. Wolfgang Lenzen and the Midwives
  3. Richard Hare and the Rabbits
  4. Conclusion
- References

1. *Introduction*

Do we have an obligation to bring happy people into existence? Do we have an obligation to conceive them? And then not to abort them? Or do we have neither of these obligations?

Someone who thinks that, other things being equal, individuals with satisfied preferences ought to exist will say that we ought to conceive and not to abort such an individual; a person who believes this can be termed a *Rabbit*. If somebody is not a Rabbit but thinks that, *if* there is an individual, she ought to have satisfied preferences, he will say that, although we have no obligation to conceive, once we have conceived, we have an obligation not to abort; a supporter of this sort of view may be called a *Midwife*. The terminology follows that of the introduction to this symposium and has mnemonic advantages: Rabbits are notoriously given to procreation, and Midwives are concerned with the step from pregnancy to birth.

Wolfgang Lenzen belongs to the camp of the Midwives, and Richard Hare to that of the Rabbits.<sup>1</sup> This paper criticizes the arguments of both – Lenzen's in section 2 and Hare's in section 3. A number of general beliefs of mine are in direct conflict with Lenzen's and Hare's theories. But since it would require at least an extra paper to argue for these beliefs first, I will, by and large, keep them out of the picture and meet Hare and Lenzen on their own ground.

---

<sup>\*</sup> I am grateful to Christoph Fehige for countless discussions on possible preferences and related topics; it is difficult to tell which arguments, in our respective writings, are his and which are mine. Thanks are also due to Christopher Abbey, Krister Bykvist, Wolfgang Lenzen, Richard Hare, Elijah Millgram, and Wlodek Rabinowicz, whose comments on drafts of this paper were very helpful. Work on this paper was part of the research project "Was zählt?"; I am indebted to the Deutsche Forschungsgemeinschaft (DFG) for supporting the project.

<sup>1</sup> See Lenzen (1991), (1995), and (1998); Hare (1975), (1988a), (1988b), and (1998).

Since the issue is complicated enough already, I have refrained from spelling out at every point precisely which variation of midwifism and which variation of rabbitism would, or would not, survive which variation of my objections. However – and I ask the reader to keep the claim in mind and check it out herself – some of my points clearly are, or imply, objections to midwifism and rabbitism in general, not just to Lenzen’s and Hare’s versions.

Four terminological remarks. Firstly, *ceteris paribus* clauses and similar modifiers will frequently be kept implicit during the argument. For instance, when I say that Hare is for procreation and against abortion, this means that he is for procreation and against abortion in standard cases – roughly speaking, where, if conception and birth took place, the individual in question would have a reasonably happy life, and nobody else thereby an unreasonably unhappy life.

Secondly, to say that a preference is *rational* is not to say that, on pain of irrationality, the individual *must* have that preference; instead, having that preference would be *compatible* with rationality.

Thirdly, by calling a life *happy* I am not implying that it is better for somebody to live that life rather than no life. Similarly, by calling something a *future good* I am not implying that it is better for somebody to live to see that item than not to live to see it. In other words, *given* that one lives, it would no doubt be good if the life were happy and contained the goods – but whether it is good that one lives a happy life rich in future goods (rather than no life at all) is a moral question which the *meaning* of “happy” and “future goods” – the way I use the words here – leaves open.

Fourthly, deviating from medical terminology, I will use the term “fetus” as an umbrella term to cover the following three classes of entities: zygotes or pre-embryos (fertilization less than a fortnight ago), embryos (weeks three to eight after fertilization) and fetuses in the strict sense of the word (weeks nine to thirty-eight).

## 2. *Wolfgang Lenzen and the Midwives*

### 2.1. *Lenzen’s Position (Before Supererogation)*

Wolfgang Lenzen believes that we have obligations not to have abortions, but no obligation to conceive in the first place. Some of the familiar qualifications are supposed to apply.<sup>2</sup>

<sup>2</sup> For instance, other people than the child itself could benefit from its existence, and conceiving it would thereby be a good thing – cf. Lenzen (1998), sect. 6.2. Vice versa, if there is a

My presentation and critique will be limited. Parts of Lenzen's theory seem to be based on views about supererogation – that is, about good actions that are not obligatory.<sup>3</sup> These are parts that I will skip. I'm interested in what's good. How much of what is good remains obligatory if we drive some wedge or other between the two is a question outside the scope of my inquiry.

This limitation is not unfair to Lenzen. If he thought that supererogation made *all* the moral difference between conceiving and not having an abortion, he would have to say that either of them is *better* than its opposite (i.e. conceiving is better than not conceiving, and not having an abortion is better than having an abortion), though only the latter (not having an abortion) is also *obligatory*. But he clearly wants to deny the symmetry in the realm of betterness, too. He believes that, while, other things being equal, it is *not* better to conceive than not to conceive,<sup>4</sup> it *is*, other things being equal, better not to have an abortion than to have an abortion. There is nothing wrong with an inquiry into this *part* of his position – the part, prior to supererogation, about what is better than what.

My concern with betterness has a terminological consequence, too. Where I use words like “obligation”, “permission”, “ought”, “right”, and “wrong”, these are just meant to refer to what is (*ceteris paribus*) better or worse; I'm not saying that the corresponding judgements would survive one theory or another of supererogation.

So what we're looking at is Lenzen's reasons for believing the following:

- Anti-Abortion: On the one hand, a world in which we make sure a fetus gets a happy life is better than one in which we abort it.
- Not-Pro-Fertilization: On the other hand, a world in which we conceive and then make sure the fertilized egg grows to see a happy life is not better than one in which we don't.<sup>5</sup>

Behind these claims there lies a more general creed. Couched in terms of betterness, it runs as follows:

---

serious threat to the health of either the pregnant woman or of the child; if the woman has been raped; or if, for financial reasons, the mother is unable to look after the child: then an abortion may be justifiable – cf. Lenzen (1991) and (1998), sect. 1.

<sup>3</sup> Cf. Lenzen (1998), sects. 2.4, 5.1, and 6.2.

<sup>4</sup> Cf., e.g., Lenzen (1998), sect. 6.2. There is strong indirect evidence, too. It would be bizarre if Lenzen wanted supererogation to do *all* the work but didn't tell us so – that is, if he thought that conception was indeed better than non-conception (just not good enough to be obligatory), but tacitly passed over this.

<sup>5</sup> For Lenzen's support of Anti-Abortion and Not-Pro-Fertilization, cf. his (1990), sects. 2.3 f., (1991), sect. 4.3, and (1998), sect. 4.



- Anti-Death: On the one hand, given an individual who exists, it is better that the individual gets a happy future than that she gets no future; this is better even if the individual does not want such a future beforehand.
- Not-Pro-New-Life: On the other hand, a world with a happy extra individual is not better than a world without that individual altogether.<sup>6</sup>

Anti-Death is supposed to support Anti-Abortion, and Not-Pro-New-Life is supposed to support Not-Pro-Fertilization.

Four things will happen. Section 2.2 will address a critical question to Anti-Death and thus, as will be shown, to Anti-Abortion as well. Section 2.3 will voice two types of doubts regarding the *conjunction* of Anti-Death and Not-Pro-New-Life or, more precisely, regarding the moral asymmetries they jointly imply. Section 2.4 will ask whether Lenzen really has deduced his specific claims (Anti-Abortion and Not-Pro-Fertilization) from his general claims (Anti-Death and Not-Pro-New-Life). The answer is no. He has not shown that his general theory about the value of life and the badness of killing, even if it were tenable, would draw a moral line between abortion and non-conception.

The four objections – one in 2.2, two in 2.3, and one in 2.4 – are mutually independent.

### 2.2. *A Problem for Anti-Death and Anti-Abortion: The Suicide Argument*

The first of my four points is one against Anti-Death and Anti-Abortion. I call it the *suicide argument*, and it runs as follows: let us imagine the case of a person *a* who does not attach any positive value to her own survival. She wants survival neither in itself nor as a means to another end of hers. She wants to commit suicide, even though she would be very happy if she didn't. What should we do? Should we help *a* to kill herself? Should we try to stop her? Or neither the one nor the other?

Of course, the problem *might* be (and in most real-life cases is) that the person in question doesn't believe, or at any rate doesn't fully grasp that she would be happy if she went on living. Suppose however that, this time, belief and full representation just aren't the problem. Person *a* fully represents to herself all the satisfaction and all the ice-cream that life has in store for her –

<sup>6</sup> For Lenzen's support of Anti-Death, cf. his (1990), sect. 2.2, (1991), sect. 3, and (1998), sect. 3; for his support of Not-Pro-New-Life, cf. his (1990), sect. 2.1, and (1998), sect. 6.2.

yet, she'd rather do without it. She has a rational preference for committing suicide.<sup>7</sup>

In such a case, I believe, morality ought to respect the preference for death. Other things being equal, the suicide ought to take place and ought not to be prevented.

Maybe my belief is little more than an intuition. But since Wolfgang Lenzen makes generous use of *his* moral intuitions,<sup>8</sup> perhaps I may be allowed to drag in one of mine. Moreover, that morality licenses the suicide at issue has at least *some* credentials from other sources. Since the suicide is *desired*, the verdict that I propose follows from the decision-theoretic standard link between a person's rational preferences and her welfare (utility, etc.).

In the extraordinary case at hand, then, death is, since it is rationally desired, *ceteris paribus* the better thing. Now let us weaken the case a little. We regard a person *b* who is perhaps not against survival, but not *for* it either. She has no preference for surviving although she knows, and fully represents to herself, that she would become happy if she didn't die. In the previous story, the rational preference *against* life generated the moral *betterness* of death. So now that the story is toned down, the rational *absence* of a preference for life will at least generate the moral verdict that *b*'s death is *not worse* than her survival.

The latter case, however, is in all relevant respects identical to that of an early fetus that would become happy if it were not aborted. Both *b* and that fetus have *no preference for life*. If, as we said, there is no moral reason to keep *b* alive, then neither will there be one to keep the fetus alive.

#### *How the Suicide Argument Survives Various Objections*

Five objections come to mind. One of them, which will be dealt with first, questions the very *possibility* of the decisive case from the suicide argument, the other four query my moral *evaluation* of it.

Objection (1) to the suicide argument says that individual *b* simply *cannot* rationally fail to prefer a happy survival. This is due to certain conceptual connections between happiness and rational preference. Happiness is by definition something that one cannot rationally decline. In other words, the following claim

<sup>7</sup> Lenzen's juvenile *J* seems to be a different case. He doesn't believe that he would become happy.

As to "rational" in this context, see the introduction to this paper.

<sup>8</sup> Cf., e.g., Lenzen (1998), sects. 2.2 ("we would be reluctant to consider", "[i]t would seem quite natural to suppose", etc.) and 5.1 (where the role of intuitions is mentioned explicitly).

(\*) Individuals want, other things being equal, to be happy

is a conceptual truth, at least if the wanting in question is supposed to be compatible with rationality.

Reply: I have no objections to conceptual links between happiness and rational preference – but precisely which ones? Upon closer inspection, claim (\*) turns out to be too strong. The connection we should put into the meaning of the word “happy” is, at best, this one:

(\*\*) Individuals want, other things being equal, to spend their life happy.<sup>9</sup>

If we subscribe to (\*\*) and not to (\*), then somebody can consistently and rationally reject the package that consists of life and happiness.<sup>10</sup> This attitude is compatible with (\*\*), just in the same way that the absence of a wish to be knocked down and taken to hospital is compatible with the wish, given that one has been knocked down, to be taken to hospital. And indeed, while the rejection of happiness for one’s life-time would be fishy, what would justify us in ruling out the rejection of the combination (i.e. of life *and* happiness) as irrational? Why would it have to be irrational, or even impossible, for somebody to find the idea of non-existence just as attractive as that of a happy existence? It may be a rare desire, but that doesn’t make it either irrational or impossible. Objection (1) is thus based on a stronger ‘requirement of rationality’ than justified.

Objection (2) to the suicide argument says that what action is good for someone does not just depend on what he wants at the time of action. Put it this way: our heroine *b* is a chain of person stages  $b_1$  to  $b_n$ ; suppose that the person stage of *b* that rationally lacks a preference for a happy survival is  $b_{1001}$ . Now, what is good for *b* consists of what is good for *all* her person stages, not just from what is good for  $b_{1001}$ . We must consult them all, not just one of them. That is what the suicide argument, by looking at  $b_{1001}$ ’s preferences alone, has failed to do; if it had done so, the result would have been different.<sup>11</sup>

Reply: Let’s take a moment to sort this out. There is no disagreement here with respect to *b*’s past person stages  $b_1$  to  $b_{1000}$ ; they have existed, they have

<sup>9</sup> Cf. Fehige (1998), sect. 1, subsect. “Orexigenesis vs. prophylaxis I”, and Singer (1993), p. 131.

<sup>10</sup> It might be objected that happiness *entails* life and that, therefore, talking about the *package* of life and happiness is misleading. This may be so. But if we decide not to be misled, the word “package” comes in handy.

<sup>11</sup> This is, or resembles, Lenzen’s objection to the Mind Readers – see, e.g., his (1998), sect. 5.2. For the term “Mind Reader”, see Fehige/Wessels (1998).

had preferences, and these should be taken into account. And if they had preferences for the existence of, say,  $b_{1005}$ , then clearly these could override the *ceteris paribus* judgement that, since  $b_{1001}$  doesn't want life, death would be all right.

This leaves us with the possible future person stages  $b_{1002}$ ,  $b_{1003}$ , etc. from the survival scenario. The objection bids us to take into account *their* happiness as well. To make sure that this happens, let us require that  $b_{1001}$  empathize with them, and that  $b_{1001}$ 's overall preference, in order to have moral authority, must give the same weight to  $b_{1001}$ 's own interests as to those of every possible later  $b_i$ .

Thus we demand empathy – a prudential, intrapersonal type thereof.<sup>12</sup> But what precisely will empathy come up with? It will certainly require  $b_{1001}$  to pity every person stage which is unhappy; the unhappier a person stage, the less good its existence.<sup>13</sup> But what about the other way round? What about the stages which would be happy? Does the Midwife suggest that we (or person stage  $b_{1001}$ ) should *pity* a person stage for not existing if its existence would be happy?

Note the deep, deep contradiction that is lying in wait for the Midwife here. Midwives insist on arguing just from the interests of objects that exist (have existed, will exist); it is a central tenet of midwifism – since without it, Midwives would be Rabbits – that the fact that an object would be happy if it existed does not make its existence, compared to its non-existence, a good thing.

The central tenet would be violated if we now took, as the objection we're dealing with suggests we do, the interests of *the possible future stages* as a moral reason for *their existence!* This is an argument a Midwife cannot use. (To employ one of the Midwife's favourite expressions, person stage  $b_{1002}$  isn't *deprived* of anything if it never exists.) Note the subtle difference: a Midwife can say (though I don't see why, and though I think the suicide argument speaks against it) that, if the future person stages are happy, it is in *the person's* interest to have them. But a Midwife *cannot* say, and hence cannot *support* the previous claim by saying, that the interests of *the possible future person stages* generate a reason for *their* existence. And this answers objection (2).

Objection (3) to the suicide argument says that it would justify us in bringing up a fetus even if we knew that, for some unavoidable reason or other, his life would be hell. The *ex-ante* absence of a preference against damnation would,

<sup>12</sup> A demand that is not unusual – see, e.g., Hare's requirement of prudence, referred to in note 56 below.

<sup>13</sup> See also the reply to objection (3).

by the reasoning of the suicide argument, show that damnation is not worse than non-existence; therefore, the suicide argument would permit the creation of a life of eternal suffering. This is a *reductio ad absurdum*.

Reply: It would be one indeed, but the suicide argument does not entail the scandalous verdict. In order to count, an individual's preference (or lack of preference) must be rational; it must be compatible with the verdict rationality would come up with if faced with the options at issue. What does this mean for damnation? Unhappiness is something that rationality advises us to avoid. Thus, if our morality is to be guided by rational preferences, we must *ascribe* to people a preference to avoid suffering. We must ascribe it even in cases in which it has no or little psychological reality – as is the case with a fetus, which has no 'real' preference whatsoever. In that sense, i.e. in the sense that includes the ascribed preferences, the fetus does *not* lack preferences against eternal suffering, and thus the suicide argument does not apply.<sup>14</sup>

But, it might be asked at this point, should we not, just the way we have ascribed preferences *against* an *unhappy* future, also ascribe preferences *for* a *happy* (as opposed to no) future?

The answer is twofold. The Midwife just cannot say that the interests of *the individual's possible future person stages* should prompt us to do so – see the reply to objection (2). So we're left with the question whether it is simply *in the individual's interest* to have, even where a corresponding preference is lacking, a happy future as opposed to no future. That no is the right answer is precisely what the suicide argument itself is out to suggest.<sup>15</sup>

Objection (4) to the suicide argument says that the step from the moral evaluation of the suicide to that of the fetus went wrong. For we should not, in general, treat an entity that cannot grasp a proposition *p* as if it had no preference for *p*.

Reply: Well, not if *pure rationality* has a different verdict about *p* – see the reply to objection (3). But otherwise, the procedure is correct and alternatives would be hard to justify. If an individual *cannot* grasp *p* and pure rationality does not recommend *p*, then it is clearly unwarranted to treat her as if she wanted *p*.<sup>16</sup>

<sup>14</sup> For a similar discussion, cf. Fehige (1998), sect. 1, subsect. "Orexigenesis vs. prophylaxis II".

<sup>15</sup> As to rational preferences and happiness (or welfare), cf. also the discussion in sect. 3.3 of this paper.

<sup>16</sup> The point becomes blatant if we think of compensation. Suppose there's an individual who has no 'real' preference for *p* (where *p* is not recommended by pure rationality) and we treat her as if she wanted *p*; something she does indeed want, although less strongly, is *q*. In a situation where she can have just one of the two goods, we would be required to give her *p*.

Objection (5) to the suicide argument says that, if we accept the suicide argument, and thus the claim that a happy future as such does not speak against death, then, apart from *other* people's sorrow and the like, very little would speak against somebody's death at all, and that is implausible.

Reply: Death would be bad *for* a person because of her past and present preferences – preferences for her survival or preferences that can only be satisfied if she survives. In the rare cases where such preferences are lacking, I don't see why death should be bad for the person. And since most people have many strong preferences of this type – *how* strong becomes obvious when we see how much pain and misery are still preferred to death! –, death will usually remain a terrible thing, even if we deny that the future as such, without an orectic anchor in the past or the present, speaks against death. We do not need Anti-Death in order to say that killing the man on the Clapham omnibus would be very bad.<sup>17</sup>

Thus, the five objections to the suicide argument do not work, and the argument remains intact. It is possible that somebody rationally wants to commit suicide even though he is fully and vividly aware of the fact that he would be happy if he went on living. He can simply prefer non-existence to a happy existence. In such a case, we ought not to prevent him from committing suicide. Similarly, we have no moral obligation to keep somebody alive who has no preference for his happy survival. A fetus who has no preferences whatsoever resembles, in the relevant respect, the latter case; therefore, we have no moral obligation to keep it alive.

Two more remarks on the scope of the argument. Firstly, the argument works no matter *which* goods a proponent of Anti-Death has in mind when he speaks of a "happy future". It doesn't matter whether what he's talking about are future experiences, future satisfied preferences, or future 'objective' goods of whatever kind. For any such goods, it is always possible to rationally prefer the end to a sequel full of them. (The same thing applies to our discussion of principles (\*) and (\*\*) in the reply to objection (1). Roughly speaking, different notions of happiness won't make a difference.)

Secondly, there are ways of weakening Anti-Death that won't help. To say that a happy future makes death bad *if this or that condition is fulfilled*, is, as can easily be seen, no way out<sup>18</sup> (unless, of course, the "if" clause requires ex-

---

In other words, we would be required to satisfy a preference she doesn't have, and doesn't by virtue of rationality have to have, at the cost of one she does have. This would be a weird notion of benefitting the individual.

<sup>17</sup> Cf. Fehige (1998), sect. 1, subsect. "Orexigenesis vs. Prophylaxis I".

<sup>18</sup> Jeff McMahan, e.g., thinks that, preferences aside, an individual's interest in future goods de-

ante preferences for survival, in which case the character of the theory is altered radically). Conditions or not, one can, without lack of empathy for one's future person stages, lack any preference for the future in question; therefore, the suicide argument will work, even against the weakened versions. By itself (especially if it hasn't been desired beforehand), a happy future – that is what the argument shows – *never* makes death bad.

2.3. *Problems with the Conjunction of Anti-Death and Not-Pro-New-Life: The Suicide Revisited and the Wretched Child*

My second and third objections, entitled “The Suicide Revisited” and “The Wretched Child”, concern the conjunction of Anti-Death and Not-Pro-New-Life – in other words, the alleged moral asymmetry of the case in which the person in question exists anyway and the case in which she does not.<sup>19</sup>

*The Suicide Revisited*

Objection no. two is again to some degree an intuitive point, just like many of Lenzen's.<sup>20</sup> Consider once more the case of the suicide *b* from the previous section (2.2): *b* has no wishes, not even implicit ones, for survival (not even for a happy future), nor for anything that her survival could help to achieve.<sup>21</sup> Furthermore, let us assume that there are no friends, relatives or any other people whose interests would be affected by *b*'s death.

(In the previous section, I advocated a particular verdict about *b*'s death or future. Now my objective is different. I want to point out that, *whatever the verdict is*, it should be the *same* in her as in a certain other case.)

Now look at a segment of happy future: a certain number of days, with a certain number of pleasant experiences, vastly outnumbering the unpleasant ones, and a certain number of satisfied preferences, vastly outnumbering the frustrated ones. Suppose that we can either give that future to *b*, or create a

---

depends on the psychological relatedness between her present and future person stages (McMahan 1998, sect. 3). The relevant modification (note the italics to come) of our suicide argument would proceed from a case in which, although aware of the fact that the possible future person stages *that are psychologically related with*  $b_{1001}$  would enjoy a happy life, person stage  $b_{1001}$  would have no preference for a happy survival.

<sup>19</sup> Technically, of course, arguments against one of the conjuncts – such as the argument that has been presented in section 2.2 above – are also arguments against the conjunction. What I'm now concerned with, however, are the problems which arise from *marrying* Anti-Death to Not-Pro-New-Life; I will show that the marriage is a mismatch – regardless of any merits or shortcomings each partner might have in its own right.

<sup>20</sup> See note 8 above.

<sup>21</sup> Note that the suicide differs from Lenzen's comatose (1998, sect. 5.2) who, I take it, was more normal in that, before the accident, he had a preference for survival.

*new* individual who will live through precisely that future; suppose also, and again improbably, that there are no morally relevant side-effects either way (for example, no benefits or costs, emotional or other, involved in killing *b* and in creating a new being) – then what should we do? Give the future to the existing person or create a new one and give it to her?

It doesn't matter. Both scenarios contain the same number of happy days, and both scenarios are equally good.<sup>22</sup> Lenzen thinks it does matter – see his claims Anti-Death and Not-Pro-New-Life. He says that we are doing a good thing if we give the happy future to the suicide (who never wanted it, not even implicitly), but not a good thing if we create a person and give it to *her*.

I would of course agree that under *normal* circumstances providing an existing person with the happy future would be better than creating a person and providing *her* with it. *Normal* people *want* a happy future and *that* would create a moral asymmetry between the existing person and the extra person (who, by virtue of her non-existence, has no such preference). Wishes of this type get frustrated if we give the future to the extra person, but not if we give it to the existing person; thus we'd better give it to the latter.

The case at issue, however, was *not* normal, and preferences for survival were *not* there. Their absence, I suggest, is the absence of any reason to favour the existing person. *It is precisely by having no preference for a happy future that the previous self has left us with no moral reason to tie the happy future to him rather than to a new bearer.* Therefore, both scenarios are equally good. If you share this intuition, you cannot support Anti-Death *and* Not-Pro-New-Life. One of them will have to go.

### *The Wretched Child*

Here is my third objection, and the second one against the conjunction of Anti-Death and Not-Pro-New-Life. Clearly, we want morality to forbid sadistic creations – in other words, the creation of persons who will be miserable throughout their life. Lenzen agrees that such creations are wrong.<sup>23</sup>

But look at the strange *group* of claims we get if we add this reasonable requirement, which Lenzen himself endorses, to his Anti-Death and Not-Pro-

<sup>22</sup> The issue I'm raising here is that of "replaceability"; cf., e.g., Parfit (1984), § 122, Singer (1993), pp. 161–5, and, in this volume, Singer, sects. 1 f., and McMahan, sect. 3.3. It is normally discussed on the basis of more realistic cases: should we kill a disabled fetus if the mother will have another try and is then likely to have a healthy child? Or if a 14-year-old girl is pregnant and can have a less happy child now or an abortion and a happier child later – what ought she to do?

<sup>23</sup> Cf., e.g., Lenzen (1998), sect. 6.2.



New-Life:<sup>24</sup>

- (A) On the one hand, a bad future is a reason not to bring a person into existence.
- (B) On the other hand, a good future is not a reason to bring a person into existence.

Whence this asymmetry? The question becomes even more puzzling if we recall:

- (C) A good future *is* a reason not to *end* a person's existence.

How do these claims go together? If a bad future makes existence a bad thing, then why does a good future – and one that is good enough to make continued existence a good thing! – not make existence a good thing?

I find this hard to answer. Not so hard, of course, if you *set out* to save the three intuitions that conception is usually not obligatory, *and* the existence of a wretched child is worth avoiding, *and* abortion is usually wrong. Indeed, these three almost force you to embrace claims (A), (B), and (C). This, however – as an unbiased look at (A), (B), and (C) suggests –, is rather an argument *against* the former trio.

I asked how claims (A), (B), and (C) go together. Note that the answer could *not* be: “While the creation of an unhappy person leads to the *existence* of somebody who suffers, and thus harms the person in question, the creation of a happy child does not benefit that child – because the child didn't exist at, and would never exist without the, creation.”

This would be morals based on a linguistic muddle – a real howler. If creating a happy person couldn't count as benefitting her because at the point of creation there was no “her” to benefit, then creating an unhappy person couldn't count as harming her because at the time of creation there was no “her” to harm. Thus we have to put a ban either on both ways of speaking (“being created can harm a person” *and* “being created can benefit a person”), or on neither of the two. *Tertium* would be cheating.

If we put a ban on both of them, (A) will have to go, and the creation of misery will be in order.<sup>25</sup> Since this route is deeply unattractive, and is not the one Lenzen chooses to take, I will not pursue it here.

<sup>24</sup> The requirement appears as (A) in the following list, paraphrases of Anti-Death and Not-Pro-New-Life as (C) and (B) respectively.

<sup>25</sup> This becomes particularly clear when we invent situations in which, after the act of procreation, there is no way of putting an end to the misery, i.e. no way of killing the person in question. The causal road to relief is blocked. In such a situation, the two-step construction Lenzen once had in mind (see sect. 2.4 of his 1990 and note 16 in his 1998) could not avoid the cruel result that producing misery is all right.

Lenzen goes in the other direction and is prepared to permit both ways of speaking: “[L]et it here be taken for granted”, he writes, “that procreation would [...] be good for a happy but bad for a wretched child.”<sup>26</sup> But if something is good *for somebody*, then clearly it is *ceteris paribus* good. *Thus Lenzen is committed to the claim that it is ceteris paribus good to bring a happy person into existence.* In that case, (B) will have to go.

We haven’t been shown – and I fail to see how one *could* show – why we should endorse the puzzlingly asymmetric trio of (A), (B), and (C).

#### 2.4. *Problems with the Inference from Anti-Death and Not-Pro-New-Life to Anti-Abortion and Not-Pro-Fertilization: Identity and Its Relatives*

We come to the fourth objection against Lenzen’s position. Having questioned in the previous two sections (2.2 and 2.3) Anti-Death itself and Anti-Death in combination with Not-Pro-New-Life, we now ask whether the conjunction of Anti-Death and Not-Pro-New-Life, tenable or not, entails that of Anti-Abortion and Not-Pro-Fertilization. Wolfgang Lenzen thinks that it does. Whether he’s right will depend on questions of identity.

*On the one hand,* it would have to be shown that a fertilized egg *is* the individual whose future we are talking about; only then can, in Anti-Death, the clause “given an individual who exists” get to work and generate Anti-Abortion, i.e. the obligation to lead the fertilized egg into a happy life and, a fortiori, not to abort it.

*On the other hand,* it would have to be shown of other items that they are *not* the individual whose future we are talking about: items such as the unfertilized egg, the sperm, and the unfused pair of an unfertilized egg and a sperm.<sup>27</sup> If one of these objects *were* the individual, then – contrary to Lenzen’s theory – that object, too, would fall under the jurisdiction of Anti-Death and would have to be led into a happy future (just like the fertilized egg, see “on the one hand”). Contrary to Lenzen’s theory, conception, too, would then be obligatory, as a step on the object’s journey to happiness, which, as has just been reported, would itself be obligatory. Only if *none* of these objects *is* the individual,

<sup>26</sup> Lenzen (1998), sect. 6.2.

<sup>27</sup> Lenzen himself is fully aware of the fact that his theory has to make a moral difference not only between a fertilized and an unfertilized egg, but also between the *pair* of sperm and egg before and after they fuse. Cf. Lenzen (1995), p. 231.

is Not-Pro-New-Life rather than Anti-Death applicable and can generate Not-Pro-Fertilization, i.e. the permission to refrain from conception.

*Thus, Anti-Death and Not-Pro-New-Life pass the buck to our theory of identity.* The *moral* asymmetry between non-conception and abortion boils down to the *identity-theoretical* asymmetry between the unfused and the fused pair of sperm and egg. The former is *not* the individual whose happy future we're talking about whereas the latter *is*.

Identity, however, is notorious – a philosophical maelstrom if ever there was one. Criteria for the identity of objects in general and of people in particular are controversial and abound with puzzles. If this is so, and if for Lenzen so much hinges on questions of identity, then why doesn't he discuss them? Where is the *theory* of identity (and where is the defense of it) that bears out his central identity-theoretical claim?<sup>28</sup> Identity is supposed to do a lot of work, but receives little attention.

This disproportion is all the more baffling since serious identity-theoretical proposals to be found in the literature clearly do *not* bear out Lenzen's central claim.

- Some theories deny that identity obeys bivalence. It need not be the case, they say, that either *a* is identical to *b* or *a* is not identical to *b*. As to the beginning of life, it may then well be that the unfertilized egg is *not* identical to the subsequent child, and that the mass of, say, 128 adhering cells *is*, but that the intermediate stages neither are nor are not.<sup>29</sup>
- Some theories say that identity is a matter of degrees, not of all or nothing.<sup>30</sup> Faced with gradual developments like that of a human life, such a theory would argue that identity emerges, and that the fertilized egg is at best 'slightly more identical' to the child than the pair of sperm and unfertilized egg.
- Some theories base identity on similarity or shared properties.<sup>31</sup> The pair of egg and sperm just before the fusion and the entity just after the fusion

<sup>28</sup> I.e. the claim that neither the unfertilized egg, the sperm nor the unfused pair of unfertilized egg and sperm are the individual whose happy future we're talking about, whereas the fertilized egg is.

<sup>29</sup> For this type of view, see, for instance, van Inwagen (1990), chs. 14 and 17–9.

<sup>30</sup> See, for instance, Hare (1981), ch. 5.4, and Lyon (1980), pp. 178–80. Cf. also Parfit (1971), (1984), part III, and (1995), esp. part II; strictly speaking, Parfit himself is not after identity, but after 'what matters in survival' – this, however, is precisely Lenzen's topic. The Buddhist view seems to be similar, cf. chs. 2 f., esp. pp. 33–40, of LaFleur (1992).

<sup>31</sup> As does Wittgenstein, see (1958), pp. 61–3; cf. also Loux (1978), pp. 124–6, Borowski (1976), and Lyon (1980).

share, intuitively speaking, such an ‘amount’ of properties that they would have to count as the same object by this type of standard. They are close neighbours in time and share most of their physical make-up, of their history, of their causal potential,<sup>32</sup> and of their position in space.

- Many theories hold our brain or our mind to be a necessary component of our identity.<sup>33</sup> They would not support Lenzen’s position, since the early fetus does not have a brain or a mind, and thus could not be the same individual as, e.g., the seven-year-old child.

All these are candidates. The answer to the decisive question – to the question “Who is who?” – is far from obvious. We have to diagnose, at the very least, a monumental lack of arguments at the heart of Lenzen’s theory.

### *Three Moves That Won’t Help*

Although Lenzen himself does not seem to build on them, I would like to mention three moves that could *not* help support his views.

(1) It would be of no use to say that the controversy among identity theories has nothing to do with morals and that, if we are looking for a *morally* relevant notion of identity (which we are), the choice of fertilization as the crucial point is just *obvious*.

It isn’t. This is why even people who are quite happy with Lenzen’s general position (Anti-Death and Not-Pro-New-Life) come up with different cut-off points (points, that is, before which killing a fetus or even an infant would, and after which it would not, be all right). Even they cannot agree what identity, in the morally relevant sense, *is*; some of them think, in contrast to Lenzen, that it requires a brain, or even consciousness.<sup>34</sup> To declare that the choice of an identity theory is a moral choice is to say that it requires a *moral* argument, not that it requires none.

(2) What about the fact that there’s space between the egg and the sperm before they fuse – doesn’t that settle the issue? No. Imagine we could cut adults

<sup>32</sup> We might be tempted to say: “Fusion makes an enormous difference in causal potential since it is *necessary* on the way to a happy person.” If this argument were valid, identity would be renewed whenever a necessary condition were met. For instance, the acquisition of consciousness, being necessary on the way to a happy person, would also create a new individual. This would ruin Lenzen’s theory since the individual who experiences the happy future would then not be the same as the fertilized egg.

<sup>33</sup> David Lewis, Anthony Quinton, and Sydney Shoemaker are among those who have proposed psychological criteria of personal identity; Thomas Nagel and John L. Mackie think that we are essentially our brains. For references, see Johnston (1987), notes 3 and 18. Cf. also in this volume McMahan, sect. 2.2: “[E]ach of us began to exist when the brain of his or her body developed the capacity to support consciousness and mental activity.”

<sup>34</sup> See, e.g., Leist (1990, ch. V), Sumner (1981, ch. IV), and McMahan (1998).

in half and put them back together again. Then while Paul is disassembled, we would still say “Look, there’s Paul cut in half”. The space between the halves does not impugn his identity. Neither, of course, does the absence of space suffice to establish identity. People touch each other without becoming one; and bacteria live in people, fetuses in mothers, etc., and are still not one with their hosts.

(3) Lenzen’s claim about identity could not be supported by adducing genetic make-up as a criterion. As a *necessary* condition for identity, genetic make-up would not show that identity starts with the fusion; at best, that it starts no earlier than with the fusion (but perhaps much later) – a result that is too weak for Lenzen’s purposes.

Could genetic make-up be a *sufficient* condition for identity? No. Firstly, the fused pair of egg and sperm can still split up and would thus be identical with two or more children – a *reductio ad absurdum*.<sup>35</sup> Secondly, we can think of brainwashings that would change our identity without changing our genetic make-up. And if ‘I’ had been raised with ‘my’ genetic make-up, but in a different culture and century and class, then I would have become a different person. The connection between genes and identity is too weak to serve as the missing link.

#### *What Fertilized Eggs Do by Themselves*

What is, according to Lenzen, the crucial difference between before and after fertilization? As we have seen, the answer that *identity* is the difference (i.e. that the individual whose future we’re talking about comes into existence at fertilization) is not obvious, to put it mildly. It needs back-up. *Lenzen needs to show us properties that constitute identity and coincide with fertilization; or at any rate properties that – identity or not – make a moral difference and coincide with fertilization.*

I will now look at Lenzen’s candidates for such properties. I will ask whether (or how, or at what moral costs) they succeed in singling out fertilization as the crucial threshold.

The morally relevant difference between the fertilized and the unfertilized egg, Lenzen tells us, is that the latter “does not by itself develop into a being who lives a life worth living. Its ‘life’ normally ends with menstruation, and as such has no more value than, say, the ‘life’ of a finger-nail or the ‘life’ of an appendix.”<sup>36</sup>

<sup>35</sup> Cf. Singer (1995), p. 94. Genetic make-up is nothing that Lenzen himself wants to rely on, cf. his (1995), p. 231.

<sup>36</sup> Lenzen (1998), sect. 4.

Now, in what sense of “by itself” is it true that an ordinary fertilized egg that turns into a happy life does so by itself, while an ordinary unfertilized egg (or an ordinary pair of an unfertilized egg and a sperm<sup>37</sup>) that turns into a happy life does not do so by itself? I suggest we check out the various readings that come to mind.

First reading of “by itself”: Let us say that something that happens does so by itself if it happens ‘normally’.

This interpretation does not support Lenzen’s claim that when an ordinary fertilized egg turns into a happy life it does so by itself. As Lenzen himself confirms, *neither* the unfertilized egg *nor* the fertilized one normally makes its way to a happy life: “[O]nly 5 out of 100 fertilized human eggs survive up to delivery, and the corresponding survival-rates for embryos and fetuses at later stages isn’t very much higher”<sup>38</sup>. Thus, the odds are both against unfertilized and fertilized eggs and won’t yield the qualitative difference.

Second reading of “by itself”: Let us try another statistical interpretation – one that refers not to the events but to what people do; let us say that something that happens does so by itself if it is true that it would happen if everybody did what is ‘normal’ for them to do.

Under this reading it might be true of some real-life fertilized eggs that turn into happy persons that they do so by themselves. But now there is more bad news for Lenzen. Firstly, the same might hold for some unfertilized eggs – again fertilization has failed to make the difference. Secondly, the whole idea of giving moral authority to the question what people *usually* do is bizarre. (To the extent that we grant that authority, we exclude the possibility of a widespread vice.) The third and most important point is an application of the second one. It is easy to see that, within Lenzen’s theory, the current interpretation of “by itself” would imply that, if most women had abortions, abortions would be all right. Since this is doubtless a verdict Lenzen would reject, this interpretation cannot be the one he has in mind.

<sup>37</sup> See note 27 above.

<sup>38</sup> Lenzen (1998), sect. 6.1; cf. also Singer (1993), ch. 6. The problem I am addressing here should not be confused with the one that Lenzen wants to solve, and can solve, by his expected utility approach (cf. his 1998, sect. 6.2). *Once we have established* that, in Lenzen’s sense, the happy future of the fetus ‘counts’ whereas that of the unfused pair of sperm and egg does not, then the expected utility approach would correctly insist that, when calculating the badness of killing a fetus, we weight its possible future happiness with the (somewhat dim) probability that the fetus, if not killed by us, would live to see it. But we are still at the step I have just italicized; expected utility doesn’t help, and wasn’t invoked by Lenzen to help, with *this* problem.

Third reading of “by itself”: Suppose we have a sufficiently clear distinction between acts and omissions. To avoid confusion with the word “act” in the wider sense (in which even omissions can be acts), I will call acts that are not omissions *active acts*, an expression that is ugly but unambiguous. Let us say that something that happens does so by itself if it would have happened even if nobody had performed any active acts.

Again, this reading does not support Lenzen’s claim that when an ordinary fertilized egg turns into a happy life it does so by itself. While pregnant, mother must do at least everything that is necessary for her own survival, then she must give birth, then the child has to be brought up. Obviously, all this involves lots of active acts; mother has to eat, for instance, and later either she or somebody else will have to feed her child – and clearly eating and feeding are active acts, not omissions.<sup>39</sup>

Fourth reading of “by itself”: When I say that something happens by itself (so somebody might try to explain their usage to us), I have in mind a certain *list of actions*. The way I use these words, something that happens does so “by itself” if it is true of the event that it would happen even if nobody performed an action from the list I have in mind.

Now everything will depend on the list. Say that “fertilizing” is on it, but that things like “eating while you’re pregnant” or “feeding after birth” are not. Then Lenzen will be all right. *On the one hand*, a fertilized egg that turns into a happy person would do so by itself. Fertilization is on the list, but fertilization is not necessary (not for the step in question, that is); mother’s eating while she’s pregnant, feeding after birth, and other such manoeuvres, are necessary for the fertilized egg to acquire a happy life but are not on the list. *On the other hand*, unfertilized eggs that turn into happy persons never do so by themselves. That would require fertilization, and fertilization is on the list.

Fine – but what is the *right* list? Any list will, in Lenzen’s framework, be equivalent to a particular moral creed; it will not be, no matter by how charitable a standard, a general principle from which the particular creed can be derived. Some lists of actions would generate Lenzen’s results, lots of other lists would not. To say that we have obligations to refrain from certain ac-

<sup>39</sup> Nothing changes if we replace mother with a machine. Building and running the machine would require active acts as well – in real life, at any rate, and the existence of science-fiction scenarios in which fetuses grow up by themselves is of no help to Lenzen’s position. Firstly, science-fiction scenarios only show that it is *logically possible* for a fetus to grow up by itself; but Lenzen thinks that in many *real-life* cases they grow up by themselves and that thus it would in many *real-life* cases be wrong to kill them. Secondly, we can also tell science-fiction stories in which the pair of sperm and unfertilized egg grows into a person by itself; hence, if science-fiction stories counted, fertilization would again fail to make the difference.

tions (which would prevent fertilized eggs from turning into happy persons), but that we have no obligation to refrain from certain other actions (which would prevent unfertilized eggs from turning into happy persons) is not giving a *reason* why abortion is wrong and non-conception is not; it's just a paraphrase.

Hence the fourth interpretation turns the expression "by itself" into a wild card; insert it into a moral theory that grants authority to what happens (or would happen) by itself and you can get practically any result you like, just by tailoring the relevant list of actions to suit your tastes. The exciting question is where the list comes from. The answer would require a more *general* notion of 'by itself', and that, in turn, throws us back on the other readings. They, however, have already failed us.

Further readings of "by itself"? What other *type* of explications of "by itself" could there be? I have no idea. I'm sorry if I've missed the decisive one. But I've done my best, and it would really be Lenzen's job to say a word or two on the concept that is supposed to bear so much weight in his theory. Moreover, I don't see where additional types of candidates could come from. There seem to be just two sources that could provide us with the basic material for an explication of "by itself": on the one hand, statistics and its cousins (standard or normal conditions, standard actions, causality, similarity, etc.); on the other hand, the difference between acts and omissions. The exercises we have been through make it reasonably clear that this material won't get us anywhere near an explication of "by itself" that will make the notion entail the difference that Lenzen assumes it entails.

#### *External Intervention, Inner Nature, Biological Constitution*

Have I been paying a malicious amount of attention to the innocent words "by itself" that appear just once in Lenzen's paper? I don't think so. The words aren't innocent, Lenzen's theory rests on them. True, he gives a few other brief hints that use different words. But it is easy to show that they are in the same boat and would succumb to the same treatment.

One example is Byrne's claim (quoted approvingly by Lenzen) that "the possibility that an ovum will become a person depends upon external intervention". I just don't get it: so does, as most parents will have noticed, the possibility that a *fertilized* egg turns into a person!<sup>40</sup>

<sup>40</sup> Cf. our above discussion of reading no. one of "by itself". What I'm saying here is true if the word "person" is used to imply that the entity will indeed experience happiness; but the word has to be used in this way if Byrne's quote is to support Lenzen's theory.



Of course, if we simply *define* “intervention” so that fertilization counts as an intervention, whereas the many active acts (not just omissions) necessary for the step from pregnancy to birth (let alone from pregnancy to, say, kindergarten) do not count as interventions, then we obtain the desired result. But this is, in another guise, the list approach we encountered as reading no. four of “by itself” – a big moral issue is placed in the lap of a terminological caprice.

A second example is Byrne’s claim (again quoted approvingly by Lenzen) that fertilization changes “the inner nature and biological constitution” of the ovum; this is in line with Lenzen’s own remark that “an unfertilized ovum *qua* its nature is not capable of making experiences”<sup>41</sup>. These references to nature are – what else *could* they be? – yet another version of the claim that the road a fertilized egg has to travel to kindergarten is in a different statistical or causal class than the road a pair of sperm and unfertilized egg has to travel to kindergarten. But it is not. Certainly, we can point to *one* active act, viz. bringing about fusion, that is necessary in the one case and no longer necessary in the other. But – as long as *statistically* the odds are against even the *fertilized egg*,<sup>42</sup> and as long as myriads of other active acts are necessary even in order to pave *its* way into a happy life<sup>43</sup> – why should, of all the active acts, *that* one catapult an object into a different statistical or causal class? Simply to say that it does (“period!”), whereas the others don’t (“period!”), is, to say the least, somewhat arbitrary.<sup>44</sup>

### *Hence the Inference Fails*

Section 2.4 has been devoted to the inference from Anti-Death and Not-Pro-New-Life to Anti-Abortion and Not-Pro-Fertilization. The upshot, I think, is clear enough. We can choose any one of various closely related terminologies (be it that of what happens “by itself”, or “normally”, or that of “nature”, or “identity”), and we can *stipulate* meanings for these terms that will *generate* a difference between pairs of sperm and egg that have fused and those that have not (differences such as: after fusion the pair has a different “nature” or

<sup>41</sup> Lenzen (1998), sect. 4.

<sup>42</sup> Cf. Lenzen (1998), sect. 6.1, and above, the discussion of reading no. one of “by itself”.

<sup>43</sup> Cf. above, the discussion of reading of no. three of “by itself”.

<sup>44</sup> Helga Kuhse and Peter Singer tell a story in which the action that brings together egg and sperm in a laboratory is, by ordinary standards, tiny and undramatic (1992, p. 59). This reminder is helpful for our context, too. But it is true *anyway* that, in vitro or not, many of the acts required to turn a fetus into a happy life have, by ordinary standards, at least the causal calibre of those required for fertilization. I mention “ordinary” standards since no doubt standards could be *defined* that put fertilization in a different class; but for that sort of manoeuvre, see above, the discussion of reading no three of “by itself”.

“identity” than before; or after fusion, but not before, the pair will attain a happy life, if at all, then “by itself” or “normally”).

But this will only work if the stipulations are tailor-made for the moral results. *General* explications of the relevant concepts (in terms of statistics, causal distance, acts and omissions, genetic make-up, brain continuity, psychological connectedness, similarity, etc.) will, as we have seen, fail to yield the difference between egg and sperm before and after fusion. If somebody wants to *argue* for the difference, he cannot therefore, as Lenzen thinks he can, *invoke* these concepts. *No general criterion will do Lenzen the favour of singling out fertilization; he will be reduced to saying that fertilization just is his criterion.*

There is, Lenzen tells us elsewhere, “eine riesige *qualitative Differenz*” between the cells before and after they fuse, and this “begründet den gravierenden moralischen Unterschied” between abortion and non-conception.<sup>45</sup> But he has simply not told us what makes the tiny physical step, which we’re all aware of, a “huge qualitative difference”. Of course the step is *necessary* on the road to personhood or a happy future – but so are a thousand others. Does Lenzen not notice that, as it stands, the claim about the difference just *amounts* to the claim that abortion and non-conception have different normative status? And that, if it *amounts* to it, it can hardly *support* it? Lenzen also says (same place as just quoted) that the huge qualitative difference is usually called “the mystery of life”. I give him that much: it sure is a mystery.

### 2.5. Wolfgang Lenzen and the Midwives: Conclusion

What, then, should we think of Lenzen’s theory? No doubt it is as carefully crafted a version of midwifism as we can wish for – a version that clearly deserves, and rewards, close study. However, close study also reveals at least three fundamental flaws.

Firstly, there are good reasons to deny Anti-Death, in other words good reasons to believe that, where an individual has no ex-ante preference against death or its consequences, death cannot be bad for that person. These reasons carry over to the denial of Lenzen’s ban on abortion. For all this, see section 2.2.

Secondly, even if Anti-Death *were* tenable, the moral asymmetry claimed by the conjunction of Anti-Death and Not-Pro-New-Life is not. The mere question whether somebody is around anyway makes, other than the question whether he is around *and wants this or that to happen* no moral difference. Two arguments to that effect were presented in section 2.3.

<sup>45</sup> Lenzen (1995), p. 232.

Thirdly, even if, contrary to what “firstly” and “secondly” have just told us, Anti-Death and Not-Pro-New-Life *were* jointly tenable, Lenzen has not shown that the two of them object to abortion but tolerate non-conception – see section 2.4. Far too much identity theory is left unexplained and undefended, so that the central identity-theoretical claim – Wolfgang Lenzen’s answer to the familiar question “When does a life begin?” – is presented to us as sheer dogma. Plug the *identity-theoretical* dogma into a moral principle sensitive to identity, and you’ll get a *moral* dogma.

### 3. *Richard Hare and the Rabbits*

#### 3.1. *Hare’s Argument*

Richard Hare believes that individuals with satisfied preferences ought to exist – and thus that we have obligations to conceive and not to abort them.<sup>46</sup> In the words of the introductions to this paper and this symposium, he is a Rabbit.

Hare has put forward what is, as far as I know, the only genuine argument in support of the Rabbits’ claim. In a nutshell: to be moral is to have analogous preferences for analogous situations; some real-life people want to have been born; hence morality requires them to have, for analogous situations (in which *other* people’s birth is at issue) analogous preferences (i.e. preferences for those people’s birth).

To see what is going on, we need a more detailed version. Consider a situation that we will call *the past existential situation S*. In *S*, one of two possible worlds had to be brought about, either the *S*-birth-world or the *S*-non-birth-world. In the *S*-birth-world, an individual exists that has far more and stronger satisfied preferences than frustrated ones and that, ex post, rationally wants to have been born. In the *S*-non-birth-world, the individual does not exist and, therefore, wants nothing. In fact, the *S*-birth-world has become actual; the individual exists. Let us call her Mary. Situation *S* is shown in figure 1; in the figure, and henceforth, “ $>_x$ ” stands for “is preferred by *x* to”.

Let us pretend, for the sake of Hare’s argument, that the question of Mary’s existence or non-existence affects nobody else’s interests. Then all we have to take into account is *her* preference in the matter, her actual ex-post preference, that is, for the *S*-birth-world over the *S*-non-birth-world.

<sup>46</sup> Some qualifications are supposed to apply. For instance, if after an abortion another child comes into existence instead of the aborted one, and is happier than the latter would have been, then an abortion may be permissible – cf. Hare (1975), sect. 10.5.

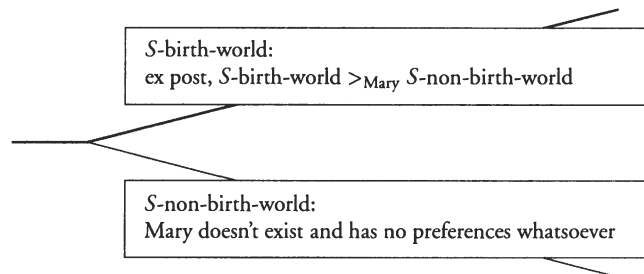


Fig. 1: The past existential situation  $S$

Hare's general theory, universal prescriptivism, says that morality requires us to have the same preferences with respect to all situations which are universal copies of each other.<sup>47</sup> Thus, if Mary has a preference for the birth-world over the non-birth world *with respect to situation  $S$* , she must, on pain of immorality, have the analogous preference *with respect to every situation that is a universal copy of  $S$* .

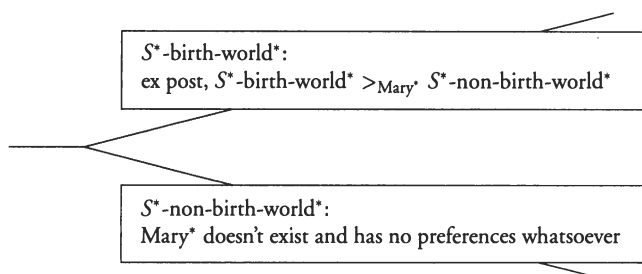
The next step is one from Mary's universal preference to a corresponding "ought"-judgement. What ought to be the case is, according to universal prescriptivism, what a universal preferrer wants to be the case. Hence, if Mary has, with respect to  $S$  and all its universal copies, a *preference* for the birth-worlds, we can say that, in  $S$  and all its universal copies, the birth-worlds *ought* to be brought about.

So that is the moral judgement about the past existential situation  $S$  and all its universal copies. Let us now turn to one such universal copy of  $S$ , the present existential situation  $S^*$ .  $S^*$  raises a moral problem right now. Somebody is currently wondering whether Mary\*, who plays in  $S^*$  exactly the part Mary played in  $S$ , ought to be born. (Throughout this paper, I will give names with asterisks to merely possible situations and the people in them. Names without asterisks refer to actual situations and to actual people.) We can represent  $S^*$  in the same way as  $S$  – see figure 2.

We already know that, in  $S$  and all its universal copies, the birth-worlds ought to be brought about. We also know that  $S^*$  is a universal copy of  $S$ . Hence, the birth-world\* of situation  $S^*$  ought to be brought about; Mary\* ought to be born.<sup>48</sup>

<sup>47</sup> A universal copy of a situation  $Y$  is a situation that resembles  $Y$  in all its universal properties; a universal property, in turn, is a property that can be specified without reference to specific individuals. Cf. Hare (1981), chs. 1.6 and 6.

<sup>48</sup> It might be objected to Hare's argument that it is never applicable to the world as it is because, in real life, a new existential situation that raises a moral problem will always differ in

Fig. 2: The present existential situation  $S^*$ 

Summing up the argument (with " $\gg_x$ " short for "preferred by individual  $x$  universally" and " $\gg_\heartsuit$ " for "better than"):

	With respect to past existential situation $S$ : $S$ -birth-world $>_{\text{Mary}}$ $S$ -non-birth-world
(Mary is a universal preferrer)	
$\Longrightarrow$	With respect to past existential situation $S$ and all its universal copies: their birth-world $\gg_{\text{Mary}}$ their non-birth-world
(universal prescriptivism)	
$\Longrightarrow$	With respect to past existential situation $S$ and all its universal copies: their birth-world $\gg_\heartsuit$ their non-birth-world
( $S^*$ is a universal copy of $S$ )	
$\Longrightarrow$	With respect to present existential situation $S^*$ : $S^*$ -birth-world* $\gg_\heartsuit$ $S^*$ -non-birth-world*

This is Hare's argument for an obligation to procreate.<sup>49</sup> It is ingenious. I will call it the *Rabbits' argument*. Its clue is the innocence of its premiss (of its first

*some* universal property from any past existential situation; we won't find an existing person whose life would be universally identical to that of the person whose creation we are considering. Hare has foreseen this objection and has added an argument to the effect that it is legitimate to replace the requirement of universal identity by the weaker requirement of relevant similarity. Let us grant that Hare can indeed meet the objection and just limit ourselves to asking whether we *would* have an obligation to create a new person whose life *would* be universally identical to that of the existing person who is happy to be alive.

<sup>49</sup> Cf. Hare (1975), (1988a), (1988b), and (1998). My presentation deviates from Hare's to some extent, but not significantly. Hare has agreed that my reconstruction captures his argument (personal communication).

line, that is). If we buy the argument, then whether we ought to bring an individual into existence can be deduced, in the end, from what ex post an *actual* person rationally wants to have happened. Compared to other questions procreation is sometimes said to confront us with (such as: what are the desires of merely possible people? How bad is it not to exist? And can we harm people who do not exist?), statements on actual rational ex-post preferences are rather unsuspecting. There's hope that, if we could deduce the morality of procreation from such statements alone, we could solve a host of major problems by *comparatively* unproblematic means.

But, I will ask, is Hare's *method* sound? Can, in general, actual rational ex-post preferences play the part Hare wants them to play? In section 3.2, I will show that, as it stands, the answer is no. If we consult actual rational ex-post preferences the way Hare suggests, the advice we get can be inconsistent. There are two ways out of the inconsistency, which will be discussed in sections 3.3 and 3.4 respectively. Both of them, however, would leave the Rabbits' argument with a large hole. Roughly speaking, in order to yield the *conclusion that people with satisfied preferences ought to exist*, the argument will need the additional *premiss that satisfied preferences ought to exist*.

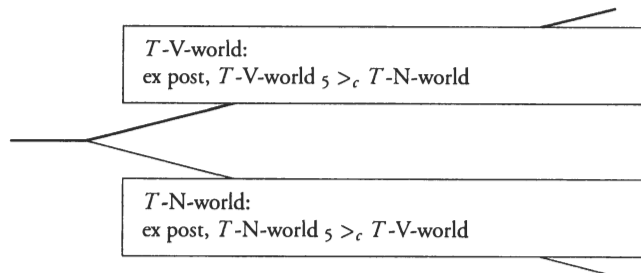
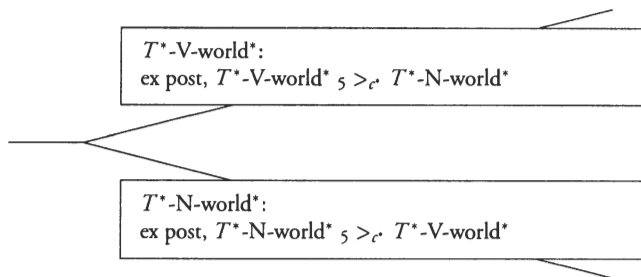
### 3.2. *A Self-Contradiction: The Vermeer-Novalis Case*

The situations  $S$  and  $S^*$  are cases of what we can call *different preferences choices* – choices which affect the number or identity of preferences.<sup>50</sup> Our question is: suppose somebody has a rational ex-post preference with respect to a different preferences choice that concerned himself; can, in general, that preference be a reliable guide for moral judgements on universally identical choices?<sup>51</sup>

Consider the following thought experiment featuring what we will call *Vermeer-Novalis situations* (for short, VN situations). Let us start with the past VN situation  $T$ . In  $T$ , it was an open question which of two possible worlds would be brought about, either the world in which an individual  $c$  turns into an expert on Vermeer and has, ex post, a rational preference of strength 5 for having become an expert on Vermeer (=: the  $T$ -V-world), or the world in which  $c$  turns into an expert on Novalis and has, ex post, a rational preference

<sup>50</sup> For the term "different preferences choice", cf. the beginning of Fehige/Wessels (1998); it is a modification of Parfit's term "different people choice" from *Reasons and Persons* (1984, sect. 120).

<sup>51</sup> The objection that Hare doesn't say that *in general* they can will be dealt with at the end of this subsection.

Fig. 3: The past VN situation  $T$ Fig. 4: The present VN situation  $T^*$ 

of strength 5 for having become an expert on Novalis (=: the  $T$ -N-world).<sup>52</sup> In fact, the  $T$ -V-world became actual;  $c$  became an expert on Vermeer. We assume that, before the decision,  $c$  had no preferences either way. If we write “ $n >_x$ ” for “is preferred by  $x$  with strength  $n$ ”, we can represent the case as usual – see figure 3.

Same procedure as before, with situations  $S$  and  $S^*$ . By adding asterisks across the board, we baptize a universal copy of  $T$ . Say that the copy,  $T^*$ , raises a moral problem right now: somebody is wondering whether  $c^*$ , who plays exactly the part in  $T^*$  that  $c$  played in  $T$ , ought to become an expert either on Vermeer or on Novalis – see figure 4.

Let us now apply Hare’s inference pattern:

With respect to past VN situation  $T$ :  
 $T$ -V-world  $>_c$   $T$ -N-world

<sup>52</sup> What is the strength of a preference? I’m not sure, but let us assume (as Hare’s system in general and some of his applications of the Rabbits’ argument in particular have to) that some such notion is available.

( $c$ is a universal preferrer)	⇒	With respect to past VN situation $T$ and all its universal copies: their V-world $\succ_c$ their N-world
(universal prescriptivism)	⇒	With respect to past VN situation $T$ and all its universal copies: their V-world $\succ_{\heartsuit}$ their N-world
( $T^*$ is a universal copy of $T$ )	⇒	With respect to present VN situation $T^*$ : $T^*$ -V-world* $\succ_{\heartsuit}$ $T^*$ -N-world*

But suppose we find another past situation,  $T'$ , with a hero  $c'$ , that differs from  $T$  just in the following respect: in  $T'$ , the  $T'$ -N-world', not the  $T'$ -V-world', has become actual – see figure 5.

Let us again apply Hare's inference pattern to the question what ought to happen in  $T^*$ , this time with the actual ex-post preference from  $T'$  as the argument's starting point.

		With respect to past VN situation $T'$ : $T'$ -N-world' $\succ_{c'}$ $T'$ -V-world'
( $c'$ is a universal preferrer)	⇒	With respect to past VN situation $T'$ and all its universal copies: their N-world $\succ_{c'}$ their V-world
(universal prescriptivism)	⇒	With respect to past VN situation $T'$ and all its universal copies: their N-world $\succ_{\heartsuit}$ their V-world
( $T^*$ is a universal copy of $T'$ )	⇒	With respect to present VN situation $T^*$ : $T^*$ -N-world* $\succ_{\heartsuit}$ $T^*$ -V-world*

The conclusion contradicts the previous one and says that the  $T^*$ -N-world\* is better than the  $T^*$ -V-world\*. *Thus, Hare's method is, as it stands, inconsistent. What someone actually rationally and ex post wants in a different preferences choice that concerned himself is no reliable guide for moral judgements on universally identical situations – not, at any rate, without further constraints.*

Hare, it might be objected, has never said that he wants to use his method for different preferences choices in general – perhaps just for different *people* choices, i.e. choices that affect the number or the identity of people who will



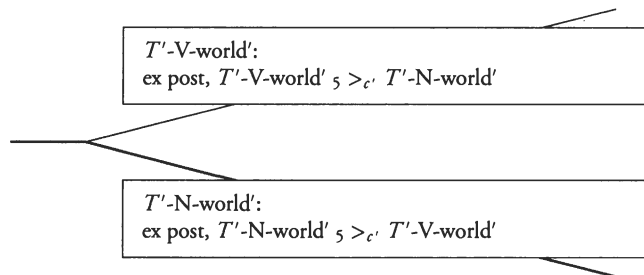


Fig. 5: The past VN situation  $T'$

ever exist.<sup>53</sup> But this move doesn't help. Firstly, Hare doesn't mention any such restriction. Secondly, consider the *spirit* of the Vermeer-Novalis case. The fact that ex-post preferences can *vary* (i.e. are not determined by the objects of choice) should make us ask how much moral authority we are willing to bestow on any one of them. Thirdly, and most importantly, we could have presented a Vermeer-Novalis case that would have been a different people choice – one in which either the person who would turn into an expert on Vermeer or *another* person who would turn into an expert on Novalis could have been brought into existence. Thus, the self-contradiction arises even in the realm of different people choices.

### 3.3. *The First Escape Route: Strengthening the Concept of a Rational Ex-Post Preference*

Can we save Hare's method from the charge of inconsistency? If we want to try to do so, we have two basic types of options. One is to strengthen the concept of what I call a *morally decisive* ex-post preference; I will pursue this route in section 3.4. The other option is to strengthen the concept of a *rational* ex-post preference – which I will try out now.

The question is: how can we get rid of the deontic contradiction in the VN case *and* retain the idea that the moral judgement about situation  $T^*$  is to be based on just what  $c$  and  $c'$  rationally wanted to have happened in  $T$  and  $T'$ ? One could argue as follows: in the VN case, something is wrong with the preferential input. The contradiction can only arise because the same person's

<sup>53</sup> For the term "different people choice", cf. note 50 above. A mutation of the objection would suggest restricting the scope of the method still further, viz. to different *number* choices. These are choices that make a difference to the number of people who will ever exist – cf., again, Parfit (1984), sect. 120. But that cannot be what Hare has in mind, since he himself applies the method to a choice that is *not* a different number choice; cf., e.g., sect. 10.5 of his (1975).

ex-post preferences differ from one world to the other. Ex-post preferences for which this holds true are perhaps *compatible* with rationality in some sense, but in order for Hare's method to work the ex-post preferences it employs have to be *dictated* by rationality.<sup>54</sup> If they are, they will no longer vary from world to world.

But how could rationality dictate them? Just one option comes to mind: We have to look at the welfare, i.e. the amount of satisfaction. We would have to claim that the following requirement is part of the meaning of "rational ex-post preference":

Metamax: If you are not better off with one combination of preferences and satisfaction than you are with another, then you cannot rationally prefer combination one to combination two.<sup>55</sup>

To see Metamax in action, look at the VN situations again. If Metamax were true, our heroes *c* and *c'* could not without being irrational (irrational by Metamax standards) have different ex-post preferences in different worlds. If they were better off in one world than in the other they would have to prefer the same world in both of them; if they were equally well off in both worlds they would have to be indifferent in each of them; and *tertium non datur*. Thus, there could not be the different metapreferences that, used as an input for Hare's method, generated the contradictory judgements in the VN case.

So this is where we are: Metamax is a means to restore Hare's method to consistency.<sup>56</sup> However, there are at least four problems with Metamax itself or with the part Metamax plays in what we can call the *Metamax Version* of Hare's method. I will be brief with problems (1)–(3), since my argument does not rest on them. But problem (4) seriously affects the Rabbits' argument, and I will discuss it in detail.

Problem (1): Metamax is at least questionable. It is simply not obvious why you should always want to want what will give you most satisfaction. Just think of various types of brainwashings which most of us would decline to undergo

<sup>54</sup> Cf. the second terminological remark in the introduction to this paper.

<sup>55</sup> The name "Metamax" echoes what the principle is up to: in some sense, rational *metapreferences* are supposed to serve the *maximum* satisfaction of the lower-order desires.

<sup>56</sup> And indeed it seems as if Hare supports Metamax. In his book *Moral Thinking*, he advocates a so-called requirement of prudence. According to this requirement, "we should always have a dominant overriding preference now that the satisfaction of our now-for-now and then-for-then preferences should be maximized" (Hare 1981, sect. 5.6). Clearly, in choices that affect the number or the identity of preferences, the requirement of prudence amounts to Metamax.

but that Metamax says cannot be rationally declined, precisely because they would give us more satisfaction by changing our preferences.<sup>57</sup>

Problem (2): Metamax makes Hare's method circular. Metamax says about any two combinations of preferences and satisfaction that, if  $x$  is not better off with combination one than with combination two, then  $x$  cannot rationally prefer combination one to combination two. In order to be able to claim that  $x$ 's preference for one combination over another can be "rational" in *this* sense, we need to know *first* that  $x$  would be better off with the combination. But if "makes  $x$  better off" means the same as "ought to be the case as far as  $x$  is concerned", then the question which combination of preferences and satisfaction makes  $x$  better off is *synonymous* with that to which the method was intended to provide an answer: what combination of preferences and satisfaction ought, for the sake of individual  $x$ , to be actualized?

Problem (3): With Metamax, Hare's method can no longer be applied to the existential situation  $S$ . In the  $S$ -non-birth-world, Mary does not exist; hence there is nobody whose well-being can be taken into account.

But let us pretend, for the sake of Hare's argument, that problems (1)–(3) can be solved. This leaves us with problem (4): whether somebody's preference for one combination of preferences and satisfaction over another is rational depends, according to Metamax, on whether it makes the person concerned better off. Thus, Metamax leaves us with the question how, in general, we can determine which combination of preferences and satisfaction is better for the preferer; and how, with respect to the existential situation  $S$  in particular, we can determine in which of the two possible worlds Mary's welfare level is higher – in the world in which she has many preferences and, as we assume, a large amount of satisfaction and just a tiny amount of frustration, or in the world in which she has no preferences, no satisfaction, and no frustration?<sup>58</sup>

Within desire-based ethics, there's just one option, and Hare would concur. We have to look at Mary's desires. Let's do so, then. As is shown in figure 6, Mary's possible desires can, depending on their actual status in the  $S$ -birth-world, be grouped into three non-empty and disjunct sets: the set of desires

<sup>57</sup> See, e.g., Nozick (1974), p. 44, Bricker (1980), sect. III, Rawls (1982), pp. 173 ff., Barry (1989), pp. 278 ff., and Goodin (1991); cf. also Williams (1973), pp. 110 ff., Wiggins (1976), sect. 4 ff. (particularly p. 105), and Railton (1984).

<sup>58</sup> Does an inquiry into the welfare of somebody who does not exist make sense? I'm not sure – see the third problem that Metamax raises. But I'm trying out an interpretation of Hare's method that has to *assume* it makes sense. I suggest that, for the sake of his argument, we grant the assumption.

Mary's desires ...

... in the <i>S</i> -birth-world	... in the <i>S</i> -non-birth-world
non-existent	n o n- e x- i-
frustrated	s t-
satisfied	e n t

Fig. 6: Mary's desires

Mary does not have, the set of desires that Mary has and that are frustrated, and the set of desires that Mary has and that are satisfied.

Now, with respect to the first group, the desires that don't exist in the *S*-birth-world, Mary is equally well off in the *S*-birth-world as in the *S*-non-birth-world. (A somewhat unusual way of putting it – but see note 58 above.) With respect to the second group, the desires frustrated in the *S*-birth-world, Mary is worse off in the *S*-birth-world than in the *S*-non-birth-world, for, *ceteris paribus*, a frustrated desire lowers one's welfare.

At this point, everything depends on the third group. So far, the *S*-non-birth-world scores better than the *S*-birth-world. Only if we gave positive weight to the acquisition of satisfied desires would we be able to reverse this judgement. Only if we subscribed to something like an

Orexigenic Axiom: *Ceteris paribus*, it is good for the preferer to have new and satisfied preferences.<sup>59</sup>

– only then could we recommend the *S*-birth-world in spite of the frustrated desires it contains.

Thus the judgement which we pass on the existential situation *S* depends on the axioms of our welfare theory. *Only if* it has an Orexigenic Axiom – an axiom, that is, according to which additional satisfied preferences are a good thing – can we say that Mary is better off with the *S*-birth-world. If our welfare theory has no Orexigenic Axiom we will think that Mary would have been

<sup>59</sup> For more on "orexigenecism", the word and the concept, cf. Fehige (1998), sect. 1.

'better off' with the *S*-non-birth-world – because in the birth-world, she has, *ex hypothesi*, more than zero frustrated preferences.<sup>60</sup>

Two remarks. One is that the diagnosis should not surprise us. Let me put it this way: the Rabbits' argument says that whether satisfied preferences ought to exist depends on whether they are welcomed by rational ex-post preferences. But then the same question arises for the satisfied rational ex-post preferences: why ought *they* to exist? We can answer this if we presuppose orexigenecism. But if we have to presuppose orexigenecism anyway we could have done so right from the beginning. We could have spared ourselves the detour via the ex-post preferences and said straightaway that satisfied preferences ought to exist.

Secondly, does it follow from my argument that one cannot rationally want to have been born? The answer depends on our notion of rationality. If we subscribe to Metamax, but not to orexigenecism, then we *cannot* rationally want to have been born. For, according to non-orexigenecism, welfare is lowered by every frustrated preference, but not raised by a satisfied one. Thus a life will never be good for the 'liver', and could therefore (according to Metamax) not be rationally preferred to the absence of life. But suppose we do not subscribe to Metamax. Then we can, even if we are non-orexigenecists, rationally want to have been born. Our rational ex-post preferences would not be dictated by our lower-order desires and their satisfaction; our ex-post preferences could rationally welcome even a combination of preferences and satisfaction that comes up with much more frustration than an ordinary life does. Life could be rationally preferred to no life. The desire to have been born would not have to be irrational.

So much for the fourth problem that Metamax has to face, and, more generally, for my first attempt to restore Hare's method to consistency. To sum up: we can get rid of the deontic contradiction in the VN case by blaming, and then taming, the preferential input. Let us make sure that the rational ex-post preferences cannot start any trouble by pointing in different directions; let us,

<sup>60</sup> You might think that, in the discussion of the fourth problem with Metamax, I should have talked about rationality, not about welfare. But note, firstly, that Metamax requires us to look at our hero's welfare. (The problem that this might be inappropriate where the hero doesn't exist has been discussed in note 58 above.) And secondly, if you prefer the rationality to the welfare jargon you are welcome. The question will then be, not whether Mary is better off with the *S*-birth-world than with the *S*-non-birth-world, but whether Mary can, without being irrational, *prefer* the *S*-birth-world to the *S*-non-birth-world. My argument would survive just as well. It would lead to the question whether *rationality* is orexigenic; and with the statement that, if it is not, then Mary cannot *rationally prefer* the *S*-birth-world to the *S*-non-birth-world.

by means of Metamax, strengthen our concept of a rational ex-post preference. Now, according to Metamax, an ex-post preference for one combination of preferences over another is rational only if the preferer is better off with the preferred combination. Thus, in order to be well-defined, the Metamax Version of Hare's method needs, as an *independent* measuring rod for the rationality of the ex-post preferences, a notion of welfare. But there are various notions of welfare, and whether the Rabbits' argument goes through depends on which notion we choose. We don't reach the conclusion that Mary's desire for birth is rational in the sense of Metamax, and that hence a happy preferer ought to be born – unless we *presuppose* that welfare is orexigenic. In other words, *if Hare wants to save his method by Metamax and wants to uphold the Rabbits' argument, then he has to find an extra argument for orexigenecism; without it, the Rabbits' argument is incomplete.*<sup>61</sup>

#### 3.4. *The Second Escape Route: Strengthening the Concept of a Morally Decisive Ex-Post Preference*

So far I have tried to get rid of the inconsistency diagnosed in section 3.2 without abandoning the idea that the moral evaluation of different preferences choices can be based solely on what a person actually, rationally, and ex post wants to have happened in a universally identical choice situation. The Rabbits' argument ran into serious difficulties that way, and I will now try another route.

It's the only other route I see that is in line with Hare's method. Let us, runs the new suggestion, base our moral judgement on an individual's ex-post preferences from *all* the possible worlds in the situation at issue; more precisely, let us *aggregate* these different ex-post preferences of one and the same individual, and then the ex-post preference that emerges from aggregation – I will term it the *morally decisive* or, for short, the *m-(ex-post) preference* – says what ought to be done (as far as the individual in question is concerned). Let us call this procedure the *Aggregational Version* of Hare's method.

The general idea behind the Aggregational Version can be described as follows: in different preferences choices, we have to regard the sets of preferences, rather than the individuals, as the "subjects of interests". Which set of preferences ought to be actualized depends on how well off each set would be if it were actualized and on how the well-being of each set is weighed against the well-being of every other set. (Talking about the well-being of sets of preferences is unusual; we normally speak of the well-being of individuals. But for

<sup>61</sup> Note that I'm not assuming here that orexigenecism is wrong. All I'm saying is that *Hare* has to assume, and does not *show*, that orexigenecism is right.

different preferences choices, where sets of preferences can differ within one and the same individual, the jargon that allows us to ascribe welfare to the *sets* can be convenient.)

In order to see how the Aggregational Version of Hare's method works let us look at the VN case from section 3.2 again. To keep things in line with Hare's system (and to avoid yet more complications), I will assume that we are utilitarians.<sup>62</sup> In situation  $T$ ,  $c$ 's ex-post preference from the  $T$ -V-world has to be weighed against his ex-post preference from the  $T$ -N-world; and ditto in situation  $T'$ :  $c'$ 's ex-post preference from the  $T'$ -N-world' has to be weighed against his ex-post preference from the  $T'$ -V-world'. Since, within each situation, the heroes' preferences are equally strong in both worlds, we should describe  $c$  and  $c'$  as being, in the morally decisive sense, indifferent, i.e. *m*-indifferent, between their V-world and their N-world; hence, the V-worlds are as good as the N-worlds for our heroes and hence, in situation  $T^*$ , the  $T^*$ -V-world\* is as good as the  $T^*$ -N-world\*. Thus, the Aggregational Version of Hare's method avoids the deontic contradiction.

#### *Transworld Interferences: The Vermeer-Novalis-Chopin Case*

By means of *intrapersonal* aggregation we can restore Hare's method to consistency. This is what the previous section has shown. But we should be careful. There are at least two problems with the Aggregational Version of Hare's method. Firstly, it leaves the Rabbits' argument with a serious hole, just as the Metamax Version does. This will be shown in the next subsection. Secondly – and this is the topic of the present subsection –, it raises the problem of *transworld interferences*.

Consider a modification of the VN situation  $T$ , the Vermeer-Novalis-Chopin situation  $U$  (for short, the VNC situation  $U$ ). In  $U$ , one of three possible worlds had to be brought about, either the world in which an individual  $d$  is an expert on Vermeer ( $U$ -V-world), or the world in which  $d$  is an expert on Novalis ( $U$ -N-world), or the world in which  $d$  is an expert on Chopin ( $U$ -C-world). In the  $U$ -V-world,  $d$ , ex post, prefers the  $U$ -V-world to the  $U$ -N-world with strength 1 and to the  $U$ -C-world with strength 2; in the  $U$ -N-world and the  $U$ -C-world,  $d$ , ex post, is indifferent between these two worlds, but he prefers both of them to the  $U$ -V-world with strength 4. Again, we assume that, before the decision,  $d$  had no preferences for one option over another. The VNC situation  $U$  is shown in figure 7.

<sup>62</sup> The remarks from this section apply, *mutatis mutandis*, to practically all social welfare functions.

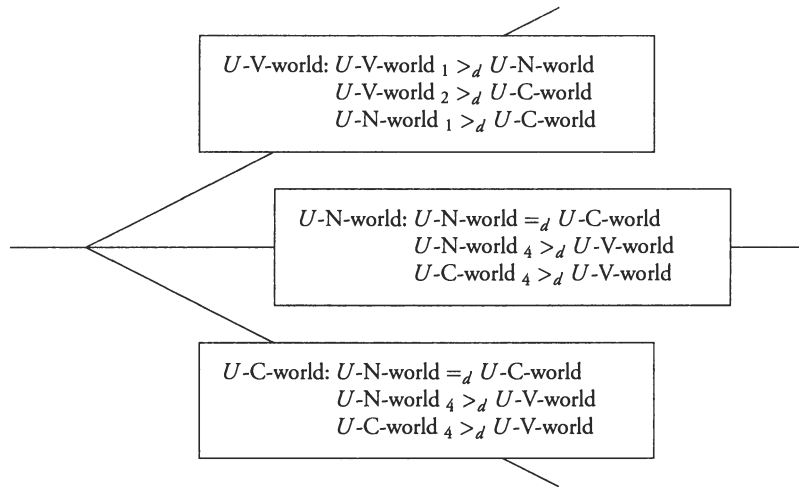


Fig. 7: The VNC situation  $U$

We now aggregate  $d$ 's ex-post preferences from all three worlds. I will continue to assume that we are utilitarians; I will write " $p_n >^m_x q$ " for " $x$   $m$ -prefers  $p$  to  $q$  with strength  $n$ "; " $p_{-n} >_x q$ " is equivalent to " $q_n >_x p$ " and " $p_0 >_x q$ " to " $q =_x p$ ". The aggregations look like this:

$U\text{-V-world:}$	$U\text{-V-world}$	$1 >_d$	$U\text{-N-world}$
$U\text{-N-world:}$	$U\text{-V-world}$	$-4 >_d$	$U\text{-N-world}$
$U\text{-C-world:}$	$U\text{-V-world}$	$-4 >_d$	$U\text{-N-world}$
morally decisive preference:	$U\text{-N-world}$	$7 >^m_d$	$U\text{-V-world}$
$U\text{-V-world:}$	$U\text{-V-world}$	$2 >_d$	$U\text{-C-world}$
$U\text{-N-world:}$	$U\text{-V-world}$	$-4 >_d$	$U\text{-C-world}$
$U\text{-C-world:}$	$U\text{-V-world}$	$-4 >_d$	$U\text{-C-world}$
morally decisive preference:	$U\text{-N-world}$	$6 >^m_d$	$U\text{-V-world}$
$U\text{-V-world:}$	$U\text{-N-world}$	$1 >_d$	$U\text{-C-world}$
$U\text{-N-world:}$	$U\text{-N-world}$	$0 >_d$	$U\text{-C-world}$
$U\text{-C-world:}$	$U\text{-N-world}$	$0 >_d$	$U\text{-C-world}$
morally decisive preference:	$U\text{-N-world}$	$1 >^m_d$	$U\text{-C-world}$

Applied to situation  $U$ , the Aggregational Version of Hare's method now yields the following evaluation: since  $d$   $m$ -prefers the  $U\text{-N-world}$  to the  $U\text{-V-world}$  with strength 7, the  $U\text{-N-world}$  to the  $U\text{-C-world}$  with strength 1,



and the *U-C*-world to the *U-V*-world with strength 6, the *U-N*-world is the best world, the *U-C*-world is second best, and the *U-V*-world is worst. But note that although in neither the *U-N*-world nor the *U-C*-world does *d* prefer one of these worlds to the other, the *U-C*-world is now said to be worse for him. *The comparative evaluation of two possible worlds would depend on preferences that nobody has in either of them.* Clearly, this is absurd.

### *The Safe Aggregational Version of Hare's Method*

As it stands, the Aggregational Version of Hare's method suffers from transworld interferences and is therefore inadequate. We are not better off with it than with the original version. This is the sad result of the previous subsection.

However, the problem of transworld interferences, too, can be solved. We just have to call upon our heroes to evaluate, in each world, only the world in which they live. In the VN situation *T*, for instance, *c* would not have to ask himself how much he prefers, in the *T-V*-world, the *T-V*-world to the *T-N*-world and how much he prefers, in the *T-N*-world, the *T-N*-world to the *T-V*-world; in each world, he would just have to ask himself how much he likes living in *that* world. He would just have to form what we can call *parochial* ex-post preferences.

So this is where we are: if we base our moral judgements about different preferences choices on *all* the ex-post preferences another person has in any of the possible worlds from a universally identical situation, then our moral judgements will not be inconsistent; and if, in addition, we require the ex-post preferences to be *parochial*, then our moral judgements will not be corrupted by transworld interferences either. We can call a repair of Hare's method that complies with both these requirements the *Safe Aggregational Version*.

Let us now return to the existential situation *S*. What happens if we apply the Safe Aggregational Version to it? Mary's story could go like this: in the *S*-birth-world, Mary wants the *S*-birth-world to have been brought about; in the *S*-non-birth-world, Mary does not exist and, therefore, wants nothing. We count the non-existence of a preference as indifference (sounds like the natural thing to do, doesn't it?). If we then weigh Mary's preference for the *S*-birth-world against her indifference towards the *S*-non-birth-world we get the conclusion in favour of procreation. Since Mary *m*-wants the *S*-birth-world to have been brought about it ought to be the case that the *S*-birth-world has been brought about. Thus, it seems as if the Safe Aggregational Version of Hare's method leaves the Rabbits' argument intact.

This time, however, there's another piece of bad news for Hare. By counting the absence of preferences (which resulted from the very absence of the

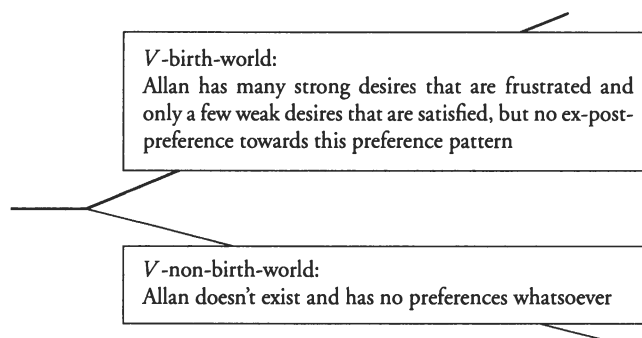


Fig. 8: The past existential situation *V*

preferrer) as indifference, the Rabbits' argument has now presupposed something like the

Indifference Principle: If a preferrer is unable to grasp a proposition *p* (as is certainly true of preferrers that do not yet exist), then morality ought to count this as indifference towards *p*.

The Indifference Principle is wrong. To see this, consider another past existential situation, the existential situation *V*, which is shown in figure 8. In *V*, one of two possible worlds had to be brought about. There is, on the one hand, the world in which Allan exists and is very badly off. He has far more and stronger frustrated preferences than satisfied ones. But Allan has no ex-post preference towards the cruel world; that is so because he can represent to himself comparatively simple states of affairs – such as being healthy, going for a walk, having friends, or eating ice-cream –, but does not have the mental capacity to grasp propositions like the one that Allan has lots of desires that are frustrated. Let us call this world the *V*-birth-world. On the other hand, there is the world in which Allan does not exist and, therefore, wants nothing (=: the *V*-non-birth-world).

The moral upshot is that, since Allan, in the *V*-birth-world, has no preference towards the *V*-birth-world and, therefore, would, according to the Indifference Principle, have to count as being indifferent towards the *V*-birth-world, there would have been, in the end, no reason not to bring Allan into existence – even though he leads a life of far more and stronger frustrated desires than satisfied ones. This is obviously absurd – and obviously something that Hare himself wouldn't want to say either.<sup>63</sup>

<sup>63</sup> The Indifference Principle, it might be objected, says what ought to be the case if a preferrer

What does Allan's case show? It shows that the Indifference Principle fails because it does not take into account that frustration is bad (even where the preferer cannot pass a judgement on it); and, more generally, because it is not sufficiently sensitive to matters settled by the very concept of welfare (or, if you like, by pure rationality – see note 60 above). No matter what preference we ascribe to someone who has no preferences with respect to his existence, and no matter how much weight we attach to it – the ascription and the weight have to be checked by an independent authority, and the independent authority is the preferer's well-being. Only if the ascription and the weight are in accord with the judgements that our notion of welfare comes up with independently, they are morally acceptable.<sup>64</sup>

For Mary's case, this means that both the ascription of a preference in the *S*-non-birth-world and its aggregation with Mary's preference from the *S*-birth-world have to be compatible with her well-being. Hence, the question arises once again (as it did in the discussion of Metamax, towards the end of sect. 3.3): how can we determine how well off Mary is? The answer depends on our notion of welfare. If it is orexigenic, then we have to say that Mary is well off in the *S*-birth world and badly off in the *S*-non-birth-world (so to speak – see note 58 above). This is so because, according to orexigenecism, satisfied preferences are good for the preferer, and she has many of them in the *S*-birth-world, but none in the *S*-non-birth-world. Thus, we would have to ascribe to Mary, in the *S*-non-birth-world, a strong preference against the *S*-non-birth-world, and we would have to make sure that an aggregation yields the conclusion that Mary *m*-wants that the *S*-birth-world has been brought about. But if welfare is *not* orexigenic, then Mary can be said to be, in the *S*-birth-world, fine, but, in the *S*-non-birth-world, very well off, even unsurpassably well. This is so because, according to non-orexigenecism, nothing can be better than the absence of preferences. Thus, we would have to ascribe to Mary, in the *S*-non-birth-world, a very strong preference for the *S*-non-birth-world, and we should make sure that her preference for the *S*-birth-world is overruled by it when we concoct her *m*-preference. In other words, if welfare is not orexigenic the Safe Aggregational Version of Hare's

---

cannot represent to himself a single *proposition p*; but the problem in Allan's case is that the hero is unable to grasp a complete *possible world*; hence, the Indifference Principle cannot be applied to the case directly. Reply: It doesn't affect the spirit of the Indifference Principle whether it is stated in terms of propositions or possible worlds. If a preferer cannot represent to himself a proposition *p*, then he certainly cannot represent to himself a possible world in which proposition *p* is true. The only reason why I chose the propositional jargon is that it is slightly less strange to talk about someone not being able to grasp a proposition than to talk about someone not being able to grasp a possible world.

<sup>64</sup> The discussion in section 2.2 of this paper bears some similarity to the present one.

method will *not* yield the conclusion that in the existential situation *S* Mary ought to have been born.<sup>65</sup>

Let me sum up section 3.4. The plan was to restore Hare's method to consistency by strengthening the concept of what I called a morally decisive ex-post preference. There are two ways of doing this. One is to go by the transworld ex-post preferences, and thus to arrive at the Aggregational Version of Hare's method. This version is prone to transworld interferences and inadequate for at least that reason.<sup>66</sup> The other way is to go by the parochial ex-post preferences. This leads to the Safe Aggregational Version of Hare's method – a version that is not subject to transworld interferences. But it still doesn't support the Rabbits' argument. In order to aggregate, Hare needs some preferential input for the non-existent preferrer – who, of course, has none. Her preferences have to be stipulated. But not any stipulation will do; nor will any aggregation; the stipulation and the aggregation have to be compatible with considerations of welfare – that was the moral of Allan's case. But there are various notions of welfare, and whether the Rabbits' argument goes through depends on which one we choose. Only if we *presuppose* that welfare is orexigenic shall we get the conclusion that a happy preferrer ought to be born. The diagnosis, then, is almost literally the same as that we reached at the end of section 3.3. *If Hare wants to replace the original version of his method by the Safe Aggregational Version and wants to uphold the Rabbits' argument, he has to find an extra argument for orexigenecism; without it, the Rabbits' argument is incomplete.*

### 3.5. Richard Hare and the Rabbits: Conclusion

Richard Hare has presented what he believes to be an argument for an obligation to procreate. I called it the Rabbits' argument. According to the Rabbits' argument, the morality of procreation can be deduced from premisses about what an actual person wants to have happened in a universally identical choice situation – see section 3.1.

The Vermeer-Novalis case from section 3.2 has shown that, as presented by Hare, the method which underlies the Rabbits' argument is inconsistent.

There are two ways out of the inconsistency. One is to strengthen the concept of a *rational* ex-post preference by requiring such a preference to

<sup>65</sup> Of course, Hare will say that his method was to *generate* the answer to the question whether welfare is orexigenic. I know this was the idea – but it doesn't work. Allan's case shows that welfare is needed for the method to work; hence the method cannot be used to *define*, or diagnose, welfare.

<sup>66</sup> I say "at least", since the version suffers also from the problem we identified, and are about to summarize, for the Safe Aggregational Version.

reflect the preferrer's well-being. This is what I did in section 3.3, by building Metamax into Hare's method. The second way, which was discussed in section 3.4, is to strengthen the concept of what I called a *morally decisive* ex-post preference by saying that, for different preferences choices, an individual's morally decisive ex-post preference has to be the *aggregate* of his ex-post preferences from all the possible worlds of this choice situation.

Both ways out of the inconsistency lead to the same result: the Rabbits' argument has a serious hole. In order to show that people with satisfied preferences ought to exist it needs, roughly speaking, the extra *premiss* that satisfied preferences ought to exist. Note that the premiss is not just large and controversial. It's so close to what the Rabbits' argument wanted to *show* that, even if we help ourselves to the premiss in order to save the argument, the argument will hardly be of use to us any longer.

Are there any other ways out of the inconsistency? The prospects are dim, I think. At any rate, the ball is back in the Rabbits' court.

#### 4. Conclusion

The conclusion of this paper is the conjunction of the conclusions from section 2 (see sect. 2.5) and from section 3 (see sect. 3.5). Wolfgang Lenzen's and Richard Hare's theories are in bad shape, and so are midwifism and rabbitism in general. To at least this extent, the alternative is likely to be true: both non-conception and abortion are morally all right.

#### References

- BARRY (1989). Brian Barry: "Utilitarianism and Preference Change", *Utilitas* 1 (1989).
- BOROWSKI (1976). E. J. Borowski: "Identity and Personal Identity", *Mind* 85 (1976).
- BRICKER (1980). Phillip Bricker: "Prudence", *The Journal of Philosophy* 77 (1980).
- FEHIGE (1998). Christoph Fehige: "A Pareto Principle for Possible People", this volume.
- FEHIGE/MEGGLE (1995). Christoph Fehige and Georg Meggle (eds.): *Zum moralischen Denken*, Frankfurt/M. 1995.
- FEHIGE/WESSELS (1998). Christoph Fehige and Ulla Wessels: "Possible Preferences", this volume.
- GOODIN (1991). Robert E. Goodin: "Actual Preferences, Actual People", *Utilitas* 3 (1991).

- HARE (1975). Richard M. Hare: "Abortion and the Golden Rule", in Hare (1993); article first publ. in 1975.
- HARE (1981). Richard M. Hare: *Moral Thinking*, Oxford 1981.
- HARE (1988A). Richard M. Hare: "A Kantian Approach to Abortion", in Hare (1993); article first publ. in 1988.
- HARE (1988B). Richard M. Hare: "Possible People", in Hare (1993); article first publ. in 1988.
- HARE (1993). Richard M. Hare: *Essays on Bioethics*, Oxford 1993.
- HARE (1998). Richard M. Hare: "Preferences of Possible People", this volume.
- HARRIS (1995). Henry Harris (ed.): *Identity*, Oxford 1995.
- VAN INWAGEN (1990). Peter van Inwagen: *Material Beings*, Ithaca 1990.
- JOHNSTON (1987). Mark Johnston: "Human Beings", *Journal of Philosophy* 84 (1987).
- KUHSE/SINGER (1992). Helga Kuhse and Peter Singer: "The Moral Status of the Embryo", in Walters/Singer (1992).
- LAFLEUR (1992). William R. Lafleur: *Liquid Life: Abortion and Buddhism in Japan*, Princeton 1992.
- LEIST (1990). Anton Leist: *Eine Frage des Lebens*, Frankfurt/M. 1990.
- LENZEN (1990). Wolfgang Lenzen: "Das Töten von Tieren und Föten", *Analyse und Kritik* 12 (1990).
- LENZEN (1991). Wolfgang Lenzen: "Wie schlimm ist es, tot zu sein? Moralphilosophische Reflexionen", in Ochsmann (1991).
- LENZEN (1995). Wolfgang Lenzen: "Hare über Abtreibung, Empfängnisverhütung und Zeugungspflicht", in Fehige/Meggle (1995).
- LENZEN (1998). Wolfgang Lenzen: "Who counts?", this volume.
- LOUX (1978). Michael J. Loux: *Substance and Attribute*, Dordrecht 1978.
- LYON (1980). Ardon Lyon: "On Remaining the Same Person", *Philosophy* 55 (1980).
- MCMAHAN (1998). Jeff McMahan: "Preferences, Death and the Ethics of Killing", this volume.
- NOZICK (1974). Robert Nozick: *Anarchy, State, and Utopia*, Oxford 1974.
- OCHSMANN (1991). Randolph Ochsmann (ed.): *Lebens-Ende*, Heidelberg 1991.
- PARFIT (1971). Derek Parfit: "Personal Identity", *Philosophical Review* 80 (1971).
- PARFIT (1984). Derek Parfit: *Reasons and Persons*, second ed., Oxford 1987; first ed. first publ. in 1984.
- PARFIT (1995). Derek Parfit: "The Unimportance of Identity", in Harris (1995).
- RAILTON (1984). Peter Railton: "Alienation, Consequentialism, and the Demands of Morality", *Philosophy and Public Affairs* 13 (1984).
- RAWLS (1982). John Rawls: "Social Unity and Primary Goods", in Sen/Williams (1982).
- SEN/WILLIAMS (1982). Amartya K. Sen and Bernard Williams (eds.): *Utilitarianism and Beyond*, Cambridge 1982.
- SINGER (1993). Peter Singer: *Practical Ethics*, second ed., Oxford 1993.
- SINGER (1995). Peter Singer: *Rethinking Life and Death*, Oxford 1995.
- SINGER (1998). Peter Singer: "Possible Preferences", this volume.

- SMART/WILLIAMS (1973). J. J. C. Smart and Bernard Williams: *Utilitarianism: For and Against*, Cambridge 1973.
- SUMNER (1981). L. W. Sumner: *Abortion and Moral Theory*, Princeton 1981.
- WALTERS/SINGER (1992). William A. W. Walters and Peter Singer (eds.): *Test Tube Babies*, Oxford 1992.
- WIGGINS (1976). David Wiggins: "Truth, Invention and the Meaning of Life", in Wiggins (1987); article first publ. in 1976.
- WIGGINS (1987). David Wiggins: *Needs, Values, Truth*, Oxford 1991; first ed. first publ. in 1987.
- WILLIAMS (1973). Bernard Williams: "A Critique of Utilitarianism", in Smart/Williams (1973).
- WITTGENSTEIN (1958). Ludwig Wittgenstein: *The Blue Book*, Oxford 1958.