

Melinda A. Roberts
David T. Wasserman
Editors

International Library of Ethics, Law, and the New Medicine 35

Harming Future Persons

*Ethics, Genetics and
the Nonidentity Problem*



Springer

Harming Future Persons

INTERNATIONAL LIBRARY OF ETHICS, LAW, AND THE NEW MEDICINE

Founding Editors

DAVID C. THOMASMA†

DAVID N. WEISSTUB, *Université de Montréal, Canada*

THOMASINE KIMBROUGH KUSHNER, *University of California,
Berkeley, U.S.A.*

Editor

DAVID N. WEISSTUB, *Université de Montréal, Canada*

Editorial Board

TERRY CARNEY, *University of Sydney, Australia*

MARCUS DÜWELL, *Utrecht University, Utrecht, the Netherlands*

SØREN HOLM, *University of Cardiff, Wales, United Kingdom*

GERRIT K. KIMSMA, *Vrije Universiteit, Amsterdam, the Netherlands*

DAVID NOVAK, *University of Toronto, Canada*

EDMUND D. PELLEGRINO, *Georgetown University, Washington D.C., U.S.A.*

DOM RENZO PEGORARO, *Fondazione Lanza and University of Padua, Italy*

DANIEL P. SULMASY, *Saint Vincent Catholic Medical Centers, New York, U.S.A.*

LAWRENCE TANCREDI, *New York University, New York, U.S.A.*

VOLUME 35

For further volumes:

<http://www.springer.com/series/6224>

Melinda A. Roberts · David T. Wasserman
Editors

Harming Future Persons

Ethics, Genetics and the Nonidentity Problem

 Springer

Editors

Melinda A. Roberts
Dept. of Philosophy and Religion
The College of New Jersey
Bliss Hall 112
Ewing, New Jersey 08628
USA
robertsm@tcnj.edu

David T. Wasserman
The Center for Ethics at Yeshiva University
Yeshiva University
2495 Amsterdam Avenue
New York City, New York 10033
USA
dwasserm@yu.edu

ISSN 1567-8008

ISBN 978-1-4020-5696-3

e-ISBN 978-1-4020-5697-0

DOI 10.1007/978-1-4020-5697-0

Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2009926947

© Springer Science+Business Media B.V. 2009

No part of this work may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording or otherwise, without written permission from the Publisher, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Acknowledgements

The editors wish most warmly to thank Derek Parfit, John Robertson and Jesper Ryberg for the invaluable moral and strategic support they provided to us in connection with this collection. The “future person” persons who have contributed to this collection as well as all future “future person” persons share an unending appreciation of Parfit’s stunning achievements. The work of Robertson, Ryberg and many others has been pivotal as well in helping philosophers and lawyers alike think through the implications of the Parfitian opus in creative and, we hope, ultimately meaningful ways.

Melinda Roberts gratefully acknowledges the support of The College of New Jersey for her work on this collection. David Wasserman gratefully acknowledges the support of the National Institutes of Health (NIH Grant R03 HG 004249) for his work on this collection.

Contents

Harming Future Persons: Introduction	xiii
Melinda A. Roberts and David T. Wasserman	

Part I Can Bringing a Person into Existence Harm That Person? Can an Act That Harms *No One* Be Wrong?

1 The Intractability of the Nonidentity Problem	3
David Heyd	

Part II If Bringing a Badly Off Person into Existence is Wrong, is *Not* Bringing a Well Off Person into Existence Also Wrong?

2 Rights and the Asymmetry Between Creating Good and Bad Lives ..	29
Ingmar Persson	
3 Asymmetries in the Morality of Causing People to Exist	49
Jeff McMahan	

Part III Must an Act Worse for People be Worse for a *Particular* Person?

4 Who Cares About Identity?	71
Nils Holtug	
5 Do Future Persons Presently Have Alternate Possible Identities?	93
Clark Wolf	
6 Rule Consequentialism and Non-identity	115
Tim Mulgan	

Part IV Is the Argument to “No Harm Done” Correct? Must an Act that Harms a Person Make that Person Worse Off?

7	Harming as Causing Harm	137
	Elizabeth Harman	
8	Wrongful Life and Procreative Decisions	155
	Bonnie Steinbock	
9	Harming and Procreating	179
	Matthew Hanser	
10	The Nonidentity Problem and the Two Envelope Problem: When is One Act Better for a Person than Another?	201
	Melinda A. Roberts	

Part V Is the Morality of Parental Reproductive Choice Special? Can Intentions and Attitudes Make an Act that Harms No One Wrong?

11	Reproduction, Partiality, and the Non-identity Problem	231
	Hallvard Lillehammer	
12	Two Varieties of “Better-For” Judgements	249
	Peter Herissone-Kelly	
13	Harms to Future People and Procreative Intentions	265
	David T. Wasserman	

Part VI Is the Person Affecting Approach Objectionable Independent of the Nonidentity Problem?

14	Can the Person Affecting Restriction Solve the Problems in Population Ethics?	289
	Gustaf Arrhenius	

Part VII What are the Implications of the Nonidentity Problem for Law and Public Policy?

15 Implications of the Nonidentity Problem for State Regulation of Reproductive Liberty 317
Philip G. Peters

16 Reparations for U.S. Slavery and Justice Over Time 333
Seana Valentine Shiffrin

Name Index 341

Subject Index 345

Contributors

Gustaf Arrhenius Department of Philosophy, Stockholm University, Swedish Collegium for Advanced Study, Sweden, Gustaf.Arrhenius@philosophy.su.se

Matthew Hanser University of California, Santa Barbara, CA, USA, hanser@philosophy.ucsb.edu

Elizabeth Harman Princeton University, Princeton, NJ, USA, eharman@princeton.edu

Peter Herissone-Kelly Philosophy Section, International School for Communities, Rights, and Inclusion, University of Central Lancashire, UK; University of Bolton, UK, pnerissone-kelly@uclan.ac.uk

David Heyd The Hebrew University of Jerusalem, Jerusalem, Israel, david.heyd@huji.ac.il

Nils Holtug Philosophy Section, Department of Media, Cognition and Communication, University of Copenhagen, Denmark, nhol@hum.ku.dk

Hallvard Lillehammer Faculty of Philosophy, University of Cambridge, Cambridge, UK, Hallvard.Lillehammer@kings.cam.ac.uk

Jeff McMahan Philosophy, Rutgers University, New Jersey, USA, McMahan@Philosophy.Rutgers.edu

Tim Mulgan University of St Andrews, Scotland, UK, tpm6@st-andrews.ac.uk

Ingmar Persson Department of Philosophy, University of Gothenburg, Gothenburg, Sweden, Ingmar.Persson@phil.gu.se

Philip G. Peters School of Law, University of Missouri, Columbia, MO, USA, petersp@missouri.edu

Melinda A. Roberts Philosophy and Religion, The College of New Jersey, Ewing, NJ, USA, robertsm@tcnj.edu

Seana Valentine Shiffirin Law and Philosophy, UCLA, Los Angeles, CA, USA, shiffirin@law.ucla.edu

Bonnie Steinbock Philosophy, The University at Albany / State University of New York, Albany, NY, USA, steinbock@albany.edu

David T. Wasserman Center for Ethics, Yeshiva University, New York, NY, USA, dwasserm@yu.edu

Clark Wolf Philosophy and Bioethics, Iowa State University, IA, USA, jwcwolf@iastate.edu

Harming Future Persons: Introduction

Melinda A. Roberts and David T. Wasserman

1 Purpose of this Collection

What are our obligations with respect to persons who have not yet, and may not ever, come into existence? Few of us believe that we can wrong those whom we leave out of existence altogether—that is, *merely possible* persons. We may think as well that the directive to be “fruitful, and multiply, and replenish the earth” does not hold up to close scrutiny.¹ How can it be wrong to decline to bring ever more people into existence? At the same time, we think we are clearly obligated to treat *future* persons—persons who don’t yet but will exist—in accordance with certain stringent standards. Bringing a person into an existence that is truly awful—not worth having—can be wrong, and so can bringing a person into an existence that is worth having when we had the alternative of bringing that same person into an existence that is substantially better. We may think as well that our obligations with respect to future persons are triggered well before the point at which those persons commence their existence. We think it would be wrong, for example, to choose today to turn the Earth of the future into a miserable place even if the victims of that choice do not yet exist. That the only persons who shall suffer are, at present, *future* persons rather than *existing* persons strikes us as no defense whatsoever against the claim that what we have done is wrong. Finally, at least many of us think that our obligations in respect of future persons are particularly stringent when those persons will come into existence as a result of our own procreative acts or choices—when, for example, they are our own offspring.

Each of these intuitive points is challenged, however, in one way or another by the powerful and compelling logic of what is known as the *nonidentity problem*. The purpose of this collection of new papers is to understand, explore and work toward the resolution of the nonidentity problem and the many puzzles it creates in moral philosophy and the law.

2 The Person-Affecting Intuition and the Nonidentity Problem

It is tempting to think that what makes an act or choice morally impermissible, or wrong, must be connected in some central way with a person's having been made *worse off*, or *harmed*, or *wronged*, by what the agent has done. We are tempted, that is, to accept the so-called *person-affecting intuition*—the idea, as Parfit puts it, that “what is bad must be *bad for someone*.”²

The person-affecting intuition, understood in one way, has a consequentialist ring to it. It seems to suggest that an act is wrong only if, by that act, the agent creates less wellbeing for some person for whom the agent could have created more wellbeing, where *wellbeing* itself is taken to be whatever it is that makes life so precious to the one who lives. However, so long as we specify that what the intuition makes important is not the creation of additional wellbeing per se—e.g., on an aggregate basis—but rather, for *each existing and future person*, preserving or enhancing wellbeing *for that same “identical” person*, the intuition might easily be endorsed by non-consequentialists and consequentialists alike.³

It is important to note that the person-affecting intuition provides only a necessary condition for when an act is wrong. It is not a complete theory. Yet it seems to capture a compelling and useful truth about moral law in a domain that offers little that is certain. And, by grounding morality in facts we seem able to grasp—by connecting an act's being wrong with its being “bad for” a person—the intuition strikes us as inherently practical. We seem able to tell when it is that an act is “bad for” a person and when it is that some other act is better. The person-affecting intuition thus seems to have the capacity to help us guide our own acts and choices as well as assess the acts and choices of others.

Yet, however compelling, the person-affecting intuition is challenged by a problem that arises uniquely in connection with *future persons*—persons, that is, who at the time the act in question is performed do not yet but will exist. The intuition implies that the future-directed acts—the acts, that is, that are not “bad for” any *existing* person—must always be “bad for” one or more members of the class of future persons if they are to be deemed wrong. According to the *nonidentity problem*, however, there is very often an obstacle to finding that a wrong act is “bad for” those persons—that it makes things worse for, or harms, or wrongs them. That obstacle arises from the simple fact that any future person's own coming into existence often remains unsettled at the time at which the wrong act is performed, and often appears to depend on the performance of the wrong act. The wrong act, in other words, itself often appears to be the deciding factor in whether one particular person comes into existence rather than a “nonidentical” other. When that happens—when we want to say the future person “owes his or her very existence” to the performance of that wrong act—it can be very hard to say just how that act makes things “bad for” or worse for, or how it harms, or wrongs, that person. Isn't, after all, the flawed existence better than none at all? And in the case where the act has no adverse consequences for anyone else—no adverse consequences, that is, for any existing

person or any other future person—the challenge of pinpointing just how that act manages to be “bad for” anyone at all seems insurmountable.

The nonidentity problem has thus been taken to establish once and for all that the person-affecting intuition is false. It shows that a “bad” act need not be “bad for” anyone, and that wrongdoing cannot fundamentally be a matter of an agent’s making things worse for, or harming, or wronging, particular persons. We can now clearly identify the greater challenge that the nonidentity problem generates for us: if morality, contrary to the person-affecting intuition, is *not* fundamentally rooted in how acts affect persons—in whether, that is, what agents do is “bad for” persons or not—then in what deeply mysterious thing *is* morality rooted?

3 The Nonidentity Cases

The nonidentity problem emerges from a series of riveting cases, each designed to show that some “bad” acts are “bad for” no one. A look at a handful of those cases, ranging from the very personal to the broadly public, suggests just how broad and how deep the problem really is.

A woman who chooses to have a child with one partner rather than another chooses more than her partner. She also “chooses” her child. And even if that child fares quite badly in life, it may seem to make no sense at all for the child to complain about the woman’s choice of partner. Had the woman instead produced a child with someone incapable of passing on the particular condition the child objects to—say, cystic fibrosis, hereditary deafness or Huntington’s chorea, or nationality, ethnicity or low socioeconomic status—*that* child would not have been made any better off than he or she in fact is. Rather, *that* child would never have existed at all. For the less well-off child and the alternative, better-off child are two, numerically distinct, *nonidentical* children. Ridding oneself of the condition means never existing at all; it does not mean having one’s lot improved.

What goes for the woman goes for the man as well. A man who chooses to have a child with an older rather than a younger woman is more likely to conceive a child with Down syndrome. Where a child is, in fact, born with Down syndrome, does it make sense to say that the father’s choice is “bad for,” or that it harms or wrongs, his child? It seems, when we look closely at how the child’s claim against the father could be articulated, that it does not. The child’s claim would be: “It’s a bad thing for me to have Down syndrome. You ought to have acted to prevent that bad thing from happening to me. And you could have done just that by producing a child with a younger woman with younger ova. Then, the suffering I now endure would have been avoided altogether.” The claim against the father collapses under its own logic. The child is asserting that the father ought to have performed act A rather than act B. But while B means a flawed existence for that child, A means none at all. For any child the father might have produced by performing A rather than B would certainly have been *nonidentical* to the child in fact produced. Since the child’s life is, we assume, clearly worth having, and given that A and B together

exhaust the father's alternatives, it becomes very difficult to articulate just how B can make things worse for, or harm, or wrong, the child. How can the *best available choice* for the child—the choice that can be demonstrated in advance to maximize wellbeing for that child—be “bad for” that child?

If this were the entire story behind the nonidentity problem, we might be left unimpressed. We might agree that the parent's choice is not “bad for” the impaired child. But we might also think that what the parent has done—the woman in choosing to risk having a child with Huntington's chorea; the man in choosing to have a child with an older woman—is not clearly wrong. What could be wrong, we might think, about the conscientiously and lovingly made choice to take the chance of producing a genetically or chromosomally impaired child—a child who will be well cared for by the parent?

But that is not the entire story behind the nonidentity problem. It is not just loving parents who make the sorts of procreative choices that lead to the existence of one child in place of another. There are unloving parents and there are loving but careless parents. And there are agents who are not parents at all but who are still engaged in the business of bringing new people into existence—engaged, that is, in procreative conduct, broadly construed. Consider, for example, the case of the doctor who has negligently failed to advise a woman that she faces a significant probability of having a child with Huntington's chorea or Down syndrome or to inform her of the availability of amniocentesis or other forms of genetic testing that would enable her to make an informed decision about continuing the pregnancy. Imagine that the child the woman produces is in fact impaired. It is fair to say that the doctor's negligence has become part of the causal sequence of acts and events ending in a particular child's coming into existence. The doctor has caused, that is, that child to come into what many of us would regard as a flawed existence. But is what the doctor has done “bad for” the child? Can the *child* legitimately claim that the doctor has harmed the child? Does the *child* have a valid claim for what, in the law, is called “wrongful life”?⁴ All that a more dutiful course of action would have accomplished would have been the existence of a distinct child, perhaps a better-off child but still a child *nonidentical* to the original. The *parents* may have a legitimate claim—for “wrongful *birth*”—against the doctor. The *parents* can perhaps make the argument that they have been harmed, or that their procreative autonomy has been compromised, by the doctor's negligence in allowing them to have a kind of child they did not want and forcing them to incur the unsought expenses and responsibilities of caring for that child. But the *child's* claim to have been harmed—according to the logic of the nonidentity problem—is itself without merit.

Other nonidentity cases challenge the person-affecting intuition in ways that may be still more persuasive. Consider, for example, the private fertility clinic that conducts a lucrative business in marketing and supplying new or risky reproductive technologies to desperate couples. Such technologies do not appear to increase the probability of having a child with Down syndrome or other common genetic or chromosomal disorders. But that does not necessarily mean they are safe. One common variety of impairment stems from multiple birth pregnancies induced by the aggressive use of ovulation-stimulating drugs. The upshot—when the woman

declines to consent to the selective reduction of fetal life or when the doctor fails to advise or even insist on it—can be the birth of several very low birth weight babies who face substantial risks of impairment, ranging from blindness and deafness to cerebral palsy and serious mental disability. Yet what chance do those identical offspring have of coming into existence *in the absence of* aggressive treatment? If the answer is little or none, it becomes difficult to argue that what has been done is “bad for” those particular offspring. Yet we may well have the sense that what has been done is surely “bad”—that it is, again, morally and perhaps legally wrong.

Or we may not. Even at this juncture, some theorists may simply bite the bullet: they may accept, within the nonidentity context, the inferences it seems we can validly derive from the person-affecting intuition. They may take the view that, at least so long as the child’s existence is worth having, the acts and choices—of the parents, the doctors, the fertility clinic or anyone else—that are “bad for” no one are neither morally nor legally wrong *even if* they lead to the existence of children who suffer badly.

Any such “no wrong done” strategy for preserving the person-affecting intuition is critically tested, however, by still more powerful forms of the nonidentity problem. Here is where the work of Derek Parfit becomes the most critical.⁵ What Parfit has demonstrated, in case after case, is how very broadly the logic of the nonidentity problem seems to extend. Thus Parfit, along with Gregory Kavka, has made a compelling case that the part of the problem that argues “no harm done” applies, not only in connection with the sort of future-directed conduct we usually think of as procreative in nature, but also in connection with a great deal of future-directed conduct that seems on its face to have nothing to do with procreation at all.⁶

If that argument is correct, then conduct identified as “harmless” by the nonidentity problem—conduct that is “bad for” no one—will include both the choices we make by the bushel on a daily basis and the large-scale, sweeping, collective choices a given community or nation may make only once in a generation or so. But it seems highly implausible that, however bleak the future they bring about, *all* such choices are morally permissible. It seems clear that, as a matter of moral law, when it comes to future-directed conduct, “laissez-faire” cannot so often be the correct posture.⁷ Our only option seems to be to reject the person-affecting intuition itself.

And the Parfitian “no harm done” argument is a persuasive one. After all, a great number of choices that are not stereotypically procreative in nature do seem on inspection to have clearly procreative *effects*. What look like “same people” choices—choices that make things better or worse for a particular future person—are really “different people” choices—choices that make things better or worse for no one at all but instead serve to bring distinct (“nonidentical”) persons into existence.

It is useful in this context to compare a perfectly ordinary, future-directed, “same people” choice that is not vulnerable to the nonidentity problem against one that is. As an example of the former sort, Parfit offers the following:

Suppose that I [for no good reason; senselessly] leave some broken glass in the undergrowth of a wood. A hundred years later this glass [foreseeably] wounds a child. My act harms

this child. If I had safely buried the glass, this child would have walked through the wood unharmed.⁸

But this ordinary wrong can be converted into a nonidentity problem by the addition of a few facts. Suppose it happens that you will conceive your first child this evening. Suppose as part of your day you make an excursion into the wood and (again, for no good reason; senselessly) leave broken glass in the undergrowth. Later on that same day, the conception takes place. Months later your child is born, and years later, when walking in the woods, that child is badly cut by that same broken glass—foreseeably so, in the sense that it was foreseeable that some child or another would be wounded in just that way. Is your act of leaving broken glass in the undergrowth “bad for” your child? Has it made things worse for, or harmed, or wronged, your child?

Had you not left the broken glass in the undergrowth and instead proceeded directly to conceiving a child, your child would not have suffered the injury later on. But, very probably, the child you then would have produced—the child who wanders into the wood and emerges unscathed—would also have been *nonidentical* to the child you in fact produced and who in fact suffers. But from your injured child’s own point of view, surely it is better to exist and suffer the injury than never to have existed at all. It seems, then, that your senseless act has not made things worse for your child, but if anything has made things better. Your child owes his or her very existence to your senseless act. How, then, can that act be “bad for” your child? At the same time, it seems clear that your act is itself “bad,” or wrong. After all, people should not leave broken glass in places where children will play.

We don’t think of the act of leaving broken glass in the wood as the kind of act that will decide who will and who won’t exist. But this example shows that perfectly ordinary acts do seem to have exactly that kind of procreative effect. The most ordinary acts, in the right circumstances, can help to determine *just what child shall be conceived and born*—just as much as the woman’s or man’s choice of partner or the doctor’s failure to discuss genetic risks with his or her patients or the fertility specialist’s decision to commence one treatment plan rather than another.

Choices of far-ranging economic and environmental policies operate on a much bigger scale than the broken glass case. They can, accordingly, generate much greater quantities of suffering—and they may accordingly strike us, if anything, as still more clearly wrong. But claims that those policies or programs harm, or are “bad for,” or wrong, future persons are just as vulnerable to the logic of the nonidentity problem.

To see that, it “may help,” as Parfit puts it, “to think about this question: how many of us could truly claim, ‘Even if railways and motor cars had never been invented, I would still have been born’?”⁹

Suppose that we are choosing between two social or economic policies. And suppose that, on one of the two policies, the standard of living would be slightly higher. . . . This effect implies another. It is not true that, whichever policy we choose, the same particular people will exist in the further future. Given the effects of two such policies on the details of our lives, it would increasingly over time be true that, on the different policies, people married different people. And, even in the same marriages, the children would increasingly over time

be conceived at different times . . . We can plausibly assume that, after one or two centuries, there would be no one living in our community who would have been born whichever policy we chose . . .¹⁰

Whatever policy we choose, the future persons who in the end exist and suffer will seem to “owe their very lives” to our choice. Had we instead done the seemingly better thing—had we taken steps to combat global warming, for example, or to conserve fossil fuels—the Earth would have been a far better place to live, but the persons who eventually exist and suffer under the one choice would not have *enjoyed* a better existence. Instead, they would never have existed at all. And that outcome surely would not have been better for them.

We again face—assuming that the selfish choice is “bad for” no existing person and for no other future person—an example of an act that is “bad for” no one at all. In many such cases, however, many of us will remain convinced that the choice itself is wrong.

Examples like the broken glass case, or Parfit’s “depletion” or “risky policy” case, or Kavka’s “slave child” or “pleasure pill” case, display the phenomenon Kavka described as the “precariousness” of existence.¹¹ The chances of any of us coming into existence would be vanishingly small had the causal chain that ends in our existence been varied in even the slightest way. But that means that there exists a vast array acts and choices that on their face have nothing to do with procreation but below the surface play a critical role in determining just what future persons we ourselves and others may one day produce. According to the nonidentity problem, those same choices cannot make things worse for, or harm, or wrong, or be otherwise “bad for,” the very future persons whom they serve to produce.

But if that is correct, the person-affecting intuition will then force us to conclude that the agents have done no wrong—that, in fact, the agents *can* do no wrong, so long as the acts themselves are future-directed and affect no existing person and the coming into existence of particular future persons is indeed “precarious.” If we continue to believe, however, about any such case that wrongs have been done—if we continue to believe that it matters morally whether the parent leaves glass in the underbrush for no reason, or takes a teratogenic pleasure pill, soon before conceiving a child, or whether we as a community throw ourselves behind one economic or environmental policy rather than another—then we must, it seems, reject the person-affecting intuition.

And it is here that we face the full power of the nonidentity problem. If we think that the future-directed act under scrutiny is wrong, how then are we to explain why that is so when the most natural and intuitive account of that fact—that it is “bad for” some future person—has been annihilated? How, in other words, are we to account for the obligations we think we have in respect of future persons?

The nonidentity problem represents a challenge not just for moral philosophy but also for law and public policy. What we say about the problem will have dramatic implications for the liability agents may have under tort law in respect of any children they negligently bring into a disadvantaged existence. It will also have implications for the scope of any constitutional guaranty that protects the privacy

and liberty interests that individuals may have in making procreative choices as they please. And it will influence our thinking about reparations for societal injustices.

We want health care providers to exercise due care in taking the family histories of couples who are planning to have children and advising couples carefully and wisely on matters relating to genetic testing. Yet many courts have implicitly accepted the nonidentity problem's "no harm done" result and have, accordingly, disallowed claims for wrongful life against the negligent provider.¹² After all, the impaired child exists only because the provider has been negligent in some way—has carelessly forgotten, for example, to recommend timely amniocentesis. It may thus seem that the child cannot have been *harmed* as a result of what the provider has done. But harm is an essential element of any claim in negligence; without harm, there is no claim.

We may think, moreover, that states should be able to regulate fertility clinics in order to reduce the chance that the offspring their services help to produce will be born impaired. Yet the nonidentity problem seems to suggest that the state's hands are constitutionally tied in any case in which the couple's procreative interests are at stake and the impaired child "owes his or her very existence" to the very practices the state now seeks to regulate. If the child has not been harmed and the couple has given their informed consent for treatment, the state would seem at a loss to show that regulation serves the "compelling aim" that it must when procreative privacy is at stake.

The question of reparations for past societal injustices raises analogous issues. To insure that the requirement to pay reparations is itself not arbitrarily imposed, we must be prepared to explain the moral basis of any such requirement. If we are persuaded by the nonidentity problem, however, that people who exist today cannot have been harmed in any morally relevant sense by those past societal injustices, then we will find it very difficult to articulate any such moral basis. And if we cannot do that, we open the door to the view that reparations are, in general, inherently unjust, at least in the case where the societal injustices happened long enough ago and were broad enough in scope to influence the identities of the future generations who suffer their effects.

4 Three Strategies for Addressing the Nonidentity Problem

The nonidentity problem has provoked an intense and wide-ranging debate among moral philosophers and legal theorists—a debate that extends from the nature of harm to the moral status of future persons to the very structure of morality and the law. In confronting the nonidentity problem, we seem to have three basic strategies to choose from. First, we can retain the person-affecting intuition and accept the result the nonidentity problem seems to entail—that no one is harmed, or made worse off, or wronged by the act or choice under scrutiny. We are then compelled also to accept that that act or choice is morally permissible. Second, we can reject the person-affecting intuition and argue that the fact that no one has been harmed,

or made worse off, or wronged does not imply that a wrong has not been done. On this approach, we will be left to conclude that, in at least some cases, what makes an act morally impermissible is grounded in something other than how that act affects any particular existing or future person. Third, we can accept the person-affecting intuition and reject the claim that the apparent victim of the questionable act has not been harmed, or made worse off, or wronged. We can take the position, in other words, that the “bad” act is, on closer inspection, indeed “bad for” someone or another. These strategies are not mutually exclusive. It is possible to argue, for example, for one class of nonidentity cases but not the other that no harm or wrong has been done.

Variations on these three strategies have emerged in earlier work on this topic, and are deployed in this present collection as well. In this Part 4, we will outline these variations. We will then preview, in Part 5 below, the particular contributions to this collection.

A. First, we can retain the person-affecting intuition and accept the basic inference that the nonidentity problem seems to support—that the acts under scrutiny in nonidentity cases do not harm, or make things worse for, or wrong any person. We must then accept that those acts, in virtue of their implications for the identities of future persons, are not wrong.¹³ On that view, morality may well impose severe constraints on our acts when those acts affect existing persons or future persons who will exist independently of what we do, but will impose no genuine constraints at all when the effects of those acts do not extend beyond the future persons who are brought into existence by those same acts.

B. Second, we can reject the person-affecting intuition and argue that the fact that no person has been harmed, or made worse off, or wronged does not imply that no wrong has been done. We can, that is, reject the person-affecting approach in favor of an account of wrongdoing that is “impersonal” at least in part. On that view, a wrong act is, in at least some cases, wrong not in virtue of whether it makes things better or worse for particular persons but rather in virtue of its failure to promote or secure certain “impersonal” values, ideals or effects. This option, while both attractive and widely deployed, faces some formidable problems. Once we say that what is important is not increasing wellbeing for particular persons, but rather increasing wellbeing in general, we raise the specter of an unconstrained obligation to procreate. After all, creating more persons can itself be one way to create more wellbeing.¹⁴ *Totalism*, for example, requires agents to maximize total wellbeing (or happiness or utility) on an *aggregate* basis. Totalism avoids the nonidentity problem. The choice to bring the less well off child into existence in place of the nonidentical better off child is wrong, according to totalism, because it fails to maximize aggregate wellbeing. But totalism also appears to make procreation obligatory whenever bringing an additional person into existence increases aggregate wellbeing—even when bringing that additional child into existence is “bad for” others who do or will exist. The plausible idea that our procreative obligations cannot be quite so stringent as that could be rooted in “commonsense morality” or a theory of liberty or simply in the intuition that our obligation is not to “make happy people” but rather to “make people happy.”¹⁵ It is, in any case, an idea we are reluctant to abandon.

Some theorists have charted a narrower impersonal course, requiring agents to bring the better off child into existence in place of the less well off child while avoiding the problematic thesis that more wellbeing in the aggregate is always better. Such an approach appeals to the intuition that we should benefit those who will benefit more.¹⁶ An example of that approach is Parfit's "principle Q," which states that, if "the same number of people would ever live, it would be worse if those who live are worse off . . . than those who would have lived."¹⁷ A variation on that approach attempts to defend the claim that a choice can be worse for someone where the person for whom things have been made worse is not strictly identical to the person for whom things might have been made better. A related approach draws on the continuing debate over cross-world identity to argue that a person-affecting account of nonidentity cases becomes plausible if we are willing to adopt a suitably coarse-grained theory of identity—one that provides, for example, that the child the agent conceives on the day he or she leaves glass in the undergrowth of the wood is the same child as the child who would have been conceived on that day by that same agent had the glass not been left in the wood. The possible relevance to the moral analysis of the many mistakes involving cross-world identity we can easily imagine agents (including ourselves!) to make has also been explored in this connection.

C. Third, we can challenge one or more of components of the nonidentity problem itself—an underlying assumption, a premise or an inference that has been put to work in constructing the argument to the result that some "bad" acts are "bad for" no one. These challenges, if successful, would allow us to retain the person-affecting approach while maintaining that at least many of the future-directed acts that appear wrong really are wrong—including leaving the glass in the undergrowth of the wood for no reason and depleting our natural resources in a way that will make the Earth of the future a miserable place. One such strategy involves close scrutiny of the inference to the result that the act in question is "bad for" no one.

That inference, of course, seems straightforward. How can an act that is critical for a particular person's very existence harm, or be "bad for," that person? In fact, however, the inference is not as obviously correct as it might seem. (i) The inference relies on a *comparative* account of harm—an account that holds that an act does not harm (is not "bad for") a person when that act does not make things *worse for* that person. An act that maximizes wellbeing for a particular person, on this view, cannot also be considered to *harm* that person. However, some theorists have argued that the comparative account of harm is mistaken. On an alternative account, a person is harmed when that person is made to suffer some discrete evil, injury or insult such as death, pain, loss or frustration. Accordingly, the fact that a person has been left at least as well off as that person would, or could, have been had he or she never existed at all does not prove "no harm done."

(ii) The "precariousness" of existence that lies at the heart of many nonidentity problems—that a person's chances of existence, had agents conducted themselves in some seemingly better way, would have been practically nil—is a phenomenon that, once pointed out, is hard to miss (though denied under the coarse-grained theories of identity mentioned above). Still, it has been argued that we can retain a comparative

account of harm (and a relatively fine-grained account of identity) but also recognize as fallacious the apparently simple transition from the idea that existence is precarious to the “no harm done” result. The idea here is that in many nonidentity cases the existence of a particular future person will be highly improbable even given the clearly wrong act, and hence that there is no basis for claiming in those cases that that wrong act is just as good for that person as any seemingly better alternative.

(iii) The person-affecting intuition can be interpreted as encompassing wrongs against persons as well as harms. On that reading, the act that *wrongs* a person can be “bad for” that person even in the case where no *harm* has been done. A number of philosophers have tried to identify ways in which acts that bring about the existence of a future person may wrong that person. Some have argued for a “birthright” to a life that is more than barely worth living; others have set various thresholds or minimum standards for such a life.¹⁸ Creating an individual whose life is expected to fall below such a minimum can be said to violate that right, independent of the issue of harm. Others argue that people have a right against the infliction of or exposure to certain noncomparative harms, a right violated in many nonidentity cases. Still others argue that future people can be wronged by the attitudes or intentions with which they are brought into existence. On this view, a procreative choice informed by disrespectful or insensitive attitudes can constitute a wrong to a particular person, even in cases in which agents could have done nothing to make things better for that person than they are.

One final point is critical at this juncture. We worry about the nonidentity problem because we want clearly to articulate our obligations in respect of future persons. We are pressed to say just what makes some future-directed conduct wrong when the most natural and intuitive account of that fact—the person-affecting account; the idea that what is “bad” is “bad for” some existing or future person—has been so compellingly challenged by the nonidentity problem. Still other theorists have argued, however, that many person-affecting accounts, independent of the nonidentity problem, are highly objectionable. They remain perilously incomplete, are not fully cogent or otherwise seem deeply problematic. We should thus keep in mind that it is an open question whether the person-affecting approach itself is worth saving.

5 Contributions to this Collection

The articles that make up this present collection explore the obligations that we have with respect to future persons and work to understand how those obligations are to be grounded within a broader moral theory.

5.1 *The Person-Affecting Intuition. Heyd*

David Heyd has elsewhere defended the view that any act that adversely affects only those future persons whose existence depends in some way on that act’s being performed cannot properly be said to be “bad for” or to harm those future persons

or to make them worse off. Heyd also retains the person-affecting intuition. The upshot is that, on Heyd's view, the choices involved in the typical nonidentity cases cannot correctly be deemed wrong. Hence he describes the "intractability" of the nonidentity problem. If an act is wrong at all, that it is wrong will be grounded, according to Heyd, in the fact that it has made or will make an existing person, or a future person who will exist whether the act under scrutiny is performed or not, worse off.¹⁹

In this new paper, Heyd extends the logic of the nonidentity problem to cases involving not just strict numerical identity but what he calls "biographical identity" as well. Thus, just as parents do nothing wrong when they bring the impaired or disadvantaged child into existence—at least, they do no wrong that can be grounded in the suffering of that particular child—parents also do nothing wrong when they raise their child under, or return their child to, a particular biographical identity. While a new "identity" may, in one sense, make things go better for that child, in another sense of identity—the biographical sense—it may do nothing more than replace that child with another child, a biographically nonidentical child, altogether.

Heyd suggests that extending the person-affecting approach in this way may help to account for the intuition that it is morally permissible, and perhaps obligatory, for the "charitable gentiles who saved Jewish children during the Holocaust" to return those children to a "Jewish environment" after the war even on the supposition that things, in some sense, would have been easier for those children if they retained their "Christian life style." Returning them would restore the biographical identities that were obscured or suspended by their concealment in gentile families—it would be good for the children they were and could still be. The very fine-grained theory of personal identity that Heyd accepts, then, along with the person-affecting intuition, suggests that there exists no moral constraint that would require the parents to raise the one child under what many might consider to be the more advantageous alternative identity.

5.2 *The Asymmetry. Persson, McMahan*

Many of us believe that it would be wrong to bring a person into a truly awful or deeply unhappy existence—an existence, that is, that is *less* than worth having. We think, in other words, that the fact that a person's life would be miserable provides a very strong moral reason *not* to bring that person into existence. At the same time, it may also seem clear to us that the bare fact that a person will come into an existence that is wonderful—or at least worth having—does not provide any moral reason at all for agents to bring that person into existence. These two positions together present us with what is called *the asymmetry*.²⁰

Ingmar Persson argues that the asymmetry itself is rooted in a certain commonsensical rights theory. According to that theory, merely possible persons cannot be rights-holders. We thus cannot violate any right that a merely possible person has by failing to bring that person into existence. In contrast, a severely diseased child may have a right that is violated when that child is brought into an existence that is less than worth having. Thinking about our moral reasons and our obligations in

terms of the child's rights thus seems to have the potential to explain the asymmetry. But there is a difficulty with such a rights-based account, according to Persson: it is forced also to attribute to the parents a sort of property right, in effect, to treat their child as they please, even as it also attributes to the child a right not to be treated in exactly that way. When we realize that the rights theory itself is fatally inconsistent and must be rejected, we should, according to Persson, find ourselves free to reject the asymmetry itself—and, in particular, the part of the asymmetry that says that we have no reason at all to bring persons who will have good lives into existence. And we can instead take the view that “we have reasons to create good lives, proportionate to their goodness, just as we have reasons against creating bad lives, proportionate to their badness.”

Having rejected the asymmetry on impersonal grounds, Persson is then in a position to address the nonidentity problem itself on impersonal grounds. Wrongdoing isn't just a matter of making things worse for a particular person. It can, instead, be a matter of creating the worse life in place of the distinct but better life.

McMahan's attempt to understand just how the asymmetry might be explained begins with a distinction among three kinds of value: what he calls “individual-affecting” value, concerning what is better or worse for an individual; what he calls “noncomparative individual” value, concerning what is good or bad for an individual, but not necessarily better or worse; and impersonal, concerning what is of value regardless of whether it is better or worse, or good or bad, for individuals. He argues that agents, within the scope of their moral decision-making, can legitimately rely on only two kinds of reasons—individual-affecting and impersonal. And he argues, moreover, that choices about whether to create new individuals can be based only on impersonal reasons since there is no individual for whose sake those choices can be made, and the outcomes of those choices cannot be better or worse for any individual. McMahan then notes two functions that “goods” and “bads” (whether individual-affecting or impersonal) can play in moral decision-making and justification—“reason giving”—providing moral reasons for action—and “cancelling”—justifying action by a favorable balance of goods to bads. He then speculates that the commonsense asymmetry derives from the assumption that, in the context of procreative choice, where the agents' reasons are limited to the impersonal, impersonal bads have both functions while impersonal goods have only a cancelling function. On that assumption, a future individual's net happiness would give us no reason to create that individual, while a future individual's misery does give us a reason *not* to create that individual unless it is cancelled by greater happiness.

In the end, however, McMahan finds that assumption, along with a host of alternative assumptions that he articulates and explores, arbitrary and ad hoc. He concludes that the most promising way to understand the asymmetry is in terms of a distinct asymmetry between harming and benefiting. But that distinction is plausible, he argues, only to the extent that it gives lesser weight, as opposed to no weight, to benefits. Discounting rather than eliminating benefits, however, is not an approach that, on its own, is capable of accounting for the asymmetry. McMahan in the end concludes, with Persson, that the asymmetry lacks any

compelling theoretical defense—again opening the door to an impersonal solution to the nonidentity problem itself.

5.3 *Identity. Holtug, Wolf, Mulgan*

The views of Persson and McMahan both permit appeal to impersonal values to explain how a choice that does not make its most plausible victim—the miserable or unhappy future person—worse off can be wrong. Nils Holtug argues that we do not need to appeal to impersonal values—or, at least, that we do not need to appeal to the value of maximizing wellbeing in the aggregate—to make that claim. By adopting a variation on a “wide” person-affecting account instead, we can, Holtug argues, explain why the act under scrutiny is wrong while explicitly taking steps to avoid the problematic notion that we are morally obligated to bring ever more additional persons into existence. Holtug thus abandons the idea that acts that “affect persons” must make things better or worse for the *same* “identical” persons. A choice may be wrong, according to Holtug, even though it does not make any particular person worse off. “What matters is that [one] child will benefit more from coming into existence than [another, possibly *nonidentical*] child.” Bringing the healthier child into existence in place of the impaired child is not a way of improving things for the impaired child. But it is, he says, a way of *affecting persons* for the better.

Holtug argues, moreover, that such an approach provides a plausible way of thinking about our reasons for performing gene therapy on the early human embryo. According to Noam Zohar, technologies that alter an individual’s genome cannot properly be regarded as “therapeutic interventions.” That is so, according to Zohar, since such technologies will often not make things better for a particular person but rather substitute a better off person for a nonidentical less well off person. Accordingly, therapeutic interventions must be “denied the special moral status of requests for therapy.”²¹ In contrast, Holtug provides a person-affecting reason—in his wide sense—for performing that therapy. The child whose genome is altered to correct an underlying defect will “benefit more” from coming into existence than will the (distinct) child who would have existed had the genome never been altered to begin with. Holtug argues, moreover, that it may not only be in the interest of the happier, healthy child to come into existence but also in the interest of the nonidentical pre-therapy fetus to be left out of existence and to have the healthy child come into existence instead. That is so, according to Holtug, because the fetus plausibly stands in the relation that “prudentially” matters (again, in Holtug’s wide person-affecting sense) to the nonidentical healthy child and thus has an interest in the benefits that accrue to that child.

Taking a position on personal identity that contrasts starkly with Heyd’s, Clark Wolf challenges the intuitive account that is presumed by both Holtug and the non-identity problem itself. Thus, the nonidentity problem takes it for granted that the disadvantaged or impaired child and the healthier, better off child who would have existed had the agents acted otherwise are “nonidentical.” Finding it implausible that one and the same child can be conceived from distinct gametes, Parfit puts

the point as follows: “It is in fact true that, if you had not been conceived within a month of [the time at which you were in fact conceived], you would never have existed.”²² In Wolf’s view, however, identity is more coarse-grained than the intuitive view suggests—and far more coarse-grained than the view described by Heyd. According to Wolf, the properties that “make us the people that we are” cannot be easily circumscribed. Certainly, the relations we may rely on to determine identity across time—for example, psychological continuity—cannot be used to determine identity across worlds. When we begin to look very closely at just what properties do determine identity, we find ourselves forced to concede that identity itself is vague in important ways. Wolf concludes that there is no reason to think that “the couple’s third child” cannot be one and the same child, whether that child turns out to be seriously genetically impaired or perfectly healthy.²³

Tim Mulgan takes a different tack. He suggests that resolving the nonidentity problem is a matter of understanding just what notions—mistaken or not—about identity ordinary agents in ordinary contexts will put to work. Because the form of rule consequentialism Mulgan defends does just that, it is able, he argues, to preserve the very aspects of commonsense morality that the nonidentity problem seems to challenge. Thus, the “ideal moral code” contemplated by rule consequentialism—the code of rules whose internalization by nearly everyone in the next generation would produce better consequences than any alternative code—will include, according to Mulgan, a host of person-affecting principles. At the same time, it will set aside certain critical facts about personal identity that we, at a more reflective level, may accept. In particular, it will treat “different people” choices as if they were “same people” choices; it will (perhaps falsely) take the impaired child to be identical to the healthier child who might have been produced instead.²⁴ Having set aside the nonidentity problem, the agent who has internalized the ideal code will then be free to appeal to ordinary, commonsensical, person-affecting principles to reach the result that he or she is obligated to bring the healthier child into existence in place of the impaired child. The strategy Mulgan suggests requires a delicately balanced ideal code—one that has the agent think carefully, but not *too* carefully, about the effects of his or her choices. In this present paper, Mulgan argues that the strategy he suggests is itself psychologically viable and that its parts, though in tension, can in fact be readily internalized by the next generation.

5.4 Harm. Harman, Steinbock, Hanser, Roberts

The nonidentity problem relies critically on the argument that the conduct under scrutiny is not “bad for,” and does not *harm* or makes things *worse for*, the person who seems its most obvious victim. Some theorists (Harman; Steinbock; see also Shiffrin, Part 5.7 below) have challenged that claim on the ground that it relies on an overly narrow, comparative (“worse for”) account of harm that is itself mistaken. Hanser agrees that a broader account of harm is in order but also argues that the fact that harm *as an effect* has been established does not entail that what the agent has done *harms* the person whose existence is flawed in some way. Still another

approach (Roberts) is to retain a comparative account of harm but argue that in many nonidentity contexts the inference from that account to the “no harm done” result is itself fallacious.

(i) *Non-comparative account of harm.* Elizabeth Harman has previously argued that an agent harms a person when the agent causes a person to be in a “bad state” such as “pain, early death, bodily damage, and deformation.”²⁵ It might seem that this view cannot account for our intuition that it is worse to deafen a baby than to conceive a baby one knows will be deaf since, on Harman’s approach, the baby in both cases has been harmed. In this present paper, however, Harman argues that the agent has an additional reason not to deafen the baby that has no application in the case where the agent conceives the congenitally deaf baby. That additional reason is simply that the first baby is made *worse off* by what the agent has done. On Harman’s view, then, our reasons come in (at least) two forms—we have reasons to allow people to be free from bad states and we also have reasons not to place people in bad states. Taking into account both sorts of person-affecting reasons—one focused on whether a person has been made worse off, the other on the particular features of the state the person is in fact in—helps us explain why some harming acts are worse than others.

Harman’s main focus in this present paper is a controversial class of cases in which the choice is not whether to bring a less well off person into existence in place of a more well off person—a prototypical nonidentity case; a case, Harman suggests, about which there is less disagreement—but instead whether to bring an *additional* person into existence in place of no one at all. Is it permissible, for example, for a couple to procreate when it is understood in advance that, due to a hereditary condition, *any* child that couple brings into existence will be deaf? Is it permissible for a couple who is enslaved to procreate when, due to an institution over which they have no control, *any* child they will bear will also be enslaved? Since in this class of cases the benefits bestowed by the procreative act will not fall to another if an alternative choice is made—the choice here is between the impaired child or none at all—those benefits do, on Harman’s view, weigh to some degree against the bad effects (deafness, or enslavement). But the fact that the life will turn out to be worth living does not on its own imply, according to Harman, that procreation is permissible. Since the procreative act still harms the child by virtue of having caused the child to be in a “bad state,” we will need to assess whether that harm can be *justified* to determine wrongdoing. That latter assessment is, according to Harman, a matter of “threshold crossing.” Since different people will have different views regarding when an applicable threshold is crossed—when, that is, a harm is so bad that it cannot be justified—“widespread, pre-theoretical disagreement” is to be expected in connection with these particularly hard cases.

Like Harman, Bonnie Steinbock rejects a purely comparative account of harm. According to Steinbock, harm can be a matter of making a particular person worse off than he or she would, or indeed could, otherwise have been. But harm can also be a matter of creating a situation in which a person has certain important interests that cannot be satisfied. Sometimes we do just that, Steinbock argues, when we bring a seriously impaired child into existence. If the child’s life falls below a “minimally decent standard”—if it does not, that is, hold a “reasonable promise of containing

the things that make human lives good,” including an ability to experience pleasure, to have relationships and to learn—then bringing that child into existence constitutes “a harm and a wrong” to that child.

Just when, however, do we reach the point at which the child’s life prospects are so dramatically reduced that the child’s interests will trump the parents’ interests in having that child? According to Steinbock, there is no reason to think that even children with serious disabilities—including Down syndrome, cystic fibrosis and spina bifida—will necessarily have lives that fall below the decent minimum standard. Whether they do or not, according to Steinbock, may depend not just on the disease itself but also on the resources the parents can draw on in caring for the child. That a miserable life is not the fault of either the disease or the parents, but is instead a function of limited resources, does not mean that the parents may make whatever procreative choice they please. The ill effects of certain diseases and disorders might combine with a limited parental ability to provide for an impaired child in a way that will create a life that falls below a “minimally decent standard.” In some cases, at least, agents thus will be required to forego reproduction altogether.

(ii) *Harming versus harm.* Matthew Hanser proposes a two-part test for whether an agent has harmed someone. First, the victim must suffer harm. And second, the agent’s action must be connected to that harm in the “right way”—it must make the agent “at least partially responsible” for that harm. If that account of harm is correct, then the fact that the future person’s bad state constitutes a *harm* will not be sufficient to show that an agent *harms* a future person in a nonidentity case. One must also show that the agent’s act makes him or her *at least partially responsible for* that bad state—that harm.

In applying the first part of the test, Hanser accepts a broad, non-comparative account of harm similar to that described by Harman and Steinbock. But Hanser weaves a comparative element into the second part of the test—with the implication that that second part will be satisfied by some nonidentity cases but not others.

In earlier work, Hanser addressed Parfit’s risky policy example. He argued that in that case future people do suffer harms, and that those who choose the policy are partially responsible for those harms.²⁶ In this present paper, Hanser argues that those same results may not hold for the sorts of nonidentity cases that involve genetic and chromosomal abnormalities. He assumes for the sake of argument that the child who is seriously impaired as a result of a genetic disease or disorder is in a state that is “bad for” that child to be in and thus suffers harm. The harder question, according to Hanser, is whether the act is related in the “right way” to that bad state. It is not enough that there exists a causal chain of acts and events connecting the act under scrutiny and the child’s being in the bad state. A determination of whether the agent has harmed the child will also require a comparison between how well off that child is and how well off that same child would have been in other circumstances. But Hanser finds it unclear how we can make the necessary comparison when the unique alternative to having the impaired child is not having that child at all. It seems, in other words, that there exists no principled basis for concluding that the parent makes the child better or worse off with respect to, e.g., any particular sensory or motor impairment. According to Hanser, then, when the only alternative available to the parent is not to create that particular child at all and the child’s life is

itself worthwhile, we cannot conclude that what the parent has done has harmed the child.

The results that Hanser reaches in connection with this class of nonidentity cases—cases in which the agent does not stand in the “right” relationship to the child’s bad state—are thus very different from those suggested by Harman and Steinbock. In many of the cases in which Harman and Steinbock will find that the parents have harmed their child, Hanser will find that they have not. In all such cases, if the choice is wrong at all, it will be wrong, according to Hanser, on grounds other than how it affects the child.

(iii) *Is the inference from the comparative account to “no harm done” valid?* Roberts, like Hanser, distinguishes between types of nonidentity problems. Roberts’ initial division is between cases in which the agents *cannot* (consistent with natural law and the choices of other agents) produce the very same person free of impairment or disadvantage and cases in which the agents *can* secure a better existence for that person. It is, Roberts concedes, hard to establish that harm (in a comparative sense) is imposed on a person when agents *cannot* have produced a better outcome for that person. (Harm to others is a different matter.) But the situation may be different, Roberts argues, in the case where agents *can* produce a better outcome for that same person and the issue is just how *likely* it is that agents will be able to do so. Examples of this latter type of nonidentity problem include Kavka’s slave child case, Parfit’s depletion case and many cases involving claims for reparations. In these cases, the argument to the “no harm done” result relies on the claim that the wrong act *improves* a person’s chances of coming into existence relative to that act’s less risky alternatives. Roberts argues, however, that there is an obstacle to establishing that claim. It is true that, within the limits of what is within the agents’ grasp prior to choice, it is *highly improbable* that the person who exists and suffers would have existed, given the less risky act. Roberts argues, however, that it is a fallacy to think that things are different for the wrong act—to think that the wrong act, in some singular way, makes it any *less* improbable that that person will come into existence. This point, Roberts argues, will have ramifications for our calculations of *expected value*. She concedes that, for any possible future person, the less risky act will generate little *expected* value—and that the *actual* value of the wrong act will be much greater than that expected value for any person who eventually does exist. But a comparison between actual against expected values cannot, according to Roberts, establish that one act makes things no worse for a person than another. Moreover, comparisons between expected values will be just as unreliable unless the underlying probabilities are both calculated on the basis of what is within the agents’ grasp prior to choice. Analogous issues are raised by the “two envelope problem.” Here as well Roberts argues that a focus on what is within the subject’s grasp prior to choice—how little or, in the case of the two envelope problem, how *much*—can help to resolve the problem. Roberts concludes that intuition is preserved in both cases: the argument to “no harm done” fails in the slave child case, and so does the argument that it is better for the subject to switch envelopes rather than not in the two envelope problem.

5.5 *A Distinctive Morality for Reproductive Decisions?* *Lillehammer, Herissone-Kelly, Wasserman*

Some theorists begin with the premise that it is simply peculiar to place the morality of reproduction, which includes the formation of some of the most intimate relationships and deepest personal attachments, under the jurisdiction of an impersonal morality. That seems particularly so in cases in which the agent is choosing his or her own future child and can bring into existence either a less well off child or a better off but nonidentical child instead.

Hallvard Lillehammer argues that it is critical to distinguish two features of the rival moralities for reproductive decisions. First, they can be either “person-involving”—that is, “concerned with benefits or harms to identifiable individuals”—or not. Features that are not person-involving in this sense are concerned instead “with good and bad states involving individuals, regardless of whether any individuals can be antecedently identified as involved in those states of affairs.” Second, they can be ethically “impartial”—that is, concerned with the maximal instantiation of some good or value—or “partial”—concerned only with its maximal instantiation in a subset of humanity, for example, the agent’s family, to which the agent has a special attachment. Aggregative forms of consequentialism happen to be both non-person involving and impartial. But the two features are not necessarily conjoined: a morality can be person-involving and impartial or non-person involving and partial. Lillehammer holds that the latter features characterize reproductive morality. While decisions about who comes into existence are non-person involving in his sense, they need not be impartial. For this reason, the impartial morality of beneficence is not appropriate for parental reproductive choices.

Herissone-Kelly argues that parental reproductive choices should not be made from an anonymous, external perspective, through “imaginative inhabitations” of the lives of possible future people. Such choices should not rest on the kind of judgment yielded by the sorts of comparisons that, for example, Holtug and others find critical (that, e.g., p’s life would be better for p than q’s life is for q). Rather, parental reproductive choices should be made from an “internal” perspective. A parent who adopts an internal perspective imaginatively inhabits the life of a single prospective child. The child’s life, assessed through that perspective, may fall short of the minimum quality of life that the parent is willing to accept in a prospective child. But in declining to have a child on the basis of a “principle of acceptable outlook,” the parent is not making an interpersonal comparison of possible lives. The minimum set by that principle is not a birthright of the child like Steinbock’s “decent minimum.” Rather, it is a variable standard set by individual parents, reflecting their own conceptions of and commitments to parenthood.

The alignment of Lillehammer’s and Herissone-Kelly’s approaches is not straightforward. Herissone-Kelly’s external perspective appears to be impartial, but its interpersonal comparisons are restricted in scope: it takes no account of the impact of the creation of different possible children. And while his internal perspective appears to be person-involving as well as partial in requiring the agent to

focus on the life of a single future child, that future child may not be “identifiable” in Lillehammer’s sense.

David Wasserman challenges the idea that whether a person is wronged by an act depends in a straightforward way on the act’s expected consequences for that person. An act that wrongs a person can be said to be “bad for” a person even if it does not make things worse for, or harm, that person. He develops the suggestion, made by Kavka, Kumar and others, that the wrong in many nonidentity cases can be understood as a kind of disrespect for the persons created.²⁷

For Wasserman, the wrong that parents may impose on their own child, independent of any question of harm, is rooted in their intentions or attitudes in creating that child. The most familiar case in which parental intentions and attitudes are widely regarded as wrong-making features of the procreative acts is that of “savior siblings”—those children created in order to provide life-saving blood or other organic material to their already-existing siblings. Although the parents have every reason to expect that their offspring will lead good lives overall, advocates of intention-sensitive accounts will not, on that ground alone, endorse what the parents have done. They will focus, instead, on whether the parents’ intentions or attitudes in respect of those children at the time of their conception are disrespectful.

Wasserman argues that the population-affecting conduct of policymakers can in some cases be evaluated by the same standards. Like parents, policymakers may have partiality-based reasons to create people with whom they have certain affinities. If they act on those reasons, and if the people they create have worthwhile lives, those people have no grounds for complaint about the unavoidable hardships in their lives. The evaluation of the policymakers’ decision thus depends in part on the attitudes and intentions they had in bringing those future people into existence to begin with.

Wasserman’s extension of an intention-sensitive account from the procreative context to the policy context—from “micro” to “macro”—finds some support in Lillehammer’s analysis of the limited role that concepts of impersonal and impartial value have in helping us understand the nonidentity problem. Yet in claiming that partiality can sometimes play a legitimate role in “the kind of large scale decision making . . . involved in public policy or population ethics,” Wasserman holds that partiality may have a moral role in domains that Lillehammer sees as characteristically impartial.

5.6 Is the Person-Affecting Approach Cogent? Arrhenius

The person-affecting approach is widely associated with the intuitive notion that agents have no independent obligation to bring additional persons into existence—and, more broadly, the notion that the act that *merely adds* an additional well-off person to a particular outcome cannot make that outcome better. Accordingly, a person-affecting constraint in some form or another can seem like an important corrective to certain impersonal forms of maximizing consequentialism, which otherwise generate implausibly stringent procreative obligations. Earlier principles that

attempt to articulate such a constraint have been criticized by John Broome and others.²⁸ That discussion has in common with the nonidentity problem the underlying idea that we must acknowledge the obligation to bring the better off person into existence in place of the less well off person: we cannot plausibly conclude that it does not matter, morally, which of those two options we choose. The discussions are, at the same time, independent in an important way. The nonidentity problem relies on facts about *identity* to make this point; Broome's argument, in contrast, can assume that the better off person is identical to the less well off person and still show that some formulations of the person-affecting constraint are inconsistent. The upshot is that even if a person-affecting approach is somehow able to deflect the nonidentity problem itself, it will still face serious conceptual difficulties. In his new paper, Arrhenius focuses on a handful of more recent attempts to articulate a person-affecting constraint in a way that avoids inconsistency. What he finds, however, is that those attempts fail to deliver what their proponents promise and that they either have counterintuitive implications of their own or are compatible with traditional impersonal theories.²⁹ Arrhenius argues, in particular, that the best among such more recent accounts raises the spectre of the Repugnant Conclusion and fails to address in any plausible way tradeoffs between the welfare of "non-uniquely" and "uniquely" realizable persons. Arrhenius concludes that the shift to a person-affecting approach will not help us solve the problems in population ethics.

5.7 Law and Public Policy. Peters, Shiffrin

Philip Peters accepts the argument that the parents' choice to bring the genetically impaired child into existence does not *harm*—in the comparative sense—the child. But he also argues that that choice is, in many instances, morally wrong. According to Peters, when the disadvantageous condition—what he calls the "injury"—is serious and can be "avoided by substitution," agents have a moral obligation to take steps to avoid that injury. His position thus has affinities with Holtug's wide person-affecting account, which holds that it is wrong to choose a child who will benefit less from existence rather than one who will benefit more. The agents must either substitute the impaired child for a healthier counterpart or refrain from bringing the impaired child into existence to begin with.³⁰

Peters then focuses on the implications that the particular moral account he offers has for both constitutional privacy law and tort law. Within tort law and, in particular, the law of negligence, a valid legal claim will include the plaintiff's assertion that he or she has been harmed as a result of the defendant's tortious conduct. It is that requirement that has led many courts to reject claims for wrongful life.³¹ The impaired child, it seems, cannot cogently claim that he or she has been harmed by the very negligence that has allowed that particular child to come into existence. On the grounds that a main function of tort law is compensatory, Peters does not challenge the dominant legal analysis of wrongful life. To put the point another way: Peters accepts that tort law and, in particular, the law of negligence are to be construed in person-affecting terms. If an act is "bad for" no one at all, it is not an act

for which the defendant can be required to pay compensatory damages. That is so, Peters suggests, despite the fact that the plaintiff can be said to have been “injured” (under Peters’ substitutional account) and despite the fact that what the defendant has done may well constitute a *moral* wrong.

John Robertson has argued that a person-affecting interpretation is similarly appropriate within the constitutional realm. According to Robertson, important procreative choices that can be demonstrated to *harm no one*—no existing and no future person—are presumptively entitled to protection as an exercise of a fundamental right under the due process clause of the Fourteenth Amendment of the U.S. Constitution.³² For the same reason the wrongful life claim in negligence should itself be dismissed by the court, it is appropriate for the court to declare unconstitutional state regulation that is restrictive of the sorts of important, highly personal, procreative choices that are inherently harmless. What “compelling state aim” could such a regulation possibly serve, when we can demonstrate, in advance, that the conduct at issue *harms no one*—no existing and no future person?

In contrast, Peters argues that a person-affecting reading of the law of negligence is not appropriately extended to the constitutional realm. He argues that the fact that courts do and should reject victimless claims in negligence should not be taken to mean that courts will be similarly unimpressed when faced with the state’s argument that regulation of analogous procreative conduct is in order. That courts cannot be expected to adopt Peters’ substitutional criterion for “injury” on the tort law side thus does not, in his view, mean that they cannot and should not adopt that criterion for purposes of determining whether constitutional standards for state regulation have been satisfied.

Peters argues that precedent exists for a court to take a broader view on what constitutes a compelling interest. He thus notes that statutes prohibiting incest are uniformly upheld as constitutional, notwithstanding the fact that, according to the nonidentity problem, any offspring produced by such a union are not harmed. Peters concludes that the state’s interest in regulating conduct that leads to “injuries” that are common and serious, and that could have been avoided by substitution, should be recognized as compelling for constitutional purposes.

What, then, are we to say about reparations? If we accept the “no harm done” implications of the nonidentity problem and think that reparations for past social injustices have the same compensatory function that the tort law has, then the validity of reparations may seem questionable. After all, if the adverse situations of some contemporary African-Americans (for example) can be traced to their forebears having been enslaved, so can, according to the nonidentity problem, their very existence. If the purpose of reparations is to “make whole” the victim—the harmed person—and if there is no victim, reparations would not seem to be owed or, indeed, even appropriate.

Seana Shiffrin, however, agrees with Harman and Steinbock that the comparative account of harm that this line of argument relies on can be challenged. According to Shiffrin, persons can be harmed, and wronged, by acts that inflict sufficiently bad conditions or cross “boundaries of respect.” Such acts cause harm even if they also happen to have beneficial effects that counterbalance the burdens they impose.

As Shiffrin has elsewhere argued, if you are hit on the head by a gold bar dropped from the sky by an agent who intends to benefit you, the bare fact that you have been enriched as well as hit on the head does not imply—even if your injuries are minor—that you have not been harmed.³³

Shiffrin extends this position to the question of reparations in the present paper. “An insult to one’s self-respect does not disappear if it accidentally results in a financial windfall. One may be harmed or wronged even if in some large accounting book one is made no worse off overall.” As to whether the institution of slavery itself transgressed boundaries of respect and imposed conditions of material deprivation—in the case of both persons who were themselves enslaved and their progeny—Shiffrin has no doubt.

6 Conclusion

The contributions to this collection agree on little. Some affirm the notion that the wrongness of an act is rooted in how it affects particular people and simply accept the inference to “no harm done” that the nonidentity problem seems to support. Others affirm the intuition while trying to find a way to avoid the result that the act under scrutiny is “bad for” no one. Still others despair of ever resolving the nonidentity problem or, indeed, of even making sense of the person-affecting approach. Despite these differences, the contributions to this collection together illustrate the richness of the nonidentity problem—and, more generally, the puzzle of articulating just what we owe to future persons. Despite, or perhaps because of, the deep fissures it exposes, the nonidentity problem emerges as an exemplary vehicle for exploring some of our deepest questions about the nature and scope of morality, the meaning of harm and the goodness or badness of the lives of persons.

Notes

1. *Genesis* 1: 27–28.
2. Parfit (1987), p. 363 (some emphasis deleted). For purposes here, the class of all *persons* may include many non-human animals and will not necessarily include all human beings. This particular statement of the person-affecting intuition provides a necessary condition on when an act is wrong. On another interpretation, the intuition concerns not the moral status of acts but rather betterness between outcomes. On that interpretation, one outcome is better than another only if there is some person for whom it is better. Broome has argued that the latter “axiological” view (which suggests that adding a new person to an outcome cannot not make that outcome better but rather is morally neutral) is attractive but (for reasons independent of the nonidentity problem) objectionable. See Broome (2004), pp. 135–136; 145–146; and 202–208. See also Parfit (1987), p. 370. While both interpretations of the intuition are explored in this collection, more attention is given to the former than the latter.
3. Thus, contemporary virtue ethicists and Kantians will surely agree that the fact that one act yields better consequences for a person than another is morally significant and legitimately bears on issues of permissibility. Rights theorists, similarly, may think that the whole point of individual rights, and the guiding principle behind how we are to go about determining

what counts as a right (or a violation thereof), is the promotion of wellbeing for persons as individuals. And consequences will be important for contractarians as well: when a particular person is wronged by an act may very well be grounded, for the contractarian, in principles that have themselves been endorsed because they make things better for persons—not in the aggregate, but one at a time—rather than worse.

4. The term “wrongful life,” while used in different ways, is perhaps most often thought to convey the idea of a life that is less than one worth living or “worth not living.” (The latter is McMahan’s expression. See McMahan (1981), p. 100.) The suggestion is that existence itself may constitute a harm—on a comparative account of harm—in view of the fact that the child is so severely impaired that he or she would have been better off never having come into existence at all. (The zero level of wellbeing implied by never existing at all would be, arguably, greater than the negative level the child is in fact made to suffer.) As a practical matter, however, legal claims for wrongful life typically do *not* involve conditions that leave the child with such grim prospects. The term “wrongful disability” is usefully introduced in Buchanan et al. (2000), pp. 222–257, to describe that less severe class of cases. The nonidentity problem, similarly, involves situations where the child’s plight is not so dire—situations, that is, in which it can more plausibly be argued that bringing the child into existence does not, in itself, harm the child. It is in that latter sense that we use the term “wrongful life” here.
5. Parfit (1987), pp. 351–379.
6. Kavka (1981), p. 93 (describing the “precariousness” of existence).
7. Kavka (1981), pp. 94–95.
8. Parfit (1987), p. 356.
9. Parfit (1987), p. 361.
10. Parfit (1987), p. 361.
11. Parfit (1987), pp. 361–366 and 371–374, and Kavka (1981), pp. 93 and 98–101.
12. See Note 4 above.
13. We might say that such acts are subject to a distinctive moral account. Heyd referred to that distinctive account as “genethics.” See Heyd (1992).
14. Or the repugnant conclusion. See Parfit (1987), pp. 381–390 and 419–441. Additional issues raised by totalism and some other aggregative forms of consequentialism are discussed in Vallentyne and Kagan (1997), pp. 5–26.
15. Narveson (1976), p. 73.
16. The well-known “Principle N” is one example of an attempt to chart such a narrower impersonal approach. See Buchanan et al. (2000), p. 249.
17. Parfit (1987), p. 360. Some narrowly impersonal positions, such as Principle N (Buchanan et al. (2000)), adopt Parfit’s same-number condition. Other positions do not require that all the selections result in the same number of people *ever* living. They thus avoid the problem of determining whether alternative choices will result in the same number of people ever existing, or, indeed, whether any pair of choices can satisfy that condition.
18. See Velleman (2007), p. 277 (discussion of child’s “birthright”). See also Reiman (2007), pp. 78ff. and p. 92 (“future people have rights irrespective of which particulars they turn out to be”), and Woodward (1986), pp. 811–826. Persson articulates but rejects a rights-based approach here.
19. See Heyd (1992).
20. McMahan (1981), p. 100.
21. Zohar (1991).
22. Parfit (1987), p. 355.
23. Velleman, as well, has recently underlined the difficulties of determining cross-world identity. See Velleman (2007), pp. 221–244. See also Hare (2007), pp. 512–523 (exploring a concept of “de dicto” betterness and a “role-affecting” approach to the nonidentity problem).
24. See Mulgan (2006), pp. 154–159.
25. Harman (2004), p. 92.

26. In particular, the harm must be related to the act in a “morally relevant sense” and in a way that makes the agent “responsible” for that outcome. Hanser (1990), pp. 54–59.
27. Kavka (1981), pp. 97 and 106; Kumar (2003).
28. See Note 2 above.
29. One such approach has been suggested by Roberts. See Roberts (2003, 1998). Still other formulations conceive the person-affecting approach as a form of “moral actualism.” See Hare (2007), pp. 498–511. For further discussion of, e.g., Narveson, see Parsons (2003), pp. 150–154.
30. Peters (2004), pp. 27–39.
31. See Note 4 above.
32. Robertson (2004), pp. 13–24. See generally Robertson (1994), esp. Chapter 2 and pp. 75–76. See also Roberts (2008).
33. See Shiffrin (1999), pp. 120–135.

References

- Broome, J. 2004. *Weighing lives* Oxford: Oxford University Press.
- Buchanan, A., D. Brock, N. Daniels and D. Wikler. 2000. *From chance to choice: Genetics and justice*. Cambridge: Cambridge University Press.
- Hanser, M. 1990. Harming future people. *Philosophy & Public Affairs* 19: 47–70.
- Hare, C. 2007. Voices from another world: Must we respect the interests of people who do not, and will never, exist? *Ethics* 117: 498–523.
- Harman, E. 2004. Can we harm and benefit in creating? *Philosophical Perspectives* 18: 89–113.
- Heyd, D. 1992. *Genethics: Moral issues in the creation of people*. Berkeley: University of California Press.
- Kavka, G. 1981. The paradox of future individuals. *Philosophy & Public Affairs* 11: 93–112.
- Kumar, R. 2003. Who can be wronged? *Philosophy & Public Affairs* 31(2): 99–118.
- McMahan, J. 1981. Problems of population choice. *Ethics* 92(1): 96–127.
- Mulgan, T. 2006. *Future people*. Oxford: Clarendon Press.
- Narveson, J. 1976. Moral problems of population. In *Ethics and population*, ed. Michael D. Bayles, 59–80. Cambridge, MA: Schenkman Publishing Co.
- Parfit, D. 1987. *Reasons and persons*. Oxford: Clarendon Press.
- Parsons, J. Axiological actualism. *Australasian Jo. of Philosophy* 80(2): 137–147.
- Peters, Jr., P. 2004. *How safe is safe enough? Obligations to the children of reproductive technology*. New York: Oxford University Press.
- Reiman, J. 2007. Being fair to future people: The non-identity problem in the original position. *Philosophy & Public Affairs* 35(1): 71–92.
- Roberts, M.A. 1998. *Child versus childmaker: Future persons and present duties in ethics and the law*. Lanham, Maryland: Rowman & Littlefield.
- Roberts, M.A. 2003. Can it ever have been better never to have existed at all? Person-based consequentialism and a new repugnant conclusion. *Journal of Applied Philosophy* 20(2): 159–185.
- Roberts, M.A. 2008. What is the wrong of wrongful disability? From chance to choice to harm to persons. *Law and Philosophy* (pub. online May 17, 2008; forthcoming in print).
- Robertson, J.A. 1994. *Children of choice: Freedom and the new reproductive technologies*. Princeton: Princeton University Press.
- Robertson, J.A. 2004. Procreative liberty and harm to offspring in assisted reproduction. *American Journal of Law and Medicine* 30: 7–40.
- Shiffrin, S. 1999. Wrongful life, procreative responsibility, and the significance of harm. *Legal Theory* 5: 117–148.

- Vallentyne, P. and S. Kagan. 1997. Infinite value and finitely additive value theory. *Journal of Philosophy* 44: 5–26.
- Velleman, D. Persons in prospect. *Philosophy & Public Affairs* 36(3): 222–288.
- Woodward, J. 1986. The non-identity problem. *Ethics* 96: 804–831.
- Zohar, N.J. 1991. Prospects for “genetic therapy”: Can a person benefit from being altered? *Bioethics* 4: 275–288.

Part I
**Can Bringing a Person into Existence
Harm That Person? Can an Act That
Harms *No One* Be Wrong?**

Chapter 1

The Intractability of the Nonidentity Problem

David Heyd

Abstract The author, in this paper and elsewhere, defends a person-affecting approach to morality, according to which an act that harms *no one* cannot be wrong, together with the argument from the nonidentity problem that any act that adversely affects only those future persons who owe their existence to that act's being performed cannot properly be said to harm those future persons. Extending the logic of the nonidentity problem to cases involving not just strict numerical identity but "biographical identity" as well, the author argues that agents do nothing wrong when they raise a child under, or return a child to, a particular biographical identity, since a new biographical identity, even if more advantageous, would not make the one child better off but instead replace the one child with another child—a biographically nonidentical child—altogether.

Keywords Biographical identity · Person-affecting approach · Nonidentity problem.

1.1 The Logical and Metaphysical Dimensions of the Problem

Ethics and metaphysics have always been bound together in a philosophically problematic way. From the first chapter of Genesis and in many of the pre-modern metaphysical systems, the very existence of the world and its inner order were explained in terms of some ultimate good. No less prevalent was the attempt to explicate the concepts of the good life and moral virtue, justice and rights, in terms of human essential nature. But with the loss of confidence in, and consensus on, an overall metaphysical picture of the world, philosophers tried to articulate moral theories with no metaphysical foundations. Moral and political constructivism, of the kind articulated by Kant and Rawls, aimed at severing the classical bond of ethics and metaphysics. But it is far from certain that even constructivism can succeed

D. Heyd (✉)

The Hebrew University of Jerusalem, Jerusalem, Israel

e-mail: david.heyd@huji.ac.il

in such complete separation. For even if a system of rights and duties, principles of justice and concepts of the good can be purely constructed by reason, some metaphysical assumptions cannot be avoided. For example, both Kant and Rawls are committed to metaphysical individualism, i.e. the non-arbitrary consideration of individuals as the basic building blocks of the normative system. The concept of the individual person cannot itself be constructed.

One sphere in which metaphysics forces itself on moral theory is what I have referred to in the past as “genethics,” namely the cluster of problems relating to the creation of new people: the determination of their very existence, their number and their identity. These problems can be divided into two categories or levels: the species and the individual. Questions such as the value of the existence of human beings (think of God’s ante-diluvian and post-diluvian reflections) or the ethics of shaping the human genome (of the kind Jürgen Habermas is concerned with in his *The Future of Human Nature*)¹ belong to the first category. The issues of family planning, sex selection and demographic policies, which decide which individuals will exist and how many, belong to the second. But on both levels of discussion, some metaphysical questions cannot be avoided: in the case of the species, what are the contours of human nature (if there is any such essential nature)? In the case of individuals, how is a human being identified as a particular person and to what extent is such individuation relevant to the morality of procreation?

The nonidentity problem is one of the most succinct metaphysical challenges to moral theory. In retrospect, it seems surprising that it was not addressed by philosophers till the 1970s. The obvious explanation is that the numerical identity of those to whom moral judgment applies was naturally taken for granted and that only once humanity has acquired far-reaching control over procreation (by means of birth control, demographic planning and genetic screening) did the problem of nonidentity impose itself on moral theory. It is to the immense credit of Derek Parfit that he was the first to take up the challenge and not only articulate the problem but also show how fundamental and inescapable it was. Unlike the identity problem of the kind Locke was thinking of in the context of his discussion of responsibility and punishment, the *nonidentity* problem raises the question of the general limits of moral judgment.

The nonidentity problem can be mapped on a two-tier structure. On the first level lies the question whether the identity of persons is at all a relevant issue to moral judgments concerning the good, the right and the just. If the answer to this question is positive, a second-level question must be addressed, namely what kind of identity is presupposed by such judgments? The first question is conceptual, relating to the logic and conditions of moral judgment. The second is metaphysical and concerns the nature of those entities that are considered the carriers of value and rights. The first, preliminary question is the focus of the deep debate between the “impersonalist” and the “person-affecting” approaches to the nature of value in general. But then, once this debate is decided, the metaphysical question of what are the objects of value judgments and who are their subjects arises. More specifically, if we adopt a person-affecting view of morality, who are the relevant “persons” (affected)? I will first discuss the first question (in Section 1.2) and then proceed to make some comments on the second (in Section 1.3).²

1.2 Four Strategies for Responding to the Nonidentity Problem

The nonidentity problem presents ethical theory with a major challenge: if the consequences of our “genetical” choices are such that the affected future people are different in number or in identity from those who would have been affected had our choice been different, can we apply our moral principles (whether utilitarian or deontological, right-based or duty-based) to these choices? Although most of our moral choices remain unaffected by the nonidentity problem,³ modern science and technology have created a long list of important decisions in which nonidentity is an intriguing theoretical obstacle. Wrongful life cases, demographic policies, intergenerational justice, genetic engineering, sex selection through PGD are all concerned with future people under problematic identity. And, as philosophers have lately noted, there are also backward-looking cases like affirmative action or compensation and apology for past crimes which raise the problem of nonidentity.⁴ Should we compensate someone for wrongs done to her ancestors when it can be proved that *she* would not have existed had the wrong not taken place?

There are four principal ways to deal with the challenge of nonidentity:

1. Denying it is a problem to begin with.
2. Aspiring to solve it in some (yet unknown) integrative moral theory in the future.
3. Attenuating it so as to make it more palatable to our moral intuitions and theories.
4. Biting the bullet, i.e. accepting all the implications of the nonidentity problem.

The first strategy characterizes the view called “impersonalism,” which holds that value is not human-dependent but an attribute of the world. The second response is associated with Parfit’s own approach and his search for “Theory X,” combining person-affecting and impersonal intuitions, both of which are impossible to give up. The third way tries to adhere to a person-affecting view by interpreting it in a wide sense or by supplementing it with impersonal features. The fourth reply to the challenge consists of embracing all the consequences and ramifications of the nonidentity problem, including those which may be less appealing to our common intuitions, and doing so by adhering to a strict person-affecting view.

I will try in Sections 1.2.1, 1.2.2 and 1.2.3 below to examine in some detail the flaws of the first three responses and discuss in Section 1.2.4 the inescapable superiority of the fourth, thus demonstrating that nonidentity matters and makes genetical choices categorically different from choices regarding present or actual people. This discussion belongs to the conceptual analysis of the nonidentity problem. In Section 1.3, I will address the metaphysical question of the kind and scope of a person’s identity and explore the way in which this can serve to support the person-affecting approach.

A preliminary methodological comment might be worth making. The nonidentity problem is *sui generis*. It is unique in the sense that analogies from other contexts of moral judgment can be of little help. Since it relates to the sphere of the very creation of subjects and objects of moral values and rights, applying theories of value and rights to the act of their creation is logically puzzling like any bootstrapping feat.

Our established notions of legal harm and benefit cannot serve as guides in the genetical sphere, since the whole point of the challenge of the nonidentity problem is to show that they *presuppose* the existence of identifiable persons.⁵ We may nevertheless think of two possible analogies from which we might derive some insight on the matter—*divine* creation and *self*-creation, which are both “genetical” in their nature. The former, which is not discussed here, is an abstract and pure theological test case which has some interesting results with which I have dealt elsewhere.⁶ The latter has to do with the unique human capacity to control one’s own identity to some degree and in that respect create oneself. It will turn out below (1.3) to be a useful analogy in the discussion of the nonidentity problem.

1.2.1 Denying the Problem

The first response on our list to the nonidentity problem is the denial of its relevance. Even if goodness, justice and freedom can be attributed to the world only through the mediation of human beings who are their subjects, there is value in promoting goodness, justice and liberty in the world, independently of human beings. This might sound strange, since goodness, justice and liberty cannot exist without human beings. But this is exactly the genetical challenge to moral theory: is there a value, or even a duty, to create “carriers” of those values, namely human beings, so as to have goodness, justice and liberty in the world? Is a human-less world any worse than a human-populated one? Impersonalists answer the question in the positive. They attribute value “to the world.” A world of million happy people is better than a world with no people at all. And it is also better than a world with half a million equally happy people, even if these are completely different people. The question of the identity of the people in two alternative worlds which are compared for their value does not arise. *Who* is made happier by some beneficent act is unimportant as long as it creates more happiness in the world than any alternative act.

There is no direct way to rebut this view of value. Philosophers have brought up various examples that seem to make this approach unattractive, if not plainly absurd, like the famous “Repugnant Conclusion” or the duty to bring children to the world whenever that serves to increase the “total” happiness. But impersonal utilitarians are unperturbed by the nonidentity problem, since for them, even if the repugnant conclusion is an embarrassment, nonidentity is not; for whenever we have to choose between two options with the same number of people created in each option, we should follow the impersonal balance of utility and ignore the (different) identity of the individuals affected. But even the somewhat less disturbing case of the duty to promote the “average” happiness creates problems for the impersonalist, since it might imply serious restrictions on reproduction and a demand that each generation become more selective in the creation of new people than its predecessors. When it comes to justice, the impersonalist encounters even a harder conflict with commonsense moral thinking. For justice seems to be an ideal *for* human beings rather than some good *tout court*. Thus, when Rawls, in his later writings, argues that the principle of justice to future generations only requires the preservation of

the justice of the basic structure of future society (rather than any particular inter-generational distribution), it is not clear what kind of value justice is. On the one hand, it seems that it is good only if there are human beings for whom it is good. On the other hand, we could shape the interests of our descendants (either by educational or by genetic means) so that they do not appreciate the value of justice (e.g. are not concerned with fairness). Would such a decision be, according to Rawls, morally permissible? And if not, would it not commit him to an impersonal view of justice?

Impersonalism dismisses the charge of nonidentity by simply denying the person-affecting nature of value (rights, justice, equality, etc.). It is the world in general which is made better by good deeds, regardless of who are the carriers or subjects of that value. Indeed, value can be attributed to the natural world independently of human beings, as some advocates of the principle of biodiversity claim, and as long as human beings are around on the planet and exercise control over it, they have a duty to preserve the variety of species and the sustainability of the environment independently of its value for human beings. But these are views which are difficult to defend, and philosophers often appeal to non-moral (aesthetic or religious) notions of shame or loss (“it would be a pity if. . .”) as substitute for the moral grounding of such duties and values.

The impersonalist must identify what makes the world better (impersonally). It could be, as classical utilitarianism suggested, positive states of mind (such as pleasant experiences). But these do not lend themselves to easy individuation (can one answer the question “how many positive experiences did you have today?”). In response, the impersonalist could suggest that rather than aggregate positive and negative psychological states, we should aim to increase the overall happiness in the world by making people happier *and* by creating happy people. But this approach assumes that there is an objective way, independent of first-personal evaluations, to assess the overall balance of the happiness of a single person as well as to compare it to the balance of happiness of another person. This difficulty forces the impersonalist back either to the weighting and aggregation of individual experiences or to the way individual persons assess the happiness of their lives from their point of view. The first possibility raises the above mentioned problem of individuation of depersonalized experiences; the second makes the impersonalist vulnerable to the nonidentity problem. Impersonal utilitarianism tries to avoid both the classical problem of personal identity (the unity behind a series of discrete experiences) and the new problem of nonidentity (the incomparability of utilities of an actual person with those of possible, non-identical alternative persons). But such avoidance comes with a price.

1.2.2 Hoping for Future Solution

Derek Parfit is at pains to salvage his fundamental impersonalist intuitions from the challenge of the nonidentity problem. An action can make the world worse without wronging any actual human being. Unwilling to give up impersonalism but

recognizing the force of the constraints of identity of the persons affected by human action, Parfit tries to do justice to both. But after having meticulously examined various options, he admits that he has failed. All he is left with is a *hope*, an aspiration to find a “Theory X” which would integrate our conflicting intuitions and resolve the tension between the Repugnant Conclusion and the nonidentity problem.⁷ The final section of *Reasons and Persons* expresses a hope (which is reminiscent of the Kantian “dialectic” hope in being only partly grounded in reason) that “non-religious ethics,” which at present is only at an early stage, would make *progress* in the future and develop moral reasoning that would provide us with such Theory X.⁸ But can we even hope for such a theory?

Parfit suggests in that final section that the difference between killing 100% of humanity and killing 99% of it is much larger (i.e. is worse) than that between killing 99% of the humanity and killing no one. This is a very bold impersonalist statement, which Parfit justifies in terms of the loss of overall happiness involved in the cessation of all future human life on the planet and the loss of the potential of progress in art and science *and* (surprisingly!) in moral reasoning. The problem is that although Parfit explicitly dissociates himself from any religious conception of ethics, his approach seems to be committed to some teleological view. For, happiness (as well as justice, artistic beauty and scientific truth) is either good for actual human beings or for the world (a kind of realization of an essential potential which it would be wrong to curtail). But once there are no human beings who can satisfy their desires and perfect their abilities, can we say, without making strong teleological assumptions, that the world would lack something that could be good? It is hard to see how progress in science, art and ethics is valuable independently of human beings for whom such progress is a genuine interest, source of satisfaction or ideal. After all, the world has no interests. Parfit, who follows Sidgwick on that matter, is wrong in arguing that the destruction of humanity is the worst conceivable crime due to “the vast reduction of the possible sum of happiness.”⁹ Voluntary collective suicide of human beings is in my view less of a “crime” (if it is a crime at all) than the deliberate murder of people. Or to put the argument against the impersonalist analysis in person-affecting terms: there are no crimes against humanity; only crimes against humans.

So it seems that the hope for a Theory X is misguided since as Parfit himself has shown pure impersonalism with no regard for person-affecting considerations leads to absurd results but person-affecting considerations involve the insurmountable problem of nonidentity. There seems to be nothing that we don’t know “yet” about the matter and which through progress in moral reasoning we would be able to discover. The nonidentity problem is not a scientific or a mathematical problem which will be solved through further research or reasoning. It seems more likely that it will be dissolved rather than solved, and that this will happen when our notions of identity change and adapt to the new forms of control we can expect to acquire in genetics and in social policy over future human beings. So although this is a matter of speculation, it seems that the “solution” will occur on the level of the metaphysics of the subject of moral judgment rather than on the level of the logic of the ascription of value.

So although Parfit was the first to recognize the acuteness of the nonidentity problem, he aligned himself with the impersonalist position, at least in the sense that he resisted the option of a fully person-affecting view of morality. But unlike the impersonalist denial of the challenge, Parfit believes that moral theory cannot be complete without addressing it.

1.2.3 Accommodating Nonidentity

We turn now to the third response. As is typical of sharp moral dilemmas, philosophers as well as lay people feel the pull of both horns. Most of us shirk the counterintuitive implications of pure impersonalism (like the duty to create a huge number of barely worthwhile human lives as long as the *overall* level of happiness in the world is promoted through their sheer numbers). But limiting moral judgment to actual, identifiable individuals makes us feel equally uneasy. Parfit recognizes this conflict but leaves us only with the hope of some future theoretical solution. But many philosophers in the past two decades have suggested solutions to the nonidentity problem. They may be categorized into two groups: the first, those who believe that we should accept combined personal and impersonal considerations in moral judgment, at least in those rare “genetical” cases; the second, those who deny that such a compromise is coherent and advocate a revised form of person-affecting theory that can accommodate at least most of the difficult challenges of the impersonalist. I shall discuss the two in turn in the following Sections 1.2.3.1 and 1.2.3.2.

1.2.3.1 Combining Person-Affecting with Impersonalist Approaches

The authors of *From Chance to Choice* admit that they do not have a full solution to the nonidentity problem and that it is a very complex issue which can be decided only in the light of a broad spectrum of questions that lie beyond genetics.¹⁰ Nevertheless they do commit themselves to the view that despite the general validity of the person-affecting approach, there are cases in which an appeal to impersonalist principles is inevitable (such as the creation of “wrongful life” in personal reproductive decisions or the so-called “different number choices” in population policies). They accept Parfit’s argument that the child in his famous example of the 14-year-old girl who decides to conceive is not harmed, and they also claim that it is not wronged (even though the mother acts wrongly). They argue, in conclusion, for a principle according to which it is wrong to create *any* child who would suffer from a serious disability if that can be avoided without a high cost to the parents and without affecting the number of future people.¹¹ The adolescent girl should wait till she can conceive *another* child, later on, who will be happier than the one she contemplates conceiving now. They concede that this is a “non-person-affecting” principle.

This, however, is at most an *ad hoc* solution, which might appease our intuitive objections to some implications of the person-affecting view, but is not theoretically

satisfying in the sense that it does not provide a principled way of relating person- and non-person-affecting principles (as the authors admit¹²). Its *ad hoc* nature is manifest in the authors' claim that the impersonalist principle is to be introduced only when the suffering or the defect of the future child is "serious." But nonidentity is not a matter of the *degree* of harm or pain but a conceptual constraint regarding the conditions for making any moral judgment. And hence the authors' evaluation that causing a handicap to an existing child is a *more* serious moral wrong than causing the birth of a similarly handicapped child remains theoretically unexplained and appears more like a compromise which tries to pacify the challenge of nonidentity.

Being similarly committed to the importance of identity in moral judgment but equally reluctant to accept the full implications of the nonidentity problem, David DeGrazia, like the authors of *From Chance to Choice*, opts for a middle way. He takes seriously the nonidentity problem and argues that although intentionally conceiving a handicapped child rather than a healthy one is wrong, it is not a wrong *to* the child.¹³ But then what kind of wrong is it? One way to describe it (which he correctly ascribes to me) is that the parents wrong themselves or society. But since DeGrazia strongly rejects that kind of explanation of the wrong, all he is left with are impersonal considerations which suppress the nonidentity problem. Consequently, like the authors of *From Chance to Choice*, DeGrazia suggests supplementing the person-affecting principles with an impersonal one: in Same People choices, apply person-affecting considerations; in Same Number choices, apply impersonal considerations. But this solution seems again to be *ad hoc*, an attempt to explain our intuitions about the wrongness of the choice of the 14-year-old girl without embracing a fully impersonalist view. DeGrazia himself admits that he does not have a theoretical explanation for the relationship between the person-affecting and the impersonal supplementary principles in ethical theory and that when it comes to Different Number Choices the right way for ethical theory to approach them "remains mysterious."¹⁴ Even if he is right in considering the person-affecting principle as having more moral weight than the impersonal, how, for example, would he compare the one-time wrongness done to an actual individual with the impersonal bad consequences created to a thousand future, non-identifiable people who are going to be born as the result of some social choice we make? Can the two be compared in the first place?

Jeff McMahan is also working within a generally person-affecting view. In the end of his detailed critical examination of Peter Singer's "replaceability argument," he reaches the conclusion that choosing between an existing (suffering) newborn and another future (happy) child can be made on person-affecting grounds. That is to say, Singer's problem of infanticide should be decided in terms of the interests of existing people (weighing the interests of the existing child in going on living and those of the parents and society) rather than on impersonal grounds of the kind Singer appeals to (according to Singer, infanticide can be justified in terms of the overall increase in the impersonal balance of happiness in the world achieved by the "replacement" of the suffering newborn with a future healthy child).¹⁵ But then McMahan is very sensitive to the "notorious" difficulty to defend the asymmetry between the duty to prevent the birth of a suffering child and the absence of a parallel

duty to cause the birth of a happy child. If consistency is sought on this matter, we will be forced to accept that the fact that a person would have a good life serves as a reason to cause that person to exist, even though this reason is weak, weaker than that assumed by Singer.¹⁶ But again, one may wonder why there should be a difference between the weight of the reasons to avoid the conception of a suffering child and those of the reasons for creating a happy child. McMahan himself asserts that he doubts whether the person-affecting and the impersonal intuitions we have can ever be reconciled. I contend that the consistent solution to the asymmetry problem is to deny that there are moral reasons (relating to the rights and interests of the future child) either to create it (happy) or not to create it (suffering).

1.2.3.2 Widening the Scope of the Person-Affecting Approach

It seems then that if pure impersonalism is incompatible with some of our fundamental moral views and if there is no way to incorporate it with person-affecting principles in one integrated theory (either in Parfit's sense of Theory X or in some compromise of the kind examined in the previous section), we are left with the option of the person-affecting approach. But the nonidentity problem presents us with a serious challenge, which threatens to undermine some equally fundamental views in the morality of procreation, demographic planning and environmental policies. Many philosophers have tried to meet the challenge without abandoning the person-affecting view or resorting to impersonalist supplements.

One strategy is to view future possible people *as if* they were all actual, or, in Parfit's terms, to consider "different people choices" as if they were "same people choices." Tim Mulgan believes that such an "as if" approach would lead to overall better reproductive choices from an impersonal point of view, thus avoiding the trap of the nonidentity of possible people which seems to lend us permission to create whomever we want. But Mulgan himself is aware that like any "as if" policy of this kind, this approach can work only if people are not aware of it! So even if this manipulative strategy is pragmatically useful, it obviously cannot serve as a theoretical response to the challenge of nonidentity. In any case Mulgan's idea indirectly proves how strong our person-affecting intuitions are and that the impersonalist consequentialist results of the kind Mulgan wishes to attain are best served by a person-affecting illusion.¹⁷ Like the attempts to integrate person-affecting and impersonalist principles, discussed above, Mulgan wishes to give moral standing to both actual and possible people. But since he believes the former have *more* moral weight than the latter, he recommends that we think "as if" all future people are actual.

Another strategy to save the person-affecting approach from the challenge of nonidentity is taking *types* of people rather than individuals as the objects of evaluation in genesis choices. It is usually agreed that harm can be done only to actual people who have undergone some loss due to the harm done *to* them and hence that wrongful conception cases cannot make strict legal sense in tort law.¹⁸ But some philosophers, like Rahul Kumar, argue that a child born in such circumstances can be said to have been *wronged*, even if not harmed. Although the child is not worse off

than he could have been, his or her respect has been violated since wronging (in contrast to harming) is a matter of the agent's character rather than of the consequence for the other party. But why is not wronging subject to the same constraint of the identity of its "victim" in exactly the same way as harming is? Kumar answers that one can be wronged as a *type* of person, so that I have similar duties to my future (unidentified) children as I have to my students (whoever they are). And society in general has such duties towards future generations, when the type of person to which these duties are owed is the rational, autonomous individual of a Scanlonian contract. Such a person-type has a legitimate expectation of respect that should not be violated even if no harm can be identified to any person-token.¹⁹

This is an interesting attempt to preserve the person-affecting view while escaping the paradoxical implications of the nonidentity problem. However, it faces some serious difficulties. First, if wronging (in contrast to harming) has to do exclusively with the character of the *agent*, how can it affect the "receiver" of the action in any way, including the violation of her respect or dignity? And if it does affect her, does it not involve making her "worse off" than she was or could have been? This suggests that we either view wronging as having the same structure as harming in the effects on another party or consider it as a matter of the agent's character and accordingly judge it in terms of the way the agent alone is affected. In both cases the person-affecting view is upheld but only on the basis of assuming the identity of an actual person (agent or victim) who is affected by the action. Secondly, Kumar's contractualist abstraction of actual individual people into idealized types of agents who have some universal properties is indeed a powerful theoretical tool in the device of a hypothetical contract, but it turns out to be based on confusion between a hypothetical contract and a contract between hypothetical people. Teachers and students can form contractual agreements like employers and employees, but the contract is not between types of people but between actual individual people, present or future, known or unknown (albeit of a certain type).

To examine this confusion more closely, consider Jeffrey Reiman's proposal for the solution of the nonidentity problem within a person-affecting view. According to his view, "future people have rights irrespective of which particulars they turn out to be."²⁰ Reiman correctly points out that Rawls' veil of ignorance hides the property of one's temporal position (namely, to which generation one belongs) in the same way as it conceals sex, race and social position. But then he proceeds to argue that in the original position what matters morally speaking are only the *properties* future individuals are going to have rather than their particular, i.e. *numerical identity*.²¹ This, I believe, is a wrong reading of Rawls' idea of the original position, or indeed of any social contract.

The social contract can only be made by *actual* people, who in order to create fair conditions for their bargaining, place themselves under a virtual veil of ignorance. This veil conceals all their particular properties (that might prejudice the way they choose the principles of justice), but *cannot* hide the fact of their very existence (or actuality). Indeed, the contract can (and should) also include future people who do not exist *yet*, but they must be *particular* individuals who are going to live anyway. Or, in other words, the hypothetical contract cannot take place between *possible*

people and cannot include principles for their own creation or the choice of their identity.²² Indeed, Reiman says that the contractors “represent all and only those people who, from this moment on, will ever exist: people who are currently living, and future people who do not yet exist but who one day will,”²³ but according to his own argument he cannot mean by that all numerically identifiable people who will actually live. For his whole point is to show how by deciding to cause the creation of a handicapped child rather than of “an individual” with better properties, we violate the rights of “the individual”. But there is no one “individual” here, whose rights are allegedly violated, but rather two numerically distinct individuals: one who is a healthy individual and *another* who is handicapped. The Rawlsian contractors are not abstract, identity-less place-holders, but particular individuals. We should not confuse the level of the background conditions of the contract (actual people trying to agree on principles of justice) with the level of the procedure of the contract itself (the veil of ignorance as a device of representation). On the first level the motive to “enter” a contract is the wish to promote *my* interests and prospects (not that of a numerically distinct individual even if she is like me in all her properties). Numerical identity is a condition for ascribing not only rights but also the interest in having a normal functioning.²⁴ Thus, in contrast to Reiman’s thesis that according to Rawls *any* future individual has rights against us, I maintain that Rawls’ contract creates rights only for actual people who are either living or are going to live anyway in the future, i.e. whose identity is fixed (in the strong numerical sense). The idea that people have the right to be born healthy or with normal capabilities is incoherent, for no contractor would choose a principle of justice which would lead to the birth of someone else in his or her stead (even if that individual would be better off). Creating children with the good properties might be a noble goal, but it can only be grounded in an impersonal (rather than contractual) conception of justice.

For Parfit, the only way around the nonidentity problem was impersonal, through a principle of beneficence. Rights, unlike overall welfare, seem to be more typically linked to metaphysically identifiable people who are the subjects of the rights. An attempt to save the person-affecting approach and avoid impersonalism lies therefore in showing that the wrongness of conceiving a child at the age of fourteen is connected with rights. James Woodward argues that since the violation of rights (again, unlike harm) does not necessarily involve a decline for the worse in one’s welfare, future people can claim that their rights have been infringed by the act of their “wrongful” creation even if their inborn and foreseeable handicap is offset by an otherwise happy and worthwhile life.²⁵ His main argument for this claim is from analogy: when an air company refuses for racist reasons to sell a ticket to a member of a minority group, the person may claim that his rights were infringed even if it turns out that he benefited from the refusal due to a crash in which all the passengers of the plane were killed. But the analogy does not work since it ignores what we referred to above as the *sui generis* nature of genesis choices. The crucial difference is that the individual discriminated against is a fully identifiable person *at whom* the offensive refusal was directed, while the possible child of the 14-year-old girl is not. The issue of the overall balance of good and bad, or the weighting of the infringement of rights vs. the benefit in welfare, is irrelevant to the logical

question of the conditions that make *any* such ascriptions possible. Similarly, it is wrong to compare the 14-year-old girl's action to that of a person making to another person a promise he cannot fulfill. For in the first case, that of creating a new person, parents do not make any promise to anyone (but only assume a future responsibility). Methodologically, the use of these analogies simply begs the question of the relevance of identity to the ascription of rights.²⁶

Another attempt (similar to Reiman's) to save the person-affecting approach from the quandary of nonidentity is through widening the scope of what is considered "person" so as to include all people affected by our acts "whoever they turn out to be." Melinda Roberts offers a sophisticated version of such a person-affecting theory. She contrasts it to my own narrow person-affecting approach that considers only actual existing people or future individuals who are going to exist "anyway" as objects of our moral duties.²⁷ According to Roberts, planting a piece of broken glass in my garden is wrong in being potentially harmful to future (yet non-existent) children, whether they are my neighbor's children (over whose creation I have no control) or my own (over whose creation I have control), i.e. whether they are (future) actual or merely possible people. But the nonidentity problem is a challenge to the person-affecting view in more specific circumstances, as Roberts seems to concede, namely when the act of planting the broken glass is *itself* the cause of the conception of a *particular* child. In such cases, known as wrongful conception, there is no way in which the particular child who gets injured could have been born without being injured by the glass. Of course, this is a bit far-fetched, but Roberts herself is willing to consider such a possibility,²⁸ and it is exactly the case of Parfit's 14-year-old girl or many of the legal complaints for wrongful conceptions. So I agree that the identity condition does not imply any knowledge of the actual identity of a future person and not even the total lack of control over his or her creation. It only applies to cases in which, in Roberts' language, a world with a particular person existing and without a certain adverse condition affecting him is *not* "accessible," or in simple words, the person would not and could not have existed without that adverse condition taking place.²⁹

Roberts' thesis is that the nonidentity problem applies only to "two-alternative" situations (a child can be either born in defect or not born at all), but not to "three-alternative" situations (a child can be born in defect, not be born at all, or be born healthy). This sounds plausible, even compelling. For example it can demonstrate why giving birth to a child in order to sell him as a slave (Kavka's case) is wrong: the same child could be born without being sold as a slave (such a possible world is "accessible").³⁰ Roberts' original point is that it even can show in person-affecting, non-Parfitian terms why a depletion policy is wrong: for there *is*, at least theoretically, an alternative in which the same child could have been born better off without the depletion policy being implemented. However, it appears that Roberts must view the 14-year-old girl as a "two-alternative" example. For, if the girl waits before conceiving the child, there is logically no way in which the *same* child could have been better off. From the point of view of nonidentity, what is the difference between this case and creating a child with a genetic defect which Roberts concedes belongs to the "two-alternative" category? Since Roberts agrees that "two-alternative" choices

are immune to moral criticism in terms of the wrongness *to* the child due to nonidentity, the disagreement about the response to Parfit's puzzles relates only to the scope of cases which belong to that category. Roberts' argument is convincing in that, due to causal possibilities regarding the coming to be of children and to metaphysical considerations relating to their individuation, the scope of the "two-alternative" choices is not as wide as we might have thought. But she shares the narrow person-affecting view (even though not easy to swallow) that conceiving a handicapped child, even if a healthy sibling could have been created in its place (by PGD), should not be considered as wronging the child.³¹

Impersonalists like Parfit believe that there is no difference between the moral standing of the child claiming damages for having been born in defect as the result of neglect during pregnancy and that of the child who is born as the result of negligent counseling before conception. Person-affecting philosophers tend to judge the cases as different, but feel uneasy about it. One way in which they justify the difference is by appealing to cases in which life is *very* bad or even not worth living. Thus, some of them hold that life as such can be good or bad for a person and that if it is bad on the whole, it would be wrong to create that person. They consider non-existence as having zero value to the person and life with certain serious defects as having negative value.³² Since zero is better than a negative value, non-existence can be said to be better for the person than living in severe handicap. The problem with this analysis is that non-existence is given *a value* (zero), although there is no one to ascribe it to. Non-existence is neither good nor bad nor neutral for anyone, since good and bad can be ascribed only to metaphysically identifiable individuals. But zero is the balance or cutting point on a scale between good and bad. We cannot say that someone who has no bank account can be considered as having a zero balance! For there is no person or bank account to which we can ascribe the value zero.³³ Again, the nonidentity problem is *sui generis* in the sense that the comparison (whatever its merits) between having a bank account with a debit and having no bank account to begin with does not serve as an analogy to the comparison between life with great suffering and no life to begin with. Indeed, we can think of a world in which some actually existing individual (including myself) does not exist. But this person cannot say that such a world would have been better for him, because no value whatsoever can be ascribed to that person.

1.2.4 Embracing the Implications of Nonidentity

We have so far shown why the attempts to either combine impersonalism and person-affecting principles or to re-interpret person-affecting principles in a way which would overcome the nonidentity problem fail. Pure impersonalism cannot be said to fail in a similar way, that is to say, it is a coherent and systematic approach to "genetical" choices and avoids the issue of nonidentity by denying its relevance to the evaluation of people's coming into existence, their numbers and their identity. But the price of impersonalism is high: it leads to the Repugnant Conclusion, it implies a duty to procreate (happy children) and it commits us to the judgment that

the world after the evolution of humanity is a better place than the world preceding it. To many people these are unpalatable consequences. The alternative to impersonalism with which we are thus left is the fourth strategy in the list of responses to the nonidentity problem, namely adhering to a strict person-affecting view.

The advocate of a strict person-affecting view is not indifferent to wrongful life cases and believes that there are moral reasons for avoiding the intentional creation of handicapped children or a reckless population policy. David Wasserman, for example, correctly points out that the moral constraints on the creation of children include the purpose for which the parents decide to have a child. Accordingly, it would be permissible to conceive a retarded child, even when another, healthy child could be created had the parents waited for a while or selected another embryo in a PGD procedure, as long as one of their *reasons* or motives included his own good. That is to say, although prospective parents cannot (logically) create a child “for its own sake”, they are expected to be motivated by the prospect of parental giving and concern for their future child’s good. Lack of sensitivity to the child’s future suffering is morally repugnant since it violates the general expectation of parents.³⁴ For Wasserman, the parents are subject to duties derived from the “role morality of parents.”³⁵ I would put it slightly differently. What counts in our moral judgment is the parents’ moral profile, the kind of people they are, rather than the objective condition of the child (or the world). And it is of course also true that if parents are completely indifferent to the welfare of their planned future child, they are liable to become bad parents and to violate their parental duties *to* the child once she is born.

To reinforce Wasserman’s approach, consider two couples: the first can *only* conceive a seriously ill child due to a permanent genetic condition from which they suffer; the second can conceive *now* a child who will be equally seriously ill, but if they postpone conception for a year, they will have a healthy child. Now, from the point of view of the child, there is no difference between the two cases, since both children are born into an equally painful life. But we do judge the parents’ choice in the two cases differently, harshly condemning the second while approving or at least sympathizing with the first. This difference indicates that the judgment of procreative choices is made with regards to existing people, usually the parents. The first couple’s choice is not merely excusable; it is morally understandable, even noble (if they are committed to take good care of the child once it is born). The second couple’s choice is perverse, even “sick,” and reflects a deformed character of people who are insensitive to pain and suffering or even derive satisfaction from it. The “positive” counterpart of that example is a case in which there are two options for giving birth to a healthy child. Our proverbial 14-year-old girl is offered two options to avoid the plight of her prematurely conceived child: wait another few years (as in Parfit’s example) or let another (more mature) woman give birth to a child “instead of you.” The latter offer sounds of course absurd, though from the point of view of the future child there is no difference between it and the first offer, since both children are going to be different from the originally planned child. The difference between the two options (which is of course significant) relates only to the girl, whose interests of satisfying motherhood can only be fulfilled by the first option.

Advocates of the strict person-affecting view, like myself, must face some consequences which are definitely counter-intuitive. They may nevertheless have reasons to “bite the bullet,” so to speak. This could mean the willingness to take moral responsibility for action (or policy) based on the strict person-affecting view with all its ramifications for future children and future generations. It could alternatively mean a skeptical view regarding the theoretical possibility of a normative justification of procreative (genetical) prohibitions. This skeptical attitude is compatible with leaving the actual moral and political choice to be guided by intuitions and public perceptions even when these are confused and inconsistent. For anyone who is not a stringent impersonalist and who is convinced by the arguments against a diluted or compromised version of the person-affecting view, the strict or narrow person-affecting analysis seems to be the lesser theoretical evil in being both consistent and doing justice to some of our fundamental intuitions.³⁶

It is important to note that conflicts of intuitions on the way to deal with the non-identity problem can arise on two levels. They can refer either to the content of the intuition itself or to the way an agreed upon intuition is explained. Thus, the person-affecting intuition that the world was not made any better by the evolution of human beings stands in direct contrast to the opposite intuition held by the impersonalist (who would also bemoan the painless and voluntary disappearance of humanity). On the other hand, the wrongness of a free and intentional choice of a 14-year-old girl to conceive a child is a shared intuition by personalists and impersonalists alike. They only disagree about the *grounds* or the explanation of that intuition. To the former, the wrongness derives from the way the decision reflects on the mother’s character, the irrationality of her act in terms of her own interests, or the burden it creates for society. For the latter, the wrongness lies in the bad consequences of the decision *tout court* (the world is made a worse place than it could have been). Unlike the direct conflict of intuitions of the first type, these differences in explanation are partly intuitive but partly theoretical. Hence there are better chances of engaging in a theoretical discussion about the second kind of conflicts than about the first.

But the appeal to intuitions in the morality of procreation (population policy, genetic engineering, etc.), although so central in the debate, is problematic. These intuitions are often confused due to the *sui generis* character of this problem and the difficulty in drawing analogies to it. Furthermore, they are not always sharp since the problem is not only theoretically unique but also historically new, and intuitions take time to form. Since much of the debate about remote examples such as Parfit’s takes place among philosophers, it should not come as a surprise that the intuitions appealed to, even if clear, are heavily colored by theoretical considerations.

1.3 Numerical, Biographical and Autobiographical Identity

If moral judgments about the creation of people must be based on person-affecting considerations, the question remains as to the identity of persons. This raises the second-tier question about the *metaphysical* conditions of the relevance of the non-identity problem. What is this “person,” the identity of which is a constraint on the

ascription of right and wrong in genetical choices? Who is the he or the she who counterfactually could have been better off had the choice been different? Logically, the person-affecting approach is not restricted to individualist conceptions of a person (as a carrier or subject of value). A collective, like a tribe or a nation, a family or a kibbutz, can be the object of duties and rights. But the nonidentity problem would equally apply to these entities: it would be incoherent to claim that it would have been better had one of these groups not existed in the first place and an alternative group created in its stead. For the same person-affecting question would arise: good for whom? However, beyond the problems of attributing the status of a moral subject to a collective, empirical circumstances make decisions on the creation or the identity (let alone the number) of future communities or collectives very rare (Moses in Egypt?).³⁷ So in the rest of this article the discussion will be limited to individual persons.

Since human beings are self-conscious and free, their identity is unique in being a combination of some general essential features and some constructed or self-constructed traits. Thus, my genetic makeup is essential to my personal identity, but being loyal to my nation might be essential “to me.” Some theorists view the distinction between sex and gender as illustrating this double nature of identity. A person’s identity begins to be formed at the moment of conception, but continues after birth through the powerful forces of socialization and education and later through the “big” choices the person makes for herself. The geneticist of the future definitely has control over the identity of people, but so do parents of young (or not so young) children, and later the individual adult who is forming and transforming her own personality.

Recent philosophical literature refers to this unique feature of human identity by distinguishing between numerical identity and narrative identity, or between personal identity in the traditional metaphysical sense (as in Leibniz or Kripke³⁸) and biographical identity. (I will use the term “biographical,” since it can refer both to the aspects of identity that are constructed by others and to those constructed by the self; furthermore, the term “narrative” in the description of a person’s life is misleading in assuming that human life is similar in its construction to that of a story or a novel). In his fine discussion of the distinction, David DeGrazia points out that essential, numerical identity (*de re*) must precede and is assumed by “narrative” identity (*de dicto*).³⁹ Numerical identity is fixed by a particular event and at a certain moment (be it conception or some time around it). Biographical identity is gradually formed throughout the person’s life, from childhood to adulthood, by parents, society and the person himself. But there is a point in constructing or self-constructing a biographical identity only *of* a numerically distinct entity (person) whose biography it is. Hence the logical precedence of the numerical over the biographical.

Narrative identity is taken (for example by DeGrazia) to relate to the way an individual forms her own identity as a person. And indeed this is an important constituent in what is important for us in continuing to be what we think we basically *are*. It is a first-person perspective on personal identity, a matter over which the subject has authority. It applies both prospectively (what kind of person I want to be) and retrospectively (how I interpret the kind of person I have been). But biographical

identity is not only first personal and there is an important additional dimension of identity between the natural or essential identity of human beings and their autobiographical or first-person identity. Take, for instance, the dilemma facing charitable gentiles who saved Jewish children during the Holocaust. Should the children be raised as Christians (like their adopting parents)? Should they, after the end of the war, be “returned” to a Jewish environment (assuming that their biological parents died)? There is no question about the numerical identity of the children on the one hand, and they have not yet acquired the power to autonomously decide their own identity on the other. But they are definitely Jewish in some deep sense.

Now the application of the nonidentity problem to this case is not as clear as in pre-conceptive decisions like the 14-year-old girl exactly because it relates to biographical rather than to numerical identity. Consider how we should solve the dilemma if we take the principle of “the best interests of the child” as our decisive guide. There is a sense in which, given the numerical identity of the child, we can say that it would be in this child’s best interest to stick to her Christian lifestyle to which she has been exposed during the war years. It would be definitely easier for her since she only vaguely remembers her Jewish origin or even forgot it altogether. Yet, there is a strong argument for judging the child’s best interests as remaining Jewish since that is the way she was treated in her early phase of life. In other words, the *person* “affected” by the decision might be considered either this (numerically identified) *child* or this (biographically identified) *Jewish* child, in a way leading to two opposite conclusions. In the former case the religious or national identity of the child is created *ex nihilo*, with no moral constraints, like in the standard pre-conceptive “genetical” choices. Due to the “nonidentity” of the pre-war Jewish child and the post-war Christian child, there are no moral constraints on the decision to preserve the Christian lifestyle of the child.

When we move from biographical to *autobiographical* identity (which comes closer to the idea of narrative identity), the application of the nonidentity problem becomes even murkier but equally instructive. For human autonomy, usually exercised in the pursuit of the ends of a person of a certain kind, might, at least in extreme cases, serve to change the kind of people we are. Conversion is a conspicuous example and so is sex (or gender) change. At least from the first-person perspective, people report that they have become different persons, following the transformative change (although they obviously realize that they have not changed “numerically”). Autobiographical identity is a matter of identification, of what I find crucially important in my life, that without which my life would be meaningless or not worth living. Culture, religion and moral character are typical examples of such identity-fixing traits (which are not essential to my numerical identity).

The nonidentity problem, accordingly, casts doubt on the logical basis of complaints about biographical (and even autobiographical) identity. My claim here is that wrongful identity complaints are no more coherent than wrongful life suits due to the nonidentity condition. I cannot regret not having been born in the eighteenth century or to different parents because it would not have been (numerically) me. But equally it would not make sense for me to criticize my parents for having been *brought up* as an Israeli, secular Jew rather than as an English Anglican priest, even

though from some abstract, impersonal perspective, the latter identity would have made my life easier or more successful according to some impersonal criterion. The reason is that it simply would not be me in some deep sense which suffices to make that comparison absurd or at least senseless from the point of view of what is good for me. Indeed, my parents could raise me (the numerically identical person) to become a different sort of person, but since the way I evaluate my good is partly informed by my biographical identity, such an alternative cannot be judged as better or worse for me.⁴⁰

In that sense of biographical identity, I am in the position to make genetical choices about my own future in the same way as my parents did for me. A decision, for example, to convert to another religion or to immigrate to another country (and adopt another culture) must be based on person-affecting considerations, namely my *present* actual values rather the good of my “future self.” For even though my numerical self can have alternative biographical identities, those cannot be fully compared to each other. How can Paul compare his life after the conversion on the way to Damascus with that preceding it: in terms of his previous Jewish values or in terms of his new Christian vision? Even if he believes there are objective (impersonal) reasons for preferring the post-conversion life, it is hard to justify that preference on person-affecting grounds. Or, to take a more mundane example, can a person regret having chosen a life-long career as a philosophy professor rather than embarking on that of a politician? After all, his present preferences are to much extent formed by the professional identity he actually formed and to have chosen otherwise cannot be considered as either better or worse for “him.” The limits of our complaint to our past selves are logically similar to those we have towards our parents for having either bequeathed us with their genetic makeup or for having formed in us a particular cultural identity.

Since biographical and “narrative” identity is subjective and a matter of degree, these examples are naturally controversial and from a third-person point of view might be considered overblown. And indeed, from an objective point of view, it would be coherent to criticize even a major choice in another’s life as a mistake in terms of *her* overall interests. Paul could thus argue that his post-conversion life better fitted his true or genuine identity (his character, dispositions, personality). But note that such a criticism would have to relate to some underlying features in one’s identity which are fixed and stable throughout the person’s life. In other words, *if* and to the extent that we take biographical identity seriously, we must concede that the evaluation of individuals’ choices during their lives are susceptible to the challenge of nonidentity in the same way as wrongful life suits are in the case of the creation of numerical identity.

The metaphysical question of the identity of human beings should be clearly distinguished from the question of the beginning of the life of a person. Numerical identity is probably formed some time after conception (for example, after the moment of possible twinning of the fertilized egg), but that does not mean that this is the point in which a human organism becomes a person (which as many philosophers have shown depends on the acquisition of some advanced mental powers, consciousness, or some other traits). Therefore, the issue of wrongful life (or

wrongful identity) should be set apart from that of abortion. It is perfectly consistent to hold that a woman has a right to abort her fetus but is prohibited from harming it. The reason is that the embryo, not being a person (yet) has no rights (including the right to life), but the future existing person will have a cause for complaint for having suffered harm during pregnancy since *he* could have been better off without suffering it.⁴¹ He is definitely numerically the same entity as the embryo from which he developed. Nonidentity does not undermine this kind of grievance. In that respect, “person-affecting” is not restricted to “persons” as human beings with full moral standing.⁴²

1.4 Conclusion

I hope to have shown in this article that the nonidentity problem is *intractable* in both the sense of being “stubborn” or difficult to handle and in the (etymological) sense of being impossible to “draw,” extract or remove. Like most genuine philosophical questions, this is exactly what makes it fascinating and deep, in both theory and practice.

Acknowledgements Faculty and Fellows of the Department of Bioethics at the NIH were generous and very helpful with numerous incisive comments, oral and written, on this paper.

Notes

1. Habermas (2003).
2. I have dealt in the past more with the first than with the second question, but most of the arguments which I critically examine in this article were raised after my book was published. See Heyd (1992).
3. If we adopt a global consequentialist view, every action in the world can easily be seen as having some effect on the timing of conception of future people, which would mean that the challenge of nonidentity is much more pervasive than we usually think.
4. Thompson (2000), pp. 470–475.
5. Admittedly, this presupposition is not held by impersonalist theories of value. It should also be noted that the argument for the uniqueness of the nonidentity problem in terms of the person-affecting theory of value is circular and that one of the best arguments for that theory itself lies in our intuition about the nonidentity problem. So there is nothing logically inconsistent in impersonalism together with denial of the challenge of nonidentity. But once the person-affecting view is accepted, it becomes clear why for conceptual reasons there are no analogies to the ethics of creation of people.
6. Heyd (1997), pp. 57–70.
7. Parfit (1984), p. 405.
8. Parfit (1984), pp. 453–454.
9. Parfit (1984), p. 454.
10. Buchanan et al. (2000), p. 247. The nonidentity problem is discussed in pp. 245–255.
11. Buchanan et al. (2000), p. 244.
12. Buchanan et al. (2000), p. 250.
13. DeGrazia (2005), p. 274.

14. DeGrazia (2005), pp. 277 and 279. Parfit holds that there is no (moral) difference between not taking certain pre-conceptive measures to ensure the health of a future child and not taking similar measures after the child is born. I contend that there is a radical difference between the two. DeGrazia suggests that there is “a *slight* difference,” thus trying to follow Parfit’s acceptance of impersonalism without denying the grip of the nonidentity problem. However, I don’t see how this kind of difference can be a matter of degree.
15. McMahan (2002), p. 357.
16. McMahan (2002), pp. 353–354.
17. Mulgan (2006), pp. 155–156. Mulgan advocates rule-utilitarianism as a middle position between impersonalism, which absurdly requires the creation of happy children, and the person-affecting view, which equally absurdly leaves us free to create a suffering child even if we can avoid it. But his solution seems to belong to the attempts at combining person-affecting and impersonalist principles of the kind discussed in the previous section. For a similar proposal that the government act *as if* future generations have rights over us, see Marc D. Davidson (forthcoming). Davidson’s reason for his proposal is that most people believe that future people can be wronged, i.e. do not think in terms of the nonidentity problem.
18. For a view according to which the wrongful creation of life can be considered as a direct harm to the individual created, see Shiffrin (1999), pp. 117–148 and Harman (2004), pp. 89–113. Harman’s argument is based on the elaborate moral weighting of reasons regarding the harms and benefits to the future person; but my claim is that any such weighting can take off the ground only once there *is* an identifiable individual as the subject of these benefits and harms. The act of giving life (of whatever quality) to a person is in itself neither a benefit nor a harm and the pain suffered by a person once she is born and throughout her life cannot be considered as caused by the parents’ act of bringing her into the world.
19. Kumar (2003), pp. 99–118. Markie, like Kumar, tries to stick to the person-affecting view, without relinquishing the grounds for wrongful life complaints, by distinguishing between being wronged and being harmed. Unlike Kumar, he appeals to the results (the existence of a handicapped child) rather than to the character of the agent. But Markie is misled by the ambiguity of “result”: the existence of such a child is a bad result “for the world (parents, society)” but not “for the child.” Markie is thus pushed back to impersonal evaluation which is exactly what he wanted to avoid. Markie (2005), pp. 290–305.
20. Reiman (2007), p. 92.
21. Reiman (2007), p. 84.
22. I have elaborated this critical interpretation of Rawls in Heyd (2009, forthcoming). Reiman quotes *A Theory of Justice*, in which Rawls indeed requires real capital accumulation from one generation to the next as part of the duties of justice, but Rawls himself abandoned that view and the intergenerational application of the difference principle in his later writings in favor of a more modest requirement to maintain the institutions of justice. Rawls (2001), p. 159.
23. Rawls (2001), p. 79. Thus, unlike Reiman, I believe it does not make sense “to think that it is in [a person’s] interest to be born with certain properties rather than others” independently of the person’s particular identity. Numerical identity is a condition for any ascription of interests to a person.
24. This is the reason for my disagreement with the recently published solution to the nonidentity problem offered by Caspar Hare, who suggests that there may be harms done to people *de dicto* even if no harm was done to anyone *de re*. Thus, according to Hare, in the same way I can harm the 35th President of the U.S., I can harm “my future child,” even though the identity of the particular person who is going to “fill” the relevant description is not fixed yet. I believe this cannot be the case since types of people (exemplifying some description) cannot have interests or rights (or be better or worse off). Only actual people can have them. One may have rights and interests *in virtue* of being of a certain type (or description), but the rights and interests relate to her *de re*. See Hare (2007), pp. 514–520.
25. Woodward (1986), pp. 804–831. Smolkin similarly believes that future (possible) people have rights, even if they cannot be said to have interests, arguing that rights are not dependent “substantively” on interests. But as he himself acknowledges, my own view is “formal” (or

- logical) rather than substantive, denying the ascribability of both interests and rights to possible people, irrespective of the *degree* of suffering or disability. Smolkin (1994), p. 319; and Smolkin (2002), pp. 202ff.
26. For a similar analogy between creating a debt which cannot be repaid and “wrongful life,” see Singer (1998), pp. 383–398.
 27. Roberts (1998), p. 18.
 28. Roberts (1998), p. 24.
 29. It is hard to judge whether the world in which this particular child exists without the broken glass by which it will be injured is “accessible” to us or not since the example is a bit artificial. Surely, the parents could delay the moment of conception by engaging in an alternative, less deleterious activity than that of planting broken glasses. And of course they could, after conception, remove the glass (as the parents in Kavka’s famous example could breach the contract and not give away their child to slavery after birth). But Parfit’s real-life examples and much of the legal history of wrongful conceptions do not suffer from this ambiguity and hence highlight the nonidentity problem more clearly.
 30. See Roberts’ article in this collection, in which she solves the slave-child paradox by first distinguishing between the assessment of the procreative choice on the basis of its *actual* value and such an assessment on the basis of its (probable) *expected* value and then showing that the source of the sense of paradox lies in the fallacious mixing of the two kinds of assessments. In the first kind of assessment, we could think of an alternative in which the same child is born without becoming a slave; and in the second, the chances (at the moment of making the choice) for the coming into being of this particular child are anyway infinitesimally low due to the biology of the procreative act.
 31. Roberts (1998), pp. 28–29.
 32. Roberts (2003), pp. 159–185, particularly p. 168.
 33. Roberts tries to solve the problem of comparability between a world in which a person has an anguished life and a world in which she does not exist by weighing “amounts”: “what we are comparing is the *amount* of well-being that Nora’s having certain properties and lacking certain others at α adds up to at α with the *amount of well-being* that Nora’s *lacking all properties* at β adds up to at β ” (p. 177). The question is, how can Nora be individuated or identified while “lacking *all* properties?” And how can all these absent properties add up to zero, as Roberts suggests, when she admits that zero is not a property of well-being that we can attribute to a non-existent person? Roberts (2003), p. 178.
 34. Wasserman (2005), pp. 132–152. Wasserman, although leaning to the person-affecting approach, offers his solution to the challenge of wrongful life complaints in terms which seem to lie somewhere between the person-affecting and the impersonal.
 35. This leaves Wasserman’s account with some measure of ambiguity. For on the one hand he says that the completely selfish parents can be accused of moral insensitivity, but that their child, born out of this insensitivity, cannot have grounds for complaint for having been born with impairment (pp. 146–147); yet on the other hand he argues that parents who conceive a child for reasons that have nothing to do with the child’s good “make themselves vulnerable to a complaint from the child for the unavoidable hardship in his life” (p. 151). The two statements are inconsistent. A purely person-affecting view, which accepts the relevance of nonidentity, can judge the act of the parents as a moral fault *only* in terms of the way it reflects on *their* character.
 36. In Heyd (1992), I discussed in detail the superiority of the person-affecting view over its impersonalist rival in terms of a “global” assessment of the merits and flaws of both as general theories of value.
 37. A similar argument would apply to animals. If they are considered as moral subjects of any kind, the nonidentity condition must equally constrain judgments about their creation in the same way as it does in the case of human beings.
 38. Parfit’s discussion of personal identity belongs to the former rather than the latter: although it is concerned with human consciousness (psychological continuity), it is not that part of the identity of persons which is constructed or self-constructed.

39. DeGrazia, (2005), pp. 28–29, p. 114 and Chapter 3. The force of DeGrazia’s hierarchy of the two kinds of identity does not depend on the particular substantive details of the metaphysics of human and personal identity on which philosophers have varied views. Hence, I do not have to take a position here on the question whether prenatal genetic interventions in an embryo should be considered as “treatment” of an identifiable individual or as the creation of a new individual. This is a separate metaphysical issue which demonstrates that biographical identity may include also prenatal events in the individual’s life (once it is numerically individuated but not yet having some fundamental features of her biographical identity, such as those which can be genetically molded in her). But I tend to agree with DeGrazia that such genetic manipulation after the third week of embryonic life does not change one’s essential (numeric) identity (pp. 263–264).
40. I want to distinguish my analysis of biographical identity from the mechanism of “adaptive preferences” originally described by Jon Elster. Adaptive preferences are one kind of response to gaps between desires and the chances of their satisfaction, i.e. the re-formation of desires in the light of the conditions which would make them satisfiable. They are considered irrational since they do not reflect the genuine preferences of the person but rather the frustration of her inability to gratify them. In contrast, self-construction of identity (or conversion) is not necessarily an *adaptive* change in one’s set of preferences but an authentic choice which is not tailor-made to fit external circumstances. Adaptive preferences are created *within* the existing biographical identity of a person. Conversions are changes *of* that identity.
41. Jeff McMahan doubts whether the (adult) person born out of a fetus that was harmed can be said to have been wronged, since he might have adapted to his condition and preferred it to not having been born at all or even to having been born without the harmful condition. See McMahan (2002), p. 301. But I think that the question of the preferences or degree of adaptability of the adult person is not what should decide the question of harm. The fact that *this* particular person could have been better off (in terms of opportunities) suffices to view the harm done to the fetus as harm to the ensuing adult person.
42. David DeGrazia eloquently makes this point by carefully distinguishing between persons and human animals and by using the term “individual-regarding” (rather than person-regarding) so as not to prejudice the issue of moral standing. See DeGrazia (2005), p. 263.

References

- Buchanan, A., D. Brock, N. Daniels, and D. Wikler. 2000. *From chance to choice: Genetics and justice*. Cambridge: Cambridge University Press.
- Davidson, M. 2008. Wrongful harm to future generations: The case of climate change. *Environmental Values* 17: 471–488.
- DeGrazia, D. 2005. *Human identity and bioethics*. Cambridge: Cambridge University Press.
- Habermas, J. 2003. *The future of human nature*. Malden, MA: Polity.
- Hare, C. 2007. Voices from another world: Must we respect the interests of the people who do not, and will never, exist? *Ethics* 117: 498–523.
- Harman, E. 2004. Can we harm and benefit in creating? *Philosophical Perspectives* 18: 89–113.
- Heyd, D. 1992. *Genethics: Moral issues in the creation of people*. Berkeley: University of California Press.
- Heyd, D. 1997. Divine creation and human procreation: Reflections on genesis in the light of genesis. In *Contingent future persons: On the ethics of deciding who will live, or not, in the future*, eds. N. Fotion and J.C. Heller, 57–70. Dordrecht: Kluwer.
- Heyd, D. 2009. A value or an obligation: Rawls on justice to future generations. In *Justice between generations*, eds. L. Meyer and A. Gosseries. London: Oxford University Press.
- Kumar, R. 2003. Who can be wronged? *Philosophy and Public Affairs* 31: 99–118.
- Markie, P.J. 2005. Wrongful conception and harmless wrongs. *Ratio* 18: 290–305.

- McMahan, J. 2002. *The ethics of killing*. Oxford: Oxford University Press.
- Mulgan, T. 2006. *Future people: A moderate consequentialist account of our obligations to future generations*. Oxford: Clarendon Press.
- Parfit, D. 1984. *Reasons and persons*. Oxford: Clarendon Press.
- Rawls, J. 2001. *Justice as fairness: A restatement*. Cambridge, MA: Harvard University Press.
- Reiman, J. 2007. Being fair to future people: The non-identity problem in the original position. *Philosophy and Public Affairs* 35: 69–92.
- Roberts, M. 1998. *Child versus childmaker: Future persons and present duties in ethics and the law*. Lanham, MD: Rowman and Littlefield.
- Roberts, M. 2003. Can it ever be better never to have existed at all? Person-based consequentialism and a new repugnant conclusion. *Journal of Applied Philosophy* 20: 159–185.
- Shiffrin, S. 1999. Wrongful life, procreative responsibility, and the significance of harm. *Legal Theory* 5: 117–148.
- Singer, P. 1998. Possible preferences. In: *Preferences*, eds. C. Fehige and U. Wessels 383–398. Berlin: De Gruyter.
- Smolkin, D. 1994. The non-identity problem and the appeal to future people's rights. *Southern Journal of Philosophy* 32: 315–329.
- Smolkin, D. 2002. Towards a rights-based solution to the non-identity problem. *Journal of Social Philosophy* 30: 194–208.
- Thompson, J. 2000. The apology paradox. *Philosophical Quarterly* 50: 470–475.
- Wasserman, D. 2005. The nonidentity problem, disability, and the role morality of prospective parents. *Ethics* 116: 132–152.
- Woodward, J. 1986. The non-identity problem. *Ethics* 96: 804–831.

Part II
**If Bringing a Badly Off Person into
Existence is Wrong, is *Not* Bringing
a Well Off Person into Existence
Also Wrong?**

Chapter 2

Rights and the Asymmetry Between Creating Good and Bad Lives

Ingmar Persson

Abstract This paper argues that common sense morality implies an asymmetry to the effect that there are stronger reasons not to cause to exist individuals whose lives will be miserable than to cause people whose lives will be worthwhile because it entails a theory of rights, according to which our general rights are negative. But it also argues that this theory of rights should be rejected in favour of a symmetrical morality which give us as much reason to see to it that worthwhile lives begin as that miserable lives do not.

Keywords Negative and positive rights · General and special rights · Asymmetry between creating good and bad lives · Moral obligation of procreation · The non-identity problem.

2.1 The Asymmetry, Common Sense Morality and Rights

Jeff McMahan claims that, according to common sense morality (CSM),

while the fact that a person's life would be worse than no life at all (or "worth not living") constitutes a strong moral reason for not bringing him into existence, the fact that a person's life would be worth living provides no (or only a relatively weak) moral reason for bringing him into existence.¹

I shall follow him in calling this intuition the Asymmetry. The Asymmetry is puzzling given a moral theory, such as utilitarianism, according to which all moral reasons are what I call *reasons of beneficence*, to the effect that we have reasons to see to it that beings are better off, reasons whose strength is proportionate to the degree of improvement. For, against the background of such a theory, it might seem that, if there is a moral reason—proportionate to the badness of the outcome—against causing a life that is worse than no life for the subject, there must also be a moral reason—proportionate to the goodness—in favour of causing a life that would be better than no life for the subject.

Although I believe that CSM includes reasons of beneficence, I believe it also includes other kinds of moral reasons, including the only kind relevant in this

I. Persson (✉)

Department of Philosophy, University of Gothenburg, Gothenburg, Sweden
e-mail: Ingmar.Persson@phil.gu.se

context, *reasons of rights*. I think that CSM entails a theory of rights which bestows upon us rights to ourselves, i.e., to our lives and limbs, and to property, rights that are *negative rather than positive*. These rights are *general* in the sense of holding against *all* others who are capable of recognizing rights. They are negative because they are to the effect only that others do *not* interfere with one's use of one's faculties or property, not that they positively aid one in one's use of these. The other side of these negative rights are negative duties or obligations of others to abstain from such interference, e.g., not to kill, mutilate, rape, imprison or steal. Call this theory the *doctrine of negative rights* (DNR).

In addition to these general, negative rights, DNR postulates *special* rights, which may be positive, and which correspond to special duties or obligations. People impose these duties on themselves by their (responsible) acts against (general) rights-holders (an obvious example is the act of promising). They are special in the sense that they are owed only to the patients of the relevant acts; thus, the corresponding rights of the patients are special by holding only against the agents in question. Special rights could well include rights to receive positive help to preserve life, limb and property, for instance, from someone who has endangered them.²

According to CSM, reasons of beneficence are normally weaker than reasons not to infringe general rights. This is what makes it wrong, for instance, to kill one innocent in order to save the lives of two or more (at least provided one is not under a special obligation to save the latter). Nevertheless, it seems reasonable to hold that reasons of beneficence can occasionally outweigh reasons of rights, so that it can be permissible, and even obligatory, to kill one if, e.g., this is the only way to save millions of lives.

In the next section I shall argue that, if we take CSM to entail DNR, we can explain why CSM implies the Asymmetry. However, I believe DNR to be untenable. This is can be argued independently of the issue of procreation raised by the Asymmetry, though I shall in Section 2.3 detail only arguments against DNR which concern these matters. I shall then in Section 2.4 contend that if we take morality to consist in nothing but reasons of beneficence, purged of the reference to rights, it will be symmetrical in the sense of affirming that we have reasons to create good lives, proportionate to their goodness, just as we have reasons against creating bad lives, proportionate to their badness. Such a morality can handle Parfit's non-identity problem. But because of its denial of the Asymmetry and some other implications regarding issues of possible people and procreation, it will strike us as counterintuitive. However, this is not an argument against this morality, since these implications strike us as counterintuitive only because we accept a theory of rights which is false.

2.2 How the Doctrine of Negative Rights Explains the Asymmetry

The hypothesis that the fact that CSM features DNR could explain why it encompasses the Asymmetry faces an obvious difficulty. This is that DNR seems clearly not applicable to non-existent, merely possible beings. Whatever the conditions

for qualifying as a rights-holder—and these are admittedly somewhat obscure—surely merely possible beings do not qualify: at least existence is required for being a rights-holder. Consequently, causing a being to exist can be as little wrong by infringing a right it had before existing as not causing a being to exist.

The claim of the rights theorist could at most be that beings acquire certain rights *if or when they begin to exist*. Michael Tooley suggests: “Every individual has a right to a life that is worth living” which is acquired when the individual begins to exist.³ Your causing a gravely deformed or handicapped individual to exist might then be wrong because it imposes upon you a special duty that you could not discharge, namely, to make the life of this unfortunate individual worth living. As remarked, according to DNR, our acts could put us under special duties, or bestow upon the patients of these acts special rights against us.

Against Tooley it could be pointed out, *ad hominem*, that it is hard to see how he could consistently make adequate use of such a right as he argues at great length that only *persons*—that is, according to his analysis, beings with a conception of themselves as persisting subjects with mental states, self-conscious beings, as we could put it—can have rights to life. Since, as Tooley would admit, no biological being is self-conscious when it begins to exist, it cannot then have a right to a life that is worth living, for surely only a being with a right to life can have a right to a life that is worth living. Nor will a human being ever acquire such a right if it is sufficiently gravely mentally handicapped never to become self-conscious, but this is precisely a case in which it could be most gravely wrong to cause a human being to begin to exist.⁴

Tooley is, however, prepared to admit that sentient beings that do not qualify as persons can have other rights, for instance, a right not to be tortured, or caused gratuitous pain, which derives from their interest or desire not to feel pain.⁵ If a being has any rights at all, such a right not to be caused to have painful experiences, and not to be deprived of pleasant experiences, without good reason, is plausibly a right that it acquires as soon as it becomes sentient. Nevertheless, according to DNR, this will have to be a negative right, and if so, it is not clear that this right could be violated by causing miserable beings to exist. For the progenitors of such beings do not literally inflict pain or suffering on their offspring as, say, torturers inflict it on their victims. Consequently, it is not clear that they violate these rights of their offspring and, so, not clear how a theory of rights like DNR could explain why it can be wrong to cause a miserable being to exist.

We have a more promising start if we could find a basis for claiming that human beings acquire rights to life and limb as soon as they become sentient.⁶ For we could then argue that, by causing their offspring to exist and be sentient, progenitors shoulder a special obligation to help them to exercise their rights, by remaining alive and using their faculties in ways that make their lives at least as good as non-existence. But this is an obligation which cannot be discharged if the offspring caused to exist is sufficiently handicapped. This would explain how these progenitors could act wrongly without relegating them to the same level as, say, torturers who violate general rights to life and limb. They would instead act wrongly by violating a special right they incur by their acts of procreation.

Now I think we can plausibly hypothesize that, according to DNR, infants do acquire rights to their life and limbs as soon as they become sentient. Consider an infant who is currently having some pleasurable experience. Imagine that it has the right not to have this experience interfered with, which Tooley cites. Then it is quite plausible to think that the infant must also have a right not to have its life and limbs interfered with to the extent that this interferes with the pleasure it is experiencing. If so, it has a negative right to life and limb, though it is not a person in Tooley's sense. This could be the right that its parents have a special obligation to help it to exercise.

People who have children are then analogous to agents who cause (general) rights-holders to be in situations in which they can no longer (as easily) keep things to which they have rights. Suppose, for instance, you endanger Vic's life and cause him suffering by pushing him into the sea. It is plausible to say that this puts you under a special obligation to try to save him, an obligation that you would not have had had you not caused his fall.⁷ It should be emphasized, however, that the positive obligation to save Vic presupposes that Vic has a (negative) right to life. It is because Vic has such a right and is transferred, by your act, from a state in which he could maintain his life to one in which he cannot that you are put under a positive obligation to try to save him.

It might be objected that the case of causing a being to exist differs crucially from the case of causing Vic to fall into the sea in that in the former you do not transfer a rights-holder from a state in which she is capable of keeping things to which she has rights to a state in which she is incapable of this. For, as remarked, merely possible beings cannot be rights-holders (or, for that matter, have any capacities). However, by causing these beings to exist, you may transfer them from a state in which it is *not* the case that they are incapable of keeping things to which they have rights—for in the state of non-existence they have no rights—to one in which this is the case. For instance, when they begin to exist, they may be incapable of maintaining their life because they are insufficiently developed to obtain food and fend off external threats. This may reasonably be held to be enough to put you under a positive obligation to help them to do so since, it might be claimed, the situation of being incapable of maintaining one's life is often worse than that of having no life to maintain. For in the former state you often positively suffer (as you are likely to do when you are pushed into the sea).

In the case of the drowning person, the content of the obligation incurred is to restore, as well as you could, the safer situation you have eliminated, by pulling the person ashore. It might be thought that the counterpart to this in the case of causing somebody to exist would be euthanasia, since the state of being dead is the closest possible equivalent to that of not having begun to exist. Although it is controversial, I think this may indeed be the best course of action if—but only if—you have happened to cause the existence of individuals with such severe handicaps that, with available assistance, their life must remain worse than non-existence. Such individuals could be presumed to be prepared to waive their right to life had they had the competence to do so. However, when you could provide them with a worthwhile

life, they would not waive their right to life, but would rather have such a life. They would then have a special right against you to be provided with it.

Even if you have not caused someone's plight, like Vic's struggling in the water, you would have, according to my construal of CSM, a reason of beneficence to assist him, though this reason is not so strong as your reason would be had you pushed Vic into the water. According to CSM, however, you do not have even a reason of beneficence to cause beings to lead good lives because, as conceived by CSM, such reasons dictate that we benefit beings by helping them to maintain things to which they have rights. So, on the assumption that DNR is part of CSM, CSM will imply a strong form of the Asymmetry on which we have *no* moral reason to create, for their sake, beings who will lead good lives, while we have strong moral reasons not to create beings whose lives will be miserable, since we shall thereby incur special obligations that we cannot discharge. We shall, however, see that, once rights are rejected, and reasons of beneficence are cleansed of their reference to rights, their range will expand and the road paved for a moral symmetry.

2.3 The Grounds and Groundlessness of Rights

Much would have to be said about the (putative) grounds of general and special rights if the above account were to serve not merely as an *explanation* of why we are intuitively inclined to believe in the Asymmetry but also as a *justification* of the truth of this belief. My ambition is, however, to show, to the contrary, that the Asymmetry, on this construal, cannot be justified because the grounds of general rights involve us in something like *paradoxes of procreation*. To do this, I need first to outline what I take to be the grounds of rights, according to CSM.

2.3.1 The Grounds of Rights

With respect to the grounds of rights, there is an interesting contrast between general and special rights: general rights are based upon facts about *the rights-holders*, whereas special rights are based upon facts about *the people under the corresponding duties*. I shall have to be all too brief about the grounds of special rights and duties, as conceived by CSM. As can be gleaned from the examples above, CSM takes an *action* from you, such as pushing Vic into the sea—no mere omission, e.g., to prevent someone else's push—to put you under a special obligation, and the patient under a special right against you. For, intuitively, we take ourselves to be more responsible for what we actively cause than for what we passively let happen.⁸ Something for which we take ourselves to have a higher degree of responsibility is required to put ourselves under something which has the greater reason-providing power of a duty.

As regards procreation, this action is normally sexual intercourse. But an expectation, let alone an intention, that a child results from a sexual act is not necessary for

the pair engaging in intercourse to acquire parental obligations as regards a resulting child. For this obligation would, by the law of many countries, be foisted upon a man who made a woman pregnant by having sex with her even if he took the precaution of using condom which, unbeknownst to him, was defective.⁹ In contrast, the interception of another free act after the sexual act seems to affect responsibility. For, as McMahan remarks: “if a woman were surreptitiously to preserve and impregnate herself with the contents of her lover’s discarded condom,” the lover would not owe the resulting child parental duties.¹⁰ If the man instead had anticipated the woman’s behaviour, then the pair would share parental duties for the child. This is so on the condition that the woman’s intervening act is free: if the man had forced or coerced her to act as she did, he alone would have had parental duties (at least if the woman had no chance of later undergoing abortion). The conception then becomes analogous to a conception due to rape. There is more to be said about the commonsensical grounds of special rights and duties, and to what extent they are justifiable, but this must be left for another occasion.

As I have already indicated, the ground of general rights must be sought in some fact about *the rights-holders*, and not in any acts of the people who are under the corresponding duties. For everyone who is capable of recognizing rights is under these duties, without having performed any particular act towards the rights-holders. But what could it be about rights-holders that has the power to impose duties upon others?

A traditional idea is that it is facts of *first appropriation or occupancy*. John Locke is a famous spokesman of this idea:

[E]very man has a property in his own person. This nobody has any right to but himself. The labour of his body and the work of his hands, we may say, are properly his. Whatsoever, then, he removes out of the state that Nature hath provided and left it in, he has mixed his labour with it, and joined to it something that is his own, and thereby makes it his property.¹¹

Locke seems here to be advancing the following two claims. (a) *The self-ownership claim*: we own, or have a (property) right to, our own bodies and the psychological characteristics that make them embody persons. True, Locke does not explicitly assert that each of us has property in their *body*, but in their *person*, which he elsewhere famously distinguishes from its body.¹² But it is hard to understand how our bodily labour could be ours if our bodies are not. (b) *The property-acquisition claim*: in virtue of the fact that we own or have a right to our bodies, we come to own hitherto unowned natural resources with which we “mix” our bodily labour, i.e., mix something that is ours. The reason we own, or have a right to, our bodies may be that we are the first ones to “appropriate” or “occupy” them in the sense of exercising voluntary control over them.¹³ We could then come to own other natural resources by being the first ones to appropriate them by means of exercising control over them through our labour. There is no need stop here to consider the peculiarities of Locke’s idea of mixing our labour with natural resources.¹⁴ Instead, let us apply the self-ownership and property-acquisition claims, vaguely understood, to procreation.

2.3.2 *The Groundlessness of General Rights*

If we own our bodies, the gametes that our bodies produce must surely be ours, too. By their sexual “labour” a couple brings their respective gametes together to form an embryo, which the woman nurtures in her body. So it would seem that, in accordance with the property-acquisition claim, a child should be the property of its parents—in particular its mother.¹⁵ Now, that we have a property right to something implies that we are permitted to use it largely as we please, even if this use benefits us less than it harms others, as long as this use does not infringe any of their rights. It would seem that a property right which did not carry this implication would not really be a *right* to anything. To have a right to something is to be in a *privileged* position as regards it, a position more advantageous than the position of others as regards it. (This is true of the so-called “claim-rights” which I here have in mind, as opposed to “liberty-rights.”) So, it would seem to follow that, since children are the property of their parents, parents have a right to use their children largely as they see fit even if this is to the detriment of the children.

This clearly conflicts with there being a parental obligation to protect one’s offspring against unjustifiable suffering, for this is an obligation to protect the offspring even when this is not in one’s own best interest. This obligation is designed to protect the offspring’s interest, while the property right is designed to protect the owner’s interest. Thus, one and the same relation, that of causing somebody to exist, appears to be taken by CSM, in virtue of its inclusion of DNR, to support two incompatible conclusions. On the one hand, it supports parental obligations, given that the offspring have rights to life and limb, but on the other hand it supports the progenitor’s property rights, given the property-acquisition claim. The reason for the latter implication is that the property-acquisition claim rules out that offspring own or have rights to themselves. But, as we have just seen, the parental obligation presupposes that offspring have rights to themselves.

In fact, it seems that the Lockean theory is even *internally* inconsistent, for the property-acquisition claim both undercuts the self-ownership claim and presupposes it. To spell out this conflict, suppose that, on the basis of the self-ownership claim,

1. Some parents P own themselves.

In virtue of (1), the property-acquisition claim and the fact that they produce their offspring out of material that belongs to nobody but themselves, it follows that:

2. P own their offspring.

Consider the fetus before it develops consciousness: it must be owned by its parents rather than by itself since it cannot be the owner of anything at this pre-conscious stage. But since the fetus’s consciousness develops out of its body, which is owned by its parents, its consciousness, too, must be owned by its parents. (Or, if consciousness is not the sort of thing which can be owned, it cannot come to own the

body in which it is developed, since this body is already possessed by the parents.) Hence, (2) is true. But P are themselves the offspring of other parents, P*. So, it would seem that, at least as long as P* are alive and have not voluntarily transferred their property rights to P, P are owned by P* rather than by themselves. But this is a *reductio ad absurdum* of the Lockean theory, for it is an argument that assumes the truth of (1), but leads to the conclusion that (1) may be false.

Hillel Steiner tries to get out of this paradox by arguing that the Lockean theory does not imply that parents *fully* own their offspring.¹⁶ This is because producing the offspring “required [the parents] to mix their labour with natural resources in the form of germ-line genetic information transmitted from . . . grandparents.”¹⁷ “Natural resources,” he maintains, are “unowned things” to which we all have equal claims.¹⁸ An ownership not being full means, according to Steiner, that it is temporary and liable to expire when the offspring attains majority.¹⁹ At that time, the offspring becomes a self-owner.

There are, however, several problems with Steiner’s proposal. First, it seems compatible with the repugnant view that parents own their children until they attain majority. Second, it is not clear why children should automatically acquire the status of self-owners upon attaining majority if their parents are alive and unwilling to transfer their property rights. As the first occupiers of the bodies of their offspring, it seems that parents would keep their property rights until they voluntarily relinquish them, or die. Third, if we do not own that of ourselves which is transmitted from elsewhere, like “germ-line genetic information,” it seems that we could not own ourselves, since everything constituting us must be transmitted from elsewhere in light of the fact that once we did not exist. So, why is not the upshot a world in which nothing is owned, and nobody has a property-right to anything, but we are all free to the use of *everything*, including human bodies, according to need and perhaps other factors?

I have presented this argument against a Lockean sort of theory of the grounds of general rights because it has to do with procreation. There are other arguments, possibly stronger ones, against this theory which are independent of this issue.²⁰ But the fact that the grounding of the general rights that DNR postulates is problematic does not mean that DNR could not be a part of CSM. Since CSM has evolved unreflectively, it may well contain crude, unrefinable, and even contradictory elements. (More specifically, I think the concept of a right is a pre-cultural product which could be traced back to behavioural dispositions that we share with our non-human ancestors, such as the propensity to defend one’s turf or prey with a special ferocity.) Since the parental obligation and a property right partially coincide in the protection of the offspring—though for different reasons—the conflict between them is probably not so conspicuous that they could not be thought to coexist in CSM. Thus, despite the argued incoherence of DNR, I may well have successfully explained why CSM leads to the Asymmetry by hypothesizing that it comprises DNR. But, of course, if this explanation of why we adhere to the Asymmetry in everyday thinking is correct, the Asymmetry has not been shown to be *justifiable*, to be part of a *true* morality. To find out what the moral truth is, we must turn to other moral reasons than reasons of rights.

2.4 The Symmetry of Reasons of Beneficence

Alongside the reasons of rights given by DNR, CSM also comprises reasons of beneficence. According to CSM, these reasons are to the effect that we have reason to help beings to keep or regain things to which they have rights and, thereby, to make them better off. This reference to rights has to be dropped if DNR is rejected. Reasons of beneficence will then boil down to the idea that we have reasons to do what benefits beings in proportion to how much they are benefited. This revision is natural, for things to which we have rights are essentially things from which we could benefit. Rights to things which could not benefit us would be pointless, since we would straightaway waive these rights. (This is, I believe, why we are not normally said to have rights to the waste products our bodies produce.)

I cannot here go into the controversial matter of what it means that something benefits, or is better for, someone.²¹ But I shall make the rather innocuous assumption that it is a necessary condition of the applicability of this notion to something that it could have consciousness at some point (thus, nothing could be better or worse for plants or inanimate things). This assumption is part of the argument that I shall now give for the claim that the revision of reasons of beneficence makes them range over merely possible beings as well as beings who sometimes exist. If so, and there is no opposition from DNR, it follows that a symmetry replaces the Asymmetry, that we could have as strong moral reasons to create worthwhile lives as against creating bad lives.²² It also follows that what Derek Parfit has called the non-identity problem presents no more of a problem than similar choices affecting existing beings.²³

Parfit's case of two medical programmes can serve to illustrate how an extension of moral reasons to merely possible beings solves this problem.²⁴ To simplify Parfit's case, suppose there is a disease which will make children somewhat handicapped if their mothers contract it when pregnant. But it is easily cured. We would think that a pregnant woman acted wrongly if she did not undergo this harmless treatment, when the alternative would be to have her child born handicapped. But what if, instead of waiting until she has been cured, a woman gets pregnant, though she knows she has contracted the disease? It seems reasonable to hold that the second woman has acted more or less as wrongly as the first one, by not waiting to have a numerically distinct normal child after she has been cured a couple of months later. But we cannot justify this judgement by claiming that the second woman acted wrongly by harming the child she in fact had—as the first woman did—since we may imagine that the handicap is so slight that her child leads a life better than non-existence. We have, however, a justification for making this judgement if we are allowed to say that she had a stronger reason to bring an alternative healthy child into existence since she would have benefited it more than she benefited the handicapped child she in fact had.

This could, however, not be the way in which CSM justifies the judgement that the second woman acts as wrongly as the first one. We should rather expect this justification to be in terms of DNR. Parfit's example seems to indicate that the special obligation CSM imposes upon us to aid the beings we have caused, or will cause, to

exist goes beyond seeing to it that their lives are at least as good as non-existence. It is rather an obligation to see to it that their lives are as good as possible, as long as this does not involve sacrifices on our own part that are too great. This obligation applies irrespective of whether these beings already exist or will exist.

2.4.1 *Does Negative Value Alone Have Moral Force?*

Even if reasons of beneficence cover possible people, we could conceivably preserve the Asymmetry. Consider so-called negative utilitarianism which construes these reasons as reasons to prevent, as far as possible, that people undergo things that are bad for them. However, negative utilitarianism seems implausible, as is shown by an argument sketched by McMahan,²⁵ on the basis of an argument originally put forward by Richard Sikora (1978). This argument turns on the observation that if what would be bad for individuals in life is a reason against conceiving them, but what would be good for them is no reason in favour of conceiving them, then, as far as those individuals are concerned, it is wrong to conceive them, however much good their lives will contain, provided that they will also contain something that is bad for them. This seems clearly absurd.

A variation of this argument appeals to the fact that there is always some probability, however slight, that were you to have a child, it would be miserable, owing to some congenital disease or disability. If this is a moral reason against conceiving a child, and the probability, however great, that its life will be good for it is no reason in favour of conceiving the child, then, as far as we take into account just the child, you would only have moral reasons against conceiving the child and no reasons in favour of it. But this is patently absurd. So, we have reason to think that, if the fact that its life risks being miserable is a reason not to conceive a being, the fact that its life has a chance of being worth living is also a reason to conceive a being.

However, David Benatar has argued that coming into existence is always harmful as long as existence contains something that is intrinsically bad.²⁶ He claims that there is an asymmetry between pleasure and pain consisting in the fact that

1. The absence of pain is good, even if that good is not enjoyed by anyone,

whereas

2. The absence of pleasure is not bad unless there is somebody for whom this absence is a deprivation.²⁷

As Benatar implies,²⁸ “bad” in (2) should not be understood as meaning intrinsically bad (for the mere absence of pleasure is not intrinsically bad even if there is somebody around to experience its absence), but should be taken in a relative or comparative sense. So, (2) means

2'. The absence of pleasure is not *worse* than its presence unless there is somebody for whom it is an absence.

By analogy, (1) should mean:

- 1'. The absence of pain is *better* than its presence even if there is not anybody for whom it is an absence.

But, curiously, Benatar resists this interpretation, writing “Avoiding the pains of existence is more than merely ‘not bad.’ It is good.”²⁹ It cannot be right, however, that the absence of pain is intrinsically good, like the presence of pleasure. For the absence of pain is compatible with the absence of pleasure, i.e., the negation of what is intrinsically good. Thus, *pace* Benatar, the absence of pain is simply “not bad,” which is enough to make this state better than a state of pain which is intrinsically bad.

With the interpretation of (1) as (1'), Benatar's asymmetry consists in the idea that the existence of subjects is required for the absence of pleasure to be worse than pleasure, but not for the absence of pain to be better than pain. It is hard to see any justification for this asymmetry, e.g., why it is any less plausible to say that a universe without sentient life is worse than a universe with sentient beings enjoying pleasure than to say that it is better than a universe with such beings undergoing pain.

In support of his alleged asymmetry between (1) and (2), Benatar claims that it has “considerable explanatory power.”³⁰ First, he asserts that it could account for McMahan's Asymmetry.³¹ This should not impress us since, as we have seen, and as Benatar grants, this Asymmetry could also be explained by an appeal to negative rights. As remarked, such an appeal could also explain why we cite potential persons' interests only as a reason for not bringing them into existence and never as a reason for bringing them into existence: it is only in the first case that they could have special rights against us.³²

In favour of his asymmetry, Benatar further claims that “only bringing people into existence can be regretted *for* the sake of the person” brought into existence, not failing to create people who would have led good lives.³³ I do not see how this claim could support Benatar's asymmetry since, if it is true, it seems equally true that we can be *glad* only for the sakes of happy people we have brought into existence, not for the sakes of the miserable ones we have not caused to exist.

Finally, Benatar also affirms that while “we are rightly sad for the inhabitants of a foreign land whose lives are characterized by suffering . . . we are not similarly sad for the happy people who, had they existed, would have populated” some uninhabited land.³⁴ But, clearly, the relevant comparison is not with our attitude to people who are in the intrinsically bad state of suffering, but with our attitude to people who have been spared suffering by not existing. Suppose we find evidence indicating that sentient life was close to developing in some very advantageous place. I do not think that it would be any less natural to be sad or regret that it did not develop than that it would be to be glad or pleased that sentient life did not develop in a place where it would have been miserable. So, I see no reason to embrace Benatar's asymmetry between (1) and (2), let alone his more extreme claim that coming into existence is always a harm.

2.4.2 *Are Reasons of Beneficence Restricted to Existing People?*

There are other construals of reasons of beneficence than ones having to do with the special moral force of negative value that might save the Asymmetry. Melinda Roberts offers a “person-based consequentialism” (PBC) which “requires, for each person who ever exists, that agents maximize that person’s well-being.”³⁵ We need not here bother about the non-aggregative aspect of PBC. Thus, for our purposes it would do as well to consider an aggregative counterpart to PBC which requires agents to see to it that the sum of well-being of all those who ever exist be as great as possible. What concerns us is simply that, on PBC, “*merely possible people*—those who could have existed, relative to a given possible future, but who do not and never will exist—do *not* count for moral purposes.”³⁶ Given the assumption, which Roberts is ready to make,³⁷ and which I shall defend below, that non-existence has neutral value, or neither positive nor negative value, PBC will imply the Asymmetry.

But can this restriction of reasons of beneficence to people who exist at some point or other really be justified? Consider a human fetus just before it acquires consciousness, perhaps a fetus sixteen weeks old. Assume that if this fetus is allowed to live on, it will develop consciousness and lead a life well worth living. Does this fetus morally count, according to PBC, so that we could be said to wrong it if we kill it for no good reason? It seems so because it certainly exists. If it is allowed to live on, it will by hypothesis develop consciousness and desires. At that stage we can certainly act wrongly to it, by causing it pointless suffering. But the very same entity which would later acquire a consciousness in virtue of which it can be wronged, a certain human animal or organism, is already in existence, sixteen weeks after conception. So, it would seem that PBC should imply that we could wrong this pre-conscious fetus, e.g., by killing it.

It might, however, be replied that when a human being or organism acquires consciousness, a new entity comes into existence—call it a (human) *person* (in the broad sense of “being with consciousness” that Roberts also employs). Some claim that *we* are identical to persons rather than human animals, i.e., that these persons are the true referents of personal pronouns as used by us. But even if this view on personal identity is true, it would be very implausible to say that it is *only* the person and not the human animal who can be wronged, and for whom life can be good or bad. The human animal no less than the person can truly be said to have consciousness and desires; hence, things can be good or bad for the human animal, and it, too, can be wronged. But the human animal normally exists in a pre-conscious state. Thus, some pre-conscious entities can be wronged.

Consider now the period prior to the formation of a human animal or organism. We could imagine a sperm and an ovum before fertilization, the zygote existing just after fertilization, or the cluster of cells existing after a few divisions, before they are integrated into something like a multi-cellular organism. Nothing can ever be good or bad for these entities, since *they* (in contrast to a multi-cellular organism which will later be formed) will never acquire consciousness; they will in the natural course of events soon cease to exist.³⁸ Thus, according to PBC, the action of seeing to it that

these entities will not form an animal or organism which will develop consciousness will not by itself wrong anyone, while we have found that the action of seeing to it that, say, a sixteen-week old pre-conscious fetus will not develop consciousness could be wrong. But it is hard to see how there could be a moral difference between these cases, for the gametes, zygote, etc., on the one hand, and the pre-conscious fetus, on the other, are morally relevantly similar: both lack consciousness, and both are involved in a natural process that will produce consciousness.

The question is then: if we can wrong a person by seeing to it that it does not begin to exist in the sense of seeing to it that a human animal does not acquire consciousness, why can we not wrong a person by seeing to it that it does not begin to exist in the fuller sense of seeing to it that a human animal which would acquire consciousness does not begin to exist? Some, though not Roberts, claim that non-existence cannot be worse for someone than (worthwhile) existence. For instance, John Broome writes “if she had never lived at all, there would have been no her for it to be worse for, so it could not have been worse for her.”³⁹ So far as I can find, there could be two sorts of reason that could be offered for the view that in order for it to be true that something is worse (or better) for someone, she must exist at some point.⁴⁰

One reason is that things which never exist are not *bone fide* particulars to which we can refer and attribute properties to. Suppose I say “My first-born child will inherit my fortune” and that years later I have a daughter. Then, in making my statement, I did not refer to her specifically, saying that *she* would be my heir. I was speaking about *whoever* would be my first-born child and not about any child in particular. But, in opposition to this objection, I would like to claim that it is possible to identify specific human beings, persons or organisms, before they begin to exist. They could be identified, for instance, as the human being that would result if a certain sperm fertilized a certain egg (assuming that we could rule out the possibility of monozygotic twinning). We could intelligibly make predictions about how good or bad life might be for this specific, but still non-existent, being who might never become actual, although such predictions are of course hazardous. This procedure is analogous to identifying a particular consciousness—and thus a particular person—as the consciousness that could develop in a particular fetus.

The other sort of reason for leaving out the non-existent is that, though we can make *some* true statements about particulars which remain forever non-existent, statements that something can be better or worse for them are not among them. Now I am prepared to grant that possession of a capacity for consciousness, desires and the like, by a subject, S, at time, t, is necessary for something to be *intrinsically good or bad* for S at t. But it does not follow that, for something to be *better* or *worse*, or in a comparative sense good or bad, for S at t, S must have a capacity for consciousness and, so, must exist, at t. For the latter claim to be true, it suffices that S at t is in a state that is indifferent or of neutral value for S, and that this state is comparable to another possible state (of existence) which would be either intrinsically good or bad for S.

One way in which a state can be neutral, or neither good nor bad, for S is by being a state in which S lacks a capacity for consciousness and all attitudes. In this state, there can be neither satisfaction nor frustration of desire, neither pleasure nor pain.

This could be a state of existence, e.g., a human organism who is in a persistent vegetative state (PVS) throughout its existence. Another possible neutral state is the state of non-existence. Both are states without a capacity for consciousness, though for different reasons. Clearly, we can judge a stretch of conscious existence to be better or worse for someone than existence in PVS. So, why should we not be able to make the same comparisons with respect to non-existence? As we have seen, the reply cannot be that we cannot refer to beings who might forever remain non-existent. It is plausible to say that to be in PVS is neither better nor worse for someone than not to exist at all. Both states are then reasonably of neutral, or of neither positive nor negative, value for the subject. Consequently, conscious life of positive intrinsic value is better for one than non-existence, just as it is better for one than is PVS, and conscious life of negative intrinsic value is worse for one than non-existence, just as it is worse for one than PVS.

It might be objected that, if something's lacking consciousness and subjective attitudes implied that existence for it was evaluatively neutral, there will be countless things—like numbers and shoes—whose existence for them is neutral. These things could crowd our moral calculations and force us to endorse absurdly irrelevant comparative judgements, such as that a beggar's life might be worse for him than existence is for his shoes, or for the number five. But there is no reason to include in our moral calculations such items as numbers and shoes, since there is nothing we can do to make existence better or worse for them, for we cannot make them conscious. Furthermore, although it is normally absurd to bother to *make* such comparative judgements as the one cited—because there is virtually no context in which the information they give is called for—it is not absurd to believe them to be *true* (indeed, so obviously true that it is pointless to assert their truth). That this is the explanation of the apparent absurdity comes out if the statements are expanded, if it is said, e.g., that the beggar's life is worse for him than existence is for his shoes, since nothing can be good or bad for them, and life is bad for the beggar. The expansion removes any misleading conversational implications about the value of existence for shoes. Moreover, it is hard to deny that we can claim to have stronger reasons against creating miserable beings than against inanimate objects because we can make such comparisons.

To conclude, there seem to be no good reasons for claiming, on the basis of reasons of beneficence, that we can wrong persons by seeing to it that they do not begin to exist in the first, thinner sense, i.e., the acquisition of consciousness sense, but not in the second, fuller sense. These cases seem morally on a par. If so, there will be a symmetry in place of the Asymmetry.

2.4.3 Consciousness and the Strength of Reasons of Beneficence

But perhaps the Asymmetry should be replaced by a symmetry to the effect that we lack moral reasons of *full* strength both against starting miserable lives and in favour of starting worthwhile lives. Perhaps reasons of beneficence grow appreciably

stronger when they deal with beings who already have the capacity of consciousness. The acquisition of a capacity for consciousness is certainly morally significant because some morally relevant forms of treatment, such as hurting, become possible only after this event. But why think that the presence of this capacity affects the strength of reasons to benefit in *all* ways?

A possible answer appeals to McMahan's idea that only conscious beings have a "time-relative interest" in future existence. A necessary condition of your having a time-relative interest in the future existence of some being is, according to McMahan, that the brain areas that support your consciousness and its consciousness are the same. Of course, this implies that both of you have consciousness, but it does not imply that you can be conscious of, or anticipate, the future experiences of this being or that it can be conscious of, or remember, your experiences. The psychological relations of anticipation, memory, future-oriented desire, etc. constitute a "psychological unity"⁴¹ which could go into the basis of time-relative interest, alongside the identity of the consciousness-sustaining brain-parts. Since this is so, time-relative interest admits of degrees in proportion to the extent of the psychological unity. McMahan claims that the presence of time-relative interest provides us with "personal" moral reasons to care about the future existence of a particular being, reasons that are stronger the stronger the psychological unity is. The reason we have to cause to exist (in either of the two senses considered in the foregoing subsection) beings who will lead good lives is only of an "impersonal" kind. McMahan writes that "it is simply better, even if it is not better for anyone, that a better-off child should exist rather than a less well-off child."⁴² However, you do not harm or wrong the better-off child if you block its existence because it cannot have a time-relative interest in starting to exist. The reason you have against this measure is impersonal rather than personal.

But, although I can see why the fact that there is a psychological unity in the shape of people anticipating their future and having preferences regarding it could provide stronger reasons to protect this future, I cannot see why the mere fact of the sameness of the brain areas underlying a current stretch of consciousness and a future one should upgrade these reasons. Compare a fetus who has had its first, brief episode of consciousness, but is now unconscious or asleep, to a fetus who is about to have its first episode of consciousness. It is hard to believe that in the first case a different, personal kind of reason to be concerned about its future has kicked in, that we could in any morally significant way harm or wrong the first fetus but not the second.

Benatar considers this argument in the process of arguing for his view for "a life to be not worth continuing, it must be worse than it need be for it not to be worth starting."⁴³ In reply to my sort of argument, he claims that "coming to exist in a morally relevant sense is a gradual process."⁴⁴ It follows that we cannot compare a state *just* before and after the coming into existence, as I have done. This gradualist claim could, however, be questioned if, as Benatar believes, coming to exist in the morally relevant sense occurs when consciousness is first acquired, since this could hardly be a gradual affair, if consciousness is something distinct from the physical. But we can afford to grant the gradualist claim since, contrary to Benatar, it weakens

rather than strengthens his case. For if there is a stage in our development during which it cannot be determined whether our existence is beginning or continuing, we must decide whether or not our further development then is worthwhile *irrespective* of whether our existence is beginning or continuing, and then this distinction cannot have moral importance. This is, however, compatible with Benatar's point that we should "permit greater life-saving sacrifices as the being's interest in existing develops,"⁴⁵ i.e., as its future-oriented preferences develop—a point I have already accepted.

But before such preferences develop, a discount in proportion to the strength of McMahan's time-relative interest seems to harbour some unpalatable implications. Consider a newborn who is experiencing some intense pleasure. Imagine that, if it continues, this experience will cause the neonate damages that will shorten its life by several years. On McMahan's theory, it seems that we could well have stronger personal reasons to let the experience continue, since the neonate will have a rather powerful time-relative interest in its ongoing experience but only a very weak time-relative interest in its distant future. This seems clearly counterintuitive. We obtain a verdict more in line with our intuitions, I submit, if we do not discount the value for the neonate of its distant future in proportion to its time-relative interest. We can form the conception of an outcome, in which an individual is alive, as being better or worse for that individual, by a certain amount, than an outcome in which it never begins to live, by adding together the value its life has for that individual at each and every moment of its duration and comparing this value sum to the neutral value of non-existence.

McMahan suggests that a reference to time-relative interest could explain our feeling that it is wrong to kill a normal infant in order to replace it by creating another infant who will lead a better life but not wrong to prevent the coming into existence of an infant who will lead a less good life and to conceive instead an infant who will lead a better life.⁴⁶ My conjecture is that it is rather our pre-reflective belief in DNR, and the ascription of a right to life to an existing infant, which explains the feeling that it is wrong to kill it in order to replace by the creation of a new infant. If so, this intuition is undermined along with the Asymmetry by the rejection of DNR. In practice, however, it is difficult to be reasonably certain that the infant who we shall bring into existence will lead a better life than the infant to be replaced, and this will be more difficult the less handicapped the latter infant is. This will mean that, in practice, the difference between CSM with the Asymmetry and the revised symmetrical morality will be comparatively slight.

Nevertheless, it should be acknowledged that, if we reject DNR, and if reasons of beneficence, correspondingly revised, come to the fore, there will in the area of reproduction be implications that are counterintuitive according to CSM. For we shall have comparatively stronger moral reasons to have children than we imagine because rights import a bias in favour of the existent which will disappear with them. But it is important to point out that it does not follow that we *ought*, or have *most* reason, as often to have children when they would lead good lives as we ought not to have children when they will have bad lives. This is because, when there is a reason of beneficence to bring someone into existence, there will often be some

alternative action in favour of which there will be at least as strong reasons, e.g., an action of aiding people already existing in an overpopulated world. So, we are permitted to perform this action instead. But when there is a reason of beneficence against bringing someone into existence, there will almost always be less strong reasons against some alternative action, even if it is only twiddling our thumbs and not positively benefiting anyone. So, we are not permitted to bring an unfortunate individual into existence rather than performing some alternative act. Again, this implies that, in practice, the implications of abandoning the Asymmetry for what we ought to do will not be as dramatic as we might think at first blush.

I cannot here pursue the question of what further facts there are which could help defuse the counterintuitive implications of the revised morality which results if we drop DNR and, so, I cannot assess how great a moral difference the revision would make in practice. Let me just emphasize that we should not expect that anything as important to CSM as DNR could go by the board without our commonsensical moral intuitions being jostled and, thus, we should not take correspondence to those intuitions as a criterion of adequacy.

Acknowledgements Many thanks to the editors for helpful comments.

Notes

1. McMahan (1981), p. 100. Cf. McMahan (2002), p. 300.
2. The distinction between general and special rights was introduced by Hart (1955). I use “patient” as the opposite of “agent,” to refer to anything with moral status that is affected by a given act.
3. Tooley (1983), p. 272. Recently, David Velleman has claimed something apparently stronger: “we are obligated to have those children to whom we can give the best start” (2008), p. 276, and hence that children have a right to be given this start. (He doubts that a “coherent concept” could be attached to the term “life worth living” (2008), p. 273.) Velleman seems to think that underlying this obligation there is an “obligation to take due consideration for the importance of human life, as the context in which personhood is realized or damaged” (2008), p. 288. This obligation also makes us “obligated to realize the value of personhood in ourselves” (2008), p. 254, and “to ensure that the human race does not go extinct” (2008), p. 254. But who has the right corresponding to this alleged obligation of ours to take due consideration for the value of human life? It would seem that (at least if we exclude God) the answer could only be: humanity, i.e., the class of all human beings. However, this will not do, for imagine that all human beings tried to waive their (alleged) right against me, avowing that they do not care one whit whether I realize my personhood or reproduce. Surely, this cannot be thought to release me from my obligation to do so (if there is such an obligation). Similarly, if all humans chose to behave in the same way as I do, there would presumably still be obligations that we all failed to discharge. Therefore I think that there is no possible right-holder and, hence, no possible right corresponding to Velleman’s fundamental obligation. But such an obligation seems to me incoherent.
4. It might also be asked how every individual could have a right to a life worth living if rights are held against people who shoulder corresponding obligations. For it would seem that some individuals are so deformed and handicapped that it is *impossible* to make their lives worth living, and it is doubtful that people can be under obligations to do what is impossible (“ought” implies “can” as the dictum goes). I just mention this objection, without endorsing it.
5. Tooley (1983), p. 100.

6. I think that, if sentient human beings have these rights, so do some sentient non-human beings, but in this paper I shall stick to human beings.
7. Cf. McMahan (2002), p. 376.
8. Cf. McMahan (2002), p. 461. I discuss the interplay between this idea of responsibility and DNR in CSM in Persson (2007).
9. Cf. McMahan (2002), p. 374.
10. McMahan (2002), p. 375.
11. Locke (1690/1990), bk II, Chapter V, Section 27.
12. Locke (1689/1975), bk II, Chapter XXVII.
13. Cf. Kamm (1992), p. 101.
14. See Waldron (1988), pp. 185–6.
15. The greater contribution of the mother receives special stress by Moller Okin in her version of this argument. Moller Okin (1989), pp. 79–88.
16. See Steiner (1994). It should be noted that Locke himself senses the difficulty and tries to reply to it (1690/1990), bk. I, Chapter VI, Sections 52–54. As Nozick observes (1974), pp. 287–9, Locke’s replies are weak. But Nozick himself does not come up with anything significantly better, though the difficulty is as much a difficulty for him.
17. Steiner (1994), p. 248.
18. Steiner (1994), p. 235.
19. Steiner (1994), pp. 248 and 275.
20. See Persson (1994) and (2005), pp. 418–21.
21. For further discussion, see Persson (2005), Chapter 10.
22. Since we are prone to feel more responsible for what we do than for what we let happen, we are likely to feel less responsible for the happy lives we fail to cause than for the miserable ones we do cause. Now, I am not discussing responsibility, but what we have moral reasons for and against doing. However, it would seem that, if we could have as strong reasons for causing happy lives as against causing miserable ones, we *should* feel as responsible for failing to cause the former as for causing the latter.
23. Parfit (1984), Chapter 16.
24. Parfit (1984), p. 367.
25. McMahan (1981), p. 120, n. 28.
26. Benatar (2006), p. 29.
27. Benatar (2006), p. 30.
28. Benatar (2006), p. 41.
29. Benatar (2006), p. 39.
30. Benatar (2006), p. 31.
31. Benatar (2006), p. 32.
32. Benatar (2006), p. 34.
33. Benatar (2006), p. 34.
34. Benatar (2006), p. 35.
35. Roberts (2002), p. 326. See also Roberts (2004).
36. Roberts (2004), p. 101.
37. Roberts (2002), p. 328, n. 27.
38. For further development of this point, see Persson (2003).
39. Broome (1999), p. 168.
40. This restates reasons that I have discussed earlier, e.g., Persson (1995), pp. 27–9.
41. McMahan (2002), p. 74.
42. McMahan (2002), p. 354.
43. Benatar (2006), p. 45.
44. Benatar (2006), p. 26.
45. Benatar (2006), p. 26.
46. McMahan (2002), pp. 353ff. I am talking about preventing the existence of the less well-off child rather than simply omitting to cause its existence in order to exclude the influence of the feeling of less responsibility for omissions than for actions.

References

- Benatar, D. 2006. *Better never to have been*. Oxford: Clarendon Press.
- Broome, J. 1999. *Ethics out of economics*. Cambridge: Cambridge University Press.
- Hart, H. 1955. Are there any natural rights? *The Philosophical Review* 64: 175–91.
- Kamm, F. 1992. *Creation and abortion*. New York: Oxford University Press.
- Locke, J. 1975. *An essay concerning human understanding*, ed. P. Nidditch, Oxford: Clarendon Press.
- Locke, J. 1990. *Two treatises of government*. London: J. M. Dent & Sons.
- McMahan, J. 1981. Problems of population theory. *Ethics* 92: 96–127.
- McMahan, J. 2002. *The ethics of killing*. New York: Oxford University Press.
- Moller Okin, S. 1989. *Justice, gender, and the family*. New York: Basic Books.
- Nozick, R. 1974. *Anarchy, state, and utopia*. New York: Basic Books.
- Parfit, D. 1984. *Reasons and persons*. Oxford: Clarendon Press.
- Persson, I. 1994. The groundlessness of natural rights. *Utilitas* 6: 9–24.
- Persson, I. 1995. Genetic therapy, identity and person-regarding reasons. *Bioethics* 9: 16–31.
- Persson, I. 2003. Two claims about potential human beings. *Bioethics* 17: 503–16.
- Persson, I. 2005. *The retreat of reason: A dilemma in the philosophy of life*. Oxford: Clarendon Press.
- Persson, I. 2007. The act-omission doctrine and negative rights. *The Journal of Value Inquiry* 41: 15–29.
- Roberts, M. 2002. A new way of doing the best we can: Person-based consequentialism and the equality problem. *Ethics* 112: 315–50.
- Roberts, M. 2004. Person-based consequentialism and the procreation obligation. In *The repugnant conclusion*, eds. J. Ryberg and T. Tännsjö. Dordrecht: Kluwer.
- Sikora, R. 1978. Is it wrong to prevent the existence of future generations? In *Obligations to future generations*, eds. R. Sikora and B. Barry. Philadelphia: Temple University Press.
- Steiner, H. 1994. *An essay on rights*. Oxford: Blackwell.
- Tooley, M. 1983. *Abortion and infanticide*. Oxford: Clarendon Press.
- Velleman, J.D. 2008. Persons in prospect. *Philosophy and Public Affairs* 36: 221–88.
- Waldron, J. 1988. *The right to private property*. Oxford: Clarendon Press.

Chapter 3

Asymmetries in the Morality of Causing People to Exist

Jeff McMahan

Abstract This paper questions the justification for the common view that there is a moral reason not to cause a person to exist if his life would be miserable, but no reason to cause a person to exist because his life would be worth living. It argues that this asymmetry presupposes an ad hoc claim about the different ways in which good and bad states in an individual's life have value. The claim that there is a moral difference between harming and benefiting is more plausible but supports only a weaker asymmetry that concedes that there is a moral reason to create lives worth living.

Keywords Procreation · Asymmetry · Nonexistence · Harming · Benefiting · Life worth living.

3.1 Introduction

Most of us accept the following two propositions.

- (1) That a person would have a life that is “worth not living”—a life in which the intrinsically bad states outweigh the good—provides a moral reason not to cause that person to exist, and indeed a reason to prevent that person from existing.
- (2) That a person would have a life worth living does not, on its own, provide a moral reason to cause that person to exist, though there is no general moral reason *not* to cause such a person to exist.

In 1981, in my first published paper, I referred to this pair of propositions as the *Asymmetry*.¹ My claim then was that although the Asymmetry is intuitively compelling, it is extraordinarily difficult to defend or justify. That will be my contention again now, 27 years later. I suppose it is some consolation that this conclusion will be bracing for those who are fond of claiming that there is no progress in philosophy.

J. McMahan (✉)
Philosophy, Rutgers University, New Jersey, USA
e-mail: McMahan@Philosophy.Rutgers.edu

3.2 Definitions and Distinctions

For brevity of exposition, I will use some rather silly shorthand terms, such as “happy people” and “miserable people.” By “happy people” I will mean people whose lives are worth living, in the sense that the good elements in their lives objectively outweigh the bad. “Miserable people” are those in whose lives the bad elements outweigh the good. I will not discuss, and will leave unlabeled, those in whose lives the good elements are neither greater nor less than the bad.

I will refer to intrinsically good states within a life simply as “goods” and to intrinsically bad states as “bads,” despite the unidiomatic use of “bad” as a noun. A good and a bad are “equivalent” when they are roughly equal in magnitude on either side of the zero point. A good and a bad are *subjectively* equivalent if the potential subject is indifferent between having both and having neither. This is, in my view, a mistaken measure of equivalence, but I mention it just to clarify the notion of equivalence.

It will be important throughout the subsequent discussion to distinguish among three types of value, or ways in which things may be valuable. When something is *better for* an individual, it has *individual-affecting value*; when something is *worse for* an individual, it has negative individual-affecting value.² Note that individual-affecting value is defined comparatively: if having a certain good is better for an individual, its absence would have been worse for that individual.

There are, however, some things that are good or bad *for* an individual in an essentially noncomparative way. For example, to be caused to exist with a life worth living seems to be good for the individual to whom it happens. There is no problem in identifying the subject of this good. Yet the alternative in which that individual would not have had that good—that is, the alternative in which she never exists—would not have been bad or worse *for her*, since nothing can be good or bad for someone who never exists. Similarly, to be caused to exist with a life that is not worth living is *bad* for the individual who comes to exist but not *worse* for her.

Some people might wish to regard continuing to exist as a noncomparative good or bad in the same way that coming into existence is. Epicureans, for example, hold that ceasing to exist can be neither better nor worse for an individual, nor even good or bad for that individual, since they claim that there is no one for whom death, which involves ceasing to exist, could be good or bad, or better or worse. They could, however, still accept that *continuing to exist* could be good or bad for an individual in an essentially noncomparative way—that is, it could be good to continue to exist if life would be worth living, even if ceasing to exist would be neither bad nor worse than continuing to exist.³

I will refer to the sort of value that may be realized by an individual’s coming into existence as *noncomparative individual-affecting value*, or *noncomparative value* for short. I will reserve the term “individual-affecting value” for things that are better or worse for individuals, on the assumption that existing cannot be better or worse than never existing, and leaving open the question whether continuing to exist can be better or worse than ceasing to exist.

The third category of value is *impersonal value*. Value is impersonal when it is neither good or bad nor better or worse *for* anyone. All impersonal value

is comparative. If something is impersonally good, considered on its own, its absence would be impersonally worse, other things being equal. If it is impersonally good, *all things considered*, its absence would be impersonally worse *simpliciter*.

The category of impersonal value is heterogeneous. For all or at least most individual-affecting values, there are corresponding impersonal values. An individual's suffering is, in itself and apart from its effects, bad in individual-affecting terms. But it is also impersonally bad. Its presence makes the world worse. That individual-affecting and impersonal values overlap in this way makes it particularly difficult to understand how the two types of value can be combined, aggregated, or weighed against one another.⁴

Some impersonal values, however, have no connection to individual well-being. Species diversity, for example, might have value even if it were not instrumentally valuable for the individual members of any species. Other impersonal values make essential reference to individual well-being and yet are distinct from, need not coincide with, and can indeed potentially conflict with individual-affecting value. Deserved harm is one example. While deserved good is both good for the deserving person and good impersonally, deserved harm is good impersonally but bad, or worse, for the person harmed. *Equality* of well-being is another example. For those who accept that equality of well-being is a genuine value that is distinct from the prioritarian view that the value of making an individual better off by some fixed amount is greater the worse off that individual is, there is value in *reducing* the well-being of the best off even if this does nothing to raise the well-being of anyone else. For the increase in equality of well-being has impersonal value even if in individual-affecting terms the effects are wholly bad (so that the individual-affecting bad effects may outweigh the impersonal good effect).⁵

Impersonal value may also make essential reference to well-being and yet not correspond to any individual-affecting value. Because impersonal value is comparative, it is impersonally better, other things being equal, if an individual whose life would not be worth living never exists than if she does exist. Similarly, it is impersonally worse if an individual whose life would be worth living never exists than if she does exist. In such cases, there may be no one for whom an individual's never existing is better or worse, or even noncomparatively good or bad.

Although I have distinguished three kinds of value—individual-affecting value, noncomparative value, and impersonal value—these kinds of value give rise to only two kinds of moral reason: individual-affecting and impersonal. One has both an individual-affecting and an impersonal reason not to cause an existing individual to suffer. One's reason to preserve or promote the diversity of species may, however, be impersonal only, and the same may be true of one's reason to increase equality of well-being (for example, when the worse off would gain less than the better off would lose) or one's reason not to cause a miserable person to exist—though of course in all these cases one's impersonal reason may be reinforced by a distinct individual-affecting reason: to preserve the lives of the existing members of a species, to raise the well-being of the worst off, and to spare existing people the burden of caring for a miserable person. Yet the individual-affecting reason and the impersonal reason are never the *same* reason.

There is, admittedly, some controversy about whether choices between causing and not causing a person to exist are governed by individual-affecting or impersonal reasons. No one, of course, doubts that individual-affecting reasons can be relevant insofar as causing a person to exist can affect the well-being of others. But are there individual-affecting reasons deriving from a concern for the well-being of the person who might be caused to exist? Suppose that because one has a certain genetic abnormality, any child one might have would be miserable—that is, would have a life in which the bads would outweigh the goods. Most of us agree that there is a strong moral reason not to cause such an individual to exist. Some philosophers think this is an individual-affecting reason, one that does *not* derive from the effects the act would have on preexisting people. But this is a mistake. It is true, of course, that *if* one does cause such an individual to exist, there will then be someone for whom one's act was *bad*, though not *worse*. Yet at the time of one's choosing between acting and not acting, there is no one whose interests would be affected by one's choice. If one were to act on the reason not to cause an individual to exist, there would never be anyone for whom that would be better. And if one were to act *against* that reason, there would never be anyone for whom one's act was worse. The fact that acting against the reason would be bad in noncomparative individual-affecting terms seems insufficient to ground an individual-affecting reason not to cause a miserable person to exist.

Although it may seem obvious that one's reason *not* to cause an individual to exist must be impersonal, since if one acts on that reason there will never be anyone for whom that would be better or worse, or even good or bad, it may nevertheless seem that one's reason to *cause* a person to exist, *if* one were to have such a reason, could be individual-affecting, since acting on that reason would be *good for* the person who would then exist, and would produce noncomparative individual-affecting value. But it seems best to classify one's reason as impersonal in this case as well. At the time of one's choice, there is no one who exists or will exist independently of that choice for whose sake one could be acting in causing him or her to exist. If one chooses to cause an individual to exist, that may be good for the individual who comes to exist, but it cannot be one's reason for acting, or one's intention in acting, to bestow that good on that individual. At most one's reason might be to create additional noncomparative individual-affecting value, which would then necessarily attach to and be good for someone. But the creation of this good would not be *better* for that individual, nor would its not being created have been worse for him or her. It seems, therefore, that any moral reason to cause or not to cause an individual to exist that derives from a concern for what that individual's well-being is best considered an impersonal rather than an individual-affecting reason.

3.3 Reasons and Values Presupposed by the Asymmetry

Having drawn certain distinctions, we can now consider what assumptions about values and reasons are necessary for the Asymmetry to be true. The first of the two propositions constitutive of the Asymmetry seems to presuppose that it is worse to

cause bads to exist by causing people to exist. Since causing bads in this way is not worse *for* any individual, it can be worse only impersonally, and the reason not to do what is worse only impersonally must itself be an impersonal reason.

The second proposition, however, seems to presuppose either that it is not better to cause goods by causing people to exist or that, even if it is better, there is no reason to do what is better in this way. It denies, in other words, either that there is impersonal value in causing goods by causing people to exist or that there is an impersonal reason to create goods that are better only impersonally.

Actually, though, what is presupposed by the two propositions that are constitutive of the Asymmetry is more complex than this. If it were worse to cause bads by causing a person to exist but not good or better in any way to cause goods in this way, it seems that there would be a reason not to cause bads by causing a person to exist but nothing other than extrinsic considerations to oppose this reason. Consideration of the effect on people of being caused to exist would ground a general reason not to cause people to exist, establishing a presumption against having children, even when a child would have a life that would be well worth living. That reason might be contingently offset in many cases by individual-affecting reasons deriving from the interests of preexisting people, making it permissible to cause a person to exist in these cases. One could justify having a child only by showing that the impersonal reason not to cause the many bads that the child's life would inevitably contain would be outweighed by the good effects that the child's existence would have on the lives of others. Procreation would be a *prima facie* objectionable, essentially selfish activity.

This is not only highly implausible but is also incompatible with the second proposition of the Asymmetry, which asserts, in effect, that there is no general reason not to cause people to exist if their lives would be worth living. So the foregoing understanding of the values and reasons that underlie the Asymmetry cannot be right. What is missing is a further distinction. It seems that the Asymmetry presupposes that goods created by causing a person to exist count in one way but not another. They do not count as reasons for causing the person to exist. But they do weigh against and cancel out corresponding bads that the person's life would contain. I will refer to these two ways in which the goods that a person's life would contain might contribute to the justification for causing that person to exist as the *reason-giving function* and the *canceling function*. What the Asymmetry presupposes is that while goods have a canceling function in procreative choices, they have no reason-giving function. They have a kind of impersonal value in that they weigh against and cancel out corresponding bads, but they do not otherwise count in favor of causing a person to exist, no matter how greatly they may outweigh the bads.⁶

Common sense intuitions seem to presuppose all of the following. Individual-affecting goods and bads have both reason-giving and canceling functions. There are reasons to do what is better for people and reasons not to do what is worse for people, and at least in many cases, the good effects of an act on a person weigh against and can cancel out equivalent bad effects in the determination of the act's permissibility. In procreative choices, bads have both impersonal reason-giving and canceling functions, but goods have only an impersonal canceling function. It is

worse in impersonal terms to create bads by causing people to exist, but the negative impersonal value of these bads can be counterbalanced and canceled by the creation of equivalent goods within the life of the same person. While goods thus have impersonal value in that they weigh against bads, they do not have impersonal reason-giving value. The goods that an individual's life would contain if he or she were caused to exist can contribute to the permissibility of causing that individual to exist by weighing against and canceling out the bads that the life would contain, but they provide no positive reason for causing the individual to exist.

These seem to be the assumptions about values and reasons that underlie the Asymmetry, but they seem strikingly ad hoc. Why should goods have both reason-giving and canceling functions in individual-affecting choices but only a canceling function in procreative choices? Why should goods have one kind of impersonal value—canceling value—but not the other—reason-giving value?

My impression is that no one has tried to answer these questions, in part because no one has asked them. But having asked them, I confess that, at the moment at least, I do not know how to answer them. In the absence of a satisfactory answer to these questions, it is tempting to look for the justification for the Asymmetry in more familiar asymmetries in common sense morality.

3.4 The Distinction Between Harming and Benefiting

One such asymmetry is that between harming and benefiting. According to this view, the reason not to cause harm is stronger than the reason to confer an equivalent benefit, if other things are equal; therefore, to harm an individual is worse, or more seriously morally objectionable, than not to benefit that individual to an equivalent degree. If causing a miserable person to exist is an instance of harming, or doing harm, while not causing a happy person to exist is an instance of not benefiting, then the Asymmetry may be just a manifestation in the area of procreation of the familiar moral difference between harming and not benefiting.

This claim is, however, more problematic than it may seem. For our ordinary concepts of harming and benefiting seem to be comparative and thus to have no application to acts of procreation. Both harming and benefiting, as ordinarily understood, involve affecting whether an individual's well-being improves or declines, or does not improve or does not decline, relative to some baseline. The usual baseline is either temporal (what an individual's well-being was *before* it was affected by an act) or counterfactual (what an individual's well-being would have been had it not been affected by that act). Because improvement and decline relative to some baseline of well-being are essentially comparative notions, harming and benefiting seem to be comparative notions as well.

Understood in this way, harming, which is an instance of *doing*, involves causing an individual's well-being to decline or preventing it from improving, while benefiting involves causing an individual's well-being to improve or preventing it from declining. Not harming is not causing a decline or preventing an improvement,

while not benefiting is not causing an improvement or not preventing a decline. The distinction between harming and not benefiting is therefore compounded from two further distinctions of presumed moral significance: that between doing and not doing, and that between a decline in well-being and an improvement in well-being.

There is another dimension to the distinction between harming and benefiting that is worth noting, which is that it may also be sensitive to where, in relation to the zero-point on the scale that measures well-being, an individual's well-being is before an act is done, where it is after an act is done, or where it would be had the act not been done. Thus, preventing an improvement in the well-being of an individual whose level of well-being is already extremely high may not count as harming. Similarly, causing an improvement in the well-being of an individual whose level of well-being is below the zero-point—that is, whose life is miserable or not worth living, counts not only as benefiting but as mitigating or reducing a harm.

Obviously, however, when one causes an individual to exist, comparisons cannot be made between that individual's well-being once she exists and her level prior to the act that caused her to exist, or the level of well-being she would have had if the act of causing her to exist had not been done. One possibility for bringing acts of procreation within the scope of the concepts of harming and benefiting is, as some philosophers have done, to assign prenatal nonexistence a value of zero, effectively locating it at the zero-point on the scale that measures well-being, both positive and negative. One could then take the zero-point as the relevant baseline, claiming that causing an individual to exist harms her to the extent that her life falls below the zero-point, or benefits her to the extent that it is above the zero-point. This would, of course, require a determination of whether the object of evaluation would be her entire life, so that one would be unable to tell whether and to what extent she had been benefited or harmed until after she had died, or whether the evaluation would consider only those aspects of her life properly attributable only to her being caused to exist. But the same problem besets any minimally sophisticated account of how causing an individual to exist can harm or benefit that individual.

There is, however, no need to adopt this strained proposal, which treats prenatal nonexistence as the evaluative equivalent of a state of existence in which the intrinsic goods neither outweigh nor are outweighed by the intrinsic bads. For there is a better way of extending or expanding the notions of harming and benefiting to apply to acts of procreation. This is to claim, plausibly, that the notions of harming and benefiting have a noncomparative dimension as well as the familiar, and dominant, comparative dimension. It seems intelligible to say that to cause a miserable person to exist is to harm that person. All that need be meant by this is that the person is harmed by being caused to be in a state that is on the whole intrinsically bad for her, because the bads of her life outweigh the goods. Similarly, even if one agrees with the Epicureans that death cannot be bad for a person, one can still coherently claim that to save a person's life is to benefit her, if it causes her to be in a state that is intrinsically good for her.

Accepting that our ordinary concepts of harming and benefiting contain an implicit noncomparative dimension in addition to a comparative dimension seems preferable to claiming, as some philosophers have done, that the concepts are best

analyzed in wholly noncomparative terms. For accounts of this latter sort have the counterintuitive implication that to kill a person, or at least to kill a person painlessly, is not to harm her. (The familiar comparative account has this same implication as well *if* it is combined with the Epicurean claim that ceasing to exist cannot be bad or worse for a person.)

Suppose that we have now identified an unproblematic sense in which to cause a person to exist can be to harm or benefit that person. This allows for the coherence of the claim that the Asymmetry is just an implication of the familiar view that harming is more seriously objectionable than not benefiting. For to cause a miserable person to exist is to harm her noncomparatively, by causing her to be in a state that is intrinsically bad for her, while not causing a happy person to exist is a peculiar way of failing to bestow a peculiar benefit on the person who would otherwise have existed.

One might object to this explanation of the moral basis of the Asymmetry on the ground that, while to cause a miserable person to exist is a clear instance of noncomparative harming, not to cause a happy person to exist is *not* an instance of not benefiting, for in the latter case there is no one who is not benefited, no one of whom it is true that she could have been benefited but was not.

The problem with this objection, however, is that there is a parallel claim about not harming that seems to undermine the claim that the moral reason not to cause a miserable person to exist is that to do so would be to harm her. Let us assume that it is unproblematic to say that to cause a miserable person to exist is an instance of harming, for there is no problem in identifying the victim of the harm. The objection is instead to the claim that not to cause a happy person to exist is an instance of *not benefiting*. For in this case there is no one who is not benefited. Yet to the extent that these claims are plausible, it must be equally plausible that to cause a happy person to exist is to benefit her, while not to cause a miserable person to exist is *not* an instance if *not harming*, for there is not one who is not harmed. But the last of these propositions seems to undermine the view that the reason one has not to cause a miserable person to exist is that to cause him to exist would be to harm him. For this presupposes that the reason not to cause a miserable person to exist is an instance of the more general reason *not to cause harm*; yet in this case not to cause the miserable person to exist would *not* be an instance of not causing harm, for there would be no one who was not harmed. In short, the claim that not causing a happy person to exist cannot be an instance of not benefiting implies that not causing a miserable person to exist also cannot be an instance of not harming. And this undermines the claim that the reason not to cause a miserable person to exist is that to cause him to exist would be to harm him and there is a moral reason not to harm him.

Let us grant, then, that not to cause a miserable person to exist is indeed an instance of not harming. We should then also accept that not to cause a happy person to exist is an instance of not benefiting. The suggestion is that, on these assumptions, the Asymmetry can be explained and defended by reference to the more general asymmetry between harming and benefiting, and in particular the general claim that harming is worse than not benefiting. But while the asymmetry between harming and benefiting provides some support for the Asymmetry, the support falls well short

of entailment. For the Asymmetry holds that there is *no* moral reason to bestow a benefit by causing an individual to exist. But the common sense asymmetry between harming and benefiting does not deny that there is a general moral reason to benefit people; it merely denies that the reason to benefit people is *as strong* as the reason not to harm people to an equivalent degree. So what the general asymmetry between harming and benefiting entails, on the assumption that causing an individual to exist can be an instance of noncomparative harming or benefiting, is not the Asymmetry but what I will call the *Weak Asymmetry*.

According to the Weak Asymmetry, the reason not to cause harm by causing a miserable person to exist is stronger, perhaps considerably stronger, than the reason to bestow an equivalent benefit by causing a happy person to exist, so that it is more seriously morally objectionable to cause harm by causing a miserable person to exist than it is not to bestow a benefit by failing to cause a happy person to exist. The Weak Asymmetry, in other words, discounts the reason-giving weight that potential goods have in procreative choices, though not all the way to zero. If the Weak Asymmetry is simply an implication of the general common sense asymmetry between harming and benefiting, the degree to which the reason-giving weight of the goods in a possible person's possible life is discounted should be the same as the degree to which benefits are discounted relative to harms in ordinary individual-affecting choices.

Ought the canceling weight of potential goods to be discounted in procreative choices as well? That is, does a good that an individual's life would contain if she were caused to exist cancel out an *equivalent* bad that the life would also contain, or does it cancel only a lesser bad? If the former—that is, if the canceling weight of goods is not discounted—it can be permissible to cause an individual to exist whose life would be just barely worth living, even if this person's existence would have no positive effect on the lives of others. If, by contrast, the canceling weight of goods is discounted along with the reason-giving weight, the goods that an individual's life would contain must exceed the bads by a certain margin for it to be permissible to cause that individual to exist. The extent to which the goods must exceed the bads is determined by the degree to which the canceling weight is discounted. There are some people whose views about causing individuals to exist may find support in the view that in procreative choices the canceling weight of potential goods is discounted along with the reason-giving weight. Frances Kamm, for example, has suggested that “we should not create persons at will unless we have good reason to believe that they can have some . . . number of years of life with some degree of health and welfare.”⁷ And David Benatar claims that “for a life to be not worth continuing, it must be worse than it need be for it not to be worth starting,” which seems to presuppose that while potential goods can cancel equivalent bads in individual-affecting choices, their canceling weight is diminished in procreative choices.⁸

Again, if the Weak Asymmetry derives from the general common sense asymmetry between harming and benefiting, we could determine whether the canceling weight of potential goods in procreative choices is discounted by seeing whether it is discounted in individual-affecting choices. And it seems that it is—at least in

many cases, though there seems to be some variation with an agent's intentions. Suppose that one could confer a benefit on a person but only by intentionally causing him to suffer an equivalent harm. And suppose that the benefit and the harm are not only objectively equivalent but subjectively equivalent as well, in that if this were a prudential choice the person himself would be indifferent between having the benefit together with the harm and having neither. Most people believe that it would nevertheless be wrong to cause the harm as a means of conferring the benefit, in the absence of the person's consent. If that is right, the good does not cancel the harm and thus its canceling weight seems to be discounted. Yet it is at least arguable that it would be permissible to bestow the benefit intentionally, foreseeing that the equivalent harm would also occur, or to cause both the benefit and the harm as side effects of action intended to produce a benefit for a different person.

It seems, therefore, that there is a presumption in procreation cases that the canceling weight of goods that a possible person's life would contain is discounted, though perhaps not to the same degree that the reason-giving weight is, according to the Weak Asymmetry. For even though the canceling weight of goods in individual-affecting choices seems to be discounted in most or all cases, there is no reason to suppose that the degree of the discounting must be the same as the degree to which the reason-giving weight of those goods is discounted relative to reason-giving weight of equivalent harms.

If the canceling weight of potential goods is discounted in procreative choices, this should be reflected in the Asymmetry as well as in the Weak Asymmetry. The Asymmetry, which I take to be the common sense view, should be revised so that it claims that in procreative choices, while the potential goods that an individual's life would contain have no reason-giving weight, they do have canceling weight, though that weight is discounted; therefore, while it is *in general* permissible to cause happy people to exist, there is a threshold on the scale that measures well-being such that if an individual's life would fall below that threshold, it is wrong to cause him to exist, even though his life would be worth living and his existence would have no effects on others.

3.5 The Individual-Affecting Symmetry View

Thus far we have considered two views of the morality of procreation: the Asymmetry and the Weak Asymmetry. I suggested that the fundamental presupposition of the Asymmetry is that, while in individual-affecting choices goods or benefits have both reason-giving weight and canceling weight, in procreative choices they have no reason-giving weight, yet retain the same canceling weight they have in individual-affecting choices. This is peculiar and requires both explanation and defense; yet to my knowledge it has never received either. An appeal to the common sense moral asymmetry between harming and benefiting leads not to the Asymmetry but only to the Weak Asymmetry, to which we will return.

At this point we should consider the alternative symmetrical views of the morality of procreation. This will help us to appreciate both the importance of being able

to defend an asymmetrical view of the morality of procreation and the problems facing any version of the Weak Asymmetry that diverges significantly from the Asymmetry.

Some philosophers have resisted the idea that there can be impersonal values, or indeed even noncomparative individual-affecting values. The only values, on their view, are individual-affecting values, and the only reasons for action are individual-affecting reasons.⁹ According to this view, an act can be wrong only if there is some individual who exists at some time for whom the act is worse—though note that this allows that an act can be worse for an individual simply by excluding an alternative act that would have benefited that individual. Because there is no one for whom never existing can be worse (or even bad), it cannot be wrong to fail to cause a happy person to exist, unless there is an individual who already exists, or who will exist independently of the choice of whether to cause a new person to exist, for whom the choice not to cause a new person to exist would be worse. This individual-affecting view thus implies the second of the two views that together constitute the Asymmetry.

But it cannot imply, and is not even compatible with, the first of these two claims. One can have an individual-affecting reason to do or not to do a certain act only if the act would be better or worse for some individual. But to cause a miserable person to exist cannot be worse for that person, since “worse for” implies a comparison with an alternative that would be better for the same individual. But the relevant alternative to causing a miserable person to exist is simply not to cause that person to exist, in which case there is never anyone for whom that alternative is better. Since there can be no one for whom never existing is better, being caused to exist cannot be worse in individual-affecting terms.

I do not claim that the *only* alternative to causing a miserable person to exist is not to cause him to exist. It is possible to individuate and refer to a particular possible person—for example, by reference to a particular gamete pair. And it is also possible that the same possible person could be caused to exist not as a miserable person but in different conditions as a happy person. But in actual cases we do not know how to do this or even how to determine whether it has happened.

Nor do I claim that for an outcome to be better or worse for an individual, there must be an alternative in which that same individual would exist in a worse or better state. I accept, for example, that it can be worse for an individual to die, even though to be dead is not to be in a worse state. But in this case there is someone for whom continued life would be good but for whom death would involve nothing at all. Because something good is better than nothing *for this person*, it is coherent to claim that continuing to live would be better for him and that death would be worse, even though death would involve his no longer existing. But what is true of ceasing to exist is not true of never existing. Even if we could individuate and refer to a possible person who would have a life worth not living if he were to exist, there will never be anyone for whom never existing is better than being caused to exist as a miserable person *if* in fact we do not cause him to exist.

The claim that ceasing to exist can be worse for an individual while never existing cannot be seems to imply that there is a significant difference between an

individual's never existing and an individual's beginning to exist and then immediately ceasing to exist. In the former case, there is no one for whom nonexistence is worse, while in the latter case there is. Yet intuitively there seems to be little or no difference between these cases.

This is a serious objection to my claim only if my claim implies that an individual's ceasing to exist immediately after beginning to exist is *significantly* worse for that individual than continuing to exist would have been. My claim might indeed imply this *if* we came into existence as fully formed persons, as Athena emerged from the head of Zeus. In that case, death immediately after beginning to exist might plausibly be regarded as a tragedy. But we instead begin to exist as barely conscious beings with no psychological connections to our own future selves. As I have argued elsewhere, death is not a significant harm to such beings.¹⁰ So while there is someone for whom death is worse even when it occurs immediately following the beginning of existence, it is worse only to a slight degree. Hence there is no morally significant difference between ceasing to exist immediately after beginning to exist and never existing at all, even though the former may be worse or better in individual-affecting terms while the latter cannot be.

Given that being caused to exist can be neither better nor worse for the individual who comes into existence, the view that morality takes an exclusively individual-affecting form implies what I will call the Individual-Affecting Symmetry view of the morality of procreation. According to this view, there is complete moral symmetry between the creation of happy people and the creation of miserable people. Apart from the effects on people who already exist or who will exist, or will exist independently of one's procreative choice, it is a matter of moral indifference whether one causes a happy person or a miserable person to exist. Where the possible interests of possible persons are concerned, there is no reason to cause a happy person to exist, though no reason not to; similarly, there is no reason to cause a miserable person to exist but also no reason not to. This last implication of the view is, however, intuitively unacceptable. A view that implies that there is no moral reason not to cause an individual to exist whose life would be filled with intrinsically bad states, uncompensated for by intrinsically good states, cannot be true.

Some philosophers who were among the earliest to respond to Parfit's challenges to the view that morality takes a wholly individual-affecting form argued that it is not intolerable if the view implies the permissibility of causing miserable people to exist, for it also implies that there would be a moral requirement to euthanize any individual whose life would be miserable as soon as he or she began to exist. This could conceivably be an adequate response if euthanasia were always possible the instant such an individual began to exist and before he or she became a person whose consent would be required for euthanasia to be permissible. But given that in practice it is sometimes possible to prevent the existence of individuals whose miserable lives would not or could not be ended until after they had already endured great suffering, it seems that we must recognize that at least in such cases there is a moral reason not to cause such a person to exist, and that this reason need not have anything to do with the interests of others.

3.6 The Antinatalist Symmetry View

As I have noted, the Asymmetry seems to presuppose that intrinsically bad states have not only negative individual-affecting value but also negative impersonal value. There is not only an individual-affecting reason not to cause an existing individual to suffer something bad but also an impersonal reason not to cause a miserable person to exist, since this would cause him to suffer uncompensated bads. It would be bad for him—a noncomparative harm—even though it would not be worse for him. It *would* be worse, but not for him: it would be worse only impersonally. Common sense morality, of which the Asymmetry is an element, accepts that intrinsically good states have individual-affecting value: there is, in general, a moral reason to benefit existing individuals, or to cause them to have intrinsic goods in their lives when this is overall better for them. But there is no reason to bestow noncomparative benefits by causing happy people to exist. Common sense morality does, however, recognize the permissibility, in general, of causing happy people to exist. In doing so, it implicitly concedes that intrinsic goods or noncomparative benefits have a kind of impersonal value in procreative choices. While they do not have impersonal reason-giving value, they do have impersonal canceling value. The intrinsically bad states that any life will inevitably contain have impersonal reason-giving weight: they count morally against procreation. This is a presupposition of the Asymmetry. So in order for procreation to be permissible, there must be other values that weigh against these bads. In most cases, the desires and interests of preexisting people weigh against them. But unless individual-affecting value has in general much greater weight than impersonal value, these considerations are seldom sufficient, on their own, to outweigh the impersonal reason not to cause an individual to exist that derives from the fact that his life will inevitably contain a great deal of suffering and other intrinsic bads. What makes procreation morally permissible in most cases is the reasonable expectation that the bads in a possible person's life will be outweighed, and significantly outweighed, by the goods. The goods that a possible person's life would contain thus have impersonal canceling value, even though they do not have impersonal reason-giving value. This is the presupposition that allows common sense morality to recognize the permissibility of causing happy people to exist (or at least happy people with lives above some threshold of well-being), given its recognition that the bads that an individual's life would contain ground an impersonal reason not to cause that individual to exist, or even a reason to prevent that individual from existing.

Yet, as I noted earlier, it seems odd to suppose that the goods an individual's life would contain have impersonal canceling value but not impersonal reason-giving value. It might be theoretically more consistent to deny that goods have any impersonal value of any kind. One might claim that while bads have four kinds of value—individual-affecting reason-giving value, individual-affecting canceling value, impersonal reason-giving value, and impersonal canceling value—goods have only two: individual-affecting reason-giving value and individual-affecting canceling value. In other words, while bads have both individual-affecting and impersonal value, goods have only individual-affecting value. They have no impersonal value of any kind.

This view seems to be implicit in many passages in David Benatar's recent book, *Better Never to Have Been*, which argues against the general moral permissibility of procreation. Benatar writes, for example, that "the absence of pain is good, even if that good is not enjoyed by anyone, whereas. . .the absence of pleasure is not bad unless there is somebody for whom this absence is a deprivation."¹¹ He then argues for these claims on the ground that they have "considerable explanatory power," citing as the first and main example of this power that they explain and justify the Asymmetry. They provide, he claims,

the best explanation for the view that while there is a duty to avoid bringing suffering people into existence, there is no duty to bring happy people into being. In other words, the reason why we think that there is a duty not to bring suffering people into existence is that the presence of this suffering would be bad (for the sufferers) and the absence of the suffering is good (even though there is nobody to enjoy the absence of suffering). In contrast to this, we think that there is no duty to bring happy people into existence because while their pleasure would be good for them, its absence would not be bad for them (given that there would be nobody who would be deprived of it).¹²

According to the view advanced in this passage, it is not a reason to cause happy people to exist that this would be good for them in noncomparative individual-affecting terms. And it is not good for them in ordinary individual-affecting terms to be caused to exist—that is, it is not better for them, since their not existing would not be worse, or even bad, for them. Nor is there an impersonal reason to cause them to exist. For it is implicit in the final sentence of this passage that if not causing them to exist would not deprive them of goods and thus would not be bad or worse *for them*, it cannot be bad or worse in *any* way relevant to our duties. The absence of the good their lives would have contained might or might not be bad impersonally, but even if it is, that is irrelevant, according to Benatar.

But matters are different, he claims, in the case of suffering. If miserable people are not caused to exist, this is good, or better than if they were caused to exist, even if there is no one for whom it is good, or better. There seems to be no way to understand this claim except as a claim about impersonal value. If it is good that suffering or miserable people do not exist, even though it is not good or better for anyone, how else can we understand the status of this good except as a good that is not *good for*—that is, except as an impersonal good?

So suppose we accept that bads have both individual-affecting and impersonal value, while goods have only individual-affecting value. If we deny not only that goods have impersonal reason-giving value but also that they have impersonal canceling value, we arrive at the view that Benatar defends: that procreation is in practice always bad in its effects on the individual who is caused to exist, so that it can be justified only by reference to the interests of others. For the bads that a life would contain count against procreation and goods do not count in favor, even to the extent of weighing against and canceling out the bads. If a life would contain a single bad, that bad is uncompensated for no matter how much good the life would contain, so it is wrong to create the life unless the benefits to others outweigh the noncomparative harm to the person whose life it would be.

We can follow Benatar in describing this view as “antinatalist.” We can thus call the view that there is a moral reason not to cause a miserable person to exist and a moral reason not to cause a happy person to exist the *Antinatalist Symmetry* view. This view is in one respect the antithesis of the Individual-Affecting Symmetry, which holds that there is no reason not to cause a miserable person to exist and no reason not to cause a happy person to exist—though it also holds that there is no reason to cause either to exist.

Although I have introduced this view by reference to some suggestive passages in his book, Benatar himself denies that the basis of his antinatalist view is that bads have impersonal value while goods do not, so that it is good if the bads that a life would contain are never created, independently of whether they would be accompanied by greater goods if they were created.¹³ He seems, in fact, to reject impersonal values and writes as if his view were supported entirely by individual-affecting considerations. It is instructive to look at what he says, as it helps to show why an asymmetrical view is difficult to defend.

Benatar writes that “we have a strong moral reason, *grounded in the interests of potential people*, to avoid creating unhappy people.”¹⁴ And in a footnote appended to this sentence, he explicitly contrasts his view with an impersonal view, noting that “the condition that the moral reason (or duty) be grounded in the interests of the potential person is an important one. Those who find plausibility in the claim that we have a reason to create happy people tend to be motivated by impersonal considerations—such as there being more happiness in the world. But these are not considerations about the interests of the potential person.”¹⁵ Yet how can the reason not to cause a miserable person to exist be grounded in the interests of the potential person when, if there were a reason to cause a happy person to exist, it would have to be grounded in impersonal considerations? Here is what Benatar says of a person, X, who never existed but would have had a pleasurable life if he had: “if the absence of pleasure. . . is ‘bad’ rather than ‘not bad’ then we should have to regret, *for X’s sake*, that X did not come into existence. But it is not regrettable.”¹⁶ Yet a parallel claim can be made about a person, Y, who never existed but who would have been miserable if he had: if the absence of suffering or misery is “good” rather than “neither good nor bad,” we should have to be *relieved, for S’s sake*, that S did not come into existence. Yet there is no one for whom we can feel relieved.

In short, if there is an individual-affecting reason not to cause a miserable person to exist, there should be an individual-affecting reason to cause a happy person to exist as well. Indeed, it may even make more sense to suppose that there can be an individual-affecting reason to cause a happy person to exist, for if one acts on that reason there will be an actual individual for whom the act was good, whereas if one acts on the reason not to cause a miserable person to exist, there will be no one for whom the act was good. Yet there cannot be individual-affecting reasons either to cause people to exist or not to cause them to exist (other, of course, than any reasons based on effects on others). The reason to cause an individual to exist and the reason not to cause an individual to exist must be impersonal. It is true, of course, that if one causes a miserable person to exist, there will be someone for whom one’s act was bad; but it is also true that if one causes a happy person to exist, there will be

someone for whom one's act was good. If one does not cause a happy person to exist, there will of course be no one for whom that is bad; but it is also true that if one does not cause a miserable person to exist, there will be no one for whom that is good. The cases are exactly alike except that one involves a life that worth living while the other involves a life that is not worth living. It is for this reason that asymmetrical views of procreation are difficult to defend.

If we accept, as we must, that there is a strong moral reason not to cause a miserable person to exist, we must accept that there are values and reasons that are impersonal. If we believe that the reason not to cause a person to exist if his life would be miserable is no weaker than the reason not to cause an existing person to suffer an equivalent amount of misery, we must accept that at least in some cases, impersonal reasons have the same strength as corresponding individual-affecting reasons.¹⁷ It is the second of the two claims that together constitute the Asymmetry that is more difficult to defend. One way to defend it is to deny that goods have impersonal value. The denial that they have any kind of impersonal value at all entails the Antinatalist Symmetry. If that view had intuitively compelling implications, we might accept it despite the fact that it denies to goods the impersonal value it assigns to bads, apparently without any theoretical justification. Yet its implications—principally the implication that there is always a strong moral presumption against procreation—are profoundly counterintuitive. We can, of course, adopt the weaker claim that goods have impersonal canceling value while lacking impersonal reason-giving value. This yields the common sense Asymmetry. But this understanding of the value of goods seems even more arbitrary and ad hoc than the suggestion that they lack impersonal value of any sort.

3.7 The Impersonal Symmetry View

A further possibility is that both bads and goods have all the same kinds of value: both individual-affecting and impersonal reason-giving value and individual-affecting and impersonal canceling value. The apparent arbitrariness of claiming that bads and goods have different kinds of value forms the basis of a rather simple argument, which I call the *Symmetry Argument*, for a view of the morality of procreation that I call the *Impersonal Symmetry* view.¹⁸ The argument is as follows.

1. To cause a miserable person to exist is bad for him, and harms him noncomparatively, even though it is not worse for him than never existing.
2. There is an impersonal moral reason not to do what would be noncomparatively bad for an individual.
3. There is therefore a reason not to cause a miserable person to exist.
4. To cause a happy person to exist is good for him, and benefits him noncomparatively, even though it is not better for him than never existing.
5. Just as there is an impersonal reason not to do what would be noncomparatively bad for an individual, so there is an impersonal reason of the same strength to do what would be equivalently noncomparatively good for an individual.
6. There is therefore a reason to cause a happy person to exist.

Claim 3 of the Symmetry Argument is the first of the two claims that constitute the Asymmetry. Claims 1 and 2 plausibly explain why 3 is true. Because claims 1 through 3 seem compelling, and claims 4 through 6 are identical except that they refer to a life that is worth living rather than to a life that is not worth living, this argument is difficult to resist. But I think we must find a way to resist it, for its implications seem quite literally unacceptable. Here are some of what seem to be implications of the Impersonal Asymmetry.

- (1) The moral reason not to cause a miserable person to exist is no stronger than the moral reason to cause an equivalently happy person to exist.
- (2) If the reason not to cause a miserable person to exist is as strong as the reason not to cause an equivalent amount of misery to an existing person, the reason to cause a happy person to exist should also be as strong as the reason to provide equivalent benefits to existing people.
- (3) If one could either produce a certain amount of good by causing a happy person to exist or prevent existing people from losing a slightly lesser amount of good, one ought, other things being equal, to cause the happy person to exist.
- (4) Grant the metaphysical assumption that is most favorable to the permissibility of abortion: that we do not begin to exist until the fetal brain acquires the capacity for consciousness, so that most abortions do not involve killing someone like you or me but instead merely prevent one of us from existing.¹⁹ Even on this assumption, if actively preventing a good is more objectionable than failing to create a good, then even early abortion is *more* objectionable morally than failing to cause a happy person to exist, which is just as objectionable as causing the existence of an equivalently miserable person.²⁰
- (5) Saving a person's life does not prevent him from suffering anything intrinsically bad; rather, it merely prevents him from losing the goods of his future life. So both saving the life of a happy person and causing a happy person to exist involve enabling a person to have intrinsic goods. From an impersonal point of view, therefore, the reason to cause a happy person to exist may be just as strong as the reason to save the life of a happy person—indeed, it is usually stronger given that in general causing a person to exist produces more good than saving a life preserves.
- (6) Because death involves only the loss of goods rather than the suffering of anything intrinsically bad, and since an entire life generally contains more good than the *remainder* of another life, it may also be worse in impersonal terms not to cause a happy person to exist than it is to kill an existing person.

These implications are concerned only with comparisons between failing to cause happy people to exist and failing to benefit existing people. But if implication 1 is correct and if the reason not to cause a miserable person to exist is as strong, or even only half as strong, as the reason not to cause an existing person to suffer an equivalent amount of misery, there must then be a further range of disturbing implications concerning comparisons between failing to cause happy people to exist and failing to prevent suffering and misery among existing people, or even causing suffering among existing people.

The implausibility of some of these implications can be mitigated by appealing to the distinctions between doing and not doing and harming and benefiting. As I pointed out earlier, these distinctions provide grounds for the claim that it is more seriously objectionable to cause a miserable person to exist than it is not to cause a happy person to exist. Yet most of the implications just cited involve equivalences between different forms of *not benefiting*. Thus, for example, just as the failure to cause a happy person to exist is an instance of not benefiting, so the failure to save a person is also an instance of not benefiting, in that it involves a failure to prevent a loss of (or decline in) positive well-being. It does, admittedly, involve a failure to prevent what in ordinary language is called a harm—death—but it does not involve a failure to prevent anything intrinsically bad. In impersonal terms, the difference between the failure to cause a happy person to exist and the failure to save a person is just the difference between the failure to create goods and the failure to prevent the loss of goods.

Another response to the unacceptability of the implications of the Impersonal Symmetry is to claim that individual-affecting values and reasons have a general priority over impersonal values and reasons. That is, one might accept that both bads and goods have impersonal value and ground impersonal reasons, but that impersonal values ground weaker reasons than corresponding individual-affecting values. For example, it might be that there is an impersonal reason to cause a happy person to exist but that this reason is weaker by some degree than the corresponding individual-affecting reason to produce an equivalent range of benefits for existing people. In general, there is a stronger reason to produce a good by benefiting an existing person than there is to produce an equivalent noncomparative benefit by causing a new person to exist.

One objection to this proposal is that it seems to be more plausible in the case of goods than in the case of bads. For as I remarked earlier, it may seem that the reason not to cause a miserable person to exist is as strong as the reason not to cause an existing person to suffer an equivalent amount of misery. Or at any rate the former reason is closer in strength to the latter than the reason to cause a happy person to exist is to the reason to benefit an existing person or existing people to an equivalent degree. One might express this suggestion by saying that the impersonal reason-giving value of bads is greater than the impersonal reason-giving value of equivalent goods.

Despite this concern, it seems that we can plausibly reject some of the most disturbing implications of the Impersonal Symmetry by appealing both to the distinction between harming and benefiting and to the claim that at least some individual-affecting reasons are stronger than corresponding impersonal reasons. This would allow us to accept that there are impersonal reasons, and thus to accept the core claims 2 and 5 of the Symmetry Argument, without embracing the Impersonal Symmetry. We can instead accept some version of the Weak Asymmetry.²¹

3.8 Conclusion

Yet because the Weak Asymmetry concedes that there is an impersonal reason to cause happy people to exist, it seems to entail weaker versions of many of the

counterintuitive implications of the Impersonal Symmetry. It seems, for example, to imply all of the following.

- (1) The reason to cause happy people to exist can in certain cases be stronger than the reason to benefit existing people by giving them lesser goods.
- (2) The reason to cause happy people to exist can in certain cases be stronger than the reason to benefit existing people by enabling them to retain or to have goods they would otherwise lose or fail to obtain.
- (3) There is a moral presumption against the permissibility of abortion on the ground that it prevents the existence of a happy person.
- (4) There is some number of happy people such that one's moral reason to cause them to exist would be stronger than and, in a case of conflict, outweigh one's reason to save the life of an existing person.

These claims, while perhaps not impossible to accept, are nevertheless very difficult to believe. It may be that the only view that captures our strongest intuitions about the morality of procreation is the Asymmetry. Yet, as I hope I have indicated in this paper, the prospects for finding a compelling theoretical defense of the Asymmetry are not promising.

Acknowledgements I am grateful to Dan Halliday, Richard Kraut, Derek Parfit, and Dominic Wilkinson for comments on an earlier draft.

Notes

1. McMahan (1981), p. 100.
2. I here modify Derek Parfit's term "person-affecting," since things may be better or worse for individuals who are not persons in the same way that they may be better or worse for persons. Those who believe that there are irreducible collective entities for which things may be better or worse might wish to have a category of "group-affecting" value as well. See Parfit (1984), p. 370.
3. For a more detailed exposition of this view, see McMahan (1988), pp. 34–5.
4. See, for example, McMahan (2001), pp. 445–75 (Section IV).
5. For a powerful defense of this view, see Temkin (2003), pp. 761–82.
6. There may be reasons to doubt whether the distinction between the reason-giving function of a value and the canceling function of that value is coherent. But it is hard to see how one could make sense of the Asymmetry if there were no such distinction.
7. Kamm (1992), p. 132.
8. Benatar (2006), p. 45. As we will see, on Benatar's own view, the canceling weight of goods is discounted all the way to zero in procreative choice.
9. See, for example, Heyd (1992). For further discussion of the individual-affecting view, see my review of Heyd's book, McMahan (1994), pp. 557–9.
10. McMahan (2002), pp. 165–85.
11. Benatar (2006), p. 30.
12. Benatar (2006), pp. 31 and 32.
13. Benatar has confirmed this to me in conversation.
14. Benatar (2006), p. 33. Emphasis added.
15. Benatar (2006).
16. Benatar (2006), p. 39. Emphasis added.

17. This leaves it open that in some cases the individual-affecting reasons are stronger. See McMahan, (2001), Section IV.
18. I first advanced this argument in my doctoral dissertation, "Problems of Population Theory" (Cambridge University, 1986), and published it subsequently in McMahan (1995), pp. 182–200.
19. I have defended this view in McMahan (2002).
20. For an argument against abortion that appeals, though not directly, to the view that goods have impersonal reason-giving value, see Hare (1975), pp. 201–22.
21. Elizabeth Harman embraces a version of the Weak Asymmetry in Harman (2004), pp. 97–8.

References

- Benatar, D. 2006. *Better never to have been: The harm of coming into existence*. Oxford: Clarendon Press.
- Hare, R. M. 1975. Abortion and the golden rule. *Philosophy and Public Affairs* 4: 201–22.
- Harman, E. 2004. Can we harm and benefit in creating? *Philosophical Perspectives* 18: 89–113.
- Heyd, D. 1992. *Genethics: Moral issues in the creation of people*. Berkeley: University of California Press.
- Kamm, F. M. 1992. *Creation and abortion*. New York: Oxford University Press.
- McMahan, J. 1981. Problems of population theory. *Ethics* 92: 96–127.
- McMahan, J. 1988. Death and the value of life. *Ethics* 99: 32–61.
- McMahan, J. 1994. Review of David Heyd, *Genethics: Moral issues in the creation of people*. *The Philosophical Review* 103: 557–9.
- McMahan, J. 1995. La moralita del causare l'esistenza di persone. *Bioetica* 2: 182–200.
- McMahan, J. 2001. Wrongful life: Paradoxes in the morality of causing people to exist (revised). In *Bioethics*, ed. J. Harris. Oxford: Oxford University Press.
- McMahan, J. 2002. *The ethics of killing: Problems at the margins of life*. New York: Oxford University Press.
- Parfit, D. 1984. *Reasons and persons*. Oxford: Oxford University Press.
- Temkin, L. 2003. Egalitarianism defended. *Ethics* 113: 764–82.

Part III
**Must an Act Worse for People
be Worse for a *Particular* Person?**

Chapter 4

Who Cares About Identity?

Nils Holtug

Abstract This paper argues that transworld identity is both morally (or at least “welfare axiologically”) and prudentially insignificant. To clarify, it does not in itself morally or prudentially matter, when comparing welfare distributions across possible worlds, whether or not the same people exist in these worlds. The moral claim is defended on the basis of a (wide) person-affecting moral principle. And the argument is made that what matters from one’s own prudential point of view is not that one comes into existence or continues to exist oneself but that an appropriate continuer/replacement does. Finally, some implications for gene-therapy are considered.

Keywords Personal identity · Non-identity problem · Person-affecting ethics · Gene-therapy.

4.1 Introduction

Identity, in various ways, is often thought to have great importance for normative theory. Some of these ways concern identity in the *qualitative* sense of the term. In this sense, two objects, *a* and *b*, are identical to the extent that they have the same properties. This is the sense in which two exactly similar cars, or identical twins, are identical. This sense of the term should be contrasted with identity in the *numerical* sense. In the numerical sense, two objects, *a* and *b*, are identical if and only if they are in fact one and the same object. Thus, while identical twins may have many properties in common, they are not one and the same person and so they are not numerically identical. My concern in this paper is with numerical identity only. Or, more precisely, I am concerned with identity in the qualitative sense only to the extent it may affect identity in the numerical sense.

N. Holtug (✉)

Philosophy Section, Department of Media, Cognition and Communication, University of Copenhagen, Denmark
e-mail: nhol@hum.ku.dk

Therefore, my discussion differs from most contemporary discussions of the normative significance of identity, since these do not concern—or at least are not plausibly understood as discussions of—identity in the numerical sense. Consider, for instance, discussions in political philosophy over the cultural and national identities of individuals or of groups. It is with respect to these discussions that Kwame Anthony Appiah speaks of “the contemporary use of ‘identity’ to refer to such features of people as their race, ethnicity, nationality, gender, religion, or sexuality”; and adds that “this use of the term reflects the conviction that each person’s identity—in the older sense of who he or she truly is—is deeply inflected by such social features.”¹ While, for instance, a person’s nationality, religion and sexuality may be extremely important for her and her sense of who she is and she may speak of a “loss of identity” or “identity crisis” if they are threatened, this is not because she would literally cease to exist if she were to lose them. And when nationalists and multiculturalists speak of the importance of (recognising) people’s national and cultural identities, the point is simply that such identities are important to people or groups because they secure for them important values such as, for instance, social cohesion, solidarity and options.²

However, as I said, my concern here is with numerical identity (and henceforth, I shall take “identity” to refer to the numerical sense of the term). I shall argue that, in various ways in which it is usually or often thought important, identity is in fact normatively insignificant. Further, I shall make such claims about both the prudential and the moral importance of identity. Here, prudence concerns a person’s self-interest, where to say that someone has a self-interest in something is, roughly, to say that this something will be *valuable* for her. More precisely, the object of a self-interest is a benefit (welfare), which is what ultimately gives prudential value to a person’s life. Thus, as I use the term “self-interest” here, a theory about self-interest is a theory about (prudential) value rather than, for instance, about (practical) rationality.³

Unlike prudence, morality includes an (in some sense) equal concern for everyone. The part of morality I shall be concerned with here is the part that concerns the goodness of distributions of individual welfare. Thus, my concern is with what we might call “welfare axiology.” Furthermore, I shall primarily be concerned with person-affecting moralities, that is, moralities that cash out the concern for welfare in terms of what is good and bad (or better and worse) *for* individuals.⁴ Such moralities should be contrasted with impersonal moralities, according to which welfare is good, period, independently of whether it renders an individual better off than she otherwise would have been. To clarify the difference between impersonal and person-affecting morality, suppose (contrary to what I shall argue in the following) that it cannot benefit or harm a person to come into existence. In that case, according to a person-affecting morality, it cannot add to the value of an outcome to cause a further happy person to exist, because coming into existence cannot be good or better for her. According to an impersonal morality, on the other hand, the welfare this person will experience may contribute to the value of the outcome.

Another distinction I need to introduce is that between identity *over time* and *transworld* identity. This is because I shall make claims about the insignificance of

both kinds of identity. Again, to clarify, identity over time concerns an object at one time being identical to an object at another time, whereas transworld identity concerns an object in one possible world being identical to an object in a different possible world.⁵ With respect to the former, we may ask whether Ludwig at the time of writing *Tractatus Logico-Philosophicus* is identical to Wittgenstein at the time of writing *Philosophische Untersuchungen*. And with respect to the latter, we may ask whether Wittgenstein is identical to the child his parents would have had if they had conceived a child ten minutes prior to when they in fact did.

Derek Parfit has famously and, I believe, plausibly argued that personal identity over time is not what matters in survival.⁶ This is a claim about the prudential significance of identity over time; a claim that Parfit further believes has important moral implications. Here, I shall mostly be concerned with the normative significance of *transworld* identity. And I shall argue that it too, in important respects, is normatively insignificant. In fact, as I pointed out above, I shall argue this with respect to both prudence and morality.

With these distinctions in place, I can present the claims I shall be making in this paper. First, I shall argue that transworld identity is morally insignificant (or, more precisely, that it is insignificant with respect to welfare axiology). My point of departure will be the so-called Non-identity Problem. Suppose a woman can either have a happy or a less happy child. Does it matter for the moral assessment thereof whether these children are identical? Intuitively, irrespective of whether they are identical, it would seem that the woman ought to have the happier child, everything else being equal. However, it does matter whether they are identical, according to the person-affecting view as that view is usually understood. We reach the result that the non-identity between the two children matters, since, on that view, an outcome cannot be worse than another if there is no one *for whom* it is worse. This is because, if the two children are numerically distinct, then it is worse for neither if the woman has the less happy child. However, I shall argue that on a more plausible person-affecting account, it is in fact morally insignificant that the children are numerically different; what matters is that the second child will benefit more from coming into existence than the former child. Thus, my argument will differ from those theorists, such as Parfit, who similarly argue that transworld identity is in important respects morally insignificant, but on the basis of an impersonal account of population ethics.⁷ I agree that transworld identity is in important respects morally insignificant, and I show that a more plausible person-affecting view can take that fact into account.

Second, I shall argue that just as identity over time is not what prudentially matters, transworld identity does not prudentially matter either. Thus, what matters from my prudential point of view is not that *I* came into existence, but rather that someone whose states stand in the right sort of relation to my states did. In order to set up this argument, I first need to consider Parfit's argument for why identity over time is not what prudentially matters.

Finally, I shall illustrate all these claims and their significance by addressing a particular case in bioethics, namely gene therapy on a conceptus, embryo and fetus. This case involves issues of both transworld identity and identity over time and I

shall consider the prudential and moral implications of my claims in relation thereto. Amongst other things, I argue that my claims about the prudential insignificance of transworld identity and identity over time have importance for our person-affecting moral assessments of this case.

4.2 The Moral Insignificance of Transworld Identity

In this section, I am concerned with the moral significance of transworld identity. More precisely, I am concerned with whether the moral comparison of welfare distributions across possible worlds depends on whether or not the same people exist in these worlds. To illustrate this issue, consider the following version of the Non-identity Problem.

Suppose that due to the presence of certain chemicals in her body, a woman will have a child with a severe disease if she decides to get pregnant now and does not take medication. Because of his disease, this child will have a life that is in various ways diminished and unfulfilling, but nevertheless worth living. If, instead, the woman gets pregnant and does take medication, her child—that same child—will live a rich and fulfilling life. Clearly, everything else being equal, it is better if the woman takes the medication.

Now consider a different case. Due to the presence of certain chemicals in her body, a woman will have a child with a severe disease if she gets pregnant now. Again, because of his disease, this child will have a life that is in various ways diminished and unfulfilling, but nevertheless worth living. If, on the other hand, she postpones her pregnancy three months, when the chemicals have left her body, her child will not be afflicted with the disease and will indeed live a rich and fulfilling life. However, since this child will be a result of a different sperm and egg than if the woman does not wait, it will be a numerically different child.⁸ In other words, if she gets pregnant now, she will have one child, *a*, who will have a low (but above zero) welfare, and if she waits, she will have another child, *b*, who will have a much higher welfare.⁹

These cases are in relevant respects similar, except for the fact that in the non-identity case the children in the two outcomes compared are numerically distinct. The question is then whether this matters for our moral assessment of the two cases. Intuitively, it seems equally bad for the two women to have a child with a serious disease when each could have had a healthy child instead. The fact that, in one of the cases, this would mean that she would have a different child does not seem to matter.

Nevertheless, why might one think that it does? Many people are drawn to the idea that morality, or at least the part of morality that is concerned with welfare, should take a person-affecting form. That is, morality (or this part) should be explained in terms of what is good and bad, or better and worse, for people.¹⁰ In one version, what the person-affecting idea amounts to is this:

The Narrow Person-affecting Principle: An outcome, O_1 , cannot be better (worse) than another outcome, O_2 , if there is no one for whom O_1 is better (worse) than O_2 .¹¹

The Narrow Person-affecting Principle has intuitive merit. How can an outcome be better, if there is no one *for whom* it is better? It also has important implications for the Non-identity Problem. It implies that the outcome in which *a* comes into existence cannot be worse than the outcome in which *b* comes into existence, unless there is someone for whom the former outcome is worse. And the former outcome is not worse for *a*, because *a* does, after all, have a life worth living and would not otherwise have existed. Nor would it seem to be worse for *b*, as *b* does not exist in this outcome and so cannot have any properties, including that of being worse off. Of course, it may be worse for the woman herself, or for other third parties, but let's leave that out. We may assume that, for some reason, the choice between *a* and *b* does not affect the woman's or anyone else's welfare. It then follows that the outcome in which *a* comes into existence cannot be worse than the outcome in which *b* comes into existence. In order for it to be worse, there would have to be someone for whom it is worse.

Thus, the Narrow Person-affecting Principle implies the significance of trans-world identity for our moral assessments. It matters that, in the Non-identity Problem, *a* and *b* are not identical. After all, in the first case of the pregnant woman presented above, this principle is quite compatible with the claim that the outcome in which the child is afflicted with a serious disease is worse than the outcome in which he is not. Since he is worse off in the outcome in which he has the disease, this outcome is worse for *him*. Thus, it is precisely because *a* and *b* are *not* identical in the non-identity case, but rather inhabit different possible worlds, that the Narrow Person-affecting Principle rules out the judgement that it is worse if *a* is caused to exist.

Nevertheless, as I pointed out above, it seems intuitively plausible that it is worse to cause *a* to exist. In fact, the Narrow Person-affecting Principle seems too restrictive in that it rules out our taking into account the welfare of *b* when assessing whether the outcome in which *a* comes into existence is worse. More generally, this principle seems implausible because when assessing whether an outcome is better or worse than another, it does not take into account the welfare of the contingent individuals who exist only in the other outcome.

To further elucidate this troubling feature, note that it is also exemplified in what I have elsewhere called the "Problem of Suffering."¹² Suppose a woman is considering having a child. Due, again, to the presence of certain chemicals in her body, this child will have a miserable life, clearly worth not living. Now suppose that, taking this into consideration, she decides not to have the child. According to the Narrow Person-affecting Principle, this outcome cannot be better, because there is no one *for whom* it is better (again, we are ignoring effects on the woman herself or on third parties). But that is a highly counterintuitive result.

For these and other reasons, some philosophers abandon person-affecting morality and opt for an impersonal account instead. For example, in the context of the Non-identity Problem, we may claim that the outcome in which *a* exists is worse than the outcome in which *b* exists because the latter outcome contains more total welfare, quite irrespective of whether these outcomes are better or worse for anyone. We will then not assess these outcomes on the basis of whether they are better or

worse for *a* and *b* but on the basis of whether they are *impersonally* better or worse or, if we like, better or worse for the “world.”

Nevertheless, it may be suggested that this is too swift. These impersonal claims may seem, in a sense, too abstract. By simply saying that it is worse if *a* exists because it is better *for the world* if *b* comes into existence instead, we lose sight of the people we wish to benefit by our acts.¹³ Likewise, simply saying that the outcome in which *b* exists has a higher total seems too abstract, unless we can tie this higher total to a particular person *for whom* this outcome would be better.

Consider also the Problem of Suffering. On an impersonal account, we may of course say that it is better if the miserable child is not caused to exist because this outcome has more total welfare. However, again this is rather abstract unless the betterness of this outcome can be explained by pointing out that if the child is caused to exist, this will be worse for him. After all, it is for *his* sake that we would want the outcome in which he does not exist to come about. But then, of course, we have abandoned an impersonal approach.

Let me therefore present an alternative person-affecting principle. Unlike the Narrow Person-affecting Principle, this principle does not attach significance to transworld identity. And unlike impersonalism, it assesses outcomes in terms of their being better and worse for people. Consider:

The Wide Person-affecting Principle: An outcome, O_1 , cannot be better (worse) than another outcome, O_2 , if there is no one for whom, were O_1 to obtain, O_1 would be better (worse) than O_2 and no one for whom, were O_2 to obtain, O_2 would be worse (better) than O_1 .

This principle allows us to reach the intuitively plausible conclusion that, in the Non-identity Problem, it is worse if the woman has *a* than if she has *b*. The outcome in which she has *a* may be worse because, had the outcome in which she has *b* obtained, this outcome would be better for *b*. Note that by claiming that if this latter outcome obtains, this is better for *b*, we are not ascribing properties to a person in a world in which he does not exist. Nevertheless, I am of course presupposing that it can be better for an individual to exist than never to exist, and I shall briefly present an argument to this effect. But first, consider also the implications of the Wide Person-affecting Principle for the Problem of Suffering. According to this principle, it may indeed be better not to have the miserable child, because had the outcome in which the woman has this child obtained, this would have been worse for the child. This presupposes that it can be worse for a person to exist than never to exist.

Note that what I need to argue is that existence can be better (worse) for a person than never existing, not (merely) that existence can be good (bad). Parfit argues that while the latter non-relational claim is defensible, the former relational claim is not. I shall now argue that this is a mistake.¹⁴

Consider, then, the claim that it can be better (worse) for an individual to come into existence than never to exist (or, for short, that existence can be better (worse) for an individual than nonexistence).¹⁵ It is not the mere state *that an individual exists*, but rather the state *that an individual exists and has such and such a life* that I claim can be better (or worse) for her than nonexistence. Furthermore, the value I speak of here is intrinsic value (that is, final or non-instrumental value). I need to

be a bit more specific. The intrinsic value of an existence need not be a value that attaches to that existence or life in its entirety; it may, for instance, attach only to specific parts (say, to the pleasurable mental states contained in it). Thus, when I say that existence is intrinsically better (worse) for an individual than nonexistence, what this means is that existence either has, or contains, more intrinsic value for her than nonexistence.

The details of the argument for this claim will depend on the particular theory of welfare assumed. The simplest version relies on a preference theory, according to which one state of affairs is better (worse) for an individual than another if he has an (appropriate) intrinsic preference for the former state over the latter. For instance, if an individual intrinsically prefers skiing at some time t to reading at t , the state of affairs that he skies at t is better for him than the state of affairs that he reads at t . Of course, the preference may have to satisfy further conditions to give rise to this sort of value, which is why I have inserted a clause about “appropriateness.”¹⁶ Assuming, then, that an individual prefers his existence (with all that it includes in terms of joys, sorrows, friendships, accomplishments etc.) to never existing, his existence is better for him. Likewise, if he prefers nonexistence, existence is worse.

Of course, I cannot simply assume such a preference theory. Therefore, let me briefly point out how the argument would proceed with alternative theories of welfare. Perhaps we hold a preference theory according to which it would be more natural to compare existence and nonexistence on the basis of the preference satisfactions contained in each of these states. Or perhaps we want to compare them on the basis of hedonism, or an objective list theory. Let us suppose that a person’s life contains a net surplus of positive value (preference-satisfactions, pleasure or items on an objective list). His nonexistence, on the other hand, contains no such values. Therefore, existence seems to be better. After all, it would seem to be better for a person to have a surplus of positive value than to have no value accrue to him. (Such an absence of value may of course be realised in either of two ways—nonexistence or a life devoid of value.)

Various semantic, logical, metaphysical and value-theoretical objections have been made to the view that existence can be better (worse) than nonexistence. Two seem to have gained most support. According to the first, it does not make sense to say about individuals who may but do not yet exist that existence can be better (worse) for them. This is because they cannot be uniquely referred to (or rigidly designated).¹⁷

However, I believe that, at least in some cases, we do in fact have access to such rigid designators. We can rigidly designate possible future individuals by specifying appropriate combinations of gametes. That is, we can refer to an individual as, say, the person who will result if this sperm fertilises this egg and develops into an individual with a certain level of psychological complexity.

Of course, on some theories of personal identity, this description will not quite suffice. It may be claimed that different psychological features could be instantiated in a particular organism and that a numerically distinct person would emerge if a sufficiently different psychology was instantiated. But first, it is not clear that different persons could emerge in this way (a point I briefly return to in Section 4.4). Second,

even if they could, this would not make it impossible to provide rigid designators for possible future individuals. We would then just have to make the reference more specific. We could refer to an individual as the individual who will exist if this sperm fertilises this egg and develops into (say) a person with *such and such a distinctive psychology*.

The other objection I want to consider here is that existence cannot be better (worse) for an individual than nonexistence because this would imply that nonexistence is worse (better) for her. Nonexistence cannot be worse because an individual cannot have this or any other property if she does not exist.¹⁸ This objection relies on the following two premises:

1. x is better (worse) for S than y , if and only if y is worse (better) for S than x .
2. An individual cannot have any properties in world W if it does not exist in W .

However, (1) and (2) are in fact quite compatible with the claim that existence can be better (worse) for S than nonexistence. Admittedly, if existence is better (worse) for S , (1) implies:

3. Nonexistence is worse (better) for S than existence.

But (3) may be true in some worlds, namely worlds in which S exists. Since S exists in these worlds she can have properties there, in accordance with (2). Thus, assuming that you have a life that is worth living, it is true in the actual world that nonexistence is worse for you than existence. It is true in virtue of the obtaining of the three-place relation “the state of affairs that *you do not exist* is worse for you than the state of affairs that *you exist*.” Obviously, the last two relata obtain in the actual world and the first relatum (the state that *you do not exist*) does not obtain there. Nevertheless, a state does not have to obtain in order to be an object in a betterness relation. If it did, the following betterness relation could not obtain: the state of affairs that *the allies win the war* is better than the state of affairs that *the Nazis win the war*.

A more plausible requirement, then, is that in order for a betterness relation to obtain, its relata must *exist*. And while the state of affairs that *you do not exist* does not obtain in the actual world, it can be sensibly claimed that it *exists* there (as an abstract entity).¹⁹ And if all three relata exist, we can claim that the triadic relation—*nonexistence* is worse for you than *existence*—obtains.²⁰

To challenge (3), we would have to hold:

4. If x is better (worse) for S than y , then x is better (worse) for S than y even if x obtains.

If nonexistence is worse for you than existence, (4) implies that nonexistence is worse for you *even* if you do not exist—something ruled out by (2). However, nothing forces us to accept (4). In the actual world, you exist, so the relation that serves as the truth-maker for (3) obtains. But if we move to a possible world in which you do not exist, this relation no longer obtains. Thus, we have a perfectly natural explanation of why (4) does not hold in such cases. The metaphysical basis for (3) is not preserved.

This metaphysical account also enables me to deal with another possible objection to the claim that existence can be better (worse) for a person than nonexistence. Above, I claimed that on some theories of welfare, the argument for this claim involves comparing welfare units (preference-satisfactions, pleasure or items on an objective list) in existence and nonexistence, and the thought that it is better for an individual to have a surplus of positive value than to have no value accrue to him. To this it may be objected that in order for nonexistence to be evaluatively comparable to existence, it must be possible to ascribe a value (say, zero value) to nonexistence. If not, the state of nonexistence cannot figure on a value scale. However, we may indeed ascribe zero value to nonexistence. Thus, in the actual world in which you exist (or, for that matter, another possible world in which you also exist), the state of your nonexistence has zero value for you.

Since, then, it can be better (worse) for a person to exist than never to exist, the Wide Person-affecting Principle is compatible with the claim that, with respect to the Non-identity Problem, it is worse if the woman has *a* than if she has the happier *b*. After all, the former outcome can be worse if either this outcome is worse for someone, if it obtains, *or* if the latter outcome is better for someone, if *this* outcome obtains. And, indeed, if the latter outcome obtains, this is better for *b*. Furthermore, with respect to the Problem of Suffering, the Wide Person-affecting Principle allows that it is better if the woman refrains from having the miserable child, because if the outcome in which she has him obtains, this will be worse for him.

So we do not need to give up person-affecting morality in order to give intuitively plausible answers regarding these problems. Furthermore, the answers we then give imply the moral (or at least welfare axiological) insignificance of transworld identity. In the Non-identity Problem, it does not matter whether *a* and *b* are identical. The outcome in which *a* exists is worse, even if the alternative outcome does not benefit him or anyone else who exists in the former outcome, but only the numerically distinct *b*. In other words, even if we are attracted to the idea that morality (or the part thereof that concerns welfare) should be explained in terms of what is good and bad, or better and worse, for individuals, we should consider transworld identity morally insignificant. What matters are the welfare levels of the people who may exist.

There are of course other person-affecting principles, besides those considered here, that have implications for the moral (in)significance of transworld identity—principles that trade on various temporal and modal distinctions between people. What these principles have in common is that, unlike the Wide Person-affecting Principle, they restrict the class of possible future people to whom our moral concern extends. However, I cannot discuss these principles here. I have elsewhere argued that none is as plausible as the Wide Person-affecting Principle.²¹

In this section, I have argued that transworld identity is morally insignificant (or at least insignificant as far as our welfare axiological principles are concerned). This, however, does not imply the *prudential* insignificance of transworld identity. To see this, consider that if *b* comes into existence, it would seem to matter a great deal from his perspective that it is he, rather than *a*, who came into existence. After all, I

have argued that existence is better for him than nonexistence. However, I now want to go further and argue that, in fact, transworld identity does not matter prudentially either.

4.3 The Prudential Insignificance of Transworld Identity

To set up my argument for the prudential insignificance of transworld identity, I first need to consider Parfit's argument for the prudential insignificance of identity over time. To say that identity over time has prudential significance, or is what matters in survival, is to say that when considering a person's survival, what makes the future him a particularly natural object of his present concern is the fact that he is *identical* to the future him. This is what makes it particularly important, from his point of view, that benefits fall in his own future rather than in someone else's. Note that, according to this claim, identity matters intrinsically (non-instrumentally) and not because identity to a future beneficiary (usually) carries with it some further relation that is in fact what prudentially matters.

Note also that what matters in the prudential sense should not be confused with what matters in the "desirability" sense.²² What matters in the desirability sense is whatever benefits or welfare ultimately consist in. Obviously, it is desirable to have such items in one's future. However, what matters in the prudential sense is the relation that one can stand in to a future person that makes it especially important, from the perspective of one's own self-interest (or an appropriate extension thereof), that such items befall *that* future person.

This last remark also brings out the relation between what prudentially matters and self-interest. If identity is what matters, what is in a person's present self-interest, with respect to future benefits, is that they accrue to a future person to whom she is identical. However, if it is not identity but a different relation that matters, what is in a person's self-interest (in an appropriately extended sense), with respect to future benefits, is that they accrue to persons to whom she stands in this different relation.²³

Now to Parfit's argument. Consider a case of brain transplantation. *a* is one of three identical triplets. His brain is removed from his head, and a cerebral hemisphere of his is inserted into each of his brothers' skulls, from which their cerebral hemispheres have been removed. Since, as it is believed to be the case in some people, *a*'s two hemispheres are equally good realizers of psychological continuity, both the resulting people, *b* and *c*, are psychologically continuous with *a*, and, we may assume, to the same degree. Furthermore, idealizing somewhat, let us assume that both of *a*'s hemispheres realize all of *a*'s psychology. So *b* and *c* both seem to remember *a*'s life, and have emotional attachments, projects and character traits that are exactly similar to his. This case is represented below as World 1 (Fig. 4.1).

As Parfit argues, we can use this case to show why identity is not what matters in survival.²⁴ However, in order to show this, we need to contrast World 1 with World 2 (Fig. 4.1). In World 2, *a*'s brain is divided by destroying one hemisphere, and the other hemisphere is then inserted into one brother's skull.

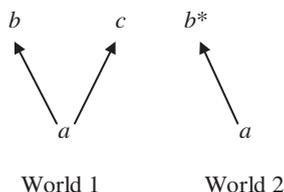


Fig. 4.1

Plausibly, if the resulting person had both of *a*'s cerebral hemispheres, then he would *be a*. He would then have *a*'s memories, plans, concerns etc. Furthermore, people seem to have survived having one cerebral hemisphere destroyed—for instance, by a stroke—although they have only one functioning hemisphere left. This judgment is supported by the fact that there may be massive psychological continuity between the pre-stroke and the post-stroke person. Combining these beliefs, it is plausible to claim that, even when only one of *a*'s hemispheres is transplanted, the resulting person is *a*. So in World 2, $a=b^*$.

What, then, about the case in which both *a*'s brothers receive a brain hemisphere of *a*'s, that is, World 1? It seems that each of *b* and *c* are as good candidates for being *a* as *b** is in World 2. But since *b* and *c* are not the same person, they cannot both be *a*. Furthermore, it would be arbitrary to say that *a* is one rather than the other. Therefore, *a* is neither. He has simply ceased to exist.

How do we get from the claims about identity in World 1 and 2 to the conclusion that identity is not what matters? Consider World 2 again. Even though what *a*'s survival amounts to is having one of his hemispheres destroyed and the other transplanted, presumably his relation to the future him, that is, *b**, contains what matters. This is because *a* is related to *b** by *b** having one of *a*'s hemispheres which continuously realizes his psychology which, of course, is also why *b** is *a*.

In World 1, *a* stands in exactly the same physical and psychological relation to each of *b* and *c* as he does to *b** in World 2. And since *a*'s relation to *b** contains what matters, so will his relation to both *b* and *c*. Therefore, *a*'s relation to *b* and *c* can contain what matters without *a* being identical to (either of) them, and so identity cannot be what matters.

If identity is not what prudentially matters, what does thus matter? For present purposes, the precise answer to this question need not really concern us. It may be continuity and connectedness of distinctive psychology (particular memories, preferences, character traits etc.),²⁵ it may be continuity of core psychology (psychological capabilities such as the capacity to remember, form intentions, reason etc.), and it may be the continued physical realization of one or both of these psychological relations.²⁶

Now, the argument for the prudential insignificance of identity over time also has immediate implications for the prudential (in)significance of transworld identity. While *c* in World 1 and *b** in World 2 are numerically distinct, *a* has equal prudential concern for them. The lack of transworld identity between *c* and *b** has no prudential significance from *a*'s point of view.

However, there is also a further, at least as spectacular respect in which transworld identity lacks prudential significance that may be discerned from consideration of the argument for the insignificance of identity over time. Just as it does not prudentially matter for a person whether she continues to exist (what matters is that an appropriate future continuer does exist), it does not prudentially matter for a person whether she comes into existence in the first place (because an appropriate replacement may be just as good). Thus, I am prudentially indifferent between a world in which I come into existence and a world in which an appropriate (but non-identical) replacement comes into existence. This second respect in which transworld identity lacks prudential significance, however, requires a bit more argument.

Consider, then, some further implications of the argument for the prudential insignificance of identity over time. Since $a \neq b$ and $a = b^*$, and since identity is transitive and necessary, it follows that $b \neq b^*$. That is, the person who comes into existence in the b -body in World 1 is numerically distinct from the person who comes into existence in the b -body in World 2 (although they have identical bodies, brains and psychologies). This is also why the argument does not violate the so-called “only x and y principle,” according to which whether y is identical to x can depend only on facts about x and y and the relations between them. This principle may seem to be violated since it may seem that all that prevents b from being identical to a is the existence of someone else, namely c (after all, in World 2, there is no competition, and so a is identical to the recipient of his (left) brain hemisphere). However, since whether or not c exists, $a \neq b$ (a is identical to b in neither World 1 nor World 2), the only x and y principle is not violated.²⁷

While the only x and y principle is not violated, it does follow that whether a continues to exist depends upon whether someone else, c , comes into existence. After all, in World 2, where c does not come into existence, a survives the transplant, whereas he does not survive in World 1. Some will hold that this is implausible for purely metaphysical reasons, but I shall not go into this here. The point is that whether or not a survives, the relation that prudentially matters for him is preserved, whether it be to b and c or to b^* . What matters to a is simply his intrinsic relation to b , b^* and c .

More importantly, for present purposes, it also follows from the argument for the prudential insignificance of identity over time that whether b comes into existence depends upon whether c comes into existence. b only comes into existence in World 1 because here c does so as well. Again, this may be considered implausible for purely metaphysical reasons, since b 's existence depends upon c 's existence even though c 's existence does not causally affect b . But, again, I shall not address the metaphysical worry here. Rather, I want to argue that even though the existence of b depends upon the existence of c , it does not prudentially matter from b 's perspective whether or not c , and so b himself, comes into existence. This is a claim about the prudential insignificance of *transworld* identity. From b 's prudential point of view, it is no worse if b^* exists than if b does.

When explaining why he finds the claim that the existence of b depends on the existence of c absurd, Harold Noonan suggests that we could say to b : “You should consider yourself fortunate that the other fellow’s brain transplant [that is, c 's brain

transplant] went so well—if it hadn't you would never have existed.”²⁸ But why claim that *b* is *fortunate*? Just as it does not matter from *a*'s point of view whether his future continuers are identical to him, as long as the relevant relations hold between him and them, it may be claimed that it does not matter from *b*'s point of view whether the person who comes into existence in the *b*-body is *b*. After all, if all the nonexistence of *b* would really amount to would be the nonexistence of *c*, then why should nonexistence matter from *b*'s perspective? So *b* might answer Noonan by saying: “Although I would not have existed if the doctors hadn't done such a good job on the other fellow, why should I care about *that*?”

When I say that “all the nonexistence of *b* would really amount to would be the nonexistence of *c*,” I do of course not mean to imply that the state of affairs that *b* does not exist and the state of affairs that *c* does not exist are one and the same state of affairs.²⁹ Obviously, they are distinct. Rather, all that *prevents* the existence of *b* in World 2 is the fact that *c* does not exist there.

Noonan holds World 1 to be better than World 2 from *b*'s prudential point of view. Or, to put it differently, he holds that *b* has a self-interest in the obtaining of World 1 rather than World 2. This is because *b* exists in the former world but not in the latter. But the existence of *b* depends in a very intimate way on the existence of *c*. This is clear when we consider that World 2 contains everything that is needed for the existence of *b* *except* the existence of *c*. Furthermore, while *c* enables the existence of *b* in World 1, this is not in virtue of any causal relations between them. Rather, it is the *mere* existence of *c* that brings *b* into existence here.

In the light of these connections between the existence of *b* and the existence of *c*, can we maintain that, since *b* exists in World 1 but not in World 2, World 1 is better from *b*'s prudential point of view? Such a claim would not be incoherent. Nevertheless, it seems to me very hard to believe. The difference between World 1 and World 2 in virtue of which *b* exists in the former but not in the latter is simply that (the causally isolated) *c* exists in the former but not in the latter. And why should the existence of *c* make a difference from *b*'s prudential point of view? Of course, it may make a difference for how *desirable* the life of the *b*-body person is, but that is a separate issue. Here, we are concerned with the prudential sense of what matters, not the desirability sense.

Some readers may be worried about these claims regarding the value of World 1 and World 2 from *b*'s prudential point of view. After all, *b* does not exist in World 2 and so does not have a prudential point of view there. So how can this world be prudentially better or worse for him than World 1?

Nevertheless, I have already developed, in Section 4.2, the metaphysical and value-theoretical apparatus necessary to substantiate my claims here. In World 1, where *b* exists, the state of affairs that *b* exists is equal in value for *b* to the state of affairs that *b** exists. And here, he has an equal self-interest in the obtaining of these two states of affairs. In other words, the fact that transworld identity does not obtain between *b* and *b** is prudentially insignificant for *b* (and, by parity of reasoning, for *b**).

What, then, *does* prudentially matter for a person with respect to existence, if it is not *her own* existence? This is another question I do not need to give a precise

answer to here. Perhaps it is the existence of a person some of whose psychological and physical states are identical to the states she herself would have had, had she existed. Thus, it is plausible to claim that b 's and b^* 's such states are identical, just after the transplant, and this may be why b^* is a suitable replacement from b 's prudential perspective. Or, perhaps it is some other (weaker) relation than identity between states that matters.

4.4 Genes and Identity

In this section, I want to consider the implications of my claims about the insignificance of identity for a particular issue to which the genetics revolution gives rise. It concerns genetic interventions on a conceptus, embryo or fetus that may be claimed to have such a significant impact that a different person will emerge than if no such intervention had taken place.

Suppose gene therapy is performed on an embryo prior to implantation. Therapy at such an early stage will have the advantage that it can be performed *in vitro* and it may cause genetic changes that are suitably integrated into the genome of the recipient such that they are transferable to his or her children, because they will occur not just in somatic cells but also in the germ-line. Obviously, such a form of gene therapy raises many moral questions such as whether the long-term effects of germ-line gene therapy can be predicted with sufficient accuracy, whether there are other and safer techniques available and whether it is a slippery slope to dubious practices.³⁰ However, I shall ignore these questions here and focus instead on the issue of whether this kind of therapy may have an impact on the numerical identity of the resulting person and if so, what this means for the moral case for performing it.

Noam J. Zohar argues that “regarding embryos in particular, persistence of genotype must generally be deemed a necessary condition for maintaining personal identity. Therefore, many proposals for [prenatal genetic intervention (PGI)] should be excluded from the notion of therapeutic intervention and thus denied the special moral status of requests for therapy.”³¹ This passage is a bit confusing since Zohar speaks of “*maintaining* personal identity” although he does not believe that the embryo is a person.³² To clarify the issue of personal identity here, we may distinguish between “person” used as a substance and as a phase sortal. As a substance sortal, it designates an essential property of whoever possesses it. Thus, a non-person cannot *become* a person. Therefore, an embryo cannot *maintain* its personal identity, because it is not (and nor can it become) a person. We can then understand the identity issue in the following way: if gene therapy is performed, this may mean that a numerically different person will come into existence than if gene therapy had not been performed. Thus, the question raised here is one of transworld identity rather than identity over time.

If, on the other hand, “person” designates a phase sortal, it refers to a property an individual may have at some time even if she does not have it always. Therefore, while the embryo cannot *maintain* its personal identity, it can *become* a person. And so the embryo may in the future be a person, provided that gene therapy does not

affect the identity of the embryo. Thus, the identity issue is here one of personal identity over time.

The reason why Zohar holds that gene “therapy” should be denied the special moral status of a request for therapy is that when identity is affected, there cannot be person-affecting reasons for performing it; “‘Therapy’ necessarily means helping a person, while PGI—though effecting improvements from an impersonal perspective—will frequently not consist in directly helping any person.”³³ However, if identity was not affected, there could be person-affecting reasons to perform gene therapy.

To keep things simple, I shall assume that gene therapy would be performed on a conceptus. Following Robert Elliot, I shall call such therapy on a conceptus “CT.”³⁴ If an instance of CT has the consequence that a person develops from the conceptus that would not have existed in the absence of CT, I shall call it “non-identity” CT, and if CT does not affect identity in this way, I shall call it “identity-preserving” CT. The person who comes into existence as the result of non-identity CT I shall call the “new person,” since this person owes her existence to the fact that CT has been performed, and the person that comes into existence if CT is identity-preserving I shall call the “original person.” Note that the original person may exist whether or not identity-preserving CT is performed.

The identity issue here of course depends on the particular theory of personal identity we assume. So let us in turn consider the implications of the views that we are essentially animals and that we are essentially persons.³⁵ If we are essentially persons, then we are essentially psychological beings, or beings with minds. If we are essentially animals (of the species *homo sapiens*), on the other hand, then we may exist even at times where we do not possess consciousness. Further, if we are essentially animals, we may either come into existence at conception or at some later point (such as at the occurrence of the so-called “primitive streak”). Suppose we come into existence at conception. CT is then performed on the sort of being that we essentially are. And so non-identity CT would in fact amount to killing the sort of individual we are.

However, it is doubtful if CT would really amount to the killing of a conceptus. As Elliot argues: where C is a particular conceptus, and “only a small proportion of C’s functional genetic material is replaced, as would be the case in gene therapy, C’s identity is not disrupted.”³⁶ Just like a specific table does not cease to be *this* very table when a small fraction of its parts are replaced, C does not cease to exist when a small fraction of *its* (genetic) parts are replaced. And if C continues to exist throughout the procedure, it is plausible that the therapy will be of the identity-preserving kind. After all, it is plausible that it will then be the very individual who in its early stages is referred to by “C” that later develops the property of being a person. No shift of persons in sight.

Of course, there could be interventions that would involve the killing of C, such as if the entire cell nucleus were replaced with a nucleus (and so genome) from a different conceptus. But this would be cloning, not gene therapy. And for technical reasons alone, presumably gene therapy would involve inserting or altering only one or at most a few genes.

Consider now the implications of the view that we are essentially persons (where “person” is therefore a substance sortal). Even if CT does not involve the killing of C, might it nevertheless be of the non-identity kind? This does seem conceptually possible. Thus, it seems conceptually possible that different persons might “break out” in the same organism, depending on what sort of psychology develops from it.

On the other hand, we also seem capable of imagining that our lives might have been psychologically very different from what they actually are, which suggests that we can imagine rather different psychological histories realized by one and the same person. Thus, suppose Tony Blair had been adopted at birth and taken to Brazil to work on a coffee plantation. Presumably, he would then be psychologically very different from what he in fact is. Now, this seems to be something that could have happened to *Tony Blair*. It seems, then, that where there is no question of there being more than one organism, significant psychological variation can be accommodated by one and the same person.³⁷

Nevertheless, some psychological theories of personal identity will clearly imply that at least some instances of CT may be of the non-identity kind. Consider, for instance, the view that we are essentially self-conscious beings.³⁸ Since some diseases may cause their bearers never to become self-conscious, a therapy that cured such a disease would imply that a person (in the thick, self-conscious sense) would exist who would not otherwise have existed. Thus, at least such instances of CT would affect identities, assuming the relevant kind of psychological theory.

However, while I believe that theories of personal identity that are so demanding with respect to the psychological complexity required for personhood are implausible, I shall not try to settle the identity issue here. For the sake of argument, I shall simply assume that different persons could originally break out in the same organism and so that non-identity CT is possible. What should we make of Zohar’s argument that there could not be a person-affecting reason to perform such therapy? Of course, assuming that the original person would have a life worth living, non-identity CT would not benefit her.

Nevertheless, non-identity CT would bring the new person into existence instead of the original person and, assuming that she would have a life worth living, this would benefit *her*. In fact, the original and the new person are in a relevantly similar position. Neither exists, either may be caused to exist, and either may benefit from so being. On the further assumption that, because of the therapy, the new person would benefit more from her existence than the original person would, there is a straightforward person-affecting reason to perform non-identity CT. This, of course, presupposes that it can benefit a person to come into existence, and I have defended this view in Section 4.2.

On the further assumption (for which I have also argued) that we should hold a *wide* person-affecting principle, we may even claim that it is *worse* to bring about the original person than to bring about the new person. It is worse to bring about the original person because, or assuming that, the new person would benefit more from coming into existence.

Let me now turn from the moral to the prudential insignificance of identity, which may in turn have further implications for the ethics of gene therapy. Consider first

the prudential insignificance of identity over time. And suppose that non-identity gene therapy is performed on a conscious fetus. Since this fetus is conscious, we may assume that it has present self-interests regarding the future.³⁹ Suppose also that the therapy “cures” a disease that significantly lowers its bearer’s welfare and greatly shortens her life. Suppose finally that while the therapy causes future psychological states to be very different from what they otherwise would have been, psychology is continuously realized throughout the procedure. On these assumptions, it is plausible that the fetus has a present self-interest in having the new person rather than the original person come into existence. After all, everything else being equal, much more welfare will be realized in the life of the new person. And there is no reason to suppose that the relation that prudentially matters between the fetus and the new person is less strong than that between the fetus and the original person. Thus, in terms of both core and distinctive psychology, and the continuous physical realization thereof, there is no reason to suppose that the original person is more strongly related than the new person to the rather crude psychology of the fetus.

Note, furthermore, that the pre-therapy fetus may have a present self-interest in the existence of the new person even if the fetus is itself identical to the original person and will cease to exist *due* to the therapy. To illustrate this point, suppose again that we are essentially self-conscious beings. Suppose also that unless non-identity therapy is performed, the child that develops from the fetus will never acquire self-consciousness. And suppose, plausibly, that the conscious pre-therapy fetus is identical to the original person (who is of course then not a person in the “thick” sense that requires self-consciousness). Suppose furthermore that the original person ceases to exist when the self-conscious person comes into existence, say, when the child reaches the age of two. Suppose finally that the life of the new person is much better than the life of the original person. While the therapy will then cause the individual that the fetus and (so) the original person is to cease to exist, the fetus and so the original person may nevertheless have a self-interest in having the new person come about. After all, the fetus may very well stand in the relation that prudentially matters to the new person, who will have the better life. Thus, there can even be a person-affecting reason to perform non-identity therapy that invokes only the interests of the fetus and of the original person.

Let me now turn to the prudential insignificance of transworld identity and, more specifically, to the claim that what prudentially matters to a person with respect to existence is not that she herself comes into existence. Suppose again that non-identity therapy is performed on a conceptus, C. At this stage, the original person does not yet exist and let us further suppose that, because of the therapy, nor will he ever exist. Suppose also that while the original person and the new person do not have the exact same psychological states, there is nevertheless a small degree of overlap. Suppose, for instance, that some of the new person’s first psychological states are exactly similar to some of the original person’s first psychological states. It would then seem plausible to claim that, since these same states break out in the numerically same organism (in the very same tissue, we may imagine), these states are in fact not only qualitatively but also numerically identical. And if the

existence of such states is what prudentially matters, then to the extent that these states overlap, some of what prudentially matters for the original person is realized by the new person. So from the original person's point of view, it may not be as bad never to exist as one might have suspected (more precisely, in the world in which the original person comes into existence, the state that the new person exists may not be worse for him than the state that he himself exists). Perhaps it may even be a good thing, given that the life of the new person is much better. And there can then be a person-affecting reason to perform non-identity therapy on a conceptus that invokes only the interests of the original person.

Therefore, even if we assume the metaphysical possibility of non-identity CT, there can be a person-affecting case for such therapy. This case will rest on the self-interests of the new person, and conceivably even on the interests of the pre-therapy fetus and the original person. Therefore, the identity issue turns out to be insignificant with respect to the (person-affecting) moral case for prenatal gene therapy.

To summarize, what I have argued is that (1) transworld identity is morally insignificant, (2) transworld identity (just like identity over time) is prudentially insignificant and (3) amongst other things, this implies that there can be person-affecting reasons to perform non-identity CT that rest on the self-interests of the new person, and even the pre-therapy fetus and the original person.⁴⁰

Notes

1. Appiah (2005), p. 65.
2. For, respectively, nationalist and multiculturalist arguments along these lines, see e.g. Miller (1995) and Kymlicka (1995).
3. Therefore, such theories should not be confused with what Parfit calls the Self-interest Theory, which is a theory about rationality. According to this theory, each person has one ultimate rational aim, namely that his life go, for him, as well as possible. Parfit (1984), p. 4.
4. Glover (1977), p. 66, Holtug (2004), Parfit (1984), p. 394.
5. Or, if you do not believe in transworld identity, in an object in one world being a counterpart of an object in a different world.
6. Parfit (1984), Part 3.
7. Parfit (1984), p. 378.
8. This follows from Kripke's plausible claim that a person would not have existed, had he not originated from the sperm and egg from which he in fact originates (Kripke 1980, p. 113).
9. Moral philosophers have discussed this problem extensively since Parfit presented his pioneering work on the subject. See, for example, Parfit (1976), (1982), (1984), Chapter 16, Heyd (1992), Chapter 4, Hanser (1990), Kavka (1982), Roberts (1998), Chapter 3, Temkin (1993), pp. 255–256, and Woodward (1986).
10. For an account of the intuitive attraction of the idea, see Holtug (2004), pp. 131–132.
11. The distinction between *narrow* and *wide* person-affecting principles derives from Parfit (1984), pp. 393–401.
12. Holtug (1998), pp. 170–171, Holtug (2004), p. 136.
13. Jonathan Glover makes a similar point in Glover (1992), p. 142.
14. Furthermore, I believe that the non-relational claim does not suffice to establish that a person may benefit from coming into existence. Consider Parfit's claim that while it cannot be *better* (worse) to come into existence, it can be *good* (bad). Parfit (1984), pp. 489–490. It may be

objected that Parfit does not succeed in defending the view that a person can benefit from coming into existence on the basis of a non-comparative judgement. This is because claims about goodness (badness) are themselves based on comparative judgements. Consider the following definition:

x is good (bad) for S , if and only if x is better (worse) for S than non- x .

This implies that in order for a state of affairs to be good for an individual, it must be better for her that it obtains than that it does not. So in order for existence to be good for an individual, it must be better for her than nonexistence, which is the comparative judgement that Parfit wants to avoid.

However, Parfit need not accept this definition. He can define “goodness” differently. More to the point, he can claim that a life that does not include any of what makes life worth living (and not living), for instance, pleasure or preference-satisfaction, is a life that has zero value. He can then define a “good life” as a life that is better than a life that has zero value, and a “bad life” as one that is worse. And he can then claim that it is good (bad) for a person to come into existence if and only if his life is a good (bad) one. This claim is based on a comparison, but with a life that has zero value, not nonexistence.

However, I do not believe that this can serve as an argument for the claim that it can benefit a person to come into existence. To defend the latter view, a comparison of existence and nonexistence is needed. After all, the alternative to coming into existence is not a *life* with zero value; it is nonexistence.

To illustrate this point, suppose I am considering whether I will benefit from going to sleep now. We might then define a good sleep as one in which I am better off than in, say, a state of unconsciousness (for example, I may be better off because I have pleasant dreams). However, suppose that the alternative to going to sleep now is not unconsciousness, but reading an excellent novel. Suppose also that reading it will give me more welfare than sleeping. While there is a sense in which it will benefit me to go to sleep (it is better than “nothing”), this is not the sense we will usually be interested in. Rather, we want to identify the best option. And to identify the best option, we will need to compare sleeping with reading the novel.

Likewise, for our prudential and moral purposes here, the relevant alternative to existence is not a life with zero value, but nonexistence. For instance, we want to be able to morally compare outcomes in which individuals come into existence and outcomes in which they do not. This does not mean that it is irrelevant to compare a particular life to a life that has no value when assessing the benefit of coming into existence. But the reason why it is relevant to compare to a life in which there is no value is that there is no value in never existing. In other words, a life in which there is no value is relevantly similar to never existing. But this claim, of course, involves a comparison of existence and nonexistence. And so what we need is a comparison of these two alternatives.

15. For further discussion of this claim, see also McMahan’s and Persson’s contributions to the volume.
16. For instance, perhaps such a preference must be rational, self-regarding, or synchronic with its object. For a list of possible conditions, see Fehige and Wessels (1998), pp. xxv–xxvi.
17. Bayles (1975–76), pp. 299–300, McKie (2001).
18. Broome (1993), p. 77, Heyd (1992), pp. 30–31, 122, Parfit (1984), p. 489.
19. Thus, it is common for modal actualists (i.e. proponents of the view that everything that exists actually exists) to distinguish between existence and obtaining. A merely possible entity is one that exists but does not obtain (is not instantiated). See Stalnaker (1979) and Plantinga (1979).
20. Wlodek Rabinowicz has suggested this account of the relevant relation (personal communication). In respect of ontological modesty, it is located between two alternative accounts of the relation. On the more modest account, statements about possible (but non-actual) states have actual states as their truth-makers—i.e., what renders them true, when true, is the way things actually are (e.g. Armstrong 1997, Chapter 10). The relevant relation can then be claimed to obtain between the state of affairs, *you exist*, *you*, and (appropriate) states of affairs serving as

truth-makers for claims about your nonexistence. This account is ontologically less extravagant in the sense that it does not claim the existence of non-obtaining states of affairs. More extravagant than both this account and that offered in the main text is one according to which the relevant relation obtains between *you* and the state of affairs, *you exist*, in one world, and the state of affairs, *you do not exist*, in another world. This account relies on a highly controversial modal realism defended by Lewis (1986).

21. Holtug (2004).
22. Unger (1990), pp. 92–97.
23. Holtug (2006); Holtug (forthcoming, Chapter 4).
24. Parfit (1971), Parfit (1984), Chapter 12. For a more recent account of why identity is not what matters, and an attempt to deal with various objections, see Parfit (1995).
25. Parfit (1984), Section 13.
26. Holtug (2006), pp. 97–103; Holtug (forthcoming, Chapter 4).
27. Garrett (1990), pp. 31–32.
28. Noonan (1989), p. 160.
29. Here, I speak of negative states of affairs. However, I do not want to commit myself to the claim that there are any such states. Thus, depending on what sort of an ontology we assume, claims about negative states of affairs may have to be reinterpreted in terms of positive states of affairs that may serve as truth-makers for claims like “*b* does not exist.” But since it is convenient to speak of negative states of affairs, I shall continue to do so.
30. I discuss these and related questions about gene therapy in Holtug (1993, 1997).
31. Zohar (1991), p. 275.
32. Zohar (1991), p. 288.
33. Zohar (1991), p. 275.
34. Elliot (1993), p. 28.
35. For further elucidation and defence of the view that we are essentially animals, also known as “animalism,” see e.g. Carter (1982), Oderberg (1997), Olson (1997), and Snowdon (1990). For elucidation and defence of the view that we are essentially persons, see e.g. McMahan (2002), Parfit (1984), Part 3, Shoemaker (1984).
36. Elliot (1993), p. 31.
37. For a more general argument for why different psychologies realized in the same organism would not amount to different persons, even if we assume a psychological account of persons, see Persson (1995).
38. Along such lines, Tooley suggests that while perhaps full self-consciousness is not a necessary condition for being a person, the possession (and perhaps exercise) of the concept of a continuing mental substance is. Tooley (1983), p. 146. And along what seem to be similar lines Kuhse and Singer write: “We must recall, however, that when we kill a new-born infant there is no *person* whose life has begun. When I think of myself as the person I now am, I realize that I did not come into existence until some time after my birth. At birth I had no sense of the future, and no experiences which I can now remember as ‘mine.’” Kuhse and Singer (1985), p. 134.
39. Some will deny that it has present self-interests that reach into the further future, e.g. because it does not possess self-consciousness. However, as I have argued elsewhere, it is implausible to claim that the present possession of self-consciousness is a precondition for present self-interests. Holtug (2006); Holtug (forthcoming, Chapter 2).
40. This article derives in part from Holtug (forthcoming). Parts of the book have already been published in journals. Thus, Section 4.2 relies in part on Holtug (2001, 2004). Furthermore, Section 4.4 touches on a few themes also discussed in Holtug and Sandøe (1996).

References

- Appiah, K.A. 2005. *The ethics of identity*. Princeton, NJ: Princeton University Press.
- Armstrong, D.M. 1997. *A world of states of affairs*. Cambridge: Cambridge University Press.
- Bayles, M.D. 1975–76. Harm to the unconceived. *Philosophy and Public Affairs* 5: 292–304.

- Broome, J. 1993. Goodness is reducible to betterness: The evil of death is the value of life. In *The good and the economical*, eds. P. Koslowski and Y. Shionoya, 70–84. Berlin: Springer-Verlag.
- Carter, W.R. 1982. Do zygotes become people? *Mind* XCI: 77–95.
- Elliot, R. 1993. Identity and the ethics of gene therapy. *Bioethics* 7: 27–40.
- Fehige, C. and U. Wessels. 1998. Preferences—An introduction. In *Preferences*, eds. C. Fehige and U. Wessels, xx–xliii. Berlin: Walter de Gruyter.
- Garrett, B. 1990. Personal identity and extrinsicness. *Danish Yearbook of Philosophy* 25: 23–40.
- Glover, J. 1977. *Causing death and saving lives*. London: Penguin Books.
- Glover, J. 1992. Future people, disability, and screening. In *Justice between age groups and generations*, eds. P. Laslett and J.S. Fishkin, 127–43. New Haven: Yale University Press.
- Hanser, M. 1990. Harming future people. *Philosophy and Public Affairs* 19: 47–70.
- Heyd, D. 1992. *Genethics. Moral issues in the creation of people*. Berkeley: University of California Press.
- Holtug, N. 1993. Human gene therapy: Down the slippery slope? *Bioethics* 7: 402–19.
- Holtug, N. 1997. Altering humans—The case for and against human gene therapy. *Cambridge Quarterly of Healthcare Ethics* 6: 157–74.
- Holtug, N. 1998. Egalitarianism and the levelling down objection. *Analysis* 58: 166–74.
- Holtug, N. 2001. On the value of coming into existence. *Journal of Ethics* 5: 361–84.
- Holtug, N. 2004. Person-affecting moralities. In *The repugnant conclusion. Essays on population ethics*, eds. J. Ryberg and T. Tännsjö, 129–61. Dordrecht: Kluwer.
- Holtug, N. 2006. Personal identity, self-interest and preferences. In *Philosophy and ethics*, ed. L.V. Siegal, 59–114. New York: Nova Science Publishers.
- Holtug, N. forthcoming. *Persons, interests and justice*. Oxford: Oxford University Press.
- Holtug, N. and P. Sandøe. 1996. Who benefits? Why personal identity does not matter in a moral evaluation of germ-line gene therapy. *Journal of Applied Philosophy* 13: 157–66.
- Kavka, G.S. 1982. The paradox of future individuals. *Philosophy and Public Affairs* 11: 93–112.
- Kripke, S. 1980. *Naming and necessity*. Oxford: Basil Blackwell.
- Kuhse, H. and P. Singer. 1985. *Should the baby live?* Oxford: Oxford University Press.
- Kymlicka, W. 1995. *Multicultural citizenship*. Oxford: Clarendon Press.
- Lewis, D. 1986. *On the plurality of worlds*. Oxford: Blackwell Publishers.
- McKie, J. 2001. Thinking about possible people: A comment on Tooley and Rachels. *Bioethics* 15: 146–56.
- McMahan, J. 2002. *The ethics of killing. Problems at the margins of life*. New York: Oxford University Press.
- Miller, D. 1995. *On nationality*. Oxford: Clarendon Press.
- Noonan, H. 1989. *Personal identity*. London: Routledge.
- Oderberg, D. 1997. Modal properties, moral status and identity. *Philosophy and Public Affairs* 26: 259–98.
- Olson, E.T. 1997. *The human animal. Personal identity without psychology*. New York: Oxford University Press.
- Parfit, D. 1971. Personal identity. *The Philosophical Review* 80: 3–27.
- Parfit, D. 1976. On doing the best for our children. In *Ethics and population*, ed. M.D. Bayles, 100–15. Cambridge, MA: Schenkman Publishing Company.
- Parfit, D. 1982. Future generations: Further problems. *Philosophy and Public Affairs* 11: 113–72.
- Parfit, D. 1984. *Reasons and persons*. Oxford: Clarendon Press.
- Parfit, D. 1995. The unimportance of identity. In *Identity*, ed. H. Harris, 13–45. Oxford: Clarendon Press.
- Persson, I. 1995. Genetic therapy, identity and person-regarding reasons. *Bioethics* 9: 16–31.
- Plantinga, A. 1979. Actualism and possible worlds. In *The possible and the actual*, ed. M.J. Loux, 253–73. Ithaca: Cornell University Press.
- Roberts, M.A. 1998. *Child versus childmaker: Future persons and present duties in ethics and the law*. Lanham, MD: Rowman & Littlefield.
- Shoemaker, S. 1984. Personal identity: A materialist's account. In *Personal identity*, eds. S. Shoemaker and R. Swinburne, 67–138. Oxford: Basil Blackwell.

- Snowdon, P.F. 1990. Persons, animals, and ourselves. In *The person and the human mind*, ed. C. Gill, 83–107. Oxford: Clarendon Press.
- Stalnaker, R.C. 1979. Possible worlds. In *The possible and the actual*, ed. M.J. Loux, 225–34. Ithaca: Cornell University Press.
- Temkin, L.S. 1993. *Inequality*. New York: Oxford University Press.
- Tooley, M. 1983. *Abortion and infanticide*. Oxford: Clarendon Press.
- Unger, P. 1990. *Identity, consciousness and value*. New York: Oxford University Press.
- Woodward, J. 1986. The non-identity problem. *Ethics* 96: 804–31.
- Zohar, N.J. 1991. Prospects for “genetic therapy”—Can a person benefit from being altered? *Bioethics* 4: 275–88.

Chapter 5

Do Future Persons Presently Have Alternate Possible Identities?

Clark Wolf

Abstract This paper argues that the nonidentity problem rests on an overly narrow conception of personal identity. The criteria for identity across possible worlds are vague and uncertain, unable to support the finely-grained judgments made in debating whether an action would harm future people. On the more plausible coarse-grained account of personal identity defended in this paper, there is no basis for denying that the very same child can be born genetically impaired or perfectly healthy. On this account of identity, the non-identity problem does not arise.

Keywords Personal identity · Definite descriptions · Possible worlds · Vagueness.

5.1 The Erewhon Hypothesis

In the novel *Erewhon*, Samuel Butler describes a fictional world where people believe in life before birth.¹ The Erewhonians believe that unborn souls constantly flutter around eligible parents, “giving them no peace either of mind or body until they have consented to take them under their protection” by giving birth to them. In order to be born, an Erewhonian soul must commit a kind of suicide, abandoning the felicitous advantages of the unborn to exchange them for the troubles and cares of life after birth. The exchange is not a good one: the unborn cannot be unhappy, and post-partum life is risky at best and tragically miserable at worst. So bad is the bargain of post-birth existence, that the very fact that one of the unborn desires to be born is taken as evidence that the individual’s mind is not sound and that the choice may thus be involuntary. Any unborn soul wishing to be born must endure an extended court proceeding to prove that the choice is free and fully informed. Those who are able to win their case must take a potion to “destroy their memory and sense of identity. They must go into the world helpless, and without a will of their own; they must draw lots for their dispositions before they go, and take them,

C. Wolf (✉)
Philosophy and Bioethics, Iowa State University, IA, USA
e-mail: jwcwolf@iastate.edu

such as they are, for better or worse.” After birth, the infant is held to be entirely responsible for any disadvantages it may endure and for all the risks it must incur. The birth parents compel the newborn to sign (by proxy) a document in which the infant accepts full responsibility and absolves its parents of all liability for any of the damaging or disappointing incidents of life.²

The beliefs of the Erewhonians were intended to seem very strange. But like other parodies, *Erewhon* attempts to place before us the image of a world that is both distant enough from our own to inspire curiosity, and close enough to elicit self-reflection. While Butler’s *Erewhonians* claimed to believe that the unborn are fairly well off, Richard Dawkins apparently has a very different view of their prospects. In a television series titled *The Root of All Evil*, he offers the following in praise of the good fortune we may thank for our very existence:

We are going to die, and that makes us the lucky ones. Most people are never going to die, because they are never going to be born. The number of people who could be here in my place outnumber the sand grains of Sahara. If you think about all the different ways in which our genes could be permuted, you and I are quite grotesquely lucky to be here.³

If we’re lucky not to be among the nonexistent, then those who never exist must be unlucky, or at least less lucky than we are. But it is difficult to know what this might mean. “Just who are these unlucky nonexistent persons?” one might ask. “Can’t we do something for them?”

Dawkins makes another common assumption that should, as I will argue, seem just as odd as the suggestion that nonexistent people are unlucky. The passage implies that we would not have existed at all if our genes had taken one of the innumerable permutations he mentions. If different gametes had met, a child with a different genetic makeup would have been born. In that case, Dawkins implies, someone else would have existed instead and we would never have existed at all. To some people, the claim that our identity depends on our genetic makeup in this way has seemed so obviously true as to need little supporting argument. But I will argue that it is *not true* that our identities depend on our genetic makeup in this way. More accurately, the claim is not *simply* true, since the concept of “identity” is ambiguous, and because there are useful and sensible conceptions of “identity” on which the claim is false. I will argue that the so called “non-identity problem” is partly an illusion. This illusion arises because we have unreflectively accepted an *Erewhonian* hypothesis about the alternative possible identities of people who do not yet exist. In this paper, I hope to provide arguments that go some way toward dispelling this pernicious illusion. But I will also offer advice about how to make the kinds of choices for which the non-identity problem is thought to arise.

5.2 Variations on a Theme by Parfit

It will be useful to consider just a few problematic moral choices that exemplify the problem in question. Readers of Derek Parfit’s *Reasons and Persons* will immediately recognize the following examples as variations on themes from that book.

Negligent Physician: Because Alph is concerned about the possibility that the child he might conceive will have avoidable birth defects, he visits a physician for tests. The tests show that if Alph conceives a child during the next year, the child will suffer from a severe birth defect. If he conceives a child after this period has passed, there is every reason to believe that his child will have no such disadvantage. In a state of drunken carelessness, the physician misplaces the results of Alph's test, but blithely judges that Alph is unlikely to pass on any defect to his children. As a result of the physician's reprehensible carelessness, Alph conceives Beth, who is born with a serious birth defect for which there is no effective treatment or accommodation. But for the physician's error, Alph would have conceived a different child who would not have had a disability. While Beth is seriously disadvantaged by her disability, her life is not so miserable that she regrets having been born.

Clearly Alph has a valid complaint against his negligent physician. But has Beth a similar case? Beth suffers from a disadvantageous condition which is the direct result of the physician's wrongful behavior. But if the physician had behaved properly, it is argued, Beth would not have existed at all. Because the physician's faulty action effectively determined Beth's identity, she cannot claim that she would have been better off but for the Physician's negligent action. Or so the non-identity problem would lead us to believe.

Depletory Policy: The U.S. President faces a decision that will determine the future of energy policy and will influence the availability of energy alternatives for many generations in the future. He could either choose policy A or policy B. Policy A will create dramatic but relatively short-term benefits for the next two or three generations, but is expected to lead to environmental disaster in the long run. Policy B will yield slightly lower benefits in the proximate future, but these benefits will be sustainable for the foreseeable future. Instead of leading to environmental disaster, Policy B would help to restore existing environmental damage. But because there is a time lag between the restoration and the resulting environmental benefits, the benefits will not improve the lives of anyone presently living.

Since future people are not a voting constituency, the President doesn't care about the people who will live in the distant future. But the President cares quite a lot about the opinion of present voters who are themselves mostly concerned with the present and more immediate future. For this reason the President chooses Policy A. As a direct result of this choice, things turn out very badly for people who live later. But the choice of Policy A has other subtle but wide-ranging implications for people's lives, and because of these changes different people are conceived and born than the people who would have been conceived and born if the President had chosen Policy B. By the time the disadvantages arrive, none of the people who suffer from these disadvantages would have existed if the President had instead chosen Policy B.

Once again, the problematic choice both causes disadvantage, and determines the identities (thus the very existence) of those who suffer the disadvantage. Those who suffer cannot claim to be worse off than they would have been if the president had acted differently. Thus it might be argued that they have no valid complaint against the President whose problematic or wrongful choice caused their suffering.

The case is not entirely hypothetical, since the choice described is very similar to choices we presently face. Many have urged that large scale public action is necessary to mitigate the harmful effects of climate change and to turn back the forces that are causing this change. But it is widely acknowledged that the benefits of such action will not be realized for many generations. If the identities of the people who would benefit from an effective climate policy would be changed by factors

stemming from the implementation of such a policy, then climate policy presents us with a practical application of the non-identity problem. And it is quite plausible to think that the identities of later generations really may be changed by any large-scale social policy like a climate mitigation initiative. Such policies will change many lives, so that different people will meet and the circumstances of later lives will be changed. Over time, the effects of small changes will be amplified, so that after several generations have passed the entire human population of the earth might be genetically different from the population that would have existed otherwise. Thus those who will suffer disadvantage if we fail to implement an effective climate policy are people who would not have been better off if a better policy had been implemented. They would not have existed at all. Or so we might be encouraged to believe.

These examples show the non-identity problem to be a practical problem that arises in certain decision contexts. But it is also a conceptual problem in which many of our standard moral concepts are implicated. The practical problems, as I will argue, stem from the underlying conceptual problems. Among other concepts, the non-identity problem raises issues for the pareto criterion, the concept of harm, Millian liberal political theory, and for person-affecting concepts and theories more broadly. The non-identity problem calls into question whether distant future persons could have rights against members of the present generation. For this reason, the problem seems to undermine the possibility that any theory of justice or right could apply between distant generations. Because of this problem, some theorists have more or less abandoned the idea of intergenerational justice altogether.⁴

The Pareto Criterion. The pareto criterion recommends any policy that is better for some and worse for no one. In the case described, those who benefit would not have existed otherwise, while those who bear the cost of climate mitigation would have been better off without it. Climate mitigation is thus worse for some (those who bear the present cost) but better for no one, since those who exist in the *generations* that will benefit would not have existed if different policies had been chosen. Since climate mitigation policies are worse for some and better for none, they are *pareto inferior* to policies that do not involve climate mitigation.

Harm. While it might seem that willful choice of policies that lead to environmental and human disaster are harmful to those who suffer, it has sometimes been argued that “harm” involves a counterfactual condition: A harms B only if A’s wrongful action makes B worse off than B would have been if A had acted as he should have, instead of as he did.⁵ Thus it would seem that those disadvantaged by the depletionary policy are not harmed by it, nor can Beth claim to have been harmed by the negligent physician whose faulty action led to her disability. Here the non-identity problem has had practical legal implications: U.S. Courts have been led by non-identity arguments to conclude that children who suffer disadvantage or disability as a result of malpractice cannot claim compensation if they would not have existed but for the malpractice in question.

Millian Liberalism. More broadly, those who find John Stuart Mill’s conception of liberalism, as defended in *On Liberty*, might reasonably find the non-identity problem disturbing. According to Mill, it is wrong to limit liberties except to prevent harm to others.⁶ Policies to mitigate damage due to climate change will limit

the liberty and impose costs on present people. But if the beneficiaries wouldn't have existed otherwise, then these policies can't be justified in terms of harm prevention: since those who will suffer the ill effects of climate change wouldn't have existed otherwise, they cannot be harmed if we don't take steps to mitigate change. It would appear, therefore, that Mill's principle would prohibit policies to reduce climate change. At least, it will prohibit such policies whenever they involve present restrictions on liberty, and when the policies in question would have sufficiently wide-ranging effects that they will determine the constituency of future populations of the earth.

Person-Affecting Principles and Person-Affecting Moral Concepts. Some philosophers have urged that moral theory should be "person affecting." By this, people sometimes mean to refer to a "person affecting principle [PAP]." There are different principles that have been given this name: Some times the PAP is associated with the view that

PAP-1: Nothing is bad (good) unless there is someone for whom it is bad (good).

At other times, the PAP is associated with a different principle:

PAP-2: It is good to make people happy, but we may be indifferent about making happy people.

These are not two ways of saying the same thing, they are different principles. But they are interestingly related, and both are implicated in the non-identity problem. According to PAP-1, the results of the *negligent physician's* carelessness are not bad if there is no one for whom they are bad. And according to description given, these results are not bad for Beth since she would not have existed otherwise. Similarly, if global warming changes the constituency of the future population of the world, then it would seem that it is not bad in the sense that it is not bad for the people who suffer from the effects of climate change.

Even those who do not accept PAP-1 may find that their own moral views raise problems in non-identity cases. For example, if one wanted to say that we have an obligation to future generations to reduce climate change, or that climate is a matter of intergenerational justice, then one might find it important to explain how our wrongful failure to implement climate policies would violate the rights of future generations. The non-identity problem makes it very difficult to see how this could be. Those who will suffer, whose rights might be supposed to be violated by our failure to implement such policies, are people who would not have existed otherwise. So climate policy cannot, one might argue, be a requirement of intergenerational justice. Considerations of rights and justice are "person affecting concepts," even though their use does not imply acceptance of PAP-1 (or PAP-2), or of a fully person-affecting morality. Thus any moral theory that employs concepts of rights and justice will have problematic implications in non-identity cases.

If one accepts the non-identity argument as it is usually presented, it seems to lead us to unfortunate conclusions when combined with a number of our common

evaluative concepts: I have argued above that non-identity cases raise problems for the pareto criterion, the concept of harm, Millian liberalism, and person-affecting principles, rights, and justice. These implications are especially counterintuitive if one thinks that identity does not, or should not matter in the way it seems to matter in these cases.

5.3 The No-Difference Thesis

Many people find these conclusions counterintuitive, but it is difficult to see a problem in the argument that leads us to them. One reason why these conclusions are counterintuitive is that we have independent reason to think that the alternative identities of future people shouldn't matter from the moral point of view. The fact that one decision results in the existence of different people than would have existed otherwise should make no morally significant difference at all. Parfit calls this the *No-Difference Thesis*, and supports this thesis with a compelling example:

The Medical Programmes. There are two rare conditions, J and K, which cannot be detected without special tests. If a pregnant woman has Condition J, this will cause the child she is carrying to have a certain handicap. A simple treatment would prevent this effect. If a woman has Condition K when she conceives a child, this will cause this child to have the same particular handicap. Condition K cannot be treated, but always disappears within two months. Suppose next that we have planned two medical programmes, but there are funds for only one; so one must be canceled. In the first programme, millions of women would be tested during pregnancy. Those found to have Condition J would be treated. In the second programme, millions of women would be tested when they intend to try to become pregnant. Those found to have condition K would be warned to postpone conception for at least two months, after which this incurable condition will have disappeared. Suppose finally that we can predict that these two programmes would achieve results in as many cases. If there is Pregnancy Testing, 1,000 children a year would be born normal rather than handicapped. If there is Preconception Testing, there would each year be born 1,000 normal children rather than 1,000 different handicapped children. (Parfit 1984, p. 367)

The only difference between the choice to fund treatment for Condition J, and the choice to fund treatment for Condition K is that in the former case, we would be benefiting people who would have existed anyway, while in the latter case we would be causing different persons to come into existence than would have existed otherwise. If we choose to fund treatment of condition J, it would seem that there are no potential complainants. That is, those who are born deformed as a consequence of our choice can not rightly say "but for your choice, I would be better off than I currently am." If we had chosen to fund condition K, those children would not have existed at all. On the other hand, if we choose to fund treatment for Condition K, the children who are born deformed because their mothers had Condition J *could* say that but for our choice they would have been better off, since they would have existed in any case. If we believed that the problem of non-identity had moral significance, this should lead us to the conclusion that we have an important moral reason to fund treatment of Condition J rather than Condition K. But most people find it obvious

that there is no moral difference between the two choices: the consequences of either choice are the same in all morally relevant respects. This is precisely what the no-difference thesis states: The fact of identity does not matter from the moral point of view.

Parfit's account of the no-difference thesis, and the supporting example he offers, are highly persuasive. For my own part, I find the case for this thesis entirely compelling. Thus for the purposes of this paper, I will assume that the no-difference thesis is *true* and will consider what conceptual resources can be marshaled to resolve the non-identity problem where the no-difference thesis is treated as a constraint. Identity may matter quite a lot in some contexts, but it does not matter in the way that would lead us to judge that one of Parfit's procedures is superior, from the moral point of view, to the other. As I will argue, the *non-identity* cases, including those described above, are problematic from the start because they stretch the concept of "identity" beyond its meaningful application.

5.4 Identity-Determining Choices and Identity-Determining Characteristics

To address the non-identity problem, I propose that we should carefully examine the concept of identity employed in the examples in which this problem arises. We need to consider more carefully what features of ourselves influence our identity, and just how they might do that. In this interest, I will introduce the idea of "identity determining choices and acts." An *identity determining choice* is a choice that determines that one person or group of people will exist instead of another person or group that might have existed. As we have seen, the non-identity problem arises in the cases above because there are identity determining choices that seem to cause disadvantages for those who come to exist as a result. In the most problematic cases, we may have a negligent party whose faulty act is the cause of impairment or serious disadvantage for someone who would not have existed otherwise. Under ordinary circumstances, this would be sufficient to support a *prima facie* complaint on behalf of the sufferer, against the faulty actor responsible for her predicament. But when the act in question is an *identity determining* act, then the actor seems to have an effective response: But for the faulty action in question, the person who is supposed to have been harmed would not have existed at all. If there are identity-determining choices, they are choices that determine that one person (or set of persons) will exist instead of another.

Which are the Identity Determining Characteristics? What are the characteristics that have this effect, determining our identities and our existence in this way? It is often assumed that one's genetic makeup determines one's identity. Indeed, something like this assumption is implicit in what Dawkins says in the passage quoted earlier: "If you think about all the different ways in which our genes could be permuted," he suggests, "you and I are quite grotesquely lucky to be here." Apparently, he and many other philosophers believe that we would have been among the much less lucky non-existent people if our genes had been permuted in the relevant way.

What is the relationship between our genes and our identities? We might consider the following claims:

Claim 1: “Any child my parents might have conceived that had the same genetic makeup as mine would have been *me*, even if other things were substantially different.”

Claim 2: “Any individual with different genetic makeup from my genetic makeup would be a different person from the person I am.”

Claim 3: “If the child my parents conceived had had a different genetic makeup from mine, that child would have been a different child—a different person—from the child I was.”

Claim 1 is obviously false. Identical twins, for example, have the same genetic makeup but they are different individuals.

Claim 2 is also false. If a person were given a genetic therapy that changed the DNA in each of his cells but left other of his characteristics unchanged, we would not regard him as having become a different person. Genetic therapy of this sort would not, for example, imply that the resultant individual no longer owned property that was owned by the person who chose to undergo the procedure, or that the person who left the operation would not be contractually bound to pay for it (since another person chose to undergo it!). If Claim 2 were true, hospitals would need to collect payment for such genetic therapies in advance.

Is there reason, then, to believe Claim 3? Many writers have argued that Claim 3, or something very much like it, is true. For example, Parfit defends the following

Time Dependence Claim [TDC]: If any particular person had not been conceived when he was in fact conceived, it is *in fact true* that he would never have existed. (Parfit 1982, p. 351)

Parfit’s *Time Dependence Claim* is not the same as Claim 3 above, but they are related. Parfit uses considerations like those articulated in Claim 3 to support the TDC. So if Claim 3 is questionable, then the TDC may be questionable as well.

Is Claim 3 true? Consider the child born to your parents on your birthday. Imagine a child exactly like you in all respects except one: the imaginary child had a different genetic makeup from your genetic makeup. The difference in the genetic makeup of the imaginary child, we can imagine, does not determine any phenotypic differences, so this child looks and acts exactly like you. The physical difference between this imaginary child your parents might have had, and the child they did have are real physical differences, but they can only be detected by a genetic test.

Would this imagined child be you, or would it be a different child? Why should we think that an irrelevant physical difference in genetic makeup would make this imagined child a different child from the child your parents actually did have? We don’t think of other minute physical differences as “identity determining” in this way: for example, if you had been born with a differently shaped nose or belly button, we would not regard these characteristics as determining that you are a different person from the person you are. There is nothing magic about our genes that

automatically implies that our “identities,” in the relevant sense of that word, are determined by our genetic makeup.

The examples given above are intended to coax us away from the assumption that our “identities” (again, in the relevant sense of the concept “identity”) are determined by our genes. But even if we are effectively coaxed, we might regard our genetic makeup as *relevant to* if not *uniquely determinative of* our identities. For example, our genetic makeup might be one among several different characteristics that combine to make you the person you are. Your genetic makeup might be sufficient to limit the range of persons you could possibly have been. This possibility is consistent with the thought that changing the genetic makeup (and nothing else) of an already existent person might leave him or her the same person after all.

5.5 Ambiguous “Identities”

In order to resolve this problem, we need to think more deeply about the way in which the properties we possess make us the people we are. Every one of us has formative experiences that have shaped and changed us, and we may even say that these experiences have influenced our *identities* in important ways. While the sense of “identity” we employ when we say this may be different from sense of “identity” that generates the non-identity problem, it is instructive to recognize that we use the term in a variety of different senses, and it is not obvious which sense will be the relevant one to use in different contexts. Even trivial experiences leave their mark on us and change us slightly. We don’t usually think of these small changes as the kinds of changes relevant to the non-identity problem, because we view ourselves as maintaining our identity through time. We are psychologically connected with our pasts: we remember these life-changing experiences and we remember what we were like before they shaped us. But notice that the sense of “identity” involved in this more classic problem of identity over time is different in relevant respects from the problem of “identity” involved in the non-identity problem.

In philosophical contexts, psychological connectedness is often considered a fundamental and perhaps a necessary condition of continued identity over time, but this part of our concept doesn’t apply to contingent future persons at all. When we say that one individual would come to exist rather than another, as the result of an identity-determining procreative choice, we do not simply mean that there is no psychological connectedness among the putatively different persons who might come into existence depending on the choice made. Except in an Erewhonian world, it is *impossible* to be psychologically connected with people who don’t exist, or for possible but non-existent people to be psychologically connected to people who will later exist. What we presumably mean is that the individuals who will exist will have different properties, and that the properties in question are sufficiently important, or essential, or constitutive. When these properties are changed, the resultant person is so fundamentally different that we should regard her (him?) as a different person entirely.

But which characteristics are so essential to our identities that we would have been different people if these characteristics had been different? We should be able to discover whether some characteristic is an identity determining characteristic by considering whether an existing person's identity would change if that characteristic were to change. Here is a principle that captures this intuitive idea:

Identity Principle [IP]: If characteristics C are identity determining for person P , then any child born with characteristics different from C would have been a different person from P (even if all other characteristics were the same).

Note that C might be an individual property, or a collection of essential properties that make P the person she is. IP captures the idea that the identities of future persons depend on the properties they will possess, not on some kind of psychological connectedness among possible persons.

5.6 Vague “Identities”

The identity principle encourages us to consider which properties of ourselves (or others) might determine that we are the person we are instead of some other person we might have been. But if these properties (whatever they are) may be possessed in greater or lesser degrees, then we need to consider the possibility that our “identities” may be vague. Consider the characteristics C that are regarded to be identity determining for a possible person P . That is, if the child is born with C , then that child will be P , but if the child is born without C , then the child will be a different person from P . It seems most plausible to think that C must be a set with multiple members, since we may think of our identities as dependent on more than one of our characteristics. In what follows, I will assume that C is a set, but I do not believe that my argument depends in any central way on this assumption. If C is a set of characteristics, then there will be many different ways in which the members of that set might be slightly perturbed, leaving P the same person, or *almost* the same person. Even if (by definition) these characteristics C are identity determining, it need not follow that even imperceptibly minute changes in C would result in the existence of a different person from P . But as we imagine increasingly radical changes in these characteristics, eventually we might judge that the changes are sufficient to determine that a different person exists from the one who might have existed. We might express one important part of this thought as follows:

Non-Identity Principle [NIP]: If C are the identity determining characteristics for P , and if $C1$ is a member of C , then P 's identity will not survive radical perturbations of $C1$.

If $C1$ were to change *enough*, then P would not exist, and a different person would exist instead. But what if the child isn't born *without* C , but with an imperceptibly slight variant on C ? We might think that there is a range of changes such that as long as characteristics C varies only slightly, within this specified range, then the identity of the child will not change. Within this range, the child born will still be P , not some other child with a different identity.

On this view, it is plausible to think that our identities may be vague: where C changes slightly, it might slightly change our identities, though not so much that we could properly be identified as entirely different people. So where C changes in this way, the claim that P is the same person P would have been may vary as well. And perhaps at some threshold level P's identity will utterly change as a result of changes in C. Where P's identity is determined by C, and C is a property or set of properties that can vary by degrees, P's identity will can be vague. Where C contains N characteristics, we might represent the range of possibilities as an N-dimensional array with P at the origin. As one moves from the origin in any direction, P's identity "fades out" as it approaches the boundary or threshold beyond which her identity will be utterly different. Consider the following statements, ID for "identity," and NI for "non-identity":

ID: P is the same person after the change in C as she was before.

NI: After the change in C, P is no longer the same person P was before the change in C.

On one view of vagueness, ID will become "less and less true" over small perturbations of property C, and "increasingly true" over increasing perturbations.⁷ Over this same range of changes, NI would become "increasingly true." And once C has changed sufficiently, ID will eventually become false and NI will become true. Some theorists are uncomfortable with the idea that truth might come in degrees, and other theories of vagueness will have slightly different implications concerning the identity of P. I must admit that this way of thinking about vague predicates seems right to me, but I understand the reasons that give many people discomfort with "degrees of truth."⁸ My argument here will in no way depend on any such controversial account of vagueness or truth, since other theories of vagueness have relevantly similar implications for identity.

When we consider the most minor perturbations in C, it is easy to think that such changes will not alter P's identity. Still, when we consider further minor changes, there may come a point when we are unsure whether P's identity has been changed. If so, then P's identity would seem to be *vague*. If "identity" is vague, we can make the concept precise by stipulating bright-line definitions that identify precisely which changes in which characteristics will change P's identity in the relevant, stipulated sense. At one extreme, we have the strictest conception of identity, according to which C will include all of P's properties and any change in C will constitute a complete change of P's identity. Less strict conceptions of "identity" will tolerate broader changes before P's "identity," in the relevant sense, is changed. For example, one conception might identify genetic makeup as an identity-determining characteristic so that a child conceived at time T1 under circumstances C1 will be different from a child conceived at T2 under circumstances C2, since different sperm and ova would meet in the two cases. Another conception of this child's (these children's) identity might give us no ground to distinguish the two possibilities as children with two different identities: For example, the child conceived under either of these circumstances might be identified as "John and Mitzi's fourth child."

Depending on the purpose we have in mind when we identify this child, it will be appropriate to use different criteria to individuate him or her. For example, if we are trying to determine whether this child can rightly inherit property under the terms of a will, her identity as “John and Mitzi’s fourth child” might be all that we need. If we have different purposes in mind, we may need to consider other features of this child to fix his or her “identity” in the relevant sense. Once again, it is worth noting that “identity” is not a univocal concept. Reflection on the vagueness of “identity” thus reveals yet another range of different conceptions of identity. When we employ the term, it will be crucial to identify what is relevant about identity in the particular context so that we will use the *right* concept of identity, and not the wrong one. There is no single concept of “Identity” that we can appropriately employ in all circumstances.

5.7 Alternative Conceptions of “Identity”

When does the non-identity problem arise? The answer will be different depending on what conception of “identity” we employ. At one extreme, we can consider the strictest conceptions of identity. On the strictest conception, object A is identical with object B just in case A and B have all properties in common. When this strict conception is applied to the identities of objects or persons across alternate possible worlds, it implies that all events that influence and change us, even in infinitesimal ways, are identity determining. Thus, on this strictest conception, each of our properties—every property we now possess or ever will possess—is identity-determining. On this conception, I would be a different person if I had stubbed my toe on the way home from work, from the person I would have been in the world where I didn’t stub my toe, even if everything else about my life and the world were the same.

Of course this strictest conception is not the one we usually employ when thinking about the problem of personal identity, nor is it the concept that is usually assumed by those who discuss the non-identity problem. It is certainly not the conception we apply in moral contexts, to identify the objects of our obligations: if I borrow books from my friendly, hopeful friends, they may become bitter and cynical if I don’t give them back. But it would be wrong for me to argue later that I didn’t owe the books to the bitter cynical people but only to the friendly hopeful ones I borrowed them from. As I decide whether or not to return their books, the facts about my friends’ identities that are relevant from the moral point of view do not include the changes that would arise in their personalities as a result of my failure to do so. What *is* relevant is that they are the ones who will suffer disadvantage if I fail to keep my obligations, that they are psychologically continuous with the persons from whom I borrowed the books, and perhaps other related characteristics. Which characteristics? Psychological connectedness is often regarded as a necessary condition for personal identity over time, we have noted above that it is clearly not relevant when we consider alternative possible future persons. It is not clear

what it could mean for one possible future person to be psychologically connected to another.

In discussions of the non-identity problem, it is frequently assumed that our identities are somehow fixed by our genetic code. But as we have seen above, we can reasonably question whether our genes determine identity in this way. In some cases where we need to identify the objects of present obligations, the relevant conception of a person's identity is not the one that naively associates identity with genetic makeup: for example, one might include provisions in a will for one's "fourth child, if any such exist." People with identical genetic makeup need not be the same person (they might be twins), and changes in genetic makeup need not imply changes in identity. Genetic differences are like any other physical difference we might have possessed. If we have no reason to regard those other physical differences as identity determining, at least in cases where they would not cause other significant differences, then we should regard genetic differences in the same terms.

What is the concept of "identity" that we should employ in non-identity cases? If we want to individuate the person in question for moral reasons—for the purpose of assigning responsibility or blame, for example, over the class of persons responsible for her or his existence—then this is not simply a problem in metaphysics. It is essentially a problem for moral theory. In this case, we need to find the relevant conception of "identity" by considering the use to which we intend to put it, and the theory we are applying when we employ it. Where the question involves responsibility, our theory concerning the identity-determining characteristics of future persons should capture what is significant about them from the *moral* point of view, and this may be quite different from other senses of identity that we use to individuate persons for other reasons, or within other theories and projects.

5.8 Future Persons as Vague but Identifiable Objects of Present Obligations

Suppose that I am stranded on a desert island and I launch a bottle containing a note that says "if you find this message and bring it to my wife in New York, she will reward you with \$10,000." To whom does "you" refer in the context of my note? It refers to whoever finds the note. (If the note is never found, my use of "you" fails to refer.)⁹

On the Daily Show, on 13 March 2008, Kristen Schaal made a video message for "the first woman president of the United States." At the time, it was reasonable to suppose that the relevant recipient of this message might be Hillary Clinton, who was still a contender for the Democratic nomination. At present, we can identify the addressee of this video message by a definite description, we have no idea which person will come to fill that definite description. If the United States never elects a woman to the office of president, then as David Velleman suggests in the passage above, it will simply turn out that Schaal's message won't come to be addressed to anyone at all.

In the examples above, we can identify the person in question with some precision: The finder of Velleman's bottle; the first woman President of the United States. Sometimes a description like this will capture everything that is relevant about an individual from the moral point of view. In fact, it might sometimes be *exactly* what is relevant. The finder of Velleman's bottle, identified only by that description, will be owed \$10,000. In a similarly ambiguous vein, I can refer to *you*: to the present reader of this paper, where "*present*" refers to your "now," the time when you are reading these words. As I wrote these words, of course, I didn't know who you would be, or when your "present reading" of these words would take place. As an author, I may hope for multiple referents, but I can still assure you that my words now refer to *you*, and not to any of the other readers. Your identity is fixed as the present reader of these words, as the person to whom they are addressed, and as the individual to whom they refer.

Shortly after the publication of Parfit's book, Douglas MacLean considered a novel criterion for identifying the future persons to whom our obligations are owed:

Perhaps we should insist on a person-affecting criterion for harm but a place-holder criterion for wrong. Thus a person is wronged by an action if he is identified by a definite description and is worse off than another action would make a person picked out by the same definite description. This analysis may... seem artificial, but... it attempts to save two intuitions. The first is that the identity problem should not matter to the moral evaluation of an act. Who the members of a future generation turn out to be should not matter to the moral assessment of our actions that determine their environment, opportunities, and quality of life. Parfit would agree. The second intuition is that those who bear the consequences of our reckless or selfish choices have a ground for complaint against us. Parfit goes to great lengths to try to undermine this intuition and to expose its incoherence.¹⁰

As I understand it, MacLean's suggestion is that we should pick out future individuals by a definite description that uniquely identifies the place they will occupy when they come into existence. As MacLean urges, this view accommodates the no difference thesis, while allowing that those who bear the consequences of our bad choices have ground for complaint. MacLean articulated this view with some precision, but did not develop it further and relegated it to a long (and fascinating) footnote. Perhaps he was skeptical that his proposal could be effectively carried through.

More recently, Jeffrey Reiman and Caspar Hare have made similar suggestions.¹¹ Reiman suggests that choices involving future individuals should be made behind a veil of ignorance that blinds us to the specific identities of those who will come into existence and experience the consequences of our choices. Hare offers decisive arguments against "moral actualism," the view that only the interests of *actual* people (and not possible) are relevant for determining the moral status of an action.¹² But Hare also articulates a conception of *de dicto betterness*, according to which we should identify the future claimants (beneficiaries or victims) of present actions according to a definite description. Hare considers an example in which Mary conceives a damaged child, Mariette, when she could, by waiting, have conceived a (different?) healthy child. Hare writes:

[Y]ou may have a feeling, as I do, that Mary's wrongdoing is in a certain way personal—in the way that there's a special kind of relationship between Mary and Mariette. Mariette has special grounds to feel aggrieved by what Mary did. But if Mary does wrong by making the world worse, then this is mysterious. After all, nobody has a special complaint against here. Everybody can complain "You have made things worse." Nobody can complain "You have made things worse for me." The *de dicto* concern account puts a finger on Mariette's special grievance. Mariette alone can say, "You failed to show appropriate *de dicto* concern for your child, and I am your child."¹³

MacLean's "place-holder criterion," Reiman's account of choice behind a veil of ignorance, and Hare's *de dicto betterness* all have the same general implication for moral choice and decision making, and for identifying the victims of present "identity determining" choices.¹⁴ Indeed, I think these different suggestions are, for the most part, different ways of articulating the same correct view. The difficulty, as MacLean noted in his early paper, is to identify "a nonarbitrary way of knowing when to apply a principle of wronging placeholders rather than persons."¹⁵

To see what non-arbitrary criterion will work best, it is well to reflect on the features that make the non-identity problem paradoxical: The problem is that the identity problem seems to show that something that should not matter from the moral point of view—the different possible identities of future persons—seems to have great significance after all. Once we recognize that the articulation of the problem employs a very specific concept of "identity," and recognize in addition that different conceptions of "identity" are appropriate in different circumstances, we have new resources to address this problem. To find a non-arbitrary criterion, we need to identify the conception of "identity" that is appropriately employed in cases where the non-identity problem seems to arise.

My suggestion is a simple one. The paradoxical features are associated with the normative and moral implications of these cases. Therefore it is our *moral* theory that should determine the conception of identity that is relevant for these contexts. In non-identity cases, our concern is to identify the victims, complainants, and perpetrators, and our normative intuitions should be an essential guide in these cases. We have good reason to accept the no-difference thesis, and to assert that the alternate possible identities of future persons are not relevant from the moral point of view. So the appropriate conception of identity must be a conception that supports the no-difference thesis. The only conception of "identity" that does this is the minimal conception of identity (or the range of minimal conceptions) that prevents us from individuating future individuals as possessing different identities in putative non-identity cases.

If the no-difference thesis is true, then distant future generations will have a valid complaint against us, members of the present generation, if we needlessly destroy the resources they will need. But our obligation is to them not because of their genetic makeup or their personalities or other characteristics that will be determining features of their "identities" once they exist. They will have "identities," in some important senses of that term, only after they come into existence. What picks them out as the objects of our present obligations (and as our victims, should we fail to meet our obligations) is simply the relationship that stands between their interests

and our choices. In this case, the “they” of “their interests” should be understood to refer to their identities as given by a general “place-holder” description, as MacLean suggests, not by their genetic makeup, or by their identities as particular Cartesian egos, or by other characteristics of personality or person they may come to possess.

The place-holder conception suggested here is minimalist and vague, but in this context it is both all we need, and exactly what we need. To see this, it helps to consider the characteristics that identify people as victims and claimants in several more conventional cases:

The Boulder Roller: Gimel is walking on a high path, rolling boulders down the mountain for fun. There is a path below, and Gimel cannot tell whether there is anyone walking on it. But his boulder-rolling imposes a serious risk of harm on anyone who might be hiking there. In this particular instance, Gimel’s activity imposes unreasonable risks on Daleth, who happens to be hiking on the path below. Her foot is crushed under one of the boulders Gimel has recklessly sent down the mountain.

Obviously Gimel has an obligation not to roll boulders down the mountain. But to whom is this obligation owed, and by what criterion should we identify the claimants? The most plausible criterion is itself somewhat indeterminate: the obligation is owed not just to Daleth, but to *anyone who might be walking on the path below*. If there is no one on the path, then Gimel is lucky and his wrongful action will not cause anyone harm. But the characteristics that we use to identify the right-bearer, in this case, do not identify claimants by their personal physical characteristics or by their “identities” in the sense of identity over time. The class of individuals who might be on the path below is large and varied. Gimel’s obligation is to any member of that class of people who might be affected by his reckless behavior.¹⁶

What picks out Daleth as the object of Gimel’s obligation is that Daleth (i) is among those whose interests are threatened by Gimel’s reckless behavior, and (ii) that Daleth is in fact at risk, since she is on the path below Gimel. It is not Daleth’s “identity” in most standard senses of that term, that pick out Daleth as an object of Gimel’s obligations, it is a much more general sense of Daleth’s identity as “a person who is put at risk by the behavior in question.” We should draw the same conclusion in other “different person choices:” the characteristics that are relevant for individuating future persons from the moral point of view do not include their genetic makeup or other specific characteristics they may possess. But we can still identify a definite description that uniquely picks out the persons to whom we have obligations: they are simply “*the class of persons whose interests will be influenced as the consequences of our present choices.*” Individuating future persons in this way effectively smudges over the confusing conceptions of identity that generate the *non-identity problem*. The putatively different possible people we might bring into existence are *all the same* from the moral point of view. Just as it would be morally wrong for me to distinguish the cynical bitter people my friends might become after my broken promises from the happy hopeful people to whom I made my promises, it is similarly wrong to distinguish between the different populations that might come into existence as a result of our present choices. To make such a distinction is simultaneously to make two different mistakes: One is a moral mistake, since the

criteria are moral criteria. But it is also a linguistic mistake, since it involves the articulation of a distinction where there is no relevant underlying difference.

While this suggestion may seem odd, it is not unfamiliar. There are many contexts in which the persons who are the subjects of obligations are appropriately identified by similarly broad criteria that apply broadly to people who fit a given description. Consider the obligation not to shoot bullets into the woods when one cannot see whether there is anyone there, or the obligation to care for one's brakes in case one needs to stop suddenly for a pedestrian. These are ordinary obligations. In both cases, we identify the claimants by a general criterion, and any individual who fits the criterion has a valid complaint if the obligation is violated. "Anyone in the woods" has a claim against the shooter; and "Any pedestrian crossing the street in front of your car" has a claim against you if you failed to keep your brakes in good repair. And "anyone on the path below" has a claim against Gimel the boulder roller.

In the case of future persons, we cannot separate "the criteria by which we identify future claimants" from "the criteria we use to identify our obligations to them." This is what it means to say that the concept of "identity," in these cases, is theory dependent. The relevant concept is the one that fits best in our moral theory, making best sense of other concepts that apply in these cases. Actions that are "identity determining" are said to change the properties or characteristics of future persons. But in non-identity cases, it is not just the persons who change. The concept of identity will also change as we apply it in different contexts.

Stable "Identities" Across Alternate Possible Characteristics. Note that some characteristics of future persons are stable across the kinds of cases most often identified as different person choices: If Mitzi and John conceive a child today, it might be identified as "Mitzi and John's fourth child." If they were to conceive a child two months from now, it would still be *their fourth child*. There is thus a definite description (*Mitzi and John's fourth child*) that uniquely identifies that individual, regardless of the difference in makeup that individual would have if *their fourth child* were conceived at different times. Before conception, there are many possible sets of properties that *their fourth child* might have: it might be male or female, it might or might not carry genes for baldness or shortness or blondness or brown eyes or any characteristic with a genetic link. As long as this child is a *person*, however, it is plausible to think that our obligations to this individual have everything to do with the characteristics that are stable (it would be *our* child whom we are responsible for bringing into existence) and little to do with other specific properties or characteristics it might possess. Like the general criterion that identifies the claimants and potential victims of Gimel's boulder rolling, the criteria that identify this individual as a person to whom we have special obligations do not in any way refer to his or her "identity" in the traditional philosophical senses of "identity." In the morally relevant sense—that is, the sense of "identity" that we should use in considering our obligations to this person whom we might bring into existence—this child is the same child regardless of its genetic makeup, sex, or many other variable characteristics he or she may or may not possess.

If the future people are badly off because we deplete the earth of the resources they will need, then in the relevant sense they are worse off than they (the future people) would have been if we had behaved more responsibly. Of course, if we had behaved more responsibly the persons who would have experienced the consequences of our choices would have had different properties and different genetic makeup. But in the relevant sense, from the moral point of view, this set of persons is the same population.

5.9 Parfit and “The Descriptive View”

In his original discussion of the non-identity problem, Derek Parfit considers and rejects a view that he calls “The Descriptive View.”¹⁷ He does not devote much space to this view, as he apparently regarded it as too obviously wrong-headed to consider seriously. Parfit’s argument is in two main parts: First, he argues for a principle he calls *The Time Dependence Claim*. Then he provides several arguments against what he calls *The Descriptive View*. Since the *Descriptive View* is a very close relative of the view I have described here, it will be important to consider Parfit’s objections to it. Parfit first considers the following claim:

The Time Dependence Claim [TD]: If any particular person had not been conceived when he was in fact conceived, it is *in fact* true that he would never have existed.¹⁸

Parfit notes that this claim is not obviously true, but he urges that it is non-controversial and “easy to believe.” He refers to the physical continuity between the ovum from which we developed and our present selves in support of this claim. Parfit also articulates a more minimal principle, which he calls “TD2.”

Time Dependence Claim 2 [TD2]: If any particular person had not been conceived within a month of the time when he was in fact conceived, he would in fact never have existed.¹⁹

These claims are both associated with Parfit’s conviction that our identities are fixed by the ovum from which we developed, and perhaps by our genetic makeup. It is sufficient to note that the position defended in this paper does indeed reject Parfit’s time dependence claim in both of its versions. It is, however, a *qualified* rejection of this claim: I have argued that the concept “identity” is ambiguous, and that different conceptions apply in different contexts, appropriate for different uses and functions. I need not (and do not) claim, therefore, that the time dependence claim is *false*. My claim is rather that there is a significant conception of “identity” for which the claim is false, and that this conception is the one that is often relevant from the moral point of view. In point of fact, this conception is relevant in many contexts where the non-identity problem arises, because the stricter conceptions of “identity” do not pick out what is relevant from the moral point of view.

This is not to deny that there are many reasonable ways to use the concept “identity” for which TD and TD2 are true. It is thus the ambiguity of the concept that leads us wrongly to think that TD and TD2 are both (i) obviously and uncontroversially true, and (ii) that their truth creates a problem for morality (the non-identity

problem itself). I have argued that the conception of “identity” that is relevant from the moral point of view typically picks out (individuates) persons according to their morally salient qualities. These qualities may be shared among the members of a large class of possible and actual persons. When they are, it is quite appropriate for us to identify actual individuals with other persons in this class, even if some of these other persons have (or would have had) radically different properties.

Parfit then discusses “the descriptive view.” He offers two different versions of this view:

The Descriptive View [DV]: Each person has several distinctive necessary properties. These are this person’s most important distinctive properties, and they do not include having grown from a particular pair of cells.²⁰

The Descriptive Name View [DNV]: Each person’s name means “the person who. . .” For us now, “Kant” means “the person who wrote the *Critique of Pure Reason*, etc.” A particular person’s necessary properties are those that would be listed when we explain the meaning of the person’s name.²¹

Parfit rejects both of these views on the ground that they have implications that are “too implausible to be worth discussing.” For example, he writes of DNV, “I am the second of my mothers’ three children. This claim implies absurdly that, if my mother had conceived no child when she in fact conceived me, I would have been my younger sister.”²²

The view I have articulated in this paper is different from the descriptive view, but is obviously related. The view I have described here does not imply that Derek Parfit would in fact have been his younger sister. But it does imply that Parfit’s identity as “the second of his mother’s three children” may be a significant aspect of his identity in some contexts. This aspect of his identity, for example, would relevantly identify him as the appropriate beneficiary of a codicil of any will if his mother mentioned a “second child” as a beneficiary. But more strongly, it also implies that that it would be appropriate for Parfit to identify himself with that role in any context where this aspect of his identity is morally significant. Thus in the *Negligent Physician* case described above, it is appropriate for Beth to identify with the non-disabled child who might have been conceived, and appropriate to recognize her as having a grievance against the physician whose wrongful behavior resulted in her disability.

In his objections to DV and DNV, Parfit fails to consider the possibility that different conceptions of “identity” might appropriately apply in different contexts. On the view I have described, different properties might distinctively identify an individual depending on the conception of “identity” we are employing in different circumstances. And this is just what we do in ordinary circumstances: If we are trying to assign authorship credit for the *Critique of Pure Reason*, for example, the relevant property for individuating the person who should receive credit is the property of being the person who wrote the *Critique*. Analogously, when we’re considering who will be harmed by our *depletory policies* the relevant property is that of being a future person who suffers from a disadvantageous condition that was caused by our policy choice.

In articulating the various versions of TD and DV, Parfit never considers a view of the kind proposed here. In fact, in spite of the vigorous attack on standard conceptions of “identity” which Parfit pursues in Part III of *Reasons and Persons*, in Part IV of the book Parfit seems to revert to an understanding of “identity” that is strikingly similar to those he rightly takes himself to have overthrown in his earlier arguments. Parfit was right to argue in Part III of *Reasons and Persons* that there is no “deep further fact” about identity that makes us the individuals we are or accounts for our persistence in time. When we consider the identities of persons who do not yet exist at all, our tenuous natural language concept of “identity” is simply stretched beyond the breaking point, applied to a context quite different from the context in which this concept grew. It seems clear that the concept of “identity” is complex, vague and ambiguous. This should make it less surprising to find that our concept can’t simply be applied, without prior analysis, to new kinds of problems.

Does the place-holder view ask the impossible? The view recommended here does imply that Parfit’s mother should have identified her prospective child in broad terms that would not distinguish between alternative possible children she might have had. From that perspective and that time, she should not have distinguished between the girl she might have conceived, and Derek Parfit himself. Does this mean that Parfit would have been that child, if things had been different? In asking Derek Parfit to identify with the child who might have existed if things had been different, does the place-holder view ask something that is impossible or unreasonable?

I hope and believe that it does not. The truth of the claim “Parfit would have been that child” turns out to depend on the conception of identity we employ when we answer it. On some conceptions it is true, and on other conceptions it is false. The situation is exactly the same with other ambiguous concepts we might employ, once we specify the concept in a precise way that distinguishes among the ambiguous meanings. Is it impossible to identify oneself with the relevant alternative people “one might have been,” on this view? Again, I think not. To do so is simply to identify the features we possess in common with these relevant alternatives, and to recognize these features as the ones that are relevant from the moral point of view.

5.10 Return to Erewhon

In Butler’s Erewhon, nonexistent future people have identities in a conventional sense: they exist before conception and are connected, in relevant ways, to the people they will be when they are born. What is peculiar about the Erewhonian world is that the novel describes a world in which our standard concept of “identity,” the conception employed in most discussions of “identity over time,” really does apply to nonexistent future persons. In Erewhon, the non-identity problem might be thought not to arise because there are continuously existing Cartesian souls who can be made worse or better off, and who will exist (though perhaps with radically different properties) in all of the different alternative possible worlds our choices might determine or select. Some people believe that our world is something like

Erewhon, and that babies are born with souls that exist in advance to be reincarnated. If they are right about this, then the non-identity problem is a mistake, since the same souls are waiting in the wings. Different futures may contain different individuals, but these individuals would be the continuation of the same reincarnated souls.

But if they are wrong—that is, if we are not reincarnated beings nor Erewhonian souls—the identity problem is *still* a mistake. The non-identity problem wrongly invites us to apply a particular conception of “identity” in contexts where that conception is singularly inappropriate and misleading. We should politely decline.²³

Notes

1. Butler (1910).
2. I would like to take this opportunity to thank Derek Parfit for the many hours of pleasure his work has provided for me. Although the view defended here implies that we should not distinguish between the author of *Reasons and Persons* and the Parfits who might have existed if things had gone differently, it does permit us to be pleased that things turned out the way they did.
The discussion of the World of the Unborn is in Chapters 18–20 of Butler’s work.
3. Dawkins (2006).
4. De Shalit (1995).
5. Feinberg (1986).
6. Feinberg (1984), Mill (1980).
7. Broome (1997, 2004); Priest (2000, 2001).
8. I felt squeamish the first time I encountered the idea that truth might come as a matter of degree, but it passed after several years of thought about the problem. I have now completely recovered my composure and I’m quite comfortable with degrees of truth. But at least one former colleague of mine regards this comfort as evidence of a moral or intellectual failing on my part. For this reason, I will not rely on this view of truth or vagueness in the argument that follows.
9. Velleman (2008), p. 237.
10. Maclean (1983), p. 196.
11. See Reiman (2007) and Hare (2007).
12. Hare (2007).
13. Hare (2007), p. 523.
14. In Wolf (1993) I defended a similar solution for the non-identity problem.
15. Maclean (1983), p. 196.
16. The set of individuals to whom Gimel has this obligation may even include some non-actual persons who might have existed and might have been walking on the path below. We need not think of non-actual persons as Erewhonian souls to consider that they might be involved in a theory of obligation in this way.
17. Parfit (1982), pp. 351–353.
18. Parfit (1982), p. 351.
19. Parfit (1982), p. 252.
20. Parfit (1982), p. 353.
21. Parfit (1982), p. 353.
22. Parfit (1982), p. 354.
23. It is worth noticing that the view defended here raises problems of its own: in particular, there is a question about the way in which the place-holder criterion would apply in the case of different-number choices, where different numbers of people will exist depending on our present choice. I believe that there is a natural way to accommodate this problem, but cannot articulate it here.

References

- Broome, J. 2004. *Weighing goods*. New York: Oxford University Press.
- Broome, J. 1997. Is incommensurability vagueness? In *Incommensurability, incomparability, and practical reason*, ed. R. Chang. Cambridge: Harvard University Press.
- Butler, S. 1910. *Erewhon*. Gutenberg Project. <http://www.gutenberg.org/etext/1906>. Accessed January 2008.
- Dawkins, R. 2006. *The root of all evil*. UK Channel 4. Alan Clements, Producer.
- De Shalit, A. 1995. *Why posterity matters: Environmental policies and future generations*. New York: Routledge.
- Feinberg, J. 1984. *Harm to others*. Oxford: Oxford University Press.
- Feinberg, J. 1986. Wrongful life and the counterfactual element in harming. *Philosophy and Social Policy* 14(1): 145–178.
- Hare, C. 2007. Voices from another world: Must we respect the interests of people who do not, and will never, exist? *Ethics* 117: 498–523.
- MacLean, D. 1983. A moral requirement for energy policies. In *Energy and the future*, eds. D. MacLean and P. Brown, 180–197. Totowa, NJ: Rowman & Littlefield.
- Mill, J. S. 1980. *On liberty*. Oxford: Clarendon Press.
- Parfit, D. 1982. Future generations: Further problems. *Philosophy and Public Affairs* 11(2): 113–72.
- Parfit, D. 1984. *Reasons and persons*. Oxford: Clarendon Press.
- Priest, G. 2000. *Logic: A very short introduction*. New York: Oxford University Press.
- Priest, G. 2001. *An introduction to non-classical logic*. Cambridge: Cambridge University Press.
- Reiman, J. 2007. Being fair to future people: The non-identity problem in the original position. *Philosophy and Public Affairs* 35(1): 69–92.
- Velleman, D. 2008. Persons in prospect. *Philosophy and Public Affairs* 36(3): 221–44.
- Wolf, C. 1993. Justice between generations. Doctoral Dissertation, University of Arizona.

Chapter 6

Rule Consequentialism and Non-identity

Tim Mulgan

Abstract This paper explores the relationship between rule consequentialism and the non-identity problem. It argues that rule consequentialism accommodates person-affecting intuitions without abandoning Parfit's no difference view. The paper also offers a new model of rule consequentialism—reinterpreting its various features as a series of departures from an act consequentialist ideal each motivated by human finitude and fallibility.

Keywords Future generations · Rule consequentialism · Reproduction · Non-identity · Person-affecting · Parfit · Hooker.

6.1 Introduction

This paper explores a rule consequentialist solution to the non-identity problem. In doing so, I will develop some themes from my recent book *Future People*—and respond to emerging criticisms of that book and of rule consequentialism in general.

My principle aim in *Future People* is to construct a new consequentialist account of the morality of our decisions regarding future people—from individual reproductive choices to global public policy priorities. *Future People* offers the first systematic rule consequentialist account of reproductive ethics, and of the significance of reproductive freedom, and also a new foundation for a liberal theory of intergenerational and international justice.

The present paper has a more limited scope than *Future People*, and also a different emphasis. Its scope is limited in two ways—I focus exclusively on moral theory, and, within moral theory, exclusively on rule consequentialism. One of my subsidiary aims in *Future People* was to motivate a return to the utilitarian tradition in political philosophy, and I regard the discussions of political philosophy and public policy as one of the main features of my book. However, as commentators have focused on the moral side of my project, and as my explorations of political

T. Mulgan (✉)
University of St Andrews, Scotland, UK
e-mail: tpm6@st-andrews.ac.uk

philosophy all rest on a foundation of moral theory, I concentrate here on explaining and defending that foundation.

In *Future People*, I have taken as my primary example the familiar objection that consequentialism is implausible because it makes unreasonable demands. Indeed, one of my aims in writing *Future People* was to bring together the two distinct literatures on obligations to future people and on the demands of morality. In the present paper, however, I focus on the non-identity problem instead. This is partly to fit the theme of the present collection. But I also have a more principled rationale. Largely due to pressure from commentators, I have come to regard my emphasis on the demandingness objection in *Future People* as, at best, misleading. While I still think there are important links between these two problems facing consequentialism, there are also important differences. The most serious problems facing any consequentialist account of future people lie at the intersection between non-identity and demandingness.

In addition to these changes of subject matter, the present paper also seeks to advance beyond *Future People*, by presenting replies to two key objections. The shift from demandingness to non-identity is one such reply. The second is my defence of a contingent morality in Sections 6.6 and 6.7. In several places, my argument in *Future People*, like any exercise in consequentialist ethics, rests on controversial empirical claims. This makes my theory appear vulnerable. Whether we regard this contingency as an objection depends on our views regarding the relationship between moral theory and empirical fact. I shall argue that the most plausible account of that relationship vindicates my approach.

6.2 Two Decisive Intuitions

Contemporary moral theory often begins with moral intuitions—judgements about particular cases or general ideals. The non-identity problem itself is significant because it generates a clash between our moral intuitions and the deliverances of some familiar modes of ethical thinking. The same is true of other puzzles in this area, such as the repugnant conclusion, the mere addition paradox, and the infinite utility puzzle. Intergenerational ethics is especially intuition-based.

I find it helpful to distinguish two kinds of intuitions: *decisive intuitions* (that any acceptable moral theory must accommodate) and *distinguishing intuitions* (that mark distinctive features of different theories). My aim in *Future People* is to develop a theory that accommodates all decisive intuitions, and also makes sense of a range of intuitions that are distinctive of a moderately radical utilitarian outlook.

If we all always agreed in our considered moral judgements, then all our intuitions would be decisive. However, such agreement is not to be found. Sometimes intuitions serve, not to confirm or refute theories, but to distinguish them. There is no definite line between decisive and distinguishing intuitions. No intuition is uncontroversially decisive, if only because there is always a niche in the philosophical marketplace for the first person who rejects it. Partisans of particular moral

theories often present an intuition as decisive, when their opponents would see it as distinctive of that particular theory.

In *Future People*, I begin with two decisive intuitions—two judgements that any acceptable moral theory must respect. These provide test cases for moral theories. In the area of future generations, this test is not trivial, as many familiar moral theories have great difficulty accommodating one or other of these intuitions.

The basic wrongness intuition. It is wrong to gratuitously create a child whose life contains nothing but suffering.

The basic liberty intuition. There is no obligation to have children, nor an obligation not to.

I focus on two simple principles that each have difficulty with one of our basic intuitions. The two principles are as follows.

The simple person-affecting principle. An action can only be wrong if some particular person is worse off than that person would have been if some other action had been performed instead.

Simple consequentialism. The right action in any situation is whatever produces the most valuable state of affairs.

In *Future People*, I use the simple person-affecting principle to illustrate the problems facing non-consequentialist accounts of future morality. These problems owe their prominence to the work of Derek Parfit.¹ Parfit distinguishes two kinds of moral choice. A *same people choice* occurs whenever our actions affect what will happen to people in the future, but not which people will come to exist. If our actions do affect who will come to exist in the future, then we are making a *different people choice*.

Parfit also further distinguishes two kinds of different people choices: *same number* (where our choice affects who exists, but not how many people exist), and *different number* (where we decide how many people ever exist). This second distinction is relevant because simple consequentialism, which seems to cope well in same number choices, faces many difficulties when we turn to different number choices.

Parfit makes three central claims.

1. Different people choices occur very frequently, and in situations where we might not expect them.
2. It is often difficult to tell, in practice, whether we are dealing with a same people choice or a different people choice.
3. Many traditional moral theories cope much better with same people choices than with different people choices. Our moral theories are designed for same people choices, and thus need to be amended to apply to different people choices.

These three claims constitute the *non-identity problem*, so called because, in a different people choice, those who will exist in one possible outcome are not (numerically) identical to those who will exist in an alternative possible outcome.

I must stress that, for the purposes of this paper, I take the non-identity problem to be the *general problem* that arises because we need to adapt some people moral theories to different people choices—and to situations where we are uncertain what type of choice we face. The non-identity problem is not a specific case, a specific intuition, or a specific objection. Nor, as I shall argue, is it a problem only for one theory or class of theories. In particular, non-identity is not only a problem for non-consequentialists.

The non-identity problem is a significant threat to anyone who endorses the simple person-affecting principle, as the latter clearly violates the basic wrongness intuition. If different actions bring different people into existence, then, whatever action we choose, we cannot afterwards locate any particular person who is worse off than he or she would otherwise have been. If we cannot compare existence with non-existence, then we can make no sense of the claim that *x* is worse off than if *x* had never existed. It follows that no simple person-affecting theory can ever condemn any creation choice, however horrific the resulting life.

The major alternative to any person-affecting approach is consequentialism.

Simple consequentialism seems untroubled by the non-identity problem. It easily accommodates the basic wrongness intuition in both different people and same people choices, as it is always wrong to produce less happiness than you might.

Unfortunately, while it does respect the basic wrongness intuition, simple consequentialism clearly violates the basic liberty intuition, as it always obliges us to do *whatever* maximises the good, and thus leaves almost no room for *any* liberty.² In any situation, either agents will be obliged to have children (to produce more happy people), or they will be obliged not to have children (because their resources would do more good if devoted to charity). Neither of these obligations is intuitively plausible.³

By contrast, the simple person-affecting principle has no difficulty with the basic liberty intuition. This is hardly surprising, as the problem with this principle is that it grants potential reproducers too much liberty, not too little.

Our two simple principles are very crude and over-simplified. Both the person-affecting approach and consequentialism have their defenders, who attempt to accommodate (or explain away) the decisive intuition that is problematic for the simplified version. The non-identity problem began life as an attack on person-affecting views. However, I believe that the situation is now largely reversed, and it is consequentialists who are on the back foot. There are two reasons for this: person-affecting theories thrive, while consequentialism has yet to put its own house in order.

The person-affecting view has many defenders, including many of the contributors to this volume.⁴ They argue that such a view can respect the basic wrongness intuition, as we *can* reasonably regard a life not worth living as worse *for that person* than non-existence. Person-affecting theorists also seek to generate stronger obligations regarding future people. One common defence is as follows. The main underlying person-affecting intuition is that an action is only wrong if someone is *wronged*. But a person can be wronged even if it is not the case that they would otherwise have been worse-off. (The classic example is when a person is prevented

from boarding a plane because of his race, and the plane goes on to crash. This person has been wronged—even though he would otherwise have died.) Applying this lesson to the non-identity problem, we can still say that a person has been wronged by an act leading to their creation, even if their life is worth living *and* they would otherwise not have existed at all.

I shall proceed on the assumption that my two basic intuitions are decisive, *and* that some extant person-affecting theories do successfully accommodate them both. Faced with these non-consequentialist alternatives, consequentialists must show *both* that they can accommodate the two basic intuitions, *and* that their theory offers something that even sophisticated person-affecting theories cannot. Before explaining the resources and advantages of my moderate consequentialism, we must ask exactly how—and why—simple consequentialism fails.

6.3 How Simple Consequentialism Fails

Simple consequentialism gets many things right. In addition to respecting the basic wrongness intuition, it also respects several other common intuitions regarding future people. I will examine two examples: gratuitous sub-maximisation and the no difference view.

The gratuitously satisficing mother. Betty has decided to have a child. She could have one in summer or in winter. A child born in winter will not suffer any serious ailments or disabilities, but he or she will have a lower quality of life than a child born in summer. Betty herself is completely indifferent when she has her child. On a whim, Betty decides to have her child in winter.

As the resulting child has a very worthwhile life, it is hard to see how any person-affecting theory could fault Betty's choice. By contrast, simple consequentialism clearly implies that Betty ought to create the child with the better life.

This is a case of blatant moral satisficing, where an agent deliberately produces a sub-optimal outcome on the grounds that it is “good enough,” even though she could have produced a significantly better outcome at absolutely no cost to herself. The rationality and morality of satisficing behaviour have been much discussed. I and others have argued elsewhere that blatant satisficing is clearly unjustified in same people choices.⁵ Why should we permit it in different people choices? If other things are completely equal, what possible justification is there for such a blatant failure to produce a person with a better life?

This tale generates intuitions that are much harder to avoid for a person-affecting theory than the basic wrongness intuition. On the other hand, these new intuitions are much less forceful. Proponents of the person-affecting approach may simply deny that Betty's choice is wrong. Indeed, they can see its verdict in this case as yet another strike against consequentialism.

I agree that this thought experiment generates no decisive intuitions. However, it does bring out a cluster of intuitions that are problematic for the person-affecting approach. It is at least plausible to believe that there is good reason to opt to create

the more valuable life over the less valuable one; that one ought to do so if other things are equal; and that the source of these reasons lies in the fact that the former option leads to a more valuable outcome—even if that outcome is better for no-one. Not everyone shares these intuitions. For those who do, however, they provide one motivation for exploring alternatives to the person-affecting approach. (I suggest in Section 6.7 that the ultimate fate of these particular intuitions may rest on our ability to reconcile them with the asymmetry intuitions explored in the rest of this section.)

The no difference view is even more controversial. Consider a variant on our previous tale.

The two mothers. Suppose two women (Debbie and Sally) have each decided to have a child. Both must choose between having a child in summer or in winter, where the child born in winter will have a lower quality of life than the child born in summer. On a whim, both decide to have their children in winter. However, due to differences in their respective medical conditions, Debbie faces a different people choice while Sally is making a same people choice.

According to simple consequentialism, there can be no moral difference between these two cases. If Sally's action is wrong, then Debbie's action must be wrong to exactly the same degree. Simple consequentialism implies the following.

The no difference view. If A and B are two situations, and if the only difference between them is that A is a different people choice and B is a same people choice, then there is no moral difference whatsoever between A and B.

Even though they aim to respect the basic wrongness intuition, person-affecting views typically reject the no difference view. I cannot think of a genuinely person-affecting moral theorist who thinks there is *no* difference between same and different people choices.⁶ If we embrace the no difference view, then this is a strike in favour of simple consequentialism. However, the no difference view is not universally endorsed. Indeed, the literature contains two extreme responses to these cases. Some hold that there is no difference between the two cases, while others claim that, while Sally's choice may well be wrong, Debbie's cannot be.⁷ The first response is most naturally combined with a consequentialist theory, while the second is obviously suited to a person-affecting theory.

I believe there is something to be said for both extremes. My aim in *Future People* is to develop and defend a middle road: while there are good reasons for Debbie to opt for a summer birth, perhaps Sally has additional reasons.

The no difference view follows automatically from a more general feature of simple consequentialism.

Impersonalism. The rightness or wrongness of actions depends entirely upon the value produced, without any regard for how that value is distributed across the lives of human beings.

The impersonalism of consequentialism is also what makes the theory notoriously demanding in famine relief cases. Simple consequentialism requires agents to place their own interests on a par with the interests of others. It leaves no leeway for favouring myself, my nearest and dearest, or my own community. In *Future People*, I treat the failure of simple consequentialism to respect the basic liberty intuition

as a particular *instance* of this excessive demandingness. Melinda Roberts has questioned this diagnosis. She suggests that the real problem for simple consequentialism in regard to future people is not a demandingness problem but instead the aggregative calculation that simple consequentialism is often associated with.⁸ I agree that my previous focus on demandingness is, at least, misleading. However, I also think there is some connection between the two problems. The intuitive problems facing simple consequentialism in future morality result from the combined impact of its commitment to demandingness and to the no difference view; and these specific commitments are both instances of simple consequentialism's deeper commitment to impersonalism. The demands of consequentialism are *especially* counterintuitive in different people choices.

Simple consequentialism must endorse the no difference view. If we reject that view, then we must reject simple consequentialism. Why might we reject the no difference view? The main reason is that this view conflicts with a range of intuitive asymmetries, such as the following.

The basic asymmetry. There is no obligation to have children, even if they would be extremely happy. But there is an obligation not to knowingly create people whose lives are not worth living.

This strong asymmetry is a very basic feature of commonsense morality.⁹ Imagine a couple who deliberately create a severely disabled child whose life contains absolutely nothing but excruciating agony—simply to explore their own capacity for other-regarding behaviour. Almost no one would find such behaviour morally acceptable. Many people also believe that it is wrong to reproduce if one cannot ensure that one's child's basic needs will be provided for. Yet almost no one thinks that a decision not to reproduce is wrong—at least, not wrong to anything like the same extent.

The intuition behind the basic asymmetry is, in part, an anti-demandingness intuition. Simple consequentialism is wrong to insist that everyone must always promote the good by always creating happy people. But demandingness alone cannot explain the intuitive difference between the two cases. To see this, consider another contrasting pair of cases.

Asymmetric demands. Suppose Mary and Martha are two affluent people in the developed world. They each face a choice between spending their money on themselves and spending it in a way that maximises the good. Mary's alternative is to donate her money to a charity that assists (already existing) disadvantaged people. Martha's alternative is to create a new happy person. Suppose each alternative produces exactly the same total value. Mary and Martha both spend their money on themselves. Has either done anything wrong? And, have they each done something *equally* wrong?

Simple consequentialism must conclude, not only that Mary and Martha each do something wrong, but also that they are *exactly equally* in the wrong. Many people will reject both claims. In particular, there is a strong intuition that Mary's action is open to moral criticism in a way that Martha's is not. Failing to benefit existing people is morally objectionable in a way (and to a degree) that failing to create new happy people is not.

Intuitively, we believe that we are (at least sometimes) morally free to depart from impersonal maximisation, even when that departure involves failing to benefit an existing person. But we also think that our freedom to depart from impersonal maximisation is distinctly *greater* when that departure involves choosing to create no one at all rather than a very happy person; or choosing to create a happy enough person instead of a happier person; or choosing to give a benefit to an existing person rather than creating a new person.

A similar asymmetry applies to other moral distinctions. Common sense regards causing harm as worse than failing to benefit. It draws a *greater* distinction when the contrast is between creating a person whose life is not worth living and failing to create a person whose life is well worth living. The former is clearly forbidden, while the latter is not blameworthy at all.

These asymmetries relate to different number choices, rather than same *number* different *people* choices. So they differ from the examples standardly used in discussions of the no difference view. However, these new asymmetries clearly bring out the underlying problem for simple consequentialism—that it cannot take account of the identity of persons.

6.4 Why Simple Consequentialism Fails

These intuitive failings can be traced (in part) to the fact that (at least in the literature on future generations) simple consequentialism is usually combined with the following account of value.

The total view. The value of a state of affairs is entirely a function of the total well-being it contains, and is unrelated to the distribution of well-being across persons.

One obvious solution is thus, not to reject simple consequentialism itself, but rather to reject the total view. Many consequentialists take this route for independent reasons—largely driven by puzzles in value theory such as the repugnant conclusion, the mere addition paradox, or the infinite utility problem.

Others reject simple consequentialism, adopting a moderate moral theory. I take this second route. I claim that my view has several advantages—or, at least, several distinctive features—when compared to other moderate views. The first is that it is compatible with the total view, and will thus appeal to anyone who wants to retain that view (which has many virtues, and many able defenders¹⁰), but to combine it with a moderate account of moral obligation.

Other moderate moral theories are compatible with the total view. Most obviously, any theory where moral obligation is independent of the values of states of affairs is consistent with any account of those values. However, my approach is distinctive in retaining from simple consequentialism both the total view and the idea that morality is ultimately all about the promotion of objective value. This distinguishes my approach from those who achieve moderation only by severing or weakening the connection between value and obligation.

A second advantage is that rule consequentialism can be combined with many alternative value theories. Indeed, I argue in *Future People* that most departures from the total view canvassed in the literature would reinforce rule consequentialism in its departures from simple consequentialism.¹¹ My theory can thus also appeal to those who reject the total view.¹²

A third advantage is that rule consequentialism builds on a departure from simple consequentialism that is already required before we turn our attention to future people. Even though it is more threatening in different people choices, the demandingness objection also arises starkly in same people choices. Consequentialists thus cannot avoid demandingness merely by altering their theory of value—as all salient alternatives coincide in same people choices. They must abandon simple consequentialism. It is thus worth asking whether the solution we devise for same people choices can also do the (related) job in different people choices. In *The Demands of Consequentialism*, I argued that rule consequentialism offers the best solution to the demandingness objection in (most) same people choices. Therefore, in *Future People*, I apply rule consequentialism to our obligations to future people.

A fourth advantage is that, depending on the details, my rule consequentialism may also be able to accommodate some distinctive intuitions, such as the intuition that gratuitous sub-maximisation is wrong.

The final advantage of my approach is that it offers a new account of the relationship between empirical facts and moral rules. This new account enables rule consequentialism to offer a compelling consequentialist justification whenever it either endorses or rejects distinguishing intuitions. We return to this advantage in Sections 6.6 and 6.7.

Because I think consequentialists are on the back foot regarding non-identity, my primary aim is constructive rather than destructive. Instead of seeking to refute rival theories, I concentrate on showing how rule consequentialism respects our two decisive intuitions.

6.5 Rule Consequentialism

Future People defends a form of rule consequentialism. Acts are assessed indirectly, in terms of an ideal code of rules. I use the following general formulation, based on the recent work of Brad Hooker.¹³

An act is wrong if and only if it is forbidden by the code of rules whose internalisation by the overwhelming majority of everyone everywhere in each new generation has maximum expected value in terms of well-being.

Two features of rule consequentialism play key roles in *Future People*.

1. To assess the costs and benefits of internalising a code of rules, we do not imagine any centrally co-ordinated mass indoctrination. Instead, we assume that moral rules are taught in the normal way—by family, teachers, and the broader culture.

2. We assess the costs of teaching a moral code to a *new* generation. We do not ask what would happen if we tried to teach the new code to a generation of adults who had already internalised a different moral code. This gives rule consequentialism a potential for radical innovation.

Rule consequentialism has been subject to many objections, and much debate, in the recent literature. I address some objections elsewhere, and offer my own solutions.¹⁴ My present focus is purely on rule consequentialism's ability to cope with non-identity intuitions. Can rule consequentialism provide an alternative to moderate person-affecting views?

We begin with our two basic intuitions. A moral code allowing agents to gratuitously create miserable people would not maximise value. Rule consequentialism thus easily respects the basic wrongness intuition. (Rule consequentialism also seems able to accommodate a prohibition on gratuitous sub-maximisation—as a rule telling agents to produce happier people (instead of people who are less happy) will produce better consequences than a rule permitting the creation of less happy people. However, I suggest in Section 6.7 that the relationship between rule consequentialism and gratuitous sub-maximisation is more complex.)

The harder task is to show that rule consequentialism respects the basic liberty intuition. This task lies at the heart of *Future People* and is the focus of most objections. My present project is to broaden the scope of the discussion: to show how rule consequentialism both avoids all the pitfalls caused by the impersonalism of simple consequentialism and accommodates the various personalised asymmetries of common-sense intuition.

6.5.1 Differentiating Rule and Simple Consequentialism

The first step is to differentiate rule consequentialism from simple consequentialism. Among many other failings, simple consequentialism cannot respect the basic liberty intuition. It is thus an unacceptable theory. Rule consequentialism can only be an acceptable theory if it diverges from simple consequentialism. The ideal code of rules cannot be identical to the rule—“Always do whatever produces the best consequences.”

To avoid the collapse into simple consequentialism, rule consequentialists seek a middle ground between overly simplistic rules and infinitely complex ones. Many contemporary formulations of rule consequentialism are driven by the need to avoid the collapse into simple consequentialism. I borrow my reply from Hooker, who introduces the distinction between “following a rule” and “accepting a rule” largely for this purpose.¹⁵

The acceptance of a rule by a population has consequences over and above compliance with that rule. Some people might accept a rule even though they do not always comply with it, while others might comply perfectly with a rule they do not accept. For instance, many people accept, on some level, more demanding principles regarding donations to charity than they can bring themselves to fully comply with,

while social or legal sanctions often produce compliance without genuine acceptance. To accept a rule involves many things other than a disposition to comply with that rule, such as the disposition to encourage others to comply, dispositions to form favourable attitudes toward others who comply, dispositions to feel guilt or shame when one breaks the rule and to condemn and resent others' breaking it, etc.

In *Future People*, I defend a form of rule consequentialism that relies heavily on the requirement that rules be accepted and internalized. This theory incorporates a clear distinction between acceptance and compliance. If the form of rule consequentialism defended in *Future People* is a coherent moral theory, then it does not collapse into simple consequentialism. If such a heavy emphasis on internalisation can itself be justified, then rule consequentialism is a distinct theory. However, this places even more pressure on my use of internalization, which I defend in Section 6.6.

The differentiation from simple consequentialism is, of course, only the beginning. The crucial question is whether rule consequentialism can use the gap between the two theories to provide an intuitive response to the non-identity problem.

6.5.2 Rule Consequentialism and Reproductive Freedom

Except for one or two brief comments, Hooker himself does not apply his theory to future generations. Indeed, I could find no detailed rule consequentialist account of either individual reproduction or inter-generational justice. One main purpose of *Future People* was to construct such an account. I argue at length that rule consequentialism does support a wide range of commonsense individual freedoms, including reproductive freedom. A crucial starting point is Hooker's observation that the question to which rule consequentialism is the answer is not "what if everyone did that?" but rather "what if everyone felt free to do that?" Hooker himself explicitly, if very briefly, applies this distinction to the morality of reproduction.¹⁶

Suppose my nephew tells me he refuses to have children. If everyone refuses to have children, the human species will die out. This would be a disastrous consequence. But it is irrelevant to the morality of my nephew's decision. What is relevant is that everyone's feeling free not to have children will not lead to the extinction of the species. Plenty of people who do not feel obligated to have children nevertheless *want* to—and, if free to do so, will. Thus, there is no need for a moral obligation to have children. Neither is there any need for a general moral obligation to have heterosexual intercourse.

I begin by establishing a strong *prima facie* case for reproductive freedom. I borrow from J. S. Mill's classical utilitarian defence of liberty, market freedom, and democracy. Given the nature of human beings, things go better overall if people are free to make significant moral decisions for themselves. Arguments against reproductive freedom are then examined and found wanting. *Future People* draws on a range of empirical evidence to argue that reproductive freedom is not a threat to the survival and flourishing of humanity. This leads to my defence of personal liberty and democratic institutions. While not infallible, they promote human happiness and offer the best safeguards for human survival.

Any calculation of the likely results of teaching a code of rules to a new generation involves great uncertainty. This uncertainty may seem a weakness of rule consequentialism—and many philosophers have argued that it is.¹⁷ But, in *Future People*, I argue that uncertainty is really a strength of rule consequentialism. The rules regarding reproductive freedom I develop in Chapter 6 of *Future People* are very general and leave considerable room for judgement in their application. I argue that, given the uncertainty of their future circumstances, it is better to teach the next generation these flexible general rules than to teach them a specific code tailored to the particular dilemmas we expect them to face in the future.

6.5.3 Rule Consequentialism and Person-Affecting Elements

My rule consequentialism has an impersonal foundation—the total view. This distinguishes it from other moderate moral theories. Simple consequentialism, when combined with the same impersonal foundation, yields a morality whose *content* is fully impersonal. To avoid an impersonal content, my rule consequentialism must include a range of obligations to particular people in its moral code. The best code that can be taught to human beings will include obligations to keep promises and to help friends, along with a range of other commonsense moral rules, such as prohibitions on murder and theft. This fit with conventional morality is often presented as a major benefit of rule consequentialism.

Accordingly, while it rejects a person-affecting *foundation* for morality, rule consequentialism need not reject all person-affecting elements within morality. The ideal code may include person-affecting rules and attitudes. Indeed, in *Future People*, I argue that it does include them. Recall that we are asked to imagine a moral code taught in the normal way in the context of a small set of interpersonal relationships. Any moral code is thus learnt via (specific) person-affecting rules. It is then natural to carry these rules (and their accompanying attitudes and moral outlook) over into the rest of our moral lives—even into different people choices.

There is a tension between these person-affecting arguments and the impersonal foundation of rule consequentialism. And there are limits on the content of the ideal code. Any code will include a general disposition to be benevolent, as the benefits of such a disposition are obvious. And *no* code will include the simple person-affecting principle. Someone who has internalized the ideal code will not plant a bomb in a forest that will explode in two centuries—even if they know that, because the act of setting the bomb will alter the identity of all future people, no particular future person will be worse off as a result of this action. We ourselves have learnt a code that (in its application to different people choices) departs from the simple person-affecting principle and produces better results than any code incorporating that principle. If we have learnt a better code than any simple person-affecting code, then no such code can be the best code humans could be taught.

On the other hand, we haven't learnt a code that goes to the other extreme. We have not internalized the no difference view. Furthermore, in *Future People*,

I argue that we *could* not internalize that view.¹⁸ The no difference view requires full impartiality. Partiality—of any kind—is only possible if we attach significance to the numerical identity of persons. Yet humans cannot internalise a fully impartial code, as such a code would be impossibly demanding.

6.6 A Contingent Morality

Having outlined the basis of my rule consequentialist response to the non-identity problem, I now turn to one common objection. My reply to this objection will lead to further elaboration of rule consequentialism.

Several reviewers of *Future People* object to my extensive use of empirical claims in defending rule consequentialism.¹⁹ In particular, they argue that I over-use the device (borrowed from Hooker) of rejecting counterintuitive rules on the basis of controversial empirical claims about what could (or could not) be taught to a population of human beings. For instance, to establish that rule consequentialism will not require agents to sacrifice all their own interests for those of future people, I claim that any population of humans would regard such a rule as unreasonably demanding—and thus that it cannot be successfully taught. Internalization is thus central to both my strategy for differentiating rule consequentialism from simple consequentialism and my attempt to justify reproductive freedom.

I aim to show that the rule consequentialist reliance on internalization costs is not under-motivated, and that rule consequentialism is not inappropriately reliant on empirical accidents. I also argue that, far from being a weakness, my reliance on empirical facts points to another advantage of my account—its ability to offer a plausible unifying story of the role of both empirical information and philosophical debate in the moral life of human beings.

I must begin by conceding that rule consequentialism's intuitive appeal *is* entirely contingent. Even in regard to the most decisive intuitions, rule consequentialism only gives the right answers because of (contingent) empirical factors. The reason for this is simple. Rule consequentialism offers a series of reasons to depart from simple consequentialism. Each of these reasons is built, ultimately, on a claim that is contingent. Things could have been very different. If they had been different, then simple consequentialism would have been the best moral code. There are thus possible worlds where rule consequentialism collapses into simple consequentialism. As simple consequentialism violates decisive intuitions such as the basic liberty intuition, it follows that it is only contingently true that rule consequentialism respects decisive intuitions. And there may be alternative moral theories that do not rest on such contingencies. (Consider a libertarian morality, where the demands of morality depend only on the agent's own voluntarily assumed obligations.) If it counts against a moral theory that it answers moral questions with contingent facts, then rule consequentialism is at a significant comparative disadvantage.

Even rule consequentialism's respect for the basic wrongness intuition is contingent, as it only endorses that intuition because the consequences of teaching a code requiring agents to take account of the interests of future people (even in different

people choices) are better than the consequences of teaching any code permitting disregard of future people. This comparative claim may seem obviously true. But it is significant to note that, however obvious, it still rests on contingent empirical features of human beings—not merely on logical features of rules, or on impersonal values. We can imagine creatures so deeply ingrained with a lack of interest in future people—or so wedded to the simple person-affecting principle—that any attempt to teach them *any* obligations in different people choices would be counter-productive. Rule consequentialism only endorses the basic wrongness intuition because we are not such creatures. Should this “contingency” worry us? I suggest that it should not.

6.6.1 *Defending Internalisation*

This brings us to my defence of rule consequentialism. Morality is for *creatures like us*. The contingent facts I appeal to in *Future People* are deep facts that make us what we are. It is a strength of rule consequentialism—not a weakness—that its moral verdicts apply only to creatures like us and only in situations (broadly) similar to our own.

Most will agree that morality should appeal to some contingent facts. But my argument doesn't just appeal to some facts. Instead, it rests very heavily on one particular set of facts—those relating to the costs of teaching rules to human communities. Why are *those facts* so important to morality?

To provide a more solid defence than I offered in *Future People*, I now seek to explain why the focus on internalization in particular is a response to a plausible rule consequentialist story about the role of morality—and not an ad hoc device introduced merely to render rule consequentialism more user-friendly.

Rule consequentialism regards morality as a code of rules to enable a community of human beings to live together in a way that promotes human well-being and human flourishing. It is important to note that this *not* an evolutionary, descriptive, or semantic claim—but a *normative* claim. I am not saying any of the following: “This is why morality evolved,” “This is what ‘morality’ means,” “This is what morality (empirically) is.” Rather, *Future People* develops a rule consequentialist suggestion as to how we might usefully answer the question: “Why is morality important to us?”

Rule consequentialism's basic question is this: What would happen if a code of rules (R) were to become the moral code for a community of human beings—*by the standard natural process*? For (perhaps deceptive) ease of presentation, in *Future People* I usually put this question in first-person plural terms for the present generation. (What would happen if *we* tried to teach R to the next generation?) But the focus is meant to be on the *teachability of the code*, not on *our ability to teach*. This interpretation is a logical extension of our motivation for abandoning simple consequentialism in the first place. If we are sympathetic to rule consequentialism at all, then it makes little sense to ask what would happen if R became a moral code for human beings *as if by magic*. Why would anyone be interested in *that*

question? Either we interpret the utilitarian tradition at an abstract level or we seek to apply it to the situation of real human beings. The former route leads to simple consequentialism, the latter to a form of rule consequentialism that asks what would happen if rules were taught to humans in the usual way.

The costs of teaching a code reflect that code's degree of fit with human beings and their situation. This provides a useful measure of the code's suitability as a moral code *for humans*. Rule consequentialism is right to place weight on such facts, and to use them to differentiate itself from simple consequentialism.

6.6.2 *Freedom and Person-Affectingness Revisited*

Having sketched a general defence of internalisation, we turn now to reconsider the two key features of my response to the non-identity problem: freedom and person-affectingness.

I begin with the rule consequentialist defence of freedom from Chapter 6 of *Future People*. This argument is very clearly not a priori, as it explicitly cites empirical studies made prominent by the work of Amartya Sen.²⁰ My central claim in *Future People* is that, as a matter of fact, given the kinds of creatures human beings turn out to be, things go better overall (in terms of human well-being broadly construed) if people are left to make major life choices (especially reproductive choices) for themselves rather than having those choices made for them. Any such argument is, of course, heavily dependent on (empirical) claims as to *how* people will exercise this freedom. The argument for reproductive freedom only goes through if we can be reasonably confident that people will not respond to such freedom in a way that leads to underpopulation or overpopulation.

Freedom is morally appropriate *for us*. But we can easily imagine creatures for whom it is not. There are possible creatures in other possible worlds whose well-being is maximized by coercion rather than choice. Insofar as it says anything about those creatures (and there is, by the way, no reason why it *should* say anything), rule consequentialism must say that, for them, the appropriate moral code will sanction (and perhaps require) widespread coercion.

All rule consequentialist arguments for moral freedom—of any kind—share this contingency. Freedom has obvious costs from an impersonal consequentialist point of view—as it leads to sub-optimal decision-making in some circumstances. (If all agents are allowed to refrain from maximising the good, then some will so refrain.) Therefore, to be included in the rule consequentialist ideal code, any freedom must have compensating benefits. These benefits arise because of contingent features of our nature—including, perhaps, the fact that we are creatures for whom freedom is an independently valuable component of well-being.

I argued above that, in addition to supporting reproductive freedom, rule consequentialism avoids the no difference view. As ever, my argument was not a priori. From a consequentialist point of view, more impersonal rules would offer the best fit with the total view. A community of rational agents who perfectly follow a no difference code would produce better results than one following a person-affecting

code. To avoid the no difference view, rule consequentialism must show that this view cannot be effectively internalized.

The empirical case here is similar to that offered by rule consequentialism regarding demandingness—only stronger. The no difference view is extremely impersonal. No code for creatures (remotely) like us can be nearly so impersonal. Perhaps we can imagine perfect utilitarian calculating machines who would be best suited to a simple consequentialist code that accords absolutely no moral significance to the numerical identity of persons. But these imaginary agents are unlike us in very morally relevant ways.

Does all this contingency undermine rule consequentialism? I think not. The “contingency” underlying my defence of freedom in *Future People* is not coincidence or accident. It reflects general and important features of human nature and the human situation. Why shouldn’t the content of an ethic for humans depend on such features? Indeed, on what else could it depend? On pure reason? On disembodied rationality? Rule consequentialism takes its inspiration from J. S. Mill and other classical utilitarians. Human nature is something we discover empirically, not something we intuit a priori. If future empirical studies overturned our views about human nature, then we would (and, rule consequentialism argues, we *should*) amend our moral views.

6.7 Rule Consequentialism and Moral Philosophy

If we allow it to appeal to contingent facts, then rule consequentialism can respect all decisive intuitions—those that, even on reflection, we cannot imagine giving up. Rule consequentialism reinterprets decisive intuitions as those we cannot imagine fitting together with *any* moral code that could be effectively internalised by a community of human beings. We simply don’t think that human beings *could* live *well* like *that*.

We must note that the notion of “effective internalisation” is itself cashed out in consequentialist terms. Rule consequentialism doesn’t deny that an individual human being or a community might train themselves (or be trained by some outside agency) to believe in simple consequentialism, or to have no regard for distant future people. What it denies is that these rules could be part of a moral code that maximises human well-being over the long-term.

Recall our distinction between decisive intuitions and distinguishing ones. Distinguishing intuitions are the sites of controversy in moral philosophy. Rule consequentialism offers an account of that controversy. A distinguishing intuition is one where we are not sure if it fits with the best ideal code or not. And, says rule consequentialism, we decide whether to embrace a controversial intuition *by* asking how well it fits with such a code.

Many people find this last claim implausible. Surely the way we decide between controversial intuitions bares little resemblance to rule consequentialist inquiry? I now seek to dissolve this objection, by bringing moral philosophy itself within the rule consequentialist framework.

Rule consequentialists do not expect to discover the ideal code in its entirety in one go.²¹ Instead, we discover that a particular rule—or a rule of some general type—is in the code. Given the way that human beings happen to be, a code that permits favouring self and nearest and dearest, forbids murder, obliges promise-keeping, and promotes beneficence will produce better results than one that does not. We know the ideal code includes these elements—even if we cannot hope to describe that code in all its details.

Decisive intuitions provide constraints—fixed points in the moral psychology of someone who has internalised the ideal code. Once these parameters are set, we explore controversial intuitions by asking how someone who had internalised these (decisive) rules would (most naturally) respond to other situations.

At this point in its inquiry, rule consequentialism welcomes a wide variety of (often inconclusive) empirical evidence, which can enlighten us on the limits and flexibility of human moral codes. If human beings have been effectively taught a code with rule R, then we at least know that codes with rule R can be taught to human beings. Rule consequentialism thus offers a sound consequentialist argument for borrowing moral rules from other cultures, so long as those rules work better than our current commonsense morality.

Another source of evidence within rule consequentialism is the progress of moral philosophy itself. For instance, suppose philosopher P develops an intuitively plausible and coherent moral theory, which links decisive intuitions together using plausible moral ideals. P's achievement then itself constitutes *prima facie* evidence that such a code makes sense as a moral view of the world—and thus *could* be a moral code for human beings. If our worry about code R is whether R can be (efficiently) internalized, then a coherent account of an intuitively plausible version of R helps alleviate that worry.

Consider a concrete example. Should we extend our prohibition on gratuitous sub-maximisation from some people choices to different people choices? For rule consequentialism, this is the question whether a person who had internalised the general norms of the ideal code would find it natural to bring certain particular cases under her person-affecting dispositions or under her general disposition to promote human well-being—or (somehow) bring such cases under both dispositions. The discovery that a certain pattern of thought makes best sense of our own moral intuitions may help us to decide how this idealised agent would react.

Throughout this paper, I have equivocated as to whether or not rule consequentialism prohibits or endorses gratuitous sub-maximisation in different people choices. But this is because I am unsure which of these attitudes best fits both with the basic wrongness intuition and with our decisive intuitions about some people choices. Considerable further exploration is necessary before we can settle this question. I do claim, however, that both consequentialists and non-consequentialists should be able to agree that rule consequentialism focuses attention on the right question here. Whatever its outcome, the rule consequentialist process promises a justification for one or other distinguishing intuition.

Rule consequentialism can thus borrow from person-affecting moral theories, as these moral theories provide evidence of the internalizability of moral codes.

However, rule consequentialism does not thereby *become* a person-affecting *theory*, as its foundation remains resolutely (and impersonally) consequentialist. Instead, in this contested region between the end-points marked by decisive intuitions, rule consequentialism offers a new way to organize *all useful moral input*.

Suppose we discovered a community of human beings who were clearly flourishing better than ourselves and whose moral code differed from our own. (Perhaps their moral philosophers have taught them to conceptualise Parfit's puzzle cases in a way that we cannot yet imagine, enabling their community to flourish across the generations in a way that ours does not.) Rule consequentialism's central claim is that we would conclude that their moral code was superior—that they had stumbled upon a better way for human beings to live. We would then adopt their moral views in controversial cases—or at least attempt to move our own views in their direction. An intuition is decisive when we cannot imagine encountering such a community. It is distinguishing (or controversial) when we can. (It follows, of course, that the judgement that a certain intuition *is* decisive can only ever be provisional. The fact that we cannot imagine encountering a more flourishing community who lack that intuition, doesn't prove that we won't.)

Such encounters need not be the stuff of exotic anthropology or bizarre science fiction. Judged in consequentialist terms, our present moral code is superior to the codes of earlier generations in many ways. There is no reason to expect our own generation to mark the end of moral progress. We may reasonably hope that future people will have better moral beliefs than ourselves. Rule consequentialism offers an account of what this claim means. It also suggests that, if we can discover what those superior future beliefs might be, we should adopt them for ourselves.

Acknowledgements For helpful comments on earlier drafts of this paper, I am grateful to the editors of this volume. In this paper, I draw freely on Mulgan (2006). For permission to do so, I am grateful to Oxford University Press. I presented an earlier version of my ideas on rule consequentialism and contingency at an international conference on "Utilitarianism: An Ethics of Experience?" held at the University of Rome in June 2007. For the invitation to attend the conference, for their warm hospitality, and for comments on my paper, I am very grateful to Eugenio Lecaldano, Gianfranco Pellegrino, and Francesco Orsi. For comments on the paper, I am grateful also to Brad Hooker and John Skorupski. I am also grateful to Rahul Kumar, Francesco Orsi, Gianfranco Pellegrino, Melinda Roberts, and Rivka Weinberg for their published comments on Mulgan (2006), which stimulated me to write this paper.

Notes

1. Parfit (1984), pp. 351–441.
2. Simple consequentialism can give us some liberty in cases where two or more outcomes are tied for being "the best" in terms of aggregate wellbeing.
3. In *Future People*, I argue that, in the actual world, simple consequentialism is more likely to oblige affluent people in developed countries *not* to reproduce, as they could invariably do more good by giving their money away. Mulgan (2006), pp. 16–20.
4. For discussions of the person-affecting approach, see, for instance, Feinberg (1986); Heyd (1992); Kumar (2003); McMahan (1998); Roberts (1998, 2002, 2003); Temkin (1993); Woodward (1986).

5. For discussion and references, see Mulgan (2001), pp. 127–44.
6. Since the view Nils Holtug suggests in his contribution to this collection would not count as “person-affecting” in my sense of that term (identity, for him, isn’t critical to a person-affecting assessment of wrongdoing), he isn’t a counterexample to my claim.
7. Parfit defends the no difference view. See Parfit (1984), pp. 366–71. The opposite view is adopted by Heyd (1992) and is implicit in many defences of the person-affecting approach.
8. Roberts (2007), p. 775.
9. We should note that the asymmetry is not uncontroversial, as demonstrated by the papers in this volume by Persson and McMahan.
10. See, for instance, Broome (2004).
11. Mulgan (2006), pp. 142–46.
12. However, there are limits to the flexibility of my account. In particular, I do not see how *rule* consequentialism—which evaluates rules collectively—can be combined with a relativised or person-affecting value theory. Rule consequentialism is thus a rival for the views of Partha Dasgupta and Melinda Roberts. Dasgupta (1993, 1994); Roberts (1998, 2002, 2003).
13. The following exposition of rule consequentialism draws freely on Mulgan (2006), pp. 130–60, which in turn is based on Hooker (2000).
14. For discussion and references, see Mulgan (2001), pp. 53–103; and Mulgan (2006), pp. 130–60.
15. Hooker (2000), pp. 75–80; Mulgan (2006), pp. 138–40. The original “collapse” objection is due to Lyons (1965).
16. Hooker (2000), p. 177.
17. See, for instance, Griffin (1996), pp. 103–7. For further references and discussion, see Mulgan (2006), pp. 150–52 and 244–53.
18. Mulgan (2006), pp. 154–59.
19. See especially Kumar (2007); and Roberts (2007). See also Orsi (2007); Pellegrino (2007); and Weinberg (2006). I reply to some of their concerns in Mulgan (2007a) and (2007b).
20. See, especially, Sen (1999), pp. 204–26.
21. Mulgan (2006), pp. 150–52.

References

- Broome, J. 2004. *Weighing lives*. Oxford: Clarendon Press.
- Dasgupta, P. 1994. Savings and fertility: Ethical issues. *Philosophy and Public Affairs* 23: 99–127.
- Dasgupta, P. 1993. *An inquiry into well-being and destitution*. Oxford: Clarendon Press.
- Feinberg, J. 1986. Wrongful life and the counterfactual element in harming. *Social Policy and Philosophy* 4: 145–178.
- Griffin, J. 1996. *Value judgement*. Oxford: Clarendon Press.
- Heyd, D. 1992. *Genethics: Moral issues in the creation of people*. University of California Press.
- Hooker, B. 2000. *Ideal code, real world: A rule-consequentialist theory of morality*. Oxford: Clarendon Press.
- Kumar, R. 2003. Who can be wronged? *Philosophy and Public Affairs* 31: 99–118.
- Kumar, R. 2007. Review of T. Mulgan, *Future people*. *Philosophical Quarterly* 57: 679–685.
- Lyons, D. 1965. *The forms and limits of utilitarianism*. Oxford: Clarendon Press.
- McMahan, J. 1998. Wrongful life: paradoxes in the morality of causing people to exist. In *Rational commitment and social justice: Essays for Gregory Kavka*, eds. J. Coleman and C. Morris. Cambridge: Cambridge University Press.
- Mulgan, T. 2001. *The demands of consequentialism*. Oxford: Clarendon Press.
- Mulgan, T. 2006. *Future people*. Oxford: Clarendon Press.
- Mulgan, T. 2007a. *Future people: una presentazione* [Precis of *Future People*]. *Filosofia e Questioni Pubbliche* 12: 135–140.
- Mulgan, T. 2007b. Riposti ai commenti [Replies to Commentators]. *Filosofia e Questioni Pubbliche*. 12: 163–185.

- Orsi, F. 2007. Dividere la moralità e connettere ragioni e valori. *Filosofia e Questioni Pubbliche* 12: 141–150.
- Parfit, D. 1984. *Reasons and Persons*. Oxford: Clarendon Press.
- Pellegrino, G. 2007. Libertà riproduttiva e obblighi generali nei confronti delle generazioni future. *Filosofia e Questioni Pubbliche* 12: 151–163.
- Roberts, M. 1998. *Child versus childmaker: Future persons and present duties in ethics and the law*. Lanham, Maryland: Rowman & Littlefield.
- Roberts, M. 2002. A new way of doing the best we can: Person-based consequentialism and the equality problem. *Ethics* 112: 315–350.
- Roberts, M. 2003. Is the person-affecting intuition paradoxical? *Theory and Decision* 55: 1–44.
- Roberts, M. 2007. Review of T. Mulgan, *Future People*. *Mind* 116: 770–775.
- Sen, A. 1999. *Development as freedom*. Oxford: Clarendon Press.
- Temkin, L. 1993. *Inequality*. Oxford: Clarendon Press.
- Weinberg, R. 2006. Review of T. Mulgan, *Future people*, *Notre Dame Philosophical Reviews* 2 (ndpr.nd.edu).
- Woodward, J. 1986. The non-identity problem. *Ethics* 96: 804–831.

Part IV

**Is the Argument to “No Harm Done”
Correct? Must an Act that *Harms*
a Person Make that Person Worse Off?**

Chapter 7

Harming as Causing Harm

Elizabeth Harman

Abstract This paper argues that non-identity actions are wrong because they *cause* harm to people. While non-identity actions also typically benefit people, failure to act would similarly benefit someone, so considerations of benefit are ineligible to justify the harm. However, in some non-identity cases, failure to act would not benefit anyone: cases where one is choosing whether to procreate at all. These are the *hard* non-identity cases. Not all “different-number” cases are hard. In some cases, we don’t know whether acting would result in more or fewer people; this paper argues that this *epistemological* factor makes acting in these cases wrong.

Keywords Creation · Harm · Benefit · Justification · Threshold · Procreation · Nonidentity.

7.1 Introduction

In this paper, I will offer a solution to the non-identity problem and defend that solution. The non-identity problem arises because some actions appear to be wrong, and they appear to be wrong in virtue of harming certain people, but those people would not have existed if the actions had not been performed, and those people have lives that are worth living. Such actions are puzzling because they do not make these people worse off than they otherwise would have been; but plausibly, one harms someone only if one makes her worse off. A solution to the non-identity problem would both explain why the actions are wrong and vindicate the appearance that the actions are wrong in virtue of harming the relevant people.¹

Consider the following case. A woman has a temporary condition such that if she conceives now, her child will be blind. However, if she waits for three months, she can conceive a child who will not be blind. Clearly, this woman should not

E. Harman (✉)
Princeton University, Princeton, NJ, USA
e-mail: eharman@princeton.edu

conceive now. It would be wrong to conceive now. Furthermore, it seems that if she does conceive now she is *harming* her future child, and *that* is why her action is wrong. However, it can be argued that she does not harm her child and thus does not act wrongly. Plausibly, one harms someone only if one makes him worse off than he would otherwise have been. But this woman's act of conceiving now does not make her child worse off than he would otherwise have been. Had she not performed this action, he would not exist. Furthermore, he has a life worth living. Either he is better off than he would otherwise have been (because existence can be compared to non-existence, and it is better to have a life worth living than not to exist) or he is not worse off because he is neither worse off nor better off (because existence cannot be compared to non-existence). Either way, he is not worse off than he would have been had she not performed this action. Thus, it seems that she does not harm him. Furthermore, it seems that the only candidate explanation of her action's wrongness is that it harms; so it seems her action is not wrong.

Another case involves the disposal of nuclear waste. We are trying to decide whether to enact a lax nuclear waste disposal policy. The differences between our current policy and the lax policy are considerable: they involve which disposal plants are built and which jobs are available, and they require different things of ordinary people in their daily lives (such as recycling policies). If we do enact the lax policy, this will be more convenient for us; but we know that several hundred years from now, many people will become sick from the poorly contained waste: many people will suffer from cancer and die young. However, we are also able to predict that *who* exists in the future will differ depending on whether we enact the policy: who has which jobs, who lives where, who meets whom—all of these things will differ in large and small ways and ultimately affect who has children together, or at least when people conceive. Enacting the policy seems to be wrong, and it seems to be wrong because it harms future people, but it does not make them worse off than they otherwise would have been.

The above two cases are cases about which there is widespread intuitive agreement. Pre-theoretically, the actions seem clearly to be wrong. Solving the puzzle would involve vindicating our intuitive beliefs that the actions are wrong, as well as our intuitive beliefs about *why* they are wrong.²

There is another kind of case that raises the non-identity problem, but about which there is not widespread agreement. Indeed, there is real intuitive disagreement about these cases. One example is the following case. Suppose that a woman has a permanent condition such that if she ever conceives a child, that child will be deaf. Is it morally permissible for her to conceive? Pre-theoretically, people disagree. Another example is the following. A man and a woman are married, in love, and want to have children together; but they are slaves in the American South, in 1800. If they conceive, their child will be a slave too. Is it morally permissible for them to conceive? Again, pre-theoretically, people disagree.

A solution to the non-identity problem would be better if it could explain why for some cases, intuitions are uniform, while for other cases—those I will call “hard cases”—there is disagreement. I will develop a solution that can explain this.

7.2 Harming as Causing Harm

The view I will offer, which provides a solution to the non-identity problem, begins with three claims about harm.³ The first is a sufficient condition on harming:

1. One harms someone if one causes him pain, mental or physical discomfort, disease, deformity, disability, or death.

(The slogan version of claim (1) is: harming is causing harm. That is, an action is a harming action if it causes an effect of harm.) More generally, the view is that an action harms someone if it causes the person to be in a bad state. Bad states are understood as states that are in themselves bad, not bad because they are worse than the state the person would otherwise have been in. (If one wants a further account of a bad state, I am willing to offer one: bad states are those states that are worse in some way than the normal healthy state for a member of one's species.⁴)

The most powerful intuitive claim in favor of the idea that one cannot harm someone unless one makes him worse off is the following: surely a doctor does not harm me when he gives me a life-saving operation. Nevertheless he causes me pain and he physically injures me, as well as doing whatever saves my life. On my view, the doctor does harm me. However, the doctor *permissibly* harms me, because the doctor prevents greater harm to me than he causes. Furthermore, seeing that the doctor harms me enables us to see why surgery is a last resort and is not to be undertaken lightly—we must be quite sure the harm is justified before proceeding.

The import of the claim that someone harms someone else is given by this claim:

2. An action that harms someone thereby has a strong moral reason against it.

Furthermore, while the prevention of worse harm to someone can justify harming her, the mere fact that a harming action also *benefits*—that is, provides positive good things to the harmed person—does not justify the harm. On my view:

3. The mere fact that a harming action also benefits the person harmed, and benefits her more than it harms her, is insufficient to justify the harm.⁵

Besides these three claims about harm, my view also makes a fourth claim about one kind of harm: harm to people who do not exist independently of the harming action. Some actions harm people who exist regardless of whether the action is performed. Other actions (including the actions in non-identity cases) harm people who would not have existed if those actions had not been performed.⁶ The fourth claim elaborates a particular way that reasons against harm are *strong* and so harming is not easily justified.

4. If an action harms someone who does not independently exist, then the fact that the action also benefits the person harmed, and benefits him more than it harms him, is *ineligible* to justify the harm if *failing to perform the action* would similarly benefit someone.

The intuitive thought behind claim (4) is the following. The fact that an action harms someone provides a strong moral reason against acting; it is a reason that tells in favor of *refraining* as opposed to *performing the action*. If the action also benefits the harmed person, but performing the action is not the only way to provide

such benefits—indeed, refraining from performing the action would provide similar benefits to someone—then considerations of benefit simply do not tell *in favor of acting* as opposed to *refraining from acting*. On my view, considerations of benefit are therefore “ineligible” to justify the harm in that they never justify the harm in such cases. (In Section 7.4, I discuss why claim (4) is restricted to harms to people who do not independently exist.)

I assume that counterfactual dependence is sufficient for causation: I assume that if an action is such that if it had not been performed, a particular event would not have occurred, then the action causes the event. While counterfactual dependence is clearly not necessary for causation, as cases of preemption illustrate, it is commonly taken to be sufficient for causation.

7.3 Solving the Non-identity Problem

A “non-identity case” is a case in which an action affects whether some people exist, and it appears (either to everyone, or to some people) that the action is wrong in virtue of harming these people, although they have lives that are worth living. I will explain how my view solves the non-identity problem by discussing four different types of non-identity cases:

- (i) same-number cases: cases in which the same number of people will exist in the future regardless of which way the agent acts
- (ii) different-number cases, more people if the action is not performed: cases in which different numbers of people will exist depending on whether the agent performs the action, and the agent knows that more people will exist if she does not perform the action
- (iii) different-number cases, unknown which scenario has more people: cases in which different numbers of people will exist depending on whether the agent performs the action, and the agent does not know whether more people will exist if she performs the action (and does not know whether more people will exist if she does not)
- (iv) different-number cases, more people if the action is performed: cases in which different numbers of people will exist depending on whether the agent performs the action, and the agent knows that more people will exist if she performs the action

I will begin with same-number cases. One example of a same-number case is the temporary condition case I described at the beginning of this paper. A woman is deciding whether to conceive now, while she has a temporary condition that would cause her child to be blind. She knows that if she does not perform this action, she will conceive in a few months, having a child who will not be blind. Another example is a case in which a fourteen-year-old girl is deciding whether to conceive now. If she conceives now, her child will face many difficulties because she is not mature enough to be a responsible mother. If she does not conceive now, she knows that she will have a child later, when she is more mature. (I stipulate that the women in these two cases know they will conceive later if they do not conceive now.)

Each of these procreative actions causes the created child to be in a bad state—the woman in the temporary condition causes her child to be blind; the teenager causes her child to suffer the various hardships involved in having an immature mother. That each action *causes* these effects follows simply from the fact that had the action not been performed, the effects would not have occurred. Because each action harms the created child, there is a strong reason against each action. Now we must ask whether the harm is justified.

While some people deny that there is any harm in non-identity cases, claiming that an action cannot harm someone if it does not make him worse off, others grant that non-identity actions harm but claim that non-identity actions do not “all things considered” harm the created persons, because they also more greatly benefit the persons than they harm them.

On my view, non-identity actions both harm and benefit the created persons. They do benefit the created persons *more* than they harm them, in that the created persons’ lives are overall worth living: the benefits compensate for the harms in this sense. However, it is a further question whether the fact that the action benefits can *justify* the action’s harming. In same-number cases, there are strong reasons against the actions in virtue of the fact that the actions harm the created persons. Can the fact that the actions also *benefit* these people justify the harming? On my view, this fact does not function simply to *cancel* the reasons against harm, because on my view the *mere fact* that a harming action also benefits the person harmed, and benefits him more than it harms him, is insufficient to justify the harm. Can this fact nevertheless justify the harming? It cannot, because it does not function as a reason to perform the action *rather than* to refrain from performing it. The reason against harming is a reason to refrain from acting: if she refrains, the agent won’t harm anyone in the way she will if she acts. However, the action’s benefiting is not like this. If the agent waits, she will have a different child later. Thus, whether she performs the action or refrains from performing it, she will benefit *someone* in the same way. So the fact that the action benefits provides no reason to act rather than refrain from acting, and on my view it does *no work at all* to justify the action’s harming.

I turn now to different-number cases in which more people will exist if the action is not performed than if it is performed. I do not think cases like this are normally discussed in connection with the non-identity problem—indeed, they are somewhat artificial—but I will discuss them because they are relevant to the next category of case. Here is an example of such a case. A woman has a temporary condition such that if she conceives now, she will have a child who will be deaf and she will also become infertile after this pregnancy; that is, the temporary condition will result in both deafness of the child she conceives now and future infertility if she conceives now. However, if she waits, she will go on to have two children. (She and her partner have a slight but considered preference for two children rather than one, and they have every reason to believe that they will be able to have two if they wait for the temporary condition to clear up.)

In this case, as for the prior type of case, there is a strong reason against the action because the action *harms* the child by causing his deafness. The action also *benefits* the child by causing all the good aspects of the child’s life. Can the fact that the action benefits justify the action’s harming? It cannot, because, as before,

the benefits do not provide a reason to conceive now *rather than not*; indeed, in this case, the agent will benefit *more* people if she does not perform the action.

Thus, as for same-number cases, in this type of different number case, there is a strong reason against performing the action *rather than not performing it*, which is that it harms in a particular way, and there is no reason to perform it *rather than not* available to justify the harming. Thus, both types of actions are wrong in virtue of harming.

I turn now to different-number cases in which the agents do not know whether more people will exist if the action is not performed than if it is performed (and do not know whether fewer people will exist if the action is not performed than if it is performed). Cases of this type are commonly discussed in the literature on the non-identity problem, and while it is often noted that these are different-number cases, it is not generally noted that they are cases in which the agents *do not know* which outcome would have more people. I think this fact is very important, as will be clear in my discussion of these cases. An example of a case of this type is the nuclear waste disposal case. In this case, the agent is able to see that it's very likely that *who* exists in the future will be affected by whether the lax nuclear waste policy is put in place. By the same mechanism, it's very likely that *how many people* exist in the future will be affected, but it is not at all clear which possibility would end up having more people in it. In this case, the agent should think it's overwhelmingly likely that how many people will exist in the future depends on whether the action is performed: she should have a very low degree of belief that this is a same number case. (While it's possible that it's a same number case, it would be a surprising coincidence if it were.) But she has no particular reason to favor the belief that *more* people will exist if she performs the action over the belief that *fewer* people will exist if she performs the action, and vice versa. So she should have a degree of belief of roughly 1/2 that this is a different-number case in which *fewer* people will exist if the action is performed than would exist if it is not performed. And she should have a degree of belief of roughly 1/2 that this is a different-number case in which *more* people will exist if the action is performed than if it is not performed.

In general, the following claim is true: if (a) an agent should have a degree of belief of roughly 1/2 that her situation is a particular way, and (b) if her situation were that way, then her action would be wrong, then (c) she should not act: she should do the morally cautious thing and refrain from acting.⁷ This claim applies to the case we have been discussing. In this case, the agent should have a degree of belief of roughly 1/2 that her action would be a non-identity action such that *fewer people* would exist if it is performed than if it is not performed. Such actions are wrong (for the reasons I gave above), so she should not act. To summarize: in different-number cases in which the agent does not know whether more people would exist if she performed her action (and does not know whether fewer people would exist if she performed her action), the actions are wrong *because* given what the agent knows about her situation, it is all too likely that her action would be wrong, and thus she should refrain from acting.⁸

Finally, I turn to cases in which the agent knows that different numbers of people will exist depending on whether she performs the action, and she knows that more

people will exist if she performs the action. Cases of this type are typically discussed in the literature on the non-identity problem; however, these cases are the *hard cases* I mentioned in the introduction. These are cases about which there is widespread pre-theoretical disagreement. For example, some people think it is permissible to procreate if one has a permanent condition such that *whenever* one procreates, one will have a deaf child; some people think this is impermissible. Some people think it is permissible to procreate if one is a slave whose child will be a slave; others think it is not.

In each of these cases, there is a strong reason against the action because it harms the created child by causing all the bad effects in the child's life—and in particular by causing the unusual bad effects (deafness in one case, the hardships of slavery in the other). Each action also substantially benefits the created child, by causing all the good things in his/her life. In these cases, it is not true that the agent will similarly benefit someone else if she refrains from acting, so the benefits do tell in favor of acting *rather than refraining from acting*. Thus, on the view I have developed so far, it is open whether the benefits justify the harming.

Note that cases of this type are very similar to cases of ordinary procreation. When deciding whether to procreate at all, one is deciding whether to perform an action which will affect *how many* people exist in the future and which will *harm* the created person by causing the bad aspects of his life. But cases of ordinary procreation are not “non-identity cases” as I have defined them, because they are not cases of actions that seem to many people to be intuitively wrong. Rather, ordinary procreation is intuitively permissible. Cases of ordinary procreation are not *hard cases*.

I will now flesh out my view and show that it explains why ordinary procreation is permissible, but why procreating in the “hard cases” *may not* be permissible. On my view, an action that harms someone thereby has a strong reason against it. The mere fact that the action also benefits the people it harms—and benefits them more than it harms them—is insufficient to justify the action. Indeed, the fact that the action benefits is only eligible as a *possible justification* if the agent would not provide similar benefits to someone if he failed to perform the action—that is, if consideration of benefits tells in favor of acting *as opposed to refraining*. Furthermore, it is crucial that the benefits be benefits *to* the persons who are harmed by the action. One cannot in general justify severely harming one person by the fact that one will also provide positive benefits to other people.⁹ That the action benefits the very people whom it harms (and benefits them more than it harms them) is a *necessary* condition of the action's being justified.

If this eligibility condition is met (refraining does not benefit in similar ways) and the necessary condition is met (the benefits are to the person who is harmed), then an action that harms someone *may* be justified. Whether it is justified is a matter of how serious the harm is, how substantial the benefits are, and perhaps how strong the other reasons for the action are (if there are any). My view is thus that whether harming is justified is a matter of threshold crossing. There is more than one threshold. Some harms are so bad that they are impossible to justify; so the harm must be below a certain threshold. And depending on how bad the harm is, there

will be a corresponding threshold that the benefits must cross to justify the harming. (On a variant of my view, other reasons for action may be relevant as well. The fact that a woman has a permanent condition such that whenever she conceives, she will have a deaf child, means that her only chance to have a child involves conceiving in these circumstances; there is a considerable sacrifice for her if she refrains. I do not believe that this consideration can play a role in justifying the harm of procreation, though there is something natural to thinking it can; I discuss this more below.)

On my view, while ordinary procreation involves harming the created person, the kind of harming involved is not so substantial that it cannot be justified. The benefits involved in ordinary procreation are so substantial that they do justify harming the person in this way. (It doesn't seem that *other reasons* to procreate play a role in making ordinary procreation permissible; it would be odd, for example, to say that having one's first child is more justified than later children because whether one has any children greatly affects one's life but whether one has more children doesn't greatly affect whether one has a good life—though of course it affects what one's life is like and the way it is good.)

When we think about ordinary procreation, it is natural to have thoughts along the lines I am expressing. It is natural, in imagining conceiving, to worry about all the hardships and bad things one's child's life will involve. It is natural—and on my view, correct—to think that these things will be compensated for by all the good things in one's child's life.¹⁰

Furthermore, on my view, there are some harms such that *nothing* could justify causing them. Suppose a woman has a permanent condition such that if she ever conceived a child, her child would have a disease that would at some point cause a year of total agony, akin to being tortured, though the child would otherwise have a good life. It is not permissible to conceive in this case, even though the child overall might well have a life worth living. Some things are so awful that one should simply refrain from causing them, and a year of total agony is among those things.

But many people think that being deaf is not so bad. It does not involve pain, and society makes a number of accommodations that make the lives of the deaf easier than they would otherwise be. Thus, to some people, causing deafness may seem to be a kind of harming that can be justified.

On my view, there is a threshold somewhere: some harms are so bad that nothing can justify them, such as the harm of living a year in total agony; other harms are not so bad that nothing can justify them, such as the harms in an ordinary life. My proposal to explain why the hard cases are hard cases is that people implicitly place the threshold between these cases in different places. Some people implicitly see deafness as not so bad, and as a kind of harm that can be justified. Others implicitly see deafness as quite bad and not as a kind of harm that can be justified. If there really is a threshold, it makes sense that there would be implicit disagreement as to exactly where it lies.¹¹

On a variant of my view, *other reasons* also play a role in justifying procreation. (On my view, they do not.) Supposing that other reasons *do* play a role in justifying procreation, it is nevertheless implausible that *certain kinds* of other reasons can justify harming. Consider a woman with a permanent condition such that if she ever

conceives, her child will suffer from a chronic disease that involves some pain but otherwise allows a full life. It is natural to think that there is a lot at stake for this woman: if she does not ever conceive, then she does not get to have a child. This seems to provide a significant reason in favor of conceiving. However, note that this woman might live in a society like the United States where adoption is possible. Thus, what's at stake for the woman is not raising a child, but simply conceiving and gestating her own child. Furthermore, if the permanent condition is genetic, then the woman could have someone else's egg donated and then she could gestate a child without passing on the disease. Supposing that either option is available to her, *what the woman gives up* in failing to conceive (with her own egg) is not *being a parent*, and not even *being a gestational mother*, but simply *being a genetic parent* to her child. I will not deny that there is value in being a genetic parent to one's child; however, I am skeptical that there is sufficient value in this relationship to warrant harming someone to achieve it. And I think similar considerations cast doubt on the significance of being a gestational parent: it seems insufficiently valuable to play a role in justifying harming someone.

A more compelling case involves a woman for whom adoption and egg donation are not available and who has a permanent condition that would cause any child she conceives to have a disease. What is at stake for this woman—being a parent at all—is a big deal. Furthermore, when we think of the question whether slaves in 1800 may permissibly procreate, we certainly imagine the case to be like this: their only chance of having children at all would be to have children that would become slaves. This makes it more likely—but by no means clear—that conceiving is permissible. Add in the further fact that birth control would not be available to them, and they would be giving up even more—that is, having sex at all—if they chose to avoid procreation.

Nevertheless, if any of these other reasons do justify procreation, then it's permissible for the parents to harm another person, the created child, in order to avoid some hardships for themselves. It's not in general permissible to harm others in order to avoid hardships for oneself—when the others would not be in any way guilty or involved in one's experiencing those hardships. So I deny that other reasons ever play a role in *justifying* procreation. (It is much more plausible that other reasons play a role in *excusing* procreation: that they explain why one wrongful act of procreation is less bad than another wrongful act of procreation, or why one agent is less blameworthy than another.¹²)

7.4 Harming and Benefiting Those Who Independently Exist

I will now turn to discussing some objections to the view I have offered. The first objection is that my view treats knowingly conceiving a deaf child as much worse than it really is. The objector points out that on my view, if one knowingly conceives a child who will be deaf, then one harms the child by causing his deafness. However, the objector points out, what one does is *less bad* than an action which *deafens* an

existing child. Yet on my view, both actions harm the child by causing the child to be deaf. It appears that on my view, both actions harm in the same way, and both are thereby equally bad.

This objection brings out something important about the nature of harming in non-identity cases, which is often overlooked. What is brought out is that harming in non-identity cases is in one important respect *less bad* than harming of people who already exist—and in general, harming of people who exist independently of the action. When one harms someone who independently exists, there are two strikes against the action:

- one causes bad things for the harmed person
- one prevents the harmed person from an existence without these bad things

These two things are so closely connected in typical cases of harming that they are not separated, but they should be separated. Non-identity cases involve only the first aspect: one harms by causing the bad things for the harmed person, but one does not prevent the person from an existence without these bad things. I think this is a way in which harming someone who independently exists is *worse* than harming someone in a non-identity case: the harm is worse for the independently-existing person, and the action is thereby worse.¹³

These considerations explain some important facts. Many people think that it is wrong to procreate if one has a permanent condition that will result in any child one conceives being deaf. Nevertheless, these people will acknowledge that someone who deafens an existing child is doing something much worse *to the child* than someone who conceives knowing the conceived child will be deaf. On my view, the first action is worse because it harms in two ways: by causing the deafness *and* by depriving the child of being hearing.

Thus, my response to the objector is that while two actions may both harm in the same way, in that they both cause deafness, one action may be worse because it also harms in another way: it deprives the child of an existence in which he can hear.

The discussion I have offered of harms to existing persons sets me starkly opposed to Derek Parfit's view of the two medical programs case.¹⁴ In this case, we are considering two different medical programs which would each reduce the number of cases of a certain disability by 1,000. One program would work by pre-conception screening of women for a temporary condition that would cause the disability in any child they conceived. The other program would work by screening pregnant women for another condition and treating them to prevent the disability from arising in their fetuses. Parfit claims that intuitively it is irrelevant which screening program we use—it is irrelevant that the latter program affects already-existing beings, while the former program affects *who* exists. On my view, that is wrong. If we use the pregnancy-screening program, we are doing more good than if we use the pre-conception screening program. The pregnancy-screening program prevents the disability in independently existing people, while the pre-conception screening prevents the existence of people who would have the disability. Thus, on my view, the pregnancy-screening program prevents worse harm.

The second objection is to claim (4):

4. If an action harms someone who does not independently exist, then the fact that the action also benefits the person harmed, and benefits him more than it harms him, is *ineligible* to justify the harm if *failing to perform the action* would similarly benefit someone.

The objector argues that claim (4) is false, as follows. Claim (4) is a restricted claim about some harms only; it is restricted to harms to people who do not independently exist. The objector complains that this restriction is unmotivated and ad hoc: either the general claim is true, in which case the restricted claim is trivially true as well; or the general claim is false, in which case the restricted claim must be false as well. Furthermore, the objector maintains, the general claim is false. There is nothing that could motivate the restricted claim if the general claim is false. Therefore, the objector concludes, the restricted claim (4) is false as well.

The general claim is this:

4*. If an action harms someone, then the fact that the action also benefits the person harmed, and benefits him more than it harms him, is *ineligible* to justify the harm if *failing to perform the action* would similarly benefit someone.

(Note that I do not endorse (4*.) The objector argues that (4*) is false as follows. Suppose that I am in a position to distribute a benefit; it can only go to one person, and I must choose who gets it. Suppose further that Ben and Sally both want it, and that getting it would have no bad effects on Ben, but if Sally gets it she will be harmed—suppose she will suffer a migraine. If the benefit is substantial enough, Sally may want me to choose her despite the fact that getting the benefit would cause her to get a migraine. But if (4*) is true, then I cannot give it to Sally: there is a strong reason against harming her, and the fact that my action would also benefit her is ineligible to justify the harm because I have an alternative that would similarly benefit someone. The objector points out that it clearly would be permissible to give the benefit to Sally (if the benefit is significant enough). The action harms Sally, but the harming is justified *in part* by the fact that it also greatly benefits her. (While the fact that Sally consents to be harmed is also part of the justification of the harm, consent alone is not a sufficient justification for harming someone.)

I think this case does establish that (4*) is false. But it helps to motivate the possibility that (4) is true while (4*) is false. What strikes us about the case is that if we don't harm Sally, then Sally *loses out* on the benefit. What moves us to think it may be permissible to harm Sally is the thought of what is bad for her about our *not* harming her. This brings out an important difference between actions that affect whether people exist and those that do not. If an action benefits someone who exists independently of the action, then there are two kinds of considerations in favor of the action:

- the action benefits the person
- if the action is not performed, the person is deprived of the benefits (that is, she exists but doesn't get some benefits she could have had)

When a non-identity action benefits someone, only the first of these two considerations is available to tell in favor of the action; the second does not apply. But

the second seems to play a crucial role in how, sometimes, the fact that we benefit someone can play a role in making harming that person permissible—and it plays a crucial role in why (4*) is false. Since the second consideration doesn't apply to non-identity actions, this suggests that the reasons (4*) is false just don't apply to (4). (Note that in discussing the second objection, I haven't offered an independent argument for the truth of (4). Rather, I've defused an objection to it.)

Consideration of these two objections brings out the following. On my view, the fact of *non-identity* in non-identity cases has been widely misunderstood. It is not as significant as some people have claimed: it does not undermine the assertion that harms occur in non-identity cases. But it is not as irrelevant as some have claimed: it does lessen harms (though not very much), and it does limit the significance of benefits: it is harder for benefits in non-identity cases to justify harm.¹⁵

7.5 Is Causing Harm Sufficient for Harming?

In this section, I will discuss another objection to the account I have offered. It is an objection to my claim that one harms someone whenever one *causes* the person to be in a bad state. This objection is an adaptation of an objection to earlier papers by Seana Shiffrin and by me raised by Matthew Hanser in his contribution to this collection.¹⁶ The objector claims that sometimes one causes someone to be in a particular bad state, but one does not harm the person, because one causes the bad state simply by *improving* a worse state. Improving a bad state, the objector claims, is not a way of harming someone. Here is an example. Suppose that a patient is legally blind; his eyesight is severely impaired, so that he cannot really make out anything visually. Say that his state of severe visual impairment is state S (for “severe”). Suppose that the only thing that can improve this patient's sight is an operation that will improve his sight so that he is somewhat impaired; he will be able to distinguish shapes and colors, but not fine details. This is a significant impairment but much better than his current state. Say that his visual impairment after the operation is his being in state M (for the “milder” impairment). The objector maintains that on my view, a doctor who performs the operation *harms* the patient by *causing* him to be in state M, which is a bad state. However, the doctor does not harm the patient at all; he simply improves his situation. There is no harm-based reason against the doctor's action to be overridden in this case. The objector grants that if there *were* a harm-based reason against acting, it would be overridden, but denies that that tells the right story about the case: this is not a case in which it is permissible to cause one harm because one is preventing (or alleviating) a greater harm. Rather, this is a case in which one simply improves someone's situation.¹⁷

I will say several things in response to this objection. First, it matters what the details are: it matters *how* the operation improves the person's state from S to M. Suppose the operation works as follows. The only way to improve the patient from being in state S is to start a causal chain that would completely cure the sight: using a laser to reshape the retina would create normal sight. However, a side effect of using

such a powerful laser would be to distort the nerves that transmit retinal images to the brain. Thus, the causal route by which S is cured is distinct from the causal route by which M is caused. In this case, I claim, the surgeon causes the patient to be in state M; furthermore, the surgeon causes the patient to be visually impaired. In this case, the surgeon does *harm* the patient by causing M, and the case is properly understood as one in which some harm is caused but that is justified because a greater harm is alleviated.

The objector would respond that he imagined the causal structure of the case differently. Suppose the operation works as follows. The patient's sight is initially so poor because his retina has a thick film of fatty tissue over it; this fatty tissue deflects light. The best that can be done is to remove *some* of the fatty tissue. If more were removed, the retina would be damaged. This is how the surgeon improves the patient's state to state M. In this case, M is caused simply as the end-state of improving from state S; M is not caused by an independent process from the process of improving S. Here, the objector's claim has force: it is natural to say that the doctor has not *harmed* the patient at all, but has simply lessened an existing harm. But the doctor does cause the patient to be in state M, a bad state. Thus the objector maintains that, on my view, the doctor harms the patient.

My response to this objection is as follows. In this version of the case, while the doctor has caused the patient to be in state M, the doctor has not caused the patient to be visually impaired. Nor does the doctor cause the patient to be disabled. Now I will clarify the view I stated earlier. In claim (1), I offered a *sufficient* condition for harming. Here is a clarified restatement of that claim:

1'. One harms someone if one causes him to be in pain, to be in mental discomfort, to be in physical discomfort, to have a disease, to be deformed, to be disabled, or to die.

In this version of the case, because the doctor does not cause the patient to be disabled, this condition is not met. Thus, I am not committed to the view that the doctor harms the patient.

Note that typically, when one causes someone to be in a particular bad state, one also causes him to be in the more general bad state of which this is a particular form. So typically this distinction is not relevant. Furthermore, in non-identity cases, the non-identity actions do cause the created people to be in the more general bad states they are in; non-identity actions do not cause particular bad states without causing the more general bad states.

At this point the objector might revise his objection as follows. Consider the following case. An angry man gets mad at someone in a bar, and strikes him with a broken bottle. Suppose that the victim was already somewhat visually impaired (suppose he was in state L, for "low visual impairment") but that the angry man's action increased his visual impairment (moving it to state M). Suppose further that the angry man worsens the victim's eyesight not by creating a new problem but simply by exacerbating an existing one. In this case, the angry man causes the victim to be in state M, but does not cause the victim to be visually impaired, and does not cause the victim to be disabled.¹⁸ So my sufficient condition for harm is not met. Yet, the objector might say, not only has the angry man clearly harmed the victim, but he

has harmed him in the very way that I am trying to capture with claim (1'): it is the same kind of harm, and it is a failing of my view if my view cannot acknowledge that. The complaint is that the previous version of the doctor case is importantly different, as to whether the agent has harmed, from the angry man case—yet my view does not see a difference between them.

What the objector is bringing out is that there is another sufficient condition for harm besides (1') (and there may be others¹⁹). While the doctor and the angry man both cause people to be in state M, and both do not cause those people to be visually impaired, there are other differences between them. We might offer this sufficient condition:

One harms someone if one *worsens* a bad condition the person was already in.

Some cases of worsening are already covered by (1')—cases in which the way the bad condition is worsened is sufficiently distinct from the original way it was bad—but some, like the case of the angry man, are not.

Perhaps there is a better way to state the new sufficient condition. We might note that, as some have argued,²⁰ some causal facts are irreducibly of the form: c caused e *rather than* f. The doctor caused his patient to be in state M rather than in state S. But the angry man caused the victim to be in state M rather than in state L. This is a difference in the causal facts about their actions. We might offer this sufficient condition:

One harms someone if one causes her to be in a particular bad state rather than in a better state.

In summary, it's true that one sometimes causes someone to be in a particular bad state without harming him. My view is that one harms if one causes someone to be in certain general bad states (listed in (1')). Furthermore, there are other sufficient conditions on harming; I have suggested two different ways we might state sufficient conditions to cover the angry man case.²¹

7.6 Conclusion

In this paper I have aimed to accomplish eight things.

I have developed a view that elaborates the claim that we have strong reasons against harming; I have offered one way in which we might see the reasons given by harming as *strong* and not easily outweighed.

I have argued for a new taxonomy of non-identity cases:

- (i) same-number cases
- (ii) different-number cases, more people if the action is not performed
- (iii) different-number cases, unknown which scenario has more people
- (iv) different-number cases, more people if the action is performed

I have argued that cases of types (i), (ii), and (iii) are very similar in what makes the actions wrong, while cases of type (iv) are quite different.

I have argued that *epistemic* considerations play a crucial role in why cases of type (iii) are wrong.

I have offered an account of why cases of type (iv) are wrong.

I have offered an explanation of why there is disagreement about cases of type (iv).

I have argued that harming independently existing persons is more morally serious than harming in non-identity cases (but not much worse), and also that benefiting independently existing persons has stronger, different reasons in favor of it than benefiting in non-identity cases.

Finally, I have responded to an objection from causing someone to be in a bad state by improving her condition.

Acknowledgements For helpful feedback on drafts of this paper, I am grateful to Tyler Doggett, Peter Graham, Alex Guerrero, Sarah McGrath, Brad Skow and audiences at University of Vermont and Ohio State University.

Notes

1. A *solution* would vindicate our pre-theoretic beliefs about these cases. Some have argued that no solution is possible, that we can only explain how non-identity actions are wrong by impersonal explanations that do not involve the claim that the actions harm anyone. For example, Parfit (1984) makes this claim.
2. I don't claim that everyone has the intuition that these cases are wrong; I simply claim that these intuitions are widespread. And, in particular, they are much more widely shared than are intuitions about what I call the "hard cases" below.
3. The view I offer in this paper is a revision of a view I presented in an earlier paper. See Harman (2004). The view is broadly the same, but differs importantly in two ways: my view of the interaction between reasons against harm and reasons to benefit has changed, and my view of what to say about "different number cases" (cases in which one's action affects how many people exist) has changed.

This paper is not meant to supercede but rather to supplement my earlier paper. In that paper, I take up many questions and argue for many conclusions which I do not address in this paper.

4. See Harman (2004).
5. In Harman (2004), I argued for (3) by considering two cases that support the claim that actions can be wrong in virtue of harming someone though they also benefit the person more than they harm him. These are not non-identity cases.
Shiffrin has a view similar to my claims (1) and (3). See Shiffrin (1999). On her view, reasons to prevent harm are much stronger than reasons to benefit. On her view, preventing someone from worse harm can justify harming her, but providing positive benefits to someone does not similarly justify harming her.
6. Note that whether a person who is harmed by an action exists independently of the action is not a matter of what *time* the person exists: if *whether* the person exists does not depend on *whether* the action is performed, the person exists independently of the action, even if the person does not yet exist at the time of the action.
7. We need the further condition that the agent knows that *not performing the action* is morally permissible. This condition is satisfied in non-identity cases.
8. Note that I'm not assuming that the agents in these cases *know* the moral fact that if this is a different-number case such that more people would exist if the action is not performed,

then acting would be wrong. The principle I'm appealing to does not require such knowledge. (Compare: suppose a person knows that if he shoots at random into the forest, there's a 50% chance he'll hit a person, because there are a lot of people wandering around the forest. Then he should not shoot. We needn't add the further claim that he knows shooting with 100% certainty of hitting someone would be wrong.)

Someone might object that a theory of what's *really* going on morally in non-identity cases should be an account of what the agents *objectively ought to do*, not of what they *subjectively ought to do*, whereas by appealing to the agents' ignorance, I am really telling a story about what the agents *subjectively ought to do*. The answer to this worry is that our "ought" claims fall along a long scale from *very subjective* to *very objective* and often what we want to explain is what an agent with "ordinary full information" should do in a particular case. But "ordinary full information" rarely contains full information about how all subtle future contingencies will work out, so these "ought" claims are not fully objective. And that is the kind of "ought" claim I am explaining.

9. In Harman (2004), I discussed a type of non-identity different-number case in which more people exist if the harming action is performed. I saw these cases as challenging for my view, but I focused on the idea that the existence of *additional happy people* besides the harmed people put pressure on us to say these actions are permissible, despite the fact that they harm. I now see that this consideration need not have troubled me: the fact that some additional happy people would exist cannot justify harming other people.
10. The following objection might be raised. If someone has decided to conceive a baby, and then is considering conceiving *now*, her alternative is conceiving *later*. But then on my view, the benefits cannot justify the harms, because the alternative would involve benefiting someone the same way. My response is as follows. When deciding *whether to conceive*, such considerations do not apply: there is no such alternative. Once one has decided to conceive, conceiving now is no better or worse than conceiving later: either way, one both harms and benefits someone and would have harmed and benefited someone in the alternative. This brings out something important: the fact that both acting and failing to act would harm someone in the same way is a way of justifying that harm.
11. My explanation of the hard cases involves the following thought: if my view is correct, then someone who had the moral facts *roughly* correct would believe that non-identity actions are wrong in the cases that aren't hard cases, but if she didn't have the moral facts *precisely* correct, then she might get the hard cases wrong. On the assumption that people tend to have the moral facts roughly correct, but don't always get them precisely correct, my view provides an explanation of disagreement in the hard cases: some people are implicitly placing the threshold in the wrong place. That my view can provide an explanation of the fact of disagreement about hard cases counts in favor of my view. I don't claim, however, that this provides anything like a decisive argument for my view. Nor do I claim that in general we're entitled to assume that people are getting the moral facts roughly correct. There is much discussion nowadays of the question: are we entitled to rely on moral intuitions about specific cases in making moral arguments? My view is that arguments that rely on specific moral claims about cases are not thereby illegitimate as arguments; any argument has undefended assumptions, or starting places, and these may be specific or general. Furthermore, I am unsympathetic to the thought that the existence of disagreement about specific claims undermines arguments that rely on those claims: these arguments have the specific claims themselves as premises, *not* the claim that the specific claims are believed by everyone.

This paper is aimed not to provide a conclusive argument for my overall view, but simply to lay out a view with a certain structure and show that the view has certain advantages. Some of these advantages are: getting the moral facts right. For example, my view implies that it is wrong for a woman to conceive if she has a temporary condition that will cause her child to be blind; and it implies that her action is wrong in virtue of harming her child. When I say that it's an advantage of my view that it gets the moral facts right, I'm relying on claims like this: what the woman does is wrong. I don't mean to be supporting that claim by the further claim that everyone agrees with that. But another part of my argument for my view—and it

is only a small part of the support for my view—is the claim that my view can explain why there is disagreement about some cases while there is agreement about other cases. It is here, and only here, that I intend my view to be supported by claims about what people believe.

12. See Harman (2004).
13. Someone might object that what I claim is true of harms to independently-existing people *in general* is only true of some harms. Consider a case of preemption: Joe punches Steve in the face, but Bill would have swung at Steve if Joe had not. The objector claims that while Joe harms Steve by *causing* him pain and injury, Joe does not *prevent* Steve from an existence without this pain and injury, because if Joe had not acted, Steve would still have suffered in the same way. My response is that the objector's claim is false: Joe does prevent Steve from an existence without this pain and injury. Steve *is prevented* from an existence without this pain and injury. Who prevents him? Joe does. (Bill, who stands idly by, certainly does not.) So, we see that *prevention* does not require counterfactual dependence: one can prevent an outcome even though the outcome would not have occurred had one not acted.
14. Parfit (1984), pp. 366–369.
15. Shiffrin points out that if one benefits someone by creating her, then if one had not acted, she “[would] not experience [the benefit’s] absence” and it’s not the case that her life will “go worse.” Shiffrin (1999), p. 134. But Shiffrin does not think a parallel consideration is relevant to harming: “Harms may differ [from benefits]—a harm that does not make a life worse does not always have less or significantly less weight than harms that do worsen lives, marking another asymmetry between harms and benefits.” Shiffrin (1999), p. 134, n. 32. By contrast, I think that the fact that harming an independently existing person deprives her of life without that harm does make such harming, at least to some degree, worse than harming by creating.

One might object to my account as follows. I claim that it’s particularly hard for benefiting by creating to justify harm in non-identity cases, because benefiting by creating is less morally significant than other benefiting. The objector points out that on my view, harming by creating is *also* less significant than other harming. Perhaps a less significant harming can be justified by a less significant benefiting. My response is that my goal in this paper is to spell out a certain kind of view—a view on which we take seriously that there are strong reasons against harming. Thus, on the view I am spelling out, there are strong reasons against all kinds of harming, including harming by creating, even though this is somewhat less bad than harming independently-existing persons.
16. See Harman (2004) and Shiffrin (1999). See also Hanser (2009), Sections 9.4–9.5.
17. Note that what’s at issue between the objector and me is not whether the ordinary word “harm” applies but, taking harm to be distinctively serious and to generate strong harm-based reasons, whether the doctor has harmed the patient. In fact, I will readily admit that on my view some actions harm though we would not typically say they harm; my claim is that these actions generate strong harm-based reasons. (For one example, if I affect when a couple conceives by keeping them late at dinner, my action is a cause of the existence of the particular child they create, thus my action harms the child (and also benefits the child). But while on my view, such actions do harm, there is no practical upshot of the fact that they harm. When we engage in everyday actions like this, some of which affect whether people come to exist, we don’t know whether performing the action is more likely to harm than not performing it, and so harm-based reasons aren’t available to us in our reasoning and don’t play a role in what we should choose to do.)
18. It’s not clear to me whether the causal facts can ever be this way: whether one can *worsen* someone’s eyesight without it’s also being true that one has caused the person to be visually impaired. But I will grant this to the objector. If it’s not possible, then my claim (1’) handles this case directly.
19. For example, some would say that I harm someone if I deprive her of something to which she is entitled (or has a right). I am not taking a stand on whether our reasons against depriving someone of a positive good—a benefit—to which she is entitled are *strong* in the way that harm-based reasons against causing pain and other bad things are strong.
20. Schaffer (2005).

21. In his paper in this volume, Hanser also makes the following objection to my Harman (2004). Suppose that two people are in danger and I can save exactly one of their lives; however, to save one of them I would have to break his arm. Hanser points out that the fact that I would benefit him (by saving his life) *is* sufficient to justify harming him. He takes this to be an objection to my claim that the mere fact that one greatly benefits someone, and benefits him more than one harms him, is not sufficient justification for harming him. However, this is not a case of both harming and benefiting someone, as I use the terms “harm” and “benefit.” What I mean by “benefit” is to provide a positive good to someone, such as a good experience. This is a case in which it is permissible to harm someone because one is *preventing* greater harm. (My view on this matter is the same as the view Hanser attributes to Shiffrin.)

References

- Hanser, M. 2009. Harm and harming: the non-identity problem revisited. In this collection.
- Harman, E. 2004. Can we harm and benefit in creating? *Philosophical Perspectives* 18: 89–113.
- Parfit, D. 1984. *Reasons and persons*. Oxford: Clarendon Press.
- Schaffer, J. 2005. Contrastive causation. *Philosophical Review* 114(3): 327–358.
- Shiffrin, S. 1999. Wrongful life, procreative responsibility, and the significance of harm. *Legal Theory* 5(2): 117–148.

Chapter 8

Wrongful Life and Procreative Decisions

Bonnie Steinbock

Abstract This paper defends and refines the claim that procreation can be wrongful. Procreation is wrongful first when the “nonexistence condition” is met: the person’s life will be filled with suffering that cannot be ameliorated or empty of all the things that make life worth living. Recognizing that this condition is rarely met, the paper then argues that it is wrong to create a person in less extreme circumstances: when the person is likely not to have a minimally decent life, one in which certain important interests cannot be satisfied. Although we must be very cautious about concluding that any particular impairment precludes a minimally decent life, there will be circumstances in which a future life is unlikely to hold a reasonable promise of containing the things that make human lives good. In these circumstances, and if reproduction is avoidable, we are required to forego reproduction altogether.

Keywords Harm · Benefit · Interests · Rights · Minimally decent life · Non-identity problem.

8.1 Introduction

Many people would agree that if a child is going to be born under very disadvantageous conditions, it would be wrong to reproduce, and indeed a wrong to that future child. However, it turns out to be surprisingly difficult to support this claim, in cases where nothing can be done to prevent the disadvantageous condition, except to prevent the child’s birth altogether. To capture this unique feature, David Heyd terms these cases “genesis problems.”¹ The precise nature of genesis problems is explained below in Section 8.3, The Philosophical Problem. I will start, however, by pointing out that genesis problems challenge some widely held intuitions, and raise the following question: when it is likely that the child will be born under adverse conditions, and has “no other way of getting born,” can concern for the welfare of the child ground an obligation to avoid reproduction?

B. Steinbock (✉)
The University at Albany/SUNY, Albany, NY, USA
e-mail: steinbock@albany.edu

By “obligation to avoid reproduction,” I mean an obligation to avoid the deliberate or intentional conception of a child. I am not addressing the question of whether abortion can ever be morally required of a woman, because that is a much more complicated issue, which turns on such issues as the moral status of the unborn and the woman’s right of self-determination. There is no consensus about either issue. For some, a fetus is morally equivalent to a born child. It would be no more permissible to abort the fetus, out of concern for its welfare, than to kill a born child. For those who regard the fetus as having a lesser moral status, abortion is permissible for a range of reasons, including the welfare of the future child. However, even for those who think that abortion can be *justified* in certain cases by concern for the welfare of the child, it does not follow that abortion would be morally *obligatory* in those cases, i.e., that the decision not to abort would be morally wrong. For example, a justification for abortion might be that the woman already has as many children as she can care for. Given a pro-choice perspective, the decision to abort would be permissible, but the opposite decision not to abort would not be immoral. These are decisions that are, for the most part, up to the women who have to make them. The right to bodily self-determination certainly includes a right not to be forced to have an abortion, but, I would argue, it also includes a right to make one’s own moral decision about abortion. This is not to say that every abortion decision is morally correct. It is possible to imagine ill-considered abortions, or abortions done for morally bad reasons, although such cases are most likely rare in real life.² It may be possible to imagine a morally bad decision to continue a pregnancy, but for the most part, such decisions are not considered immoral, and it would be unusual, to say the least, to claim that a pregnant woman has an obligation to kill the fetus, out of concern for its future well-being. By contrast, it is not at all odd or unusual to suggest that starting a pregnancy in disadvantageous circumstances would be irresponsible and indeed unfair to the future child. It is this judgment about conceiving a child that I examine in this paper, where the decision to procreate is clearly intentional and voluntary. This will have implications for the scope and limits of procreative liberty, and for ethical judgments about risk in assisted reproduction.

For example, the main objection to reproductive cloning in the National Advisory Bioethics Commission’s report was an unacceptable level of risk of serious defects in offspring.³ The question of risk to offspring also comes up in the President’s Council on Bioethics’ White Paper on alternative sources of stem cells. One proposal is based on an analogy with pre-implantation genetic diagnosis (PGD). It suggests that one or two stem cells could be removed from an embryo without damaging the embryo. The embryo then could be implanted to start a pregnancy. Although PGD is widely regarded as safe, the President’s Council rejected the proposal, primarily on the ground that in the absence of long-term safety studies, it is not possible to determine conclusively that embryo biopsy is safe for the future child.⁴

Objections to reproduction under adverse conditions can be based on emotional as well as physical harm. For example, critics of postmenopausal motherhood have argued that a woman who has a child in her fifties or even sixties might not be able to be an adequate parent. How, they ask, will she handle a rambunctious two-year-old

or a rebellious teen? Moreover, women who have babies in their sixties may not live to see their children grow up. Art Caplan has suggested imposing age limits for infertility treatment,⁵ to avoid creating children who will be orphaned at a young age. In addition, he argues that even if these children are not orphaned before adulthood, they are likely to be prematurely burdened with the care of an elderly parent. Is it fair to the children, he asks, to have such old mothers?

8.2 How Bad Is It?

Much of the debate centers on empirical questions that relate to the badness of the harmful condition, including how severe the harm is, and how likely is it to occur. But these are not the only questions, since we should also ask whether the harm can be lessened, even if not prevented entirely. Having an older mother is not ideal for children, perhaps, but surely it is not a tragedy either. Today, it is not uncommon for women in their forties to have babies. Perhaps in ten years, motherhood in one's fifties will not be a rarity. As women live (and are healthy) into their eighties and even nineties, the "orphan objection" may no longer be an objection to postmenopausal motherhood. In addition, many older women have had to take on the job of raising their grandchildren, and have done a pretty good job. If they can be good child rearers, it would seem that postmenopausal women can be good (or good enough) mothers as well.

8.3 The Philosophical Problem

The truth is, we often do not have very reliable evidence about the impact of these various technologies and arrangements on offspring. Sometimes objections are based on "gut reactions" and a fear of what's unusual, rather than solid empirical evidence. Sometimes they are based on stereotypical thinking and prejudice: gay and lesbian couples have been prevented from using infertility services and adoption, because of an unfounded belief that they cannot be good parents.⁶ We should be very careful in attempting to assess the empirical questions regarding the impact on offspring, to avoid unnecessarily depriving individuals of the right to have children.

However, there is a deeper philosophical issue raised by all of these examples because the technology or arrangement that results in the child's being born in a harmful or disadvantageous condition is at the same time the condition of the child's being born at all. Thus, the examples under discussion differ importantly from other examples of prenatal harming, where something can be done to prevent the harm to the child. For example, a pregnant woman can reduce the risk of prematurity or low birth weight (which are associated with various health risks) in her baby by not smoking or drinking alcohol. She can lessen the risk her child will have a neural tube disorder by getting enough folic acid in her diet. In fact, she can do this even before she gets pregnant. The fact that the child does not yet exist is not

the relevant factor. What is important is that the harm can be prevented. The child who is harmed when the harm could have been prevented has, in most cases, a legitimate grievance against the individual or individuals who could have prevented the harm. By contrast, in genesis cases, nothing can be done to prevent the harm to this child. It's life with the disadvantage or no life at all. And that makes the question of whether bringing the child into existence in a harmful condition is "unfair to the child" a much more difficult one.

Genesis problems are particularly vexing because our intuitions often go one way, while the arguments seem to go another. Another motive for examining genesis problems is that they have profound implications for ethical theory, in particular, the explanation of why wrong acts are wrong. On one plausible ethical view, acts that are wrong must be wrong for someone. Moral principles, on this view, must concern the interests of individuals; they must be "person-affecting."⁷ Genesis problems pose a challenge to this assumption because they seem to provide examples of wrong acts that are not a wrong or a harm *to* anyone.

8.4 Preventing Births to Protect Children

8.4.1 Robertson's Analysis

John Robertson, perhaps the best-known advocate for procreative liberty, has argued that banning risky procreative technologies or arrangements *out of concern for the welfare of offspring* makes no sense.⁸ As Robertson puts it, "But for the technique in question, the child never would have been born. *Whatever psychological or social problems arise, they hardly rise to the level of severe handicap or disability that would make the child's very existence a net burden, and hence a wrongful life.*"⁹ However well-meaning, the attempt to protect children by preventing their births is illogical.

The extent to which this view differs from conventional thinking cannot be overstated. Virtually every professional society or national commission or oversight group that has considered the matter takes for granted that expected impact on offspring must be taken into consideration in determining the permissibility of a reproductive treatment or arrangement. The British Human Fertilization & Embryology Act of 1990 explicitly provides that a "woman shall not be provided with treatment services unless *account has been taken of the welfare of any child who may be born as a result of the treatment* (including the need of that child for a father), and of any other child who may be affected by the birth."¹⁰ It is not clear precisely what is meant by "taking account of the welfare of the child": for example, how severe or likely the harm would have to be to deny treatment services. However, the mention of "the need of that child for a father" suggests that all sorts of social factors should be considered, and could justify denial of treatment. By contrast, on Robertson's account, the procreative liberty of individuals can be limited only when the predicted harm would constitute a "wrongful life." If the child has a life

that is, on balance, worth living, from the child's own perspective, despite whatever disadvantages it has, then, Robertson maintains, its life logically cannot be regarded as a harm or wrong to the child.

I am not concerned with the tort of wrongful life here, or the question of whether such cases can fit under traditional understandings of tort law.¹¹ The primary issue in wrongful life cases is whether the infant plaintiff deserves *compensation*,¹² whereas my concern in this paper is procreative responsibility, and whether there is an obligation not to have a child in the first place. Still, wrongful life cases are useful for understanding how birth in very disadvantageous conditions can plausibly be seen as a harm or a wrong to the child.

8.4.2 *Feinberg's Counterfactual Analysis of Harming*

The idea that children can be harmed by being born may seem incoherent. For to say that the child has been harmed by being born is to say that the child has been made worse off. But how can someone be made *worse* off by coming to exist? Nonexistence is not a better condition to be in; it is no condition at all. This suggests that it is impossible to harm someone by causing him to exist.

Joel Feinberg suggests that this conclusion comes from failing to distinguish between two interpretations of what it is to make someone worse off. On one interpretation, which Feinberg calls the "worsening condition," to make someone "worse off" is to make him worse off than he *was*. Clearly, the worsening condition cannot be satisfied in the wrongful life situation. No one can be worse off *than he was* before he existed, since this suggests comparing the existing individual with himself before he existed, which is absurd. However, to make him worse off can also be interpreted counterfactually where it means "worse off than he would have been." The counterfactual claim is that the child would have been better off not coming into existence, or "better off unborn." Before explaining what this might mean, let us consider an objection to the counterfactual analysis of harming.

8.4.3 *Harris's Objection*

John Harris rejects the counterfactual analysis as unnecessarily complicated. To be harmed, according to Harris, is simply to be put in a condition that is harmful. He writes, "I would want to claim that a harmed condition obtains wherever someone is in a disabling or hurtful condition, even though that condition is only marginally disabling and even though it is not possible for that particular individual to avoid the condition in question."¹³ To harm someone, on Harris's account, is just to be responsible, causally and morally, for the person's being in that harmed condition.

Harris's account seems counterintuitive, as is revealed in the following pair of examples. In the first example, a woman who is a smoker continues to smoke during her pregnancy, despite knowing the risk of causing asthma in her child. If the child is born asthmatic, she has harmed her child, in a straightforward sense of

“harm”; the child has been made worse off by the woman’s behavior. He or she could have been born without asthma, if the woman had stopped smoking during pregnancy.¹⁴ Contrast this case with a woman who has asthma because of a genetic predisposition.¹⁵ She decides to have a child, hoping that her child will not inherit her genetic predisposition, but well aware that any child she has might be asthmatic. On Harris’s analysis, if she has a child who has asthma, she too has *harmed* her child, and just as much as the woman who continues to smoke during pregnancy. She is causally responsible for the child’s being born with asthma, as well as morally responsible, in the sense that she knew of the risk and (let us assume) chose to keep smoking when she could have stopped.

Harris’s analysis, which equates the two cases, seems quite wrong. The difference is that the smoker could have prevented her baby being born with asthma, while the non-smoking asthmatic could not. The only way she could prevent the birth of a child with asthma would be to avoid having a child at all. Not only does Harris’s analysis fail to distinguish between the two examples, but it has the counterintuitive implication that virtually all of us harm our children, because all of us pass on genes associated with disadvantageous, though not disastrous, conditions, such as nearsightedness, acne, or allergies. To have a child, on this view, is to harm him or her.¹⁶

Responding to this objection, Harris agrees that his account makes all parents *causally* responsible for the harms they genetically transmit, but maintains that this sense of “responsible” is trivial. Parents are not *morally* responsible for the harms they cause “unless they were, first, aware that they were likely to transmit those harms and, second, aware of a better alternative child, or a better possible alternative child, and could, realistically, have produced that child instead.”¹⁷

I do not dispute Harris’s claim about moral responsibility for harm. However, while I agree that one may be morally responsible for having a child in a harmful condition if one could have substituted a different child (see below Section 8.7.2), it is hard to see how the possibility of substitution can be a condition of causal responsibility for harm. How does the fact that one could have avoided the harmful condition, by bringing a *different* child into existence make it the case that if one fails to make the substitution, one has harmed *this* child? How does the existence of a better option affect whether this child is harmed by being born? The existence of a better option, and the failure to adopt it, may be part of an argument that one has behaved irresponsibly or wrongly, but it is hard to see what role it plays in an argument that the child born with a disadvantageous condition, who could not have been born in a better condition, has been harmed.

To show that the child has been harmed, we need a counterfactual analysis like Feinberg’s, which aims at explaining how it is possible for someone to be “better off unborn.”¹⁸ The first step is to ask what this means. Many people find the expression “better off unborn” baffling. They maintain that since we cannot compare the child’s impaired condition with non-existence, it makes no sense to say of any individual that he or she would be better off never having come into existence. To address this conundrum, Feinberg suggests that we think about a comparable claim, “better off dead.” The phrase “better off dead” does not express the absurd idea that

non-existence is a better condition for someone to be in than life; non-existence is not a condition of a person at all. Rather, the phrase expresses the idea that sometimes the burdens of life outweigh the benefits—that from the individual’s own perspective, life is not worth living.

While it is fairly easy to understand what this means in the case of a competent adult, it is trickier to be able to claim, with any certainty, that a never-competent infant would be “better off dead.” Infants lack the cognitive ability to have complex preferences, such as “I prefer death to life under such-and-such conditions,” making it extremely difficult to say, on behalf of an infant, that he is better off dead or that it is better *for him* to die. John Robertson provides the provocative example of a child who is profoundly retarded, nonambulatory, blind, deaf, and who will spend a brief life in a crib on the back wards of a state institution. (This last is a social factor which could be ameliorated and so is not really relevant, but I leave it in as it is part of his example.) Robertson says that although you and I might find such a life horrible to contemplate, and might prefer death, the child, who has known no other existence, might disagree. “Life and life alone, whatever its limitations, might be of sufficient worth to him.”¹⁹

It appears that two conditions are necessary for us to be able to assert with any confidence that an infant is better off dead: excruciating physical pain and such a brief life span that the child is unable to develop any compensating abilities. Such conditions are very unusual. Most newborns who have serious disabling conditions do not have lives that will be pure torture. Nevertheless, there are some cases in which it is possible accurately to predict that they will. Given a choice between a brief life filled with nothing but severe and unrelievable pain, we can say, sadly, that the baby would be better off dead. If the baby would be better off dead, it seems that the baby would be better off unborn. We should, however, examine the notion of what it is to be “better off unborn” a bit more closely.

8.5 The Nonexistence Condition

8.5.1 A Test for Harm

As noted above, the claim “better off unborn” does not refer to any preference the impaired infant actually has. Instead, this claim is one that would be made on behalf of the infant plaintiffs by proxy choosers who act as advocates for the infants, concerned to promote their overall welfare. The proxy choosers are not to substitute their *own* views of what makes life worth living. They are not to think about the conditions under which *they* would prefer nonexistence. Rather, they are to view things, as much as possible, from the children’s perspective.²⁰ The judgment that these children would be “better off unborn” is warranted if all the children’s interests (whatever they might be) are inexorably doomed to defeat by their incurable condition. “Thus,” Feinberg says, “it would be irrational—contrary to what reason decrees—for a representative and protector of those interests to prefer the continuance of that condition to nonexistence.”²¹ Let us call this standard *the nonexistence*

condition. If the nonexistence condition is fulfilled, the child has been harmed, and therefore wronged, by birth.

The question I want to consider now is whether a child can be said to be harmed or wronged by birth only if the nonexistence condition is satisfied. Consider the following example. After years of trying to have a child, an infertile couple resorts to IVF and is able to have a much-loved child, Junior. Unfortunately, Junior turns out to have an inherited disorder that causes a massive failure of bone marrow cell production, and can lead to leukemia. Junior is healthy at present, but he probably will need a bone marrow transplant in the future, and possibly a kidney transplant as well. As it happens, the couple has several leftover embryos in storage and one is both disease-free and a perfect tissue match. The couple hires a surrogate to bring the embryo to term, with the idea that the child will be a source of bone marrow for Junior. They do not neglect or abuse “Donor” (as they name him). They just do not feel about him as they do about Junior. Indeed, they consciously suppress any tender feelings toward Donor since that might inhibit them in using him as a source of organs for Junior, should the need arise. Unlike real-life cases,²² where children have been conceived as “savior siblings,” but also loved for themselves as members of the family, this couple never intended to love Donor. If the couple did not want Donor as anything but a source of spare parts, they should not have had him in the first place. What they did was wrong, and moreover, a wrong *to* poor Donor.

On Robertson’s analysis, however, it seems that Donor has not been harmed or wronged. If Donor were to complain of his lonely, loveless existence, his parents could point out that, had they not needed the bone marrow for Junior, Donor would not be here today. He’d still be a frozen embryo. Admittedly, Donor’s life is pretty bad, but he does not want to die, nor would he prefer never to have been born. Since the nonexistence condition is not met, Donor has not been harmed or wronged by birth. That, I would argue, is completely implausible. The fact that Donor does not long for death or regret having been born surely does not get his parents off the moral hook. But is this example a counter-example to Robertson? Not necessarily. For it may be possible to treat these examples so that they fall under the sorts of ordinary person-affecting principles that Robertson himself seems to rely on. This approach is taken by Melinda Roberts.²³

8.5.2 A Person-Affecting Solution: The Third Option Approach

Like Robertson, Melinda Roberts believes that the nonexistence condition must be fulfilled for existence to be a harm or wrong to a child who has no other way of being born. However, she thinks that Robertson mistakenly assumes that if the progenitors *would not have reproduced* except under the conditions as presented, the child’s only alternatives are life with the disadvantage or no life at all. His mistake is to focus exclusively on what the couple *would have done*, instead of on what they *could have done*. If they could have brought the child to birth in a better condition, then they *have* harmed and wronged him. Perhaps this strategy can be used with the example

of Donor. It could be argued that Donor's options are not only either (1) life as an unloved child or (2) nonexistence. There's a third alternative: life as a loved and wanted child. The couple's failure to choose this third option makes Donor worse off than he *could have been* and explains why he has been harmed and wronged.

Roberts' analysis is consistent with the robust moral intuition that Donor's parents acted in an appalling fashion. It is not clear, however, that her strategy for deriving this conclusion is successful. It depends on whether a third option was in fact available to the Donor's parents. Certainly they could have *treated* Donor better. They could have treated him the same as Junior. But could they have *loved* Donor? Saying "they should have had him and loved him" sounds a little like the parental admonition to "eat it *and* like it," which has at least the ring of paradox. However, in my example, this third option is not realistically available, since the whole point of having Donor is as a source of spare parts for Junior. His parents deliberately avoid developing tender feelings for Donor so that such feelings will not get in the way if Junior needs his organs. Thus, it seems that a "third option" is not available to Donor's parents—they could not have had him and loved him in the circumstances—and therefore, on Roberts' analysis, they did nothing wrong. This, however, is morally outrageous. Having Donor as a source of spare parts for Junior is wrong, even if the resulting child does not want to die, even if he regards his life as on balance, worth living.

8.5.3 *The Decent Minimum Standard*

A more plausible criterion for "rightful" birth than the nonexistence condition is one in which life is actually a benefit to the child, as opposed to a life that is wretched, although still worth living. For life to be a positive benefit, certain minimal conditions must be satisfied, and therefore we can call this criterion for responsible procreation the "decent minimum standard." A decent minimum is reached only if life holds a reasonable promise of containing the things that make human lives good: an ability to experience pleasure, to learn, to have relationships with others. If someone's life will be inevitably and irremediably bereft of many of these goods, then we do that person no favor by bringing him or her into existence; indeed, knowingly and voluntarily to conceive a child under such conditions is a harm and a wrong to the person. This aspect of the decent minimum standard focuses on the child's capacities for a good human life. In addition, the ability to be a good enough parent is also part of the decent minimum. I maintain that it is wrong, irresponsible procreation, to have a child if one knows that one lacks either the ability to love the child or the capacity to care properly for him or her.²⁴

It might be argued that building into a decent minimum the ability to love one's children is implausible, since many people have parents who did not love them, but who still have lives well worth living. Indeed, in some cases of artists, writers, or actors, what makes their lives well worth living is a talent that derives precisely from having been deprived of parental love. If children can have lives that are well worth living, despite inadequate parenting, how is it possible to maintain that their

parents wronged them by having them? In response, I would argue that, while it is possible that a child may flourish, even without parental love, the foreseeable result is that the child will have a very diminished life. Most children need a secure basis of (at least) mother-love to be psychologically healthy. Without that basis, they are likely to have profound feelings of unworthiness and self-loathing, which make it difficult or impossible for them to have healthy relationships with others. No one should impose that on a child right from the start, not if it can be avoided. Where the only way to avoid giving a child that kind of diminished life is to avoid his or her birth, that is the responsible choice to make.²⁵

The intuition behind the decent minimum standard is that children have a right to something more than lives that are barely worth living, and individuals have a corresponding obligation not to have children under sufficiently awful conditions. Granted, no one can guarantee that one's child will be happy, and the mere possibility of hardship and burden does not make procreation wrong. At the same time, if individuals know, or should know, that they cannot provide their children with minimally good lives, then they should refrain from procreation, where this is possible.²⁶

The idea motivating this principle is that becoming a parent is not solely, or even primarily, a right. It is also, and primarily, an awesome responsibility. Prospective parents must think not simply of their own reproductive interests, but also of the welfare of their offspring, and this means thinking about the kinds of lives their children are likely to have. To bring a child into the world knowing that a decent minimum cannot be achieved is wrong; indeed, it is a wrong to the child.

To say that birth is a wrong to, or unfair to, the child suggests that the child has an interest in not being born. But how should we understand this interest? We certainly can ascribe to individuals an interest in not having lives that it would be contrary to reason to prefer, i.e., lives that meet the nonexistence condition. What I am suggesting here is that we can also say that people have an interest in not having lives that fall above the nonexistence condition, if they fall below the decent minimum standard. That is, it is not only lives that amount to pure torture that it is reasonable to reject, from the perspective of preconception existence, but also lives that are not minimally decent. If this is right, then it is possible to ascribe to individuals an interest in minimally decent lives, and to say that if they are knowingly brought into the world when this interest cannot be met, they have been wronged.

The question remains, how bad is too bad? At what point would it be wrong, unfair to the child, to bring him or her into the world? This is an issue on which reasonable people can disagree, at least about cases in the middle. We can expect considerable (though probably not unanimous) consensus about lives that fall well below a decent minimum (e.g., Robertson's deaf, blind, paralyzed, and profoundly retarded child), as well as consensus about lives that, despite certain disadvantages, fall well above (e.g., having asthma or being very nearsighted). However, we can expect disagreement about cases in-between, conditions such as Down syndrome, cystic fibrosis (CF), spina bifida, achondroplasia. On the one hand, prospective parents should be realistic about the burdens and limits such conditions may impose. Their desire to have a biological child is not the only relevant factor. On the other

hand, there are many individuals with serious disabilities who have lives that are well worth living. *Having a disability, even a serious one, does not entail life below a decent minimum.* Prospective parents who can provide their child with a life well worth living, despite a disease or disability, are not morally required to abstain from procreation. What they cannot do is claim that, on the grounds that the nonexistence condition will not be met, the prospect of a harmful condition is *morally irrelevant* to their decision.

Someone might ask why we should not raise the standard and say that prospective parents have an obligation to give their children lives that are, not just at the decent minimum level, but something much better than that? After all, we think that parents ought to make sacrifices for their children's health, education, and general welfare. They are poor excuses for parents if they settle for a decent minimum once the children are born; why not say that procreation is wrong unless offspring can be reasonably expected to have very good lives, lives considerably above a decent minimum?

It would be morally *permissible* to avoid having children under adverse conditions, on my view, since there is no moral obligation to have children at all. Refraining from having children is not something that needs justification. The question is whether individuals have an obligation to forego reproduction altogether, if the child is likely to experience physical or psychological harm. It seems to me that the higher one sets the bar, the less plausible it is that there is such an obligation. Why are individuals morally required to give up their dream of becoming parents, especially if they can be wonderful parents, simply because the child is likely to have more than the usual set of problems?

The nonexistence condition is arguably the right standard for ending someone's life, precisely because it is so restrictive.²⁷ Euthanasia—for example, killing an infant with serious defects—would only be justified if we had very good reason to think that the child's life would be unbearable. However, the nonexistence condition does not seem to be a reasonable standard for *bringing someone into existence*. That is, there seems to be an asymmetry between ending and starting a person's existence that is relevant to the morality of procreative decisions.

8.5.4 The Asymmetry Between Ending and Starting Lives

Existence makes a difference. We need not concern ourselves here with the thorny question of *when* a human being comes into existence, whether at conception or sometime during pregnancy, at birth, or sometime after birth. *Whenever* an individual comes into existence, the point I am making here is that ceasing-to-exist and never-coming-to-exist are not the same thing. Cynthia Cohen explains the difference this way:

Death is terrible, in part, because it prevents us from having future goods that we would have had if we had remained alive. The awfulness of death is also grounded in the fact that it robs us of those goods that we already have. Preconception nonexistence, on the other hand, does not involve the loss of life's goods, nor does it deprive us of goods that we

already possess. There is no loss incurred by possible children who are not brought into the world, for there is no actual “we” who could suffer such a loss at this point.²⁸

There is a natural and widespread reluctance to relinquish life, even under the worst conditions. There is a tendency to “cling to life” even when its burdens are great and its benefits marginal. Because we know that people generally prefer to go on living, even when life is filled with suffering, the standard for saying that a person’s life is not worth living, or that he or she would be better off dead, should be set quite high. The question we must ask is, “Is life better than death for this individual?” From a pre-conception standpoint, however, the standard is different, because we need not consider the tendency to cling to life, and to want to go on existing, even under the most miserable of conditions. Death is bad for people who want to go on living, but never-existing is not bad for anyone. We do not grieve for the limitless numbers of people who never were born, and this is not callousness on our part. There literally is no one to be sorry for! Of course, an infertile couple can be anguished about not having a child, but this sorrow, as real and intense as it may be, is not the same as grieving for a child who died. In grieving for a child who dies, there is the thought of *the child’s* loss of life, as well as one’s own grief in losing the child.

What is the implication of this asymmetry for procreative decision-making? It is this: If we are going to bring people into existence, we should be reasonably sure that existence will *benefit* them, and this is not the case if their lives are barely worth living. Harmful conditions which would never justify terminating the life of an infant can be excellent reasons for not having a child in the first place. The standard, then, for bringing people into the world should be higher than the standard for ushering them out. People have a right not to be brought into the world, where their births can be avoided, unless they can be reasonably assured of a decent minimum of the goods that life has to offer.

8.6 The Human Rights Approach

Like Feinberg and me, David Archard argues that children have a birthright to a life that is above a certain threshold.²⁹ However, instead of talking about basic interests which are doomed to defeat, Archard sets “the threshold of a minimally acceptable life as one in which the child has the reasonable prospect of enjoying a good number of those rights possessed by all children,” as outlined by the United Nations Convention on the Rights of the Child.³⁰ Archard does not specify what those rights are, which is perhaps not surprising, since there are dozens, including the right to know and be cared for by his or her parents, the right to education, the right to rest and leisure, the right to be protected from economic exploitation, and the right to a standard of living adequate for the child’s physical, mental, spiritual, moral and social development.

Some have taken the position that it is meaningless to claim something as a right when there is little or no prospect of the duty logically implied by the right

being fulfilled. However, I reject this Benthamite approach.³¹ The claim is not that children are already protected by these rights, but that they should be. They are, in Feinberg's term, "manifesto rights." Manifesto rights express moral ideals and aspirations. When we say that hungry children have a right to be fed, we are not necessarily pointing to anyone in particular and saying, "You have the obligation to feed these children." Certainly if a country is too poor to provide its citizens the basic necessities of life, its officials cannot be said to have an obligation to do so. Rather, we use manifesto rights to express the moral claim that a world in which children starve is morally unacceptable and should be changed.

Archard thinks that children who cannot enjoy a good number of the Rights of the Child should not get born, and that those who deliberately conceive them under these conditions wrong them.³² However, this suggests that the very poor have a moral obligation not to procreate at all, a claim John Harris calls "astonishing."³³ I agree. The very poor are already victims of injustice because of their economic situation. To maintain that they have no right to "marry and found a family"—which, I might point out, is also a fundamental human right³⁴—is doubly unjust. Instead of seeing the Rights of the Child as a minimal condition for morally permissible procreation, we should see it for what it is: an ideal. To say that all children ought to have these rights is not to say that those who are less fortunate should not be born at all. It is to say that all of us have an obligation to work toward improving the conditions under which too many children live.

So far, we have been considering situations in which the harm or disadvantage can be avoided only by foregoing reproduction altogether. The moral situation is quite different if it is possible to avoid the harmful condition and still become a parent, for example, by delaying conception. In such a case, the harm is avoided by having a different child. For this reason, it is referred to by Derek Parfit and others as "the non-identity problem."

8.7 The Non-Identity Problem

8.7.1 *The Fourteen-Year-Old Girl*³⁵

Suppose a fourteen-year-old girl decides to get pregnant. Of course, most pregnancies of fourteen-year-old girls are not planned. They happen because very young girls often do not have access to or take responsibility for birth control, or are in denial about the possibility of becoming pregnant. Often pregnancies occur as the result of relationships with men who are considerably older, where even the sexual relationship itself may have been imposed. In such circumstances, it is far from clear that the girl is fully responsible for becoming pregnant. Calling her responsible for the pregnancy, or blaming her for having a child, may seem harsh or misplaced. So let us imagine the unlikely scenario where the pregnancy is deliberate. She has not been pressured to have sex, and she has access to contraception. She decides not to use contraception, in order to have a baby, because of the prestige this would

confer in her middle school. Let us imagine (even more unlikely!) that she attempts to justify her plan to her mother. She says, "What's wrong with wanting a baby? Why are you so opposed?"

Some mothers might respond that it is just wrong to have sex outside marriage. But suppose this mother is not especially conservative about matters of sex. She would like her daughter to wait until she is older to have sex, but even if she cannot persuade her to wait to have sex, she certainly will try to persuade her to delay having a baby. One reason focuses on the hardships her daughter will likely experience. Pregnancy imposes severe health strains on, and is more likely to have complications in, girls who are not yet full-grown themselves. Moreover, she may have to drop out of school to care for the baby, thus limiting her opportunities. In any event, having a baby will deprive her of her adolescence, of spending time with friends, going out at night, etc. If these self-regarding reasons do not persuade her daughter she ought to wait, she should think about the effect on her parents who will probably end up shouldering a great deal of the burden of child care long before they are ready to be grandparents.

Finally, her mother will undoubtedly talk about the impact on the child of being born to such a young mother. Babies born to very young girls are more likely to be very low birth weight, and to be at greater risk of complications such as infection, respiratory distress syndrome, neurological problems, gastrointestinal problems, and sudden infant death syndrome.³⁶ The child is more likely to grow up in poverty, without a father, and may suffer the associated disadvantages, such as truancy, trouble with the law, and increased risk of drug and alcohol abuse. In addition, few girls of fourteen have the experience and maturity to be good mothers. Her mother might say, "How can you have a baby? For heaven's sake, you couldn't take proper care of your cat, remember? You never remembered to feed it or change the kitty litter. If you don't care about your own future or our feelings, think of your baby. Having a baby at your age is just not fair to the baby."

The girl might respond by refuting the charges of irresponsibility and inability to care for the child. "That was two years ago!" she might say. "You never give me credit for being responsible." Perhaps she has a point. At least in other cultures and other times, fourteen-year-old girls can be good mothers. Shakespeare's Juliet was deemed quite old enough to become a mother, and the same is true in many developing countries, where five-year-olds are entrusted with the care of their younger siblings.³⁷

But suppose that the girl does not defend her decision by claiming greater responsibility and maturity than her mother supposes she has. Instead she denies that this is morally relevant. For while she is emotionally immature, she is very bright, and has read Parfit, Harris, and Robertson. She says, "You're undoubtedly right. Given my youth and immaturity, I probably will not be the ideal mother. I might even be neglectful. I agree that I'd be a better mother to a child I'd have when I'm older and more mature. But *so what?* I can't do better by *that* baby. If I wait until I'm twenty and a better mother, I'll have a *different* baby. How can you say that my having a baby now would be unfair to the child?" She has a point. It is extremely unlikely that the nonexistence condition would be met, or even that the child's life would fall

below a decent minimum. How, then, can we justify the strong intuition that almost everyone has that it would be wrong deliberately to conceive a child at such a tender age, and not just wrong, but a wrong to the child?

James Woodward explains the unfairness in terms of failing to live up to the parental obligations the girl will have to her child.

Alma [the name he gives the 14-year-old girl] will—perhaps unavoidably—fail to give the child the love and affection it requires, or will fail to appreciate the importance of giving it certain kinds of training and education, or will be impatient with the child’s demands and physically or emotionally abuse it, or will neglect her child’s nutritional or medical needs. I claim that, when this is the case, Alma is roughly in the position of someone who is considering making a promise about an extremely serious matter which she has good reason to expect she will be unable fully to keep. . . . If Alma has her child and fails to meet the duties and obligations she owes to her child, the child has a complaint against her, based on a wrong done to the child.³⁸

Woodward’s claim that Alma wrongs her child is based on her failure to live up to certain duties and obligations she has to her child, and not on the claim that the child, once born, will be miserable or prefer nonexistence. This expresses the idea that it is possible to wrong someone, or treat that person unfairly, even if he or she is, on balance, better off as a result. His approach has considerable appeal, but it raises the question just what one’s obligations and duties toward a future child are. To put it another way, how good a parent would one have to be, or predict that one would be, to avoid failing in one’s obligations?

Where the choice is between having a child and not having any child, as in the case of the postmenopausal or HIV-positive woman, we might set the bar relatively low, that is, at the decent minimum standard. Where the choice is between having *this* child and a later child to whom one would be a better parent, one might argue for setting the bar somewhat higher. That is, one is not morally required to delay procreation until one can provide the best possible care or be the best parent one could be. That seems unduly perfectionist, as well as practically unrealizable. (If I have a child at age twenty-five, I’ll be physically more energetic than I would be at forty, but I might have more experience and patience at forty. When will I be “the best parent I could be?”) A more plausible view is the more modest obligation not to have a child until one will be able to be a “good enough” parent.³⁹

A difficulty with Woodward’s explanation of the wrong to the child is that it is limited to cases where the prospective parent will be unable to fulfill her obligations and duties to the child. What if the prospective parent could be a good enough parent, but has the option of having a different child in better circumstances? Is there an obligation to have the “better-off” child, and if so, can this be explained in terms of a rights-violation, or unfairness to the child that gets born? This is very problematic, as we see in the following pair of examples, which I have adapted from Derek Parfit and Dan Brock⁴⁰

Angela and Betty

Angela is pregnant. Her doctor discovers that she has a condition that will result in mild retardation in her baby. The doctor prescribes a medication that will prevent the retardation. But Angela does not want to take the medication, because a side effect of the medication is

that it can cause mild acne. So she does not take it and, as predicted, her baby is born mildly retarded.

Betty wants to get pregnant. However, she is on medication that has the following side effect: if she gets pregnant while on the medication, her baby will be born mildly retarded. Going off the medication is not a feasible option, as it would adversely affect her health as well as her fertility. Fortunately, she only needs to take the medication for a few months. Her doctor advises her to wait to get pregnant until she is off the medication. But Betty does not want to wait. She plans to visit her family during her summer vacation, and so she wants to have the baby in June at the latest. She gets pregnant right away and has a baby in June who, as predicted, is born mildly retarded.

Most people would regard both Angela and Betty as having acted wrongly. I certainly do. Both give birth to a mildly retarded child, when this easily could have been prevented, and for reasons that are morally trivial. Morally, there seems to be no difference between what Angela does and what Betty does. Those who agree accept the “No-Difference View.”⁴¹

However, as I argued earlier, there is a difference in the two cases, a difference that ordinarily would affect our judgments of wrongdoing. The difference is that Angela, but not Betty, has harmed her baby. By not taking the prescribed medication, Angela has caused her baby to be born retarded, when he could have been born with normal intelligence. She has caused him to be worse off than he otherwise would have been, which is the ordinary straightforward conception of harming. But the same is not true of Betty. She has not made her baby worse off than he would have been, or could have been. There is no way that the child she had could have been born with normal intelligence. There was nothing Betty could do to make *him* mentally normal. Admittedly, by waiting until she was off the medication, Betty could have avoided having a child who was mildly retarded. John Harris thinks this is enough to say that Betty has harmed her child, but it is hard to see why. Nothing Betty did or could have done could have prevented mental retardation in the child born in June. Waiting would have enabled her to have a child with normal intelligence, but it would have been a *different child*, one conceived from a different egg and a different sperm.

Some disability rights advocates would argue that *neither* Betty nor Angela harms her child because they reject the idea that it is possible to harm a child by causing or allowing him to be born retarded. This is because they dispute the view of disabilities generally as medical problems or as inherently disadvantageous. Instead, they believe that disabilities, including mental retardation, are largely socially constructed, and become a disadvantage, or a handicap, when the world is not organized to facilitate the abilities of the “differently abled.” Mental retardation is not a harmful or disadvantageous condition unless society chooses to make it so.

This socio-political model of disability has some truth in it. It is possible to make changes in society to enable people with certain disabilities to have access to a range of opportunities from which they were previously barred. Wheelchair ramps are a good example. At the same time, not all disabilities are alike, and the claim that disability is completely, or even mostly, a social construction is surely an exaggeration. While society can do a lot to improve the opportunities of those with developmental disabilities, there will always be opportunities foreclosed to them because of their disability. This is a reason to view mental retardation as a harmful condition,

and one to be avoided, if possible. This is entirely consistent with recognizing the worth of individuals who have disabilities, including developmental ones, and their contributions to their families, friends, and the world around them.⁴²

If mild mental retardation could be seen as making the child's life fall below a decent minimum, we could argue that Betty, as much as Angela, harms her child. But I deliberately chose mild mental retardation because it falls above that standard. Individuals who are mildly retarded can go to school, make friends, get jobs, and generally have lives that are well worth living, even if limited in various ways. This being the case, we cannot say that Betty has harmed her baby. Nor will Betty fail to fulfill her obligations to the child, like the 14-year-old girl would. So Betty cannot be seen as wronging the child. On what ground, then, can we say that Betty acts wrongly?

Melinda Roberts rejects the view that Betty has done anything wrong; that is, she rejects the No-Difference View. Betty, Roberts says, has not harmed her baby; in fact, given the absence of any third, better alternative for the baby, she's done the best she could by him. This enables Roberts to retain the person-affecting restriction (PAR) in her theory of morality, but at the high price of a completely implausible judgment.

Is there a way to retain the PAR while maintaining that Betty acts wrongly, indeed, just as wrongly as Angela? David Wasserman thinks this is possible. He writes:

For me, the intuitive difference between choosing to have a child with a given impairment rather than 1) no child or 2) a child without that impairment is best explained by the fact that the parent has a good reason for the choice in 1)—it's the only child she can have—but no obvious reason in 2)—why not wait? Once a reason is supplied in 2), e.g., the mother wants to let her ailing parents get acquainted with their first grandchild, which they will not be able to do if she waits the year necessary to avoid an impairment—the question is whether that is a good enough reason, which may be debatable. But if it is, it justifies rather than excuses her decision to have a child sooner—neither the child she has nor anyone else is wronged, nor does the mother act wrongly in any sense, by acting on a decision made for reasons that are respectful of the future child and compatible with the kind of relationship she seeks to establish with it.⁴³

Does this explanation accord with the view that Betty acts wrongly? First, note that in this example, the waiting period is one month, not one year. That is important because while it might impose a significant burden to wait a whole year before having a child, it is hard to see how waiting one month could impose a significant burden, thus giving the prospective mother a good (or good enough) reason not to wait. But second, and more important, the appeal to "good enough reasons" demonstrates the need for impersonal reasons in the morality of beneficence. To Wasserman's question, "Why not wait?" Betty has an answer. She plans to visit her family during her summer vacation, and so she wants to have the baby in June at the latest. Ordinarily, that would be a good enough reason for not wishing to delay conception. No one would blame Betty or think the worse of her for timing her birth to fit into her summer plans. So the question is, why isn't this a "good enough" reason here? Once we acknowledge that Betty hasn't harmed or wronged anyone by having the child with mild mental retardation, the demand for a better reason cannot

be justified in person-affecting terms. It seems that either we have to give up the judgment that Betty's act is wrong—or give up the PAR.

In giving up the PAR, we acknowledge that not all wrong acts are bad for someone. We give up the requirement that there must always be a victim of a wrongful act. We could instead adopt the following principle: other things being equal, it is wrong to have a child in a harmful condition if it is possible to have a different child without the harmful condition. Philip Peters calls this the principle of avoiding harm by substitution, or the substitution principle, for short.

8.7.2 The Substitution Principle

The substitution principle says that when individuals have a choice, they “should choose to bear the child who is likely to suffer the least.”⁴⁴ While I accept the intuition that the moral requirement is to avoid “gratuitous suffering,” that is, suffering that could have been avoided, to insist that individuals must choose the child who will “suffer the least” appears unduly perfectionist. It appears to require individuals to have the healthiest, happiest children they possibly could have, and to make procreation, which falls short of this ideal, morally wrong. I suggest instead this modification of the substitution principle:

Individuals who face reproductive decisions are morally required not to bring into the world children who will experience serious suffering or limited opportunity or serious loss of happiness, if this outcome can be avoided, without imposing substantial burdens or costs or loss of benefits on themselves or others, by bringing into the world different individuals who will be spared these disadvantages.⁴⁵

This principle is an impersonal principle. It is not person-affecting in that the failure to substitute does not harm any individual, or make anyone worse off, even in the appropriate counterfactual sense. There is no victim of a failure to substitute. And yet, as Peters reminds us, there is a sense in which the substitution principle is person-affecting: namely, that it is based on the badness of avoidable human suffering and limited opportunity. Concern to prevent human suffering can be seen as person-affecting, in a sense, because, as Dan Brock notes, “suffering and limited opportunity must be experienced by some person—they cannot exist in disembodied form. . . .”⁴⁶ Jonathan Glover makes a similar point when he says that comparative impersonal principles, that is, those that compare amounts of suffering in the world, are “rooted in people and their lives, rather than derived from mere abstract rules.”⁴⁷ This makes the incorporation of comparative impersonal principles into our morality more palatable than it otherwise might be.

8.7.3 Avoiding Harm by Substitution in the Real World

The examples of Angela and Betty are philosophically interesting, but highly artificial. Most people faced with a risk of disability in their offspring cannot avoid it simply by delaying conception for a few months. The details matter in determining whether one would be violating the substitution principle.

Consider prenatal testing and selective abortion. This can be viewed as a substitution method, since if the fetus is found to have a genetic or chromosomal disorder, an abortion gives the couple the chance to “try again” in a future pregnancy to have another child who will not have the disease. This is, of course, unacceptable to those who oppose abortion and regard fetuses as the moral equivalent of born children. Even those who are generally pro-choice are likely to find abortion of a wanted child, especially if this is done in the second or third trimester, psychologically and morally troubling. Because abortion can impose emotional burdens on the procreating woman or the couple, it is not required by the substitution principle, which only requires substitution if it can be accomplished without the imposition of substantial burdens. Moreover, undergoing amniocentesis increases the risk of miscarriage, which is another perfectly good reason for being unwilling to undergo it.

Another method of substitution is provided by preimplantation genetic diagnosis (PGD) and embryo selection. In PGD, embryos are created *in vitro*. A single cell is removed from each embryo and tested for genetic disease. Affected embryos are discarded, and only those that are disease-free are selected to be implanted in the uterus. This affects the identity of who will get born. Is there an obligation to undergo PGD by individuals at high risk of transmitting a genetic disease? I do not think that there is. For those who regard preimplantation embryos as having the moral status of human persons, PGD is no more morally acceptable than prenatal testing and selective abortion. If they have no obligation to abort, it is hard to see why they would have an obligation to discard embryos. Moreover, PGD requires IVF, which is expensive, often ineffective, and imposes both burdens and physical risks on the woman. For this reason, it is not required by the substitution principle.

The substitution principle might have applicability to certain procreative decisions, for example, the number of embryos to be implanted in an IVF cycle. In the United Kingdom, it was originally proposed that the number of embryos that may be transferred in any one cycle be limited to two, in order to reduce the incidence of multiple births, which have an increased risk of disability in the offspring. However, since this was likely to reduce live birth rates for older women, the policy that was ultimately adopted was a maximum of two embryos per cycle for women under forty, and three for women over forty. The Human Embryo and Fertilization Authority characterized this as “a reasonable balance between our overriding objective of reducing multiple births with the need to maximize a women’s chance of having a healthy singleton baby.”⁴⁸

In the United States, there is no central authority determining how many embryos can be transferred in any one cycle, although there are guidelines which specify that no more than two embryos should be transferred in women under the age of thirty-five, and no more than five in women over forty.⁴⁹ The American Society for Reproductive Medicine has announced that infertility treatment is moving closer to the goal of single embryo transfer, which “results in fewer multiple pregnancies, by far (although monozygotic twinning is possible) and when performed in the appropriate patient population results in cumulative pregnancy rates as good as those achieved with multiple embryo transfer.”⁵⁰ However, some couples express a desire to have two or more embryos transferred because having twins enables them

to “complete” their families with one round of treatment. They may be willing to take the increased risk of disability to accomplish this goal. However, their interests and goals are not the only relevant factors. Here, the substitution principle would seem to impose on couples a moral obligation to transfer only one or two embryos, as long as this gives them a reasonable chance of having a healthy singleton.

8.8 Conclusion

Genesis problems are challenging, but not insoluble. In particular, they do not require us to discard the common-sense intuition that the welfare of offspring is always a morally relevant consideration in procreative decision-making. It is wrong to have children who cannot have minimally decent lives, although reasonable people can disagree about what constitutes a decent minimum. Furthermore, such judgments should be based on a realistic assessment of the facts, not stereotypical thinking. In particular, it is important to remember that people can have lives that are well worth living, despite disabling conditions or poverty. Nevertheless, there are times when procreation is wrong, even though no one is harmed or wronged by birth. To explain these cases, we need to supplement a morality of person-affecting reasons with a comparative impersonal principle: the principle of substitution. This will explain some of the difficult cases, although it is often not easy to say when someone has an *obligation* to substitute. The morality of procreation, and the obligation to avoid procreation, is based partly on an objective assessment of the likely quality of the future child’s life, but also on the reasons, intentions and attitudes of those who would have children.

Acknowledgements I owe a debt of gratitude to more people than I can mention for helpful discussions of this topic over the years. In addition to those individuals mentioned in the endnotes, I would like especially to thank John Arras and Jonathan Glover.

Notes

1. Heyd (1992).
2. I create examples of morally problematic abortions in Steinbock (1999), pp. 245–267.
3. National Bioethics Advisory Commission (1997), p. ii.
4. The President’s Council on Bioethics (2005), p. 57.
5. Goldberg (2005), p. 14.
6. Most studies have found that children reared by gay or lesbian parents are psychologically normal, and that their parents’ homosexuality has not adversely affected their personality development, their gender identity, social interactions, or sexual preference. See Ethics Committee of the American Society for Reproductive Medicine (2006).
7. I use the term “person-affecting” because it is prevalent in the literature. I do not mean to imply that interests are limited to persons, or that morality concerns only the interests of persons. “Person-affecting” is simply more graceful than “interested-individual-affecting.”
8. There might be other reasons to prohibit such technologies or arrangements, including the interests of prospective parents or society at large. The issue here is whether the interests of the child justify preventing his or her birth.
9. Robertson (1994), p.122. Emphasis added.

10. Morgan (1991), p.194. Emphasis added.
11. I have addressed this issue in Steinbock (1986), pp. 15–20; Steinbock (1992), pp. 114–125; and Steinbock and McClamrock (1994), pp. 15–21.
12. Philip G. Peters argues that wrongful life cases belong only partly with the law of torts, which is relevant only to the establishment of negligence. It is fair to require the defendant to bear some of the burden of the extra expenses occasioned by the child's impaired condition because it was his or her negligence that has resulted in the child's being born. However, the payments should be seen, not as damages for harm, but rather as child support, which belongs to family law. By taking the justification of payments out of tort law, the question of whether the child has been harmed by birth—a requirement in torts—is avoided. Rather, the justification is that the negligent defendant has a responsibility to the child to help provide the resources that will enable the child to reach his or her full potential. See Peters (1992), pp. 397–454.
13. Harris (1992), p. 88.
14. There are two caveats here. First, particular causal claims, as opposed to statistical causal claims, are notoriously difficult to establish. Studies may establish a causal link between smoking in pregnancy and asthma in offspring, but that does not entitle us to say, in any particular case, that had the mother not smoked during pregnancy, this child would have been born asthma-free. Perhaps the child would have had asthma anyway. A genetic predisposition might have caused the asthma, even in the absence of maternal smoking. Or the child might have been exposed to smoke in utero from the father, or there could have been other environmental factors that could have induced asthma. Second, even if a causal connection can be established, the ascription of moral responsibility depends on whether the woman “could have stopped smoking.” Since smoking is often addictive, the woman's failure to stop may be less than fully voluntary, and her moral responsibility for the asthma might therefore be diminished. However, while the ascription of causal and moral responsibility is complex, it is possible to imagine a case in which it can be determined that the mother's smoking during pregnancy harmed her baby, and that this is a harm for which she is morally responsible.
15. I owe my colleague, Rachel Cohon, thanks for this example.
16. Although I find this conclusion to be totally counterintuitive, it is straightforwardly embraced by David Benatar in Benatar (2006).
17. Harris (2007), p. 94.
18. Actually, I prefer the expression “better off unconceived” to avoid the question of whether abortion would be in the child's best interest. However, Feinberg uses “better off unborn” so I will follow his usage here.
19. Robertson (1974–1975), p. 254.
20. This requirement is also placed on infants' surrogate decision makers in President's Commission on Ethical Problems in Medicine and Biomedical and Behavioral Research (1982).
21. Feinberg (1987), p. 164.
22. See, for example, the cover story in *Time Magazine* about Marissa and Anissa Ayala, “The Gift of Life,” June 17, 1991 <<http://www.time.com/time/covers/0,16641,19910617,00.html>> (accessed July 17, 2008). See also Steinbock (2008).
23. Melinda Roberts calls cases in which there are three alternatives, one of which is better than existence under a disadvantageous cases, “type 3-alt” cases. See Roberts (1998), pp. 92–96
24. Individuals with severe intellectual impairments may want to have a baby, unaware that they are not capable of caring for and raising a child. The fact that they are not aware of their own limitations in this respect does not entitle them to procreate. It falls to those responsible for them to prevent them from having children, both for their sake and for the sake of the child.
25. For an excellent treatment of the right of children to be loved, see Liao (2006), pp. 420–440. Liao focuses on the right of existing children to be loved, and how this human right has policy implications. He does not proclaim a duty on the part of individuals to refrain from procreation where they cannot or will not love the child, but it seems to me consistent with his view.
26. Kamm makes a similar point: “let us suppose that we should not create persons at will unless we have good reason to believe that they can have some—just how many is deliberately left

- open—number of years of life with some degree of health and welfare, and let us call these things that they should have the *minima*.” Kamm (1992), p. 132.
27. This position was taken in President’s Commission on Ethical Problems in Medicine and Biomedical and Behavioral Research (1982).
 28. Cohen (1997), p. 33.
 29. Archard (2004), pp. 403–420.
 30. Archard, *op. cit.*
 31. Steinbock (1998), pp. 13–14.
 32. Steinbock (1998), pp. 13–14.
 33. Harris (1992), p. 91.
 34. See the Universal Declaration of Human Rights: 1948–1998 <<http://www.un.org/Overview/rights.html>> (accessed July 28, 2005), and the Charter of Fundamental Rights of the European Union, Article 9 <http://www.europarl.eu.int/comparl/libe/elsj/charter/art09/default_en.htm> (accessed July 28, 2005).
 35. I have adapted this example from Parfit (1986), pp. 358–361.
 36. See Lucile Packard Children’s Hospital at Stanford, High Risk Newborn: Very Low Birthweight <<http://www.lpch.org/DiseaseHealthInfo/HealthLibrary/hrnewborn/vlbw.html>> (accessed August 1, 2005).
 37. Another possibility is to conceive this example as a “third option” kind of case. Melinda Roberts suggests that the alternatives are not life as a neglected child or no life at all. The child could be born and not be neglected if all the other agents who might affect the child’s life—his father, his parents, the young girl’s parents, and the community at large—help out. And if they do not, the blame for the child’s lack of well-being is not to be laid solely at her door. Should the girl refrain from having the child to prevent the others from wronging the child? Roberts thinks this question need not be resolved. She writes, “*Some agent* has wronged the child, according to personalism [her interpretation of the person-affecting restriction or PAR], and has, correspondingly, done something wrong. Someone remains, morally, on the hook for the wrong that has, by hypothesis, been done the child. To avoid the charge of an unconscionably loose moral standard—the charge that lies at the root of the fourteen-year-old girl objection—it seems that this result is all that is really required.” Roberts (1998), p. 111.
 38. Woodward (1986), p. 815.
 39. Bettelheim (1987).
 40. Parfit (1976); Brock (1995).
 41. Parfit (1986), p. 367. By the No-Difference View, I mean simply the claim that there is no *moral* difference between what Angela does and what Betty does. One does not act more wrongly than the other. The more generalized version of the No-Difference View holds that the wrongness of both acts must have the *same explanation*. Thus, if person-affecting reasons cannot explain the wrongness of Betty’s act, it cannot explain the wrongness of Angela’s act either. This leads Parfit to reject person-affecting reasons altogether in the area of morality concerned with beneficence and human well-being. For an excellent critique of the generalized version of the No-Difference View, see Jeff McMahan (2001). I agree with McMahan that both kinds of reasons, person-affecting and impersonal, are necessary in moral discourse, and that neither can be reduced to the other.
 42. I argued for this in Steinbock (2000), pp. 108–123.
 43. Personal communication from David Wasserman.
 44. Peters (1989), p. 515; Peters (2004), especially Chapter 4; Peters (2009).
 45. This is a simplification of a principle offered in Buchanan et al. (2000), p.249.
 46. Brock (1995), p. 399.
 47. Glover (1992), p. 142.
 48. See Human Fertilisation & Embryology Authority, Chair’s Letter CH(04)01a. <<http://www.hfea.gov.uk/HFEAGuidance/ChairsLettersArchive/2003–2004/CH0401a>> (accessed August 2, 2005).

49. See, for example, the Practice Committee of the Society for Assisted Reproductive Technology and the American Society for Reproductive Medicine (2004), pp. 773–774.
50. See the American Society for Reproductive Medicine Press Release on Single Embryo Transfer <http://www.asrm.org/Media/Press/single_embryo.html> (accessed August 2, 2005).

References

- American Society for Reproductive Medicine. Press release on single embryo transfer. 2 Aug. 2005. <http://www.asrm.org/media/press/single_embryo.html>.
- Archard, D. 2004. Wrongful life. *Philosophy* 79: 403–20.
- Benatar, D. 2006. *Better never to have been: The harm of coming into existence*. New York: Oxford University Press.
- Bettelheim, B. 1987. *A good enough parent*. New York: Alfred A Knopf, Inc.
- Brock, D. W. 1995. The non-identity problem and genetic harms: The case of wrongful handicaps. *Bioethics* 9(3/4): 269–275.
- Buchanan, A., D. Brock, N. Daniels, and D. Wikler, 2000. *From chance to choice: Genetics and justice*. Cambridge: Cambridge University Press.
- Charter of Fundamental Rights of the European Union, Article 9. 28 July 2005 <http://www.europarl.eu.int/comparl/libe/elsj/charter/art09/default_en.htm>.
- Cohen, C. B. 1997. The morality of knowingly conceiving children with serious conditions: An expanded “wrongful” life standard. In *Contingent future persons*, eds. N. Fotion and J. C. Heller. New York: Kluwer Academic.
- Ethics Committee of the American Society for Reproductive Medicine. 2006. Ethics committee report: Access to fertility treatment by gays, lesbians, and unmarried persons. *Fertility and Sterility* 86(5): 1333–1335.
- Feinberg, J. 1987. Wrongful life and the counterfactual element of harming. *Social Philosophy & Policy* 4: 164.
- The gift of life. 1991. *Time Magazine*. <<http://www.time.com/time/covers/0,16641,19910617,00.html>>.
- Glover, J. 1992. Future people, disability, and screening. In *Justice between age groups and generations*, eds. P. Laslett and J. Fishkin, 127–43. New Haven: Yale University Press.
- Goldberg, C. 2005. Experts debate age limits for childbearing; committee grapples with needs of kids vs. those of parents. *Chicago Tribune* 8 Apr. 2005: 14.
- Harris, J. 1992. The wrong of wrongful life. In *Wonderwoman and superman: The ethics of human biotechnology*, ed. J. Harris, 91. Oxford: Oxford University Press.
- Harris, J. 2007. *Enhancing evolution: The ethical case for making better people*. Princeton: Princeton University Press.
- Heyd, D. 1992. *Genethics: Moral issues in the creation of people*. California: University of California Press.
- Human Fertilization & Embryology Authority, Chair’s letter CH(04)01a. 2 August 2005 <<http://hfea.gov.uk/hfeaguidance/chairslettersarchive/2003–2004/ch0401a>>.
- Kamm, F. 1992. *Creation and abortion*. Oxford: Oxford University Press.
- Liao, M. 2006. The right of children to be loved. *The Journal of Political Philosophy* 14(4): 420–40.
- Lucile Packard Children’s Hospital at Stanford. Very low birthweight. 1 August 2005 <<http://www.lpch.org/diseasehealthinfo/healthlibrary/hrnewborn/vlbw.html>>.
- McMahan, J. 2001. Wrongful life: Paradoxes in the morality of causing people to exist. In *Bioethics*, ed. J. Harris. Oxford: Oxford University Press.
- Morgan, D., and R. G. Lee. 1991. *Blackstone’s guide to the Human Fertilization & Embryology Act 1990*.
- Parfit, D. 1976. On doing the best for our children. In *Ethics and population*, ed. M. Bayles. Cambridge, MA: Schenkman Company.

- Parfit, D. 1986. *Reasons and persons*. Oxford: Oxford University Press.
- Peters, P. G. 1989. Protecting the unconceived: Nonexistence, avoidability, and reproductive technology. *Arizona Law Review* 31: 487–548.
- Peters, P. G. 1992. Rethinking wrongful life: Bridging the boundary between tort and family law. *Tulane Law Review* 67(2): 397–454.
- Peters, P. G. 2004. *How safe is safe enough? Obligations to the children of reproductive technology*. New York: Oxford University Press.
- Peters, P. G. 2009. Implications of the nonidentity problem for state regulation of reproductive liberty. In this collection.
- The Practice Committee of the Society for Assisted Reproductive Technology, and the American Society for Reproductive Medicine. 2004. Guidelines on the number of embryos transferred. *Fertility and Sterility* 82: 773–4.
- Roberts, M. 1998. *Child versus childmaker: Future persons and present duties in ethics and the law*. Lanham, Maryland: Rowman & Littlefield.
- Robertson, J. A. 1974–1975. Involuntary euthanasia of defective newborns: A legal analysis. *Stanford Law Review* 27: 254.
- Robertson, J. A. 1994. *Children of choice: Freedom and the new reproductive technologies*. Princeton, NJ: Princeton University Press.
- Steinbock, B. 1986. The logical case for “wrongful life.” *Hastings Center Report* 16(2): 15–20.
- Steinbock, B. 1992. *Life before birth: The moral and legal status of embryos and fetuses*. Oxford: Oxford University Press.
- Steinbock, B. 1998. Rethinking the right to reproduce. *Working Paper Series Number 98.05* Harvard Center for Population and Development Studies: 13–14.
- Steinbock, B. 1999. Why most abortions are not wrong. In *Bioethics for medical education, Advances in bioethics*; v. 5, eds. R. B. Edwards and E. Bittar, 245–267. Stamford, CT: JAI Press.
- Steinbock, B. 2000. Disability, prenatal testing, and selective abortion. In *Prenatal testing and disability rights*, eds. E. Parens and A. Asch, 108–23. Washington, D.C.: Georgetown University Press.
- Steinbock, B. 2008. Using preimplantation genetic diagnosis to save a sibling: The story of Molly and Adam Nash. In *Ethical issues in modern medicine: Contemporary readings in bioethics*, eds. B. Steinbock, A. J. London and J. D. Arras. 8th ed. New York: McGraw Hill.
- Steinbock, B. and R. McClamrock. 1994. When is birth unfair to the child? *Hastings Center Report* 24(6): 15–21.
- Universal Declaration of Human Rights: 1948–1998. 28 July 2005 <<http://www.un.org/overview/rights.html>>.
- Woodward, J. 1986. The non-identity problem. *Ethics* 96: 804–831.

Other References

- National Bioethics Advisory Commission. 1997. *Cloning human beings: Report and recommendations v. 1, 2*. Washington, D.C.
- The President’s Council on Bioethics. 2005. *Alternative sources of human pluripotent stem cells: A white paper*, 57. Washington, D.C.
- The President’s Commission on Ethical Problems in Medicine and Biomedical and Behavioral Research. 1982. *Deciding to forego life-sustaining treatment: A report on the ethical and legal issues in treatment decisions*. Washington, D.C.: U.S. Government Printing Office.

Chapter 9

Harming and Procreating

Matthew Hanser

Abstract Suppose that A harms B if and only if he stands in some appropriate relation to B's suffering of some harm. Such an account of harming can explain how it is that those who choose Derek Parfit's risky energy policy harm the future people later exposed to radiation, but can it also yield the result that parents who conceive a disabled child, when they could easily have conceived a different, non-disabled child instead, thereby harm the child they actually conceive?

Keywords Harm · Non-identity problem · Wrongful life · Derek Parfit · Seana Shiffrin · Elizabeth Harman.

9.1

Derek Parfit argues that a choice does not harm someone in a morally relevant sense if that person has a life worth living and would not have existed had the choice not been made.¹ This creates a puzzle regarding choices that affect the identities of future people. Consider two examples:²

The Risky Policy. We must choose between two long-term energy policies, one of which involves the burial of radioactive waste in an area that we know will become earthquake prone in the distant future. We choose this policy, and three centuries later an earthquake releases radiation that injures or kills thousands of people. Because this policy also raises the standard of living, the people injured or killed by the leaked radiation would never have existed had we chosen a safer policy. (Standards of living have trivial but cumulative effects on people's daily lives, eventually affecting when and with whom people have sex. A choice that affects the standard of living thus affects the identities of future people.)

The Disabled Child. A couple knows that because the man carries a genetic defect any child they conceive in the ordinary way will suffer from a fairly serious disability. They also know that their doctor can isolate a sperm from the man, genetically alter it to eliminate the defect, combine it *in vitro* with one of the woman's eggs, and implant the resulting zygote in the woman's uterus. Although they could easily afford this procedure (perhaps it is covered by

M. Hanser (✉)
University of California, Santa Barbara, CA, USA
e-mail: hanser@philosophy.ucsb.edu

their insurance), they choose to conceive a child unaided. Had they opted for the *in vitro* procedure they would almost certainly have had a different child, since the sperm and egg used in that procedure would almost certainly have been distinct from the ones that gave rise to their actual child.

If Parfit is right, our choice of the risky policy does not harm the people injured or killed by radiation, since they have lives that are on balance worth living and would not have existed had we chosen a safer policy. Likewise, the couple's choice to conceive unaided does not harm their child, since he too has a life worth living and would not have existed had they selected the *in vitro* procedure. Why, then, are these choices morally objectionable? Parfit calls this the *non-identity problem*, since it arises because the people who live short, unhealthy, or disabled lives as a result of these choices are distinct from the people who would have lived longer, healthier, or disability-free lives had the choices not been made. Parfit argues that we cannot give a *narrow person-affecting* account of what's objectionable about these choices—we cannot give an account that appeals only to the choices' (supposed) bad effects on people who actually ever live. In each case the objection must be either (i) that the choice's effects on the people it brings into existence are not as good as the effects that some other choice would have had on the people *it* brought into existence, or (ii) that the choice produces a worse overall state of affairs, impersonally conceived, than some other choice would have produced.

In an earlier paper (1990) I argued that a choice *can* be objectionable owing to its bad effects on someone who has a life worth living and who would not have existed had the choice not been made. I argued further that on a correct analysis of what it is for a choice to harm someone, the choice of the risky policy *does* harm the people injured or killed by leaked radiation. But I was reluctant to extend this analysis to cases like that of the disabled child.³

More recently, Seana Shiffrin (1999) and Elizabeth Harman (2004) have presented accounts of harming structurally quite similar to my own. But they take their accounts to apply quite straightforwardly to procreative choices: neither would hesitate to say that the parents of the disabled child harm him by bringing him into existence. So who is right here? Was my reluctance to extend my account of harming to cases like that of the disabled child misplaced, or is there a genuine problem here, the force of which Shiffrin and Harman have failed to appreciate? That is the question I wish to explore in this paper.

Here's another way of framing the issue. I've described two cases, that of the risky policy and that of the disabled child. Parfit thinks that in each case there is an obstacle (the same obstacle) to saying that the choice harms someone. Shiffrin and Harman think that in neither case is there an obstacle to saying this. I think there is no obstacle in the case of the risky policy but that there may well be one in the case of the disabled child. (Obviously I must disagree with Parfit about what this obstacle is—it can't simply be that the child has a life worth living and would not otherwise have existed.) For the purposes of this paper I shall assume that I am right and Parfit wrong about the risky policy; I don't want to reargue that here. The question addressed in this paper is whether there is some obstacle to saying that the parents harm their child *other* than the one (supposedly) identified by Parfit.

9.2

Let's begin with Parfit's obstacle and why I think it is illusory. The obstacle is created by the claim that an agent does not harm someone by acting in a certain way if that person has a life worth living and would not have existed had the agent acted differently. (Parfit prefers to speak of a choice's harming someone rather than of an agent's harming someone by acting in a certain way, but I do not think that anything hangs on this.) Parfit defends this claim by appealing to a certain general account of harming. Put roughly, the account is this:

A harms B by acting in a certain way if and only if B would have been better off had A acted in some other way.

For present purposes, the "only if" direction is the one that matters: A harms B by acting in a certain way *only if* B would have been better off had A acted in some other way. This is why Parfit thinks we do not harm the people injured or killed in the nuclear catastrophe. These people have lives worth living and would not have existed at all had we chosen some other policy. They consequently would not have been better off had we chosen some other policy.

There are a number of problems with this account of harming. For one thing, it should probably be revised to say that A harms B by acting in a certain way if and only if B would have been better off *in some respect* had A acted in some other way. An agent can harm someone in a morally relevant sense even though he also benefits him, and even though the benefit outweighs the harm. In such a case the agent's action is still *pro tanto* objectionable; it may even be objectionable all things considered. (*Pro tanto* objectionable acts are ones that stand in need of justification; in the absence of justification they are objectionable all things considered.) But this revision does not suffice to undermine Parfit's claim about the risky policy, for there is arguably *no* respect in which the people exposed to radiation would have fared better had we chosen some other policy. In order for there to be a respect in which someone is better off in one scenario than in another, he must have some level or other of well-being in each scenario. And in order for a person to have a level of well-being in a scenario, he must exist in that scenario. (A person who does not exist does not have a "neutral" level of well-being.)

The real problem with Parfit's account of harming is deeper than this. The account is defective because it attempts to analyze the notion of harm all in one go, so to speak. We should instead adopt a two-step account. The first step is to analyze the notion of *harming* (i.e. of *doing harm*) in terms of the prior notion of someone's *suffering harm*; the second step is then to say what it is for someone to suffer harm. Schematically:

Step One: A harms B by acting in a certain way if and only if A's acting in that way results in his standing in relation R to some harm suffered by B.

Step Two: B suffers harm if and only if ____.

To transform this account-schema into a substantive account of harming we must do two things: specify relation R, and fill in the blank in the account of what it is to

suffer harm.⁴ In my earlier paper I did not offer a substantive account of harming. I gave relation R a name—I said that A harms B if and only if his actions make him at least partially *responsible* for a harm suffered by B—but I did not attempt to spell out the conditions for an agent's being at least partially responsible for a harm.⁵ Nor did I offer an account of what it is for someone to suffer harm. Rather, I argued that on *any* reasonable account of what it is to suffer harm, the people exposed to radiation thereby suffer harm, and that on *any* reasonable account of what it is for an agent to be responsible for someone's suffering harm, we who adopt the risky policy are responsible for the harms these people suffer. The harms they suffer, after all, are the predictable result of the release of radiation; and the release of radiation is the predictable result of our burying radioactive waste in an area which we know will eventually become earthquake-prone. On any reasonable way of fleshing out the two-step account of harming, then, we who adopt the risky policy harm the people later exposed to radiation.

9.3

Since our choice of the risky policy predictably harms people, it is *pro tanto* objectionable. But it does not follow that the choice is objectionable all things considered. To secure this stronger result we need to show that the choice is not an instance of *justified* harming. And a case can be made that it is. Despite the harms they suffer, life is on balance good for the people exposed to radiation. Since they would not have enjoyed this good but for our choice of the risky policy, we benefit them by choosing this policy. And since the benefit we bestow upon them is greater than the harm we cause them to suffer, our choice is not objectionable all things considered. Or so the reasoning goes.

Shiffrin, Harman and I all reject this line of reasoning, but our arguments are different, and it is worth pausing briefly to examine the differences. In my earlier paper I raised two objections.⁶ (In that paper I did not, and in this section I do not, question the assumption that being conceived can come to one as a benefit.) First, I expressed doubt about the assumption that if a choice brings someone more benefit than harm, then the harm-based *pro tanto* objection to the choice is automatically overridden. But instead of developing and defending this doubt, I concentrated on a second objection. I argued that by choosing the risky policy we do not, in a morally relevant sense, benefit the people who would not otherwise have existed. My claim was really that we cannot use the benefits they receive to *justify* our choice. We cannot do this because only *intended* benefits can be used to justify harming; and I argued that when we choose the risky policy we cannot plausibly be taken to intend the benefits received by those who will later owe their existence to our choice. But while I think it is defensible that only intended benefits can be used to justify acts of harming, this is a claim that many would reject. Dialectically speaking, then, I was perhaps unwise to base my argument upon it.

Harman's argument is simpler.⁷ She argues, in effect, that since by stipulation we will benefit an equal number of future people whichever policy we choose,

the benefit-based reasons in favor of choosing the competing policies cancel out, leaving the harm-based *pro tanto* objection to choosing the risky policy intact. This argument is unsatisfying. First, it can be deployed only with respect to same number choices—choices where the same number of people will ever live whichever choice is made. If a choice will affect the number of future people, then the benefit-based reasons favoring the competing options do not completely cancel.⁸ More importantly, the argument does not accurately reflect the way justification works. Suppose that I can save one or another of two people from death, but not both. And suppose that while I can save the first person without harming him, I can save the second only at the cost of breaking his arm. Clearly it would be permissible for me to save the second person. The benefit I can bestow upon him only by breaking his arm *justifies* me in breaking his arm, which is to say that it justifies me in harming him. This justification is not undermined by the fact that I could have bestowed an equivalent benefit upon someone else without doing harm. If Harman's reasoning were sound, saving the second person by breaking his arm would be impermissible: the benefit-based reasons in favor of saving each of the two people would cancel out, leaving the harm-based *pro tanto* objection to harming the second person intact.

Shiffrin has perhaps the most persuasive argument why the harm-based *pro tanto* objection to choosing the risky policy is not overridden.⁹ She distinguishes two kinds of benefit. Someone receives what we might call a *preventative* benefit when he is prevented from suffering harm. (For Shiffrin, to suffer harm is to be in a certain sort of noncomparatively bad state. I shall say more about this in the next section.) Someone receives a *pure* benefit, by contrast, when he receives a positive good—a good that does not consist in the absence or prevention of something bad. When someone is prevented from falling down a manhole, he receives a preventative benefit; when someone who is comfortably well off wins the lottery, he receives a pure benefit. Shiffrin argues that it is generally permissible to harm someone, even without his consent, in order to prevent him from suffering a greater harm. It is permissible, that is, to harm someone, even without his consent, in order to bestow upon him a more-than-compensating preventative benefit. That is why it is permissible to break a person's arm (even without his consent) when that is the only way to save his life. But Shiffrin argues that it is not generally permissible to harm someone, without his consent, merely to bestow upon him a more-than-compensating pure benefit. We cannot, for example, break someone's arm, without his consent, in order to help him win the lottery. Now if someone with a life worth living receives a benefit by being brought into existence, that benefit is pure, not preventative. (Coming into existence does not prevent one from suffering the "harm" of nonexistence.) We consequently cannot justify harming someone by pointing out that he would never have existed had we acted differently.¹⁰ (As a number of authors have argued, that the person would not later regret our action is neither here nor there.¹¹)

Shiffrin's argument applies both to same number choices and to different number choices. It also applies both to our choice of the risky policy and to the parents' choice to create a disabled child. Suppose her argument is sound. Then *if* the parents harm their disabled child by bringing him into existence, they cannot claim that the

harm-based *pro tanto* objection to their choice is overridden by the fact that he would not otherwise have received the benefit of existence. Consequently, if there is an asymmetry between our choice of the risky policy and the parents' choice to create a disabled child, it must consist in this: that while we harm people by choosing the risky policy, the parents do *not* harm their disabled child by bringing him into existence.

By contrast, if my argument is sound and Shiffrin's is not, there are potentially two asymmetries between these choices: perhaps the parents do not harm their child by bringing him into existence; and if they do, perhaps that harming is justified. It would be justified if existence were a benefit that the parents *intended* to bestow upon their child.¹² But I shall set the question of justification aside for the remainder of this paper. I shall concentrate solely on the question whether the parents of the disabled child harm him by bringing him into existence.

9.4

Recall that according to the two-step account of harming, A harms B by acting in a certain way if and only if A's acting in that way places him in relation R to some harm suffered by B. If there is an obstacle to saying that the parents harm their disabled child, then, it is either because there is an obstacle to saying that their child suffers harm or because there is an obstacle to saying that the parents' action places them in relation R to whatever harm he suffers. Now I claimed in Section 9.2 that on any reasonable account of what it is to suffer harm, the people exposed to leaked radiation suffer it, and that on any reasonable account of relation R, we who choose the risky policy stand in relation R to the harms they suffer. That is why Shiffrin, Harman and I can agree that the choice of the risky policy harms people without having first to agree upon how exactly the two-step account is to be fleshed out. But things are different in the case of the disabled child. Whether the child suffers harm depends upon which particular account of harm we adopt; and whether the parents stand in relation R to that harm, assuming that it exists, depends upon how exactly we understand relation R. In this section I shall consider whether the disabled child suffers harm. In subsequent sections I shall address the question whether the parents' actions connect them in the right way to that harm, assuming that it exists.

Of course like everyone else, the disabled child can be expected to suffer numerous harms over the course of his life. He'll suffer harm, for example, when he's beaten up by a playground bully in the fourth grade. I shall consider later whether parents, simply in virtue bringing their children into existence, thereby become at least partially responsible for all the harms their children ever suffer. But the example of the disabled child is meant raise a narrower question: do the parents harm their child by bringing him into existence *with a disability*? The question to be addressed in the present section is thus whether the disabled child suffers harm simply in having, or in coming into existence with, a disability.

The most widely accepted account of harm (call it the *standard* account) relies upon a counterfactual comparison:

A person suffers harm if and only if there occurs an event such that he would have been better off in some respect had it not occurred.

This account of harm must be distinguished from Parfit's superficially similar account of *harming*. According to Parfit, A harms B by acting in a certain way if and only if B would have been better off (in some respect) had A acted differently. The risky policy can be used to illustrate the difference between the two accounts. According to Parfit's account of harming, we do not harm anyone by choosing the risky policy, because the people exposed to radiation would not have been better off (in any respect) had we acted differently. But according to the standard account of harm, the people exposed to radiation do suffer harm: they would have lived longer, healthier lives had the event of their exposure to radiation not occurred.¹³

So let's apply the standard account of harm to the case of the disabled child. First we must distinguish between, on the one hand, the *actions* that may or may not place the parents in relation R to some harm suffered by their child, and, on the other hand, the *event* (whatever it might be) that may or may not come to their child as a harm. Perhaps this event is the child's conception. Of course we sometimes speak as if a child's conception were an action, but strictly speaking it's a "mere" event, consisting (let us suppose) in the joining together of two gametes. This event may be caused by actions performed by the parents, but it is not itself an action. Does the event of the disabled child's conception come to him as a harm? According to the standard account, it does not. Had the child not been conceived, he would not have lived a life free of disability—he would not have been in a better state with respect to the ability in question. Rather, he would not have existed at all. And we may imagine the details of the example to be such that there is *no* event such that had it not occurred, that very same child would have been born free of disability. According to the standard account, then, the child suffers no harm with respect to his disability.

But what if we accept some other account of harm? Shiffrin, who argues vigorously against comparative models of harm, proposes that harms be identified with "certain absolute, noncomparative conditions (e.g. a list of evils like broken limbs, disabilities, episodes of pain, significant losses, death) . . ."¹⁴ To suffer harm, according to Shiffrin, is not to be in a comparatively bad condition; it is to be in a certain sort of *absolutely* bad condition. It is not to be worse off than one was before, or than one otherwise would have been; it is simply to be *badly* off in a certain way. And she offers an account of what unifies the various items on her list: harms are "conditions that generate a significant chasm or conflict between one's will and one's experience, one's life more broadly understood, or one's circumstances."¹⁵ The details of her account, however, needn't concern us. The important point is that on her account, the child's disability is itself a harm. That he has always had this disability and that he could not have existed without it are irrelevant to the question whether he suffers harm in having it.

Harman agrees that disabilities are harms even for those who have always had them and who could not have existed without them. According to Harman, states of ill health, bodily damage and bodily deformity count as harms because it is worse to be in one of these states than to be in a healthy, normal state. The states on her list are comparatively bad, but the states to which they compare unfavorably are those of a healthy, properly functioning organism of the relevant species. It doesn't matter whether the individual suffering harm was ever in a position to attain this ideal. On Harman's account as well, then, the disabled child suffers harm simply in being disabled.¹⁶

I do not accept the standard account of harm, but nor do I accept Shiffrin's or Harman's account. On my view, harms paradigmatically consist in losses of "basic goods." The power of sight, for example, is a basic good for a human, so *losing* the power of sight is a harm. Merely *lacking* the power of sight, by contrast, is not a harm.¹⁷ If this is right, then someone who comes into existence lacking the power of sight has not thereby suffered harm: one can't lose what one has never had. Likewise with other disabilities. So here there is indeed a difference between my view and the views of Shiffrin and Harman. They think the disabled child suffers harm simply in being disabled, and this leaves room for the possibility that his parents harm him by bringing him into existence. On my view, the disabled child does not suffer harm simply in being disabled.¹⁸ Nor does he suffer harm in coming into existence with a disability. There is thus no room to say that the parents harm him by bringing him into existence with a disability.

But this difference in our views is more verbal than real.¹⁹ While I do not think the child suffers harm in being disabled, I do think his disability is *bad* for him—his disabled state is one that it is bad for him to be in. And I think that if his parents' actions place them in relation R to his disability, then their actions are *pro tanto* objectionable.²⁰ Shiffrin, Harman and I thus agree that if the parents' actions relate them in the right way to their child's disabled state, their actions are *pro tanto* objectionable. We differ only in whether we would in that case describe the parents as having *harmed* their child. For the sake of argument, then, I shall for the remainder of this paper accept Shiffrin's and Harman's description. I shall grant that disabilities are harms.

9.5

Shiffrin and Harman hold that a correct understanding of harming will reveal that the parents' actions satisfy the very same sufficient condition for harming that paradigm ordinary acts of act of harming satisfy. (By "ordinary" acts of harming I mean ones where the victim would have existed even if the action had not been performed.) The parents' actions may differ from ordinary acts of harming in certain respects, but the differences are irrelevant: it's what they all have in common that makes them harmings. That, at any rate, is the idea. Now according to the two-step analysis of harming, A harms B by acting in a certain way if and only if A's acting in that way places him in relation R to some harm suffered by B. I have granted, for the sake

of argument, that the disabled child suffers harm in being disabled. So finally we arrive at the crux of the matter. Do the parents' actions place them in relation R to their child's disability?

In order to answer this question, we need to know what conditions are sufficient for the obtaining of relation R. Shiffrin tells us that her argument begins with "the (contested) assumption . . . that being created can harm a person," and she explains in a footnote that the basic idea behind this assumption is that "by being the direct cause of a person's being in a situation that intrinsically delivers harm, one harms that person. . . ." ²¹ Likewise, Harman writes that "[a]n action harms a person if the action causes pain, early death, bodily damage, or deformity to her . . ." ²² Both, then, suggest that *causation* of harm is sufficient for harming. And both treat it as obvious that the parents of the disabled child cause his disability.

According to what is probably the most popular account, causation is the ancestral of the *counterfactual dependence* relation, with the relata being distinct, actually occurring events: event *e* depends counterfactually upon event *c* if and only if *e* would not have occurred had *c* not occurred; and *c* causes *e* if and only if the two events are connected by a chain of such counterfactual dependencies. ²³ (The chain might of course consist of a single link.) But this understanding of causation will not suit Shiffrin's and Harman's purposes, since they are interested in the causation of states, not the causation of events. (For both Shiffrin and Harman, harms are bad states or conditions.) Let us assume that a state is a thing's possession of a property. We may adapt the counterfactual account of causation accordingly:

β's being *G* depends counterfactually upon *α*'s being *F* if and only if *β* would not have been *G* had *α* not been *F*;

α's being *F* causes *β*'s being *G* if and only if the two states are linked by a chain of such counterfactual dependencies.

This account yields the desired result that the parents' bringing their child into existence causes him to be disabled: had they not brought him into existence, he would not have been disabled.

The parents harm their child in the morally relevant sense if (a) this account of causation is correct and (b) causation of harm is sufficient to generate a harm-based *pro tanto* moral objection. ²⁴ We could attempt to evaluate these two claims separately, but this would involve us in difficult, disputed questions about the nature of causation. ²⁵ I thus propose collapsing the two claims and evaluating the result directly. The collapsed claim is this:

If B's suffering some harm is linked, via a chain of counterfactual dependencies, to A's acting in a certain way, then A's acting in that way is *pro tanto* objectionable.

I shall argue that this claim is false. Perhaps this means that causation is not the ancestral of counterfactual dependence; perhaps it means that causation of harm is not sufficient to generate a *pro tanto* moral objection; perhaps it means both. We needn't decide.

The harm we've been discussing so far has been that of a child's being disabled. But being disabled is a *determinable* property. There are many dimensions along

which someone can be disabled, each corresponding to an aspect of proper functioning; and along each dimension there are a spectrum of ability levels, ranging from fully normal functioning (or even better than normal functioning) to a complete lack of functioning. Even a single power, such as sight, can be assessed along multiple dimensions. Let us choose a particular dimension of functioning for our example: the ability to discriminate detail visually at a distance. At one end of the ability range is vision that is 20/20 or better. Moving away from this ideal we find increasing degrees of shortsightedness—vision that is 20/30, 20/40, and so on. Having precisely 20/200 vision is a fully *determinate* disability.²⁶

Now suppose a doctor performs an operation to improve his shortsighted patient's vision. Before the operation the patient's vision is 20/250; afterwards it is 20/200. According to Shiffrin's and Harman's accounts of harm, someone with 20/200 vision suffers harm; so even after the operation, the patient suffers harm. But the doctor obviously does not harm his patient by performing the operation. There is no harm-based *pro tanto* objection to the doctor's action. In order for an action to be *pro tanto* objectionable, then, it is *not* enough that had the action not been performed, some person would not have been in the particular state of harm he's actually in. Nor is it enough that the person would instead have been in another state from along the relevant ability spectrum.²⁷ What matters is whether he would have been in a *better* state from along that spectrum had the action not been performed. If the doctor were to make his patient's sight *worse*, he would harm him. But since he actually *improves* his patient's sight along the relevant dimension, he benefits rather than harms him.²⁸

Shiffrin and Harman could, I suppose, insist that the doctor *does* harm his patient by performing the operation, and then argue that this harming is *justified*. The doctor causes his patient to suffer the harm of having precisely 20/200 vision, but he also prevents him from suffering (or continuing to suffer) the worse harm of having precisely 20/250 vision. The doctor thus inflicts a lesser harm upon his patient in order to prevent him from suffering a greater harm—a paradigm case of justified harming! But this is not a plausible analysis of the example. To be sure, there is such a thing as inflicting a lesser harm for the sake of preventing a greater one. Suppose that the only way a doctor can cure his patient of a disease causing permanent paralysis is to give him a drug that will, as a side effect, make him shortsighted. Here the doctor clearly harms his patient by giving him the drug. He damages his patient's sight, which we may suppose would otherwise have been perfect. But this harming is justified. The doctor inflicts harm along one dimension of functioning in order to prevent greater harm along another dimension. In the operation example, however, only one dimension of functioning is involved. The doctor simply decreases his patient's degree of shortsightedness. That he was unable to eliminate the shortsightedness completely—that he was unable to improve his patient's vision all the way to 20/20—is not a "cost" or unfortunate side effect of the operation, to be weighed against the benefit. In this example the doctor *simply* benefits his patient; he does not harm him *at all*.

In denying that the doctor harms his patient, I am not rejecting noncomparative accounts of harm. I agree that the patient's state after the operation is still one that it

is bad for him to be in. For the sake of argument I have even granted that in having 20/200 vision the patient suffers harm. What I am denying is that the doctor's action *relates* him to this harm in a way that makes his action *pro tanto* objectionable. Nor am I abandoning the two-step analysis of harming. I am not returning to Parfit's position that an agent harms someone only if he makes him worse off than he would otherwise have been. To determine whether the doctor harms his patient we do not compare levels of well-being; we compare degrees of *harm* along a single dimension of functioning. It is significant, however, that it has turned out to be impossible to avoid comparisons in giving an account of harming. If we eliminate them from our account of what it is to suffer harm, they reemerge in our account of what it is to stand in relation R to harm.

The claim we wished to evaluate was:

If B's suffering some harm is linked, via a chain of counterfactual dependencies, to A's acting in a certain way, then A's acting in that way is *pro tanto* objectionable.

We have seen that this is false. In order to determine whether an agent stands in relation R to a person's having some determinate disability, it is not enough to know that the person would not have had that particular, determinate disability had the agent acted differently. We must also know whether he would have been in a better or a worse state of functioning along the relevant dimension had the agent acted differently. The comparison is crucial: if the person would have been in an even worse state of functioning had the agent acted differently, the agent benefits rather than harms him.

If the claim I've rejected has seemed true, I suspect this is because people writing on the issue have tended to focus upon certain *determinable* disabilities. While determinate disabilities occupy single points along ability spectra, determinable disabilities disjunctively encompass whole ranges of ability levels. Someone is *mildly shortsighted*, for example, just in case his determinate level of visual acuity falls below 20/20 but above (let us suppose) 20/60. Since this range includes neither end of the relevant ability spectrum, the fact that a mildly shortsighted person would not have been mildly shortsighted had a certain agent acted differently is compatible both with its being the case that his vision would have been better had the agent acted differently and with its being the case that his vision would have been worse had the agent acted differently; and of course if the subject's vision would have been worse had the agent acted differently, the agent benefits rather than harms him. But things are different when the range of ability levels disjunctively encompassed by the determinable disability includes the low end of a spectrum. One such disability is being legally blind. Someone is legally blind just in case his determinate level of visual acuity falls below a certain threshold. If someone who is legally blind would not have been legally blind had a certain agent acted differently, this *rules out* the possibility that his vision would have been even worse had the agent acted differently. It consequently rules out the possibility that the agent benefits rather than harms him with respect to that dimension of functioning.²⁹ Even more extreme is the determinable harm of being disabled *simpliciter*. Someone is disabled *simpliciter*

just in case there is some dimension or other along which his level of functioning falls below the threshold of proper functioning. As long as someone exists, the only alternative to his being disabled *simpliciter* is his functioning properly in every respect.

This does not mean that how alternative levels of functioning compare is irrelevant to the question whether an agent stands in relation R to someone's being legally blind, or to his being disabled *simpliciter*. It just means that so long as the victim was going to exist either way, the outcome of the comparison is not in doubt: if the subject would still have existed but would not have been legally blind (or disabled) had the agent acted differently, then the subject would have functioned *better* along the relevant dimension had the agent acted differently; and so the agent harms him by acting as he does. There is no reason to think that whether B would have functioned better had A acted differently has any less bearing on whether A stands in relation R to B's being legally blind, or disabled *simpliciter*, than it does on whether A stands in relation R to B's being mildly shortsighted, or having precisely 20/200 vision. It would be absurd to suggest that a comparison between levels of functioning is relevant to harming only when the harm in question happens not to encompass the low end of an ability spectrum.

9.6

I have argued that the mere fact that B would not have suffered some harm had A acted differently is not enough to make A's act one of harming. It matters what the alternative was for B. Now in the procreation example the disabled child would not have existed at all had his parents not created him. Might the fact that *this* was the alternative for their child be enough to make their action one of harming?

The argument of the previous section does not rule this out. We've seen that if B, who has a certain determinate disability, would have been in an even worse state with respect to that ability had A acted differently, then A does not harm him, but rather benefits him by acting as he does. And we may presumably add that if B would have been in exactly the same state of disability had A acted differently, then (setting aside worries about preemption) A neither harms nor benefits him. So perhaps A harms B as long as neither of these conditions obtains—perhaps A harms B if B is in a disabled state and would have been in neither that same state nor a worse one (along the same dimension) had A acted differently. This sufficient condition for harming could in turn be satisfied in either of two ways: by its being the case that B would have been in a better state (along the relevant dimension) had A acted differently or by its being the case that B would not have existed at all had A acted differently.

The resulting view would be this. Suppose that B has some determinate disability. Then:

1. A *benefits* (and does not harm) B with respect to the relevant dimension of functioning if B's state of functioning along that dimension would have been even worse had A acted differently;

2. A *neither benefits nor harms* B with respect to the relevant dimension of functioning if B's state of functioning along that dimension would have been neither better nor worse had A acted differently;
3. A *harms* B with respect to the relevant dimension of functioning if *either* (i) B's state of functioning along that dimension would have been better had A acted differently, *or* (ii) B would not have existed at all had A acted differently.

I don't think this view is acceptable. The aim of the approach to the non-identity problem under discussion in this paper is to show that non-identity cases *aren't special at all*, that the acts of the parents and of those who choose the risky policy satisfy the *very same* sufficient condition for harming that paradigm ordinary acts of harming satisfy. The putative sufficient condition considered and rejected in the previous section met this *desideratum*. The condition stated in (3), by contrast, is disjunctive: it has separate clauses for identity and non-identity cases. And the non-identity disjunct—clause (ii)—seems completely ad hoc. All the other conditions specified in the view—the condition for benefiting, the condition for neither benefiting nor harming and the condition for harming in identity cases—rest upon *comparisons* between actual and counterfactual states of functioning. Why should there be a second, *noncomparative* condition for harming? This condition appears to have been tacked on simply to give the desired result in non-identity cases.

There is a further respect in which clause (ii) seems ad hoc. Consider a sequence of cases: first, an agent causes someone to have 20/200 vision rather than 20/190 vision; next, he causes someone to have 20/200 vision rather than 20/180 vision; and so on, the alternative each time approaching closer to 20/20. As the sequence continues, the acts of harming become progressively more serious. Eventually we reach the most serious case of all, the one in which the victim would have had perfect eyesight had the agent acted differently.³⁰ Now according to the view characterized by (1)–(3), equally serious (presumably) would be the harming done if the person with 20/200 vision would not have existed at all had the agent acted differently. But why should we assimilate nonexistence to perfect vision when determining whether the agent harms the person who ends up with 20/200 vision? Why not assimilate nonexistence to a complete lack of sight? Surely we could with equal justice replace (1) with:

- 1'. A *benefits* (and does not harm) B with respect to the relevant dimension of functioning if *either* (i) B's state of functioning along that dimension would have been even worse had A acted differently, *or* (ii) B's actual state of functioning along that dimension is at least better than a complete lack of functioning and B would not have existed at all had the agent acted differently.

My opponent says of the child with 20/200 vision: "Had his parents not created him, he would not have suffered the harm of being shortsighted. It's not just that he wouldn't have suffered the determinate harm of having 20/200 vision. He would not have suffered harm of *any* degree along the relevant dimension. There is consequently a respect in which his parents harm him by creating him." But we can with equal justice say: "Although it is bad for the child that he has 20/200 vision, at least he can see to *some* degree; had his parents not created him, he would not have been

able to see to *any* degree. His parents consequently benefit him by creating him. It's not just that they benefit him on balance, since his life is good on balance. They benefit him even with respect to his power of sight."³¹

Now in fact I don't think we should say either of these things. Nonexistence should be assimilated to neither end of the ability spectrum. Comparisons are necessary for determinations of harming, and while 20/200 vision can be compared with other degrees of visual acuity, it cannot meaningfully be compared with nonexistence.³² I suspect my opponent's point, that the disabled child would not have suffered harm of any degree along the relevant dimension had his parents acted differently, gets its rhetorical force from cases where comparisons *are* possible—from cases where the victim would have existed even if the agent had acted differently. In such cases the victim would have been in a good state, and hence in a *better* state, along the relevant dimension had the agent acted differently.

Alternatively, perhaps the reason some have thought it sufficient for harming that the person who suffers harm would not have existed had the agent acted differently is that they have thought it sufficient for harming that the person in question would not have suffered that particular, determinate harm had the agent acted differently; and satisfaction of the former condition entails satisfaction of the latter one. But as we saw in the previous section, the latter condition is not sufficient for harming. We require *independent* grounds for thinking that the former condition is sufficient for harming, and so far we have found none.

9.7

In Section 9.5 we considered a condition satisfied both by the parents of the disabled child and by the agents of most ordinary acts of harming: had they acted differently, certain people would not have suffered the harms they actually suffer. Unfortunately, satisfaction of this condition turned out to be insufficient for harming. In Section 9.6 we considered a condition satisfied both by the parents of the disabled child and by the agents in other non-identity cases, but not by the agents of ordinary acts of harming: had they acted differently, certain people who suffer harms would never have existed. But the suggestion that satisfaction of this condition suffices for harming turned out to be unmotivated. Now let's consider a condition the satisfaction of which *does* suffice for harming. The question will be whether it is satisfied by the parents of the disabled child.

Our reflections on the eye operation example suggest the following sufficient condition for harming:

A harms B by acting in a certain way if A's acting in that way rather than in some other way helps explain why B is in a certain bad state (with respect to a given dimension of functioning) *rather than in a better state* (with respect to that same dimension of functioning).³³

This condition appeals to what we might call "contrastive" explanation. For present purposes, we may understand contrastive explanation as the ancestral of

contrastive counterfactual dependence: β 's being G rather than G' depends counterfactually upon α 's being F rather than F' if and only if α is F rather than F' , β is G rather than G' , and if α had been F' rather than F , β would have been G' rather than G .³⁴

Now it is not true that if the parents of the disabled child had acted differently, their actual child would have been in a better state of functioning. They consequently do not satisfy this sufficient condition for harming via a *single-linked* chain of counterfactual dependencies. But we must consider whether their performing some action rather than another helps explain, via a *multi-linked* chain of counterfactual dependencies, why their child is in a certain bad state (with respect to a given dimension of functioning) rather than a better one. This possibility must be taken seriously, for it is only in virtue of such a multi-linked chain that we who choose the risky policy harm those who die in the ensuing nuclear catastrophe. After all, it is not true that if we had chosen some other policy, the people actually killed by leaked radiation would have lived to ripe old ages. There is no single-linked explanation here either. Working backwards, the explanatory chain runs as follows: the people's dying prematurely rather than living to ripe old ages depends counterfactually upon the earthquake zone's having radioactive waste buried in it rather than not having radioactive waste buried in it; and the earthquake zone's having radioactive waste buried in it rather than not having radioactive waste buried in it depends counterfactually upon our having chosen the risky policy rather than some other policy.

Can we construct such a multi-linked chain in the case of the disabled child? Let us try, again working backwards. The child would have functioned properly had the sperm from which he partially originated not been genetically defective. (We will suppose, perhaps implausibly but for the sake of argument, that this genetic defect was not identity determining.) The child's being disabled rather than fully functional, then, depends counterfactually upon the sperm's being defective rather than non-defective. And the sperm's being defective rather than non-defective in turn depends counterfactually upon the father's being a carrier of the defect rather than a non-carrier. No doubt we could continue tracing this chain of counterfactual dependencies back in time. The problem is that no *action* of the father's need appear in the chain. There need be nothing the father did, for example, such that the sperm's being defective rather than non-defective depends counterfactually upon his having done that rather than something else. Likewise, there need be nothing the father did such that his being a carrier of the genetic defect rather than a non-carrier depends counterfactually upon his having done that rather than something else. And so on.

We can of course imagine a version of the example in which the father's having acted in one way rather than another *does* figure into the relevant explanatory chain. His being a carrier of the defect rather than a non-carrier, for example, might depend counterfactually upon his having taken a certain drug in his youth rather than abstaining. In that case he *would* have harmed the child he later helped create. But there is no reason why the child's being disabled rather than fully functional *must* be explained by either parent's having at some point acted in one way rather than another.³⁵

9.8

Let us assume, then, that no action of either parent figures into the chain of counterfactual dependencies explaining why their child is disabled rather than fully functional. Mightn't their act of creating him figure into the explanation in some other way? Had they not created him, this explanatory chain would never have been forged, for in that case there would have been no child whose condition needed explaining. By creating the child, then, the couple gave the fact that the man was a carrier of a genetic defect rather than a non-carrier the *opportunity*, so to speak, to explain someone's being disabled rather than fully functional. (I continue to assume, for the sake of argument, that the genetic defect is not identity-determining.) The parents knew that if they created a child, the man's being a carrier of the defect rather than a non-carrier would lead to that child's being disabled rather than fully functional. Mightn't that be enough to make their act one of harming?³⁶

I don't think so. Suppose that B has 20/200 vision. A arranges for C to perform an operation on B; C is the only doctor in the world capable of performing this sort of operation. Unfortunately, C is also a misogynist, and although he would never admit it, he never does his best work on women. C performs the operation on B, improving her vision to 20/150. Had he not been a misogynist, he would have improved it to 20/100. B's having 20/150 rather than 20/100 vision after the operation is thus explained by C's being a misogynist rather than a non-misogynist. And by arranging for C to perform the operation, A gave C's misogyny the opportunity, so to speak, to explain B's having 20/150 rather than 20/100 vision. But A's action is not even *pro tanto* objectionable. A does not harm B by arranging for C to perform the operation. (C arguably harms B by not performing the operation as well as he should have done, but this feature of the example is not essential. We could imagine instead that the operation's effectiveness was reduced by some temporary environmental condition. Unfortunately, the operation could not be delayed.)

In this example, of course, B would have been even more shortsighted had A not arranged for the operation, whereas in the procreation example the child would not have existed at all had his parents not created him. But we have yet to see a reason for thinking that this difference *makes* a difference. It would be ad hoc to suggest that while there is no *pro tanto* objection to giving a condition or circumstance the opportunity to explain someone's being disabled rather than fully functional when that person would otherwise have been in an even worse state of functioning, there *is* a *pro tanto* objection when either (i) the person would otherwise have been in a better state of functioning or (ii) the person would otherwise not have existed at all.

9.9

Here's another possibility. That his parents created him rather than doing something else does not explain why the child is disabled rather than fully functional, but it does explain why he is existent rather than nonexistent, alive rather than not alive.

His parents thus stand in relation R to his existence or his life. But perhaps, in virtue of standing in relation R to his existence or his life, they indirectly also stand in relation R to each of the elements (states, episodes, etc.) making up that existence or life. Some of these elements are good, others are bad; and among the bad elements is his disability. Because his parents *create* him—give him life—they stand in relation R to everything in his life, including his disability. And so even though his life is on balance good, they harm him by creating him.

I think this proposal's plausibility rests upon equivocations. When we speak of life, we sometimes mean the sort of thing that a person can live or lead: a long life or a short one, a happy life or a sad one, an easy life or a hard one, a boring life or an exciting one. In this sense, a person's life is something like his history. But we can also mean the bare property of being alive. Likewise, when we speak of existence, we sometimes mean the sort of thing that can be pleasant, carefree, meager or wretched, but at other times we mean the bare property of existing. Now the parents' creating their child rather than doing something else does *not* explain why their child has one sort of life or existence rather than another. There is thus no reason to think that the parents stand in relation R to their child's life or existence in the first sense. (It won't do to say that the child would not have had the sort of life or existence he actually has if he had not been created—we've already seen that that's not sufficient for the holding of relation R.) All we're entitled to say is that the parents stand in relation R to their child's having the bare property of being alive, or of existing: their creating him rather than doing something else explains why he is alive rather than not alive, existent rather than nonexistent. The bare property of being alive, or of existing, however, does not have both good and bad "elements." I thus see no argument from the parents' standing in relation R to their child's being alive, or existing, to their standing in relation R to the various good and bad elements that comprise their child's life or existence.

9.10

I agree that being disabled is bad for the child, and that this fact must figure somehow into a correct account of why his parents' choice to create him is *pro tanto* objectionable. One way it could figure in is this: perhaps the parents produce a worse outcome by creating a disabled child (or one who suffers harm) than they would have produced by creating a child free of disability (or one who didn't suffer harm). But for a variety of reasons I would prefer a less impersonal explanation of why the parents' choice is *pro tanto* objectionable. Of course if the parents were to *harm* their child by creating him, that would provide a satisfyingly person-affecting explanation, but for reasons I have attempted to set out in this paper I have serious doubts about the availability of this explanation. Even if we broaden the concept of *harming* for moral purposes, as I think we should, the concept must still pick out a relation between agents and harms that is intuitively sufficient to generate a harm-based *pro tanto* objection, and I have yet to find convincing grounds for

thinking that such a relation holds between the parents of the disabled child and their child's disability.

Acknowledgements I am grateful to David Wasserman and Melinda Roberts for their helpful comments.

Notes

1. Parfit (1984), pp. 372–374. Parfit distinguishes between “morally relevant” and “ordinary” senses of harming. Using the word in its ordinary sense, it might be infelicitous to say that an agent harms someone if his action is only a remote cause of that person's suffering harm, but Parfit points out that such an action might nonetheless be morally objectionable owing to its role in the production of that harm. I agree that for purposes of moral theory we should employ a somewhat broader notion of harming. Henceforth I shall usually drop the modifying phrase “in a morally relevant sense,” taking it as understood. Parfit has a second reason for distinguishing morally relevant from ordinary harming: an action that harms someone in the ordinary sense needn't be morally objectionable. I find this reason unconvincing. An act of *justified* ordinary harming is not morally objectionable all things considered, but it is still, in a morally relevant sense, an act of harming. That's why it stands in need of justification. The relevant question is whether ordinary harmings ever require no justification at all; and I do not find Parfit's examples convincing on this score. See Parfit (1984), pp. 69–71.
2. The first example is a simplified version of one found in Parfit (1984), pp. 371–372. The second example comes from McMahan (1998), p. 208.
3. Nor did I defend what is sometimes called the person-affecting *restriction*: roughly, that an action cannot be wrong unless there is someone whom it harms, wrongs, or causes to be worse off.
4. When I say that Shiffrin, Harman and I offer structurally similar accounts of harming, I mean that our accounts share this two-step structure. We all think that A harms B if and only if he's connected in the right way to a harm suffered by B.
5. By “responsible for harm” I did not mean “at fault for, or to blame for, harm.” I meant that the harm is attributable to one's agency in such a way that one can legitimately be asked to defend one's role in bringing it about. If one's conduct makes one at least partially responsible for a harm, one's conduct stands in need of justification. The important question is what non-moral relation between one's action and a harm can make one at least partially responsible for that harm.
6. Hanser (1990), pp. 60–62.
7. Harman (2004), pp. 92–93.
8. Harman grants that if the risky policy would result in a larger future population, the benefit-based reason in favor of adopting it would outweigh the benefit-based reason in favor of adopting the competing policy. But she argues that the moral reason against harming is so serious that adopting the risky policy would still be impermissible. See Harman (2004), pp. 102–103.
9. Shiffrin (1999), pp. 119–133. Her actual argument concerns procreative choices, but it applies equally well to the choice of the risky policy.
10. Strictly speaking, Shiffrin does not argue that it is *impermissible* to harm someone (without his consent) in order to bestow upon him a greater pure benefit. Her article concerns the question whether the agents of such acts *owe compensation* to those they have harmed; her conclusion is that they do. By contrast, an agent who unavoidably harms someone in the course of saving his life has no obligation to compensate the recipient of his beneficence.

Shiffrin also argues that the force of the reason in favor of bestowing a pure benefit is

greatly reduced when no one will be left in a comparatively bad state if one fails to bestow that benefit. If a child is not conceived, his life does not go worse than it would have gone had he received the benefit of existence. The fact that a certain choice is a necessary condition for someone's receiving the benefit of existence consequently has limited justificatory force. See Shiffrin (1999), pp. 134–135.

11. See, for example, Woodward (1986), pp. 822–825, and Harman (2004), pp. 98–101.
12. This is close to being the view of David Wasserman (2005).
13. According to the standard method for evaluating counterfactuals, a given counterfactual is true if and only if its consequent is true in the “closest” possible world—the world most similar to the actual world—in which its antecedent is true. The world most similar to our own in which the people are not exposed to radiation is presumably one in which the earthquake does not occur, but whose history leading up to that point is otherwise the same. It is *not* a world in which we chose some other energy policy three centuries earlier. David Heyd accuses me of equivocating when I say both that the people who die in the nuclear disaster would have been better off had they not been exposed to radiation and that they would not have been better off had we had chosen a different policy. Heyd (1992), p. 113. But when I say this I am not, as he supposes, illicitly describing a single event first as the people's exposure to radiation and then as our choice of the risky policy. Rather, I am describing two different events which occur three hundred years apart.
14. Shiffrin (1999), p. 123.
15. *Ibid.*
16. Harman (2004), pp. 96–97.
17. I defend this understanding of harm, and criticize competing conceptions, in Hanser (2008). I should stress that my view is not a version of the temporal comparison view. I do not hold that suffering harm is a matter of being worse off in some respect than one was at an earlier time.
18. I remain open to the possibility that the child's disability is a *harmful* condition—a condition that gives rise to harms. But it is not itself a harm.
19. I do not mean that nothing of significance hangs on which account of harm is correct. I mean only that my disagreement with Shiffrin and Harman over whether disabilities are harms does not affect the question whether we can give a narrow person-affecting account of why it's wrong (assuming that it is wrong) to create a disabled child, when one could easily create a non-disabled child instead.
20. In my earlier paper I argued that if an action makes its agent responsible for someone's being in a bad state, then even if the victim cannot be said to suffer harm as a result of the action, the action is *pro tanto* objectionable on narrow person-affecting grounds. See Hanser (1990), pp. 64–65.
21. Shiffrin (1999), p. 119 and fn. 8. She continues that “it is sufficient for this article, though, to assume that if one is the direct cause of a person's being in a situation that intrinsically delivers harm, then one is responsible for that person's suffering harm, even if one does not *harm* her.” I take this last phrase to mean “even if one does not harm her *in the ordinary sense.*”
22. Harman (2004), p. 93. And regarding a case like that of the disabled child, she writes that “conceiving harms the child because it causes the child to be in a particular kind of bad state.” *Ibid.*, p. 94.
23. This account derives from Lewis (1973).
24. Some might object that causation of harm generates a *pro tanto* objection only when the harm is foreseeable. Since the parents could foresee that their child would be disabled, I shall set this worry aside.
25. For an overview of the pros and cons of various counterfactual analyses of causation, see the papers collected in Collins et al. (2004).
26. Given the possibility of corrective lenses, shortsightedness, unless extremely severe, is not a serious disability, and I'm not sure to what extent Shiffrin and Harman would consider it a harm. But this is not important. I use shortsightedness as an example only because there is a familiar scale for measuring its degrees.

27. This might, however, be sufficient for causation. Had the doctor not performed the operation, the patient's vision would have been 20/250 instead of 20/200. Perhaps it follows that the doctor (or the operation he performed) *caused* the patient to have precisely 20/200 vision. But again, for our purposes this doesn't matter.
28. In this argument I have assumed, along with Harman, that states of less than proper functioning are harms. For Shiffrin, harms are (roughly) states that create a cleavage between a person's experience and his will. The argument could easily be recast in these terms: an action that decreases the degree of cleavage (along some dimension) between a person's experience and his will, without eliminating that cleavage entirely, leaves him in a state of harm but does not harm him.
29. The same point holds for any determinate disability occupying the low end of an ability spectrum—being completely blind, for example. If someone who is completely blind would not have been completely blind had a certain agent acted differently, then it can't be that his vision would have been even worse had the agent acted differently.
30. For convenience I ignore the possibility of vision that is better than 20/20.
31. Would this latter line of thought be as plausible if the harm in question were an episode of pain rather than a disability? One might think not, reasoning as follows. Nonexistence is more plausibly assimilated to the *absence* of something than to its presence; that is why nonexistence is more plausibly assimilated to the complete absence of an ability (i.e. to the *worst* state with respect to the ability) than to its complete presence. But pain is bad in itself—it is not merely the absence of something good. It is thus more plausible to assimilate nonexistence to the complete absence of pain (i.e. to the *best* state with respect to pain) than to a state of maximum pain. I am unconvinced. Episodes of pain are episodes of unpleasant *consciousness*; the end-points of the relevant spectrum are consequently states of pain-free consciousness and states of maximally painful consciousness. Nonexistence can be plausibly assimilated to neither end-point.
32. It is possible to get around this difficulty when the subject's *overall* quality of life is at issue. Suppose a certain sort of life is so bad that it would be better to die now than spend the rest of one's days living such a life. Strictly speaking the terms of comparison here are alternative possible lives—it's better to have a shorter such life than a longer one—but speaking loosely we might say that it's better not to exist at all than to live such a life. Unfortunately, this maneuver cannot be used to defend the claim that the parents of the disabled child harm him by creating him. It simply isn't true that it would be better for the child if he were to die now rather than continue living with 20/200 vision. It isn't even true that *at least so far as his ability to discriminate detail visually is concerned*, it would be better for him if he were to die now rather than continue living with 20/200 vision (if indeed this thought even makes sense). Nor do I think it would be better for a person, at least so far as *pain* is concerned, if he were to die now rather than spend the rest of his life experiencing mild pain (if indeed this thought even makes sense).
33. Since I don't think that harms are bad states, I do not myself accept this as a (partial) account of harming. But I do think that it captures a sufficient condition for an action's being *pro tanto* objectionable owing to its role in explaining someone's being in a bad state.
34. I don't really think that explanation should be cashed out in terms of counterfactual dependence, for the usual reason: in cases of "preemptive" explanation, α 's being F rather than F' explains β 's being G rather than G' even though β would have been G anyway, owing to some other cause, had α been F' . But if we set cases of preemption aside, counterfactual dependence will serve well enough. (I shall not here attempt to determine whether a thing's being F rather than F' amounts to anything more than its being F and not F' .)
35. Suppose a doctor awakens a patient from an otherwise permanent coma, knowing that the patient will suffer mild pain for the remainder of his days. Does the doctor harm the patient? There are two versions of the case to consider. If the pain results from an injury caused by the waking procedure, then I think the doctor does harm the patient. The patient's being in a painful rather than a pain-free conscious state depends counterfactually upon his suffering rather than not suffering the injury; and his suffering rather than not suffering the injury

depends counterfactually upon the doctor's performing rather than not performing the procedure. (This is, of course, a case of justified harming.) But if the pain results from a condition that the patient had even before going into a coma, I am inclined to say that the doctor does not harm the patient. There is nothing the doctor does such that his doing that rather than something else explains the patient's being in a painful rather than a pain-free conscious state upon waking from his coma.

36. Perhaps Harman is suggesting something along these lines when she writes that "[i]f some factor for which we are not responsible will cause a bad result if we act in a certain way, this fact can provide a reason against the action." Harman (2004), p. 95, emphasis omitted.

References

- Collins, J., Hall, N., and Paul, L.A., eds. 2004. *Causation and counterfactuals*. Cambridge, Massachusetts: The MIT Press.
- Hanser, M. 1990. Harming future people. *Philosophy & Public Affairs* 19: 47–70.
- Hanser, M. 2008. The metaphysics of harm. *Philosophy and Phenomenological Research* 77: 421–450.
- Harman, E. 2004. Can we harm and benefit in creating? *Philosophical Perspectives* 18: 89–113.
- Heyd, D. 1992. *Genethics*. Berkeley and Los Angeles: University of California Press.
- Lewis, D. 1973. Causation. *Journal of Philosophy* 70: 556–567. Reprinted in his *Philosophical papers, volume 2*. Oxford: Oxford University Press.
- McMahan, J. 1998. Wrongful life: Paradoxes in the morality of causing people to exist. In *Rational commitment and social justice*, eds. J. Coleman and C. Morris, 208–247. Cambridge: Cambridge University Press.
- Parfit, D. 1984. *Reasons and persons*. Oxford: Clarendon Press.
- Shiffrin, S. 1999. Wrongful life, procreative responsibility, and the significance of harm. *Legal Theory* 5: 117–148.
- Wasserman, D. 2005. The nonidentity problem, disability, and the role morality of prospective parents. *Ethics* 116: 132–152.
- Woodward, J. 1986. The non-identity problem. *Ethics* 96: 804–831.

Chapter 10

The Nonidentity Problem and the Two Envelope Problem: When is One Act Better for a Person than Another?

Melinda A. Roberts

Abstract The nonidentity problem and the two envelope problem have in common an ongoing resistance to intuitive analysis. They also share certain structural features. I argue that the two problems proceed under the same error, imagining subjects to draw haphazardly from a potpourri of actual and expected values to generate results about betterness and harm rather than, as we naturally do and always should, drawing in a more discriminating way from a more orderly array. When we play by the same set of rules in calculating the values that we then compare, we, in particular, become able to discern (1) *harm* in just the cases in respect of which one important type of nonidentity problem has long been thought to show “no harm done” and (2) *no harm done* in the two envelope problem, which purports to show “harm done” when the subject refuses endlessly to switch from one envelope to another and back again.

Keywords Parfit · Kavka · Slave child example · Harm · Risk · Probability · Expected value · Nonidentity problem · Two envelope problem.

10.1 Parallel Problems

The nonidentity problem and the two-envelope problem may seem like games. But they are hard games, perennially resisting intuitive resolution. Moreover, the nonidentity problem is widely considered to have destroyed any hope that moral theory can be grounded in the highly intuitive “person-affecting,” or “person-based,” approach to ethics (“PBA”). According to PBA, the moral status of a given act is determined by facts that are *person-based* in nature. Included in PBA is, for example, the core idea that “what is bad must be bad for someone.”¹ That core idea—the “person-based intuition” (“PBI”)—expresses nothing more than a simple necessary condition on wrongdoing: the act that is “bad for” no existing and no future person cannot be morally wrong. PBA, in contrast, includes both PBI

Melinda A. Roberts (✉)
The College of New Jersey, Ewing, NJ, USA
e-mail: robertsm@tcnj.edu

and additional necessary and certain sufficient conditions on wrongdoing as well, with the principles that constitute PBA united in their effort to connect wrongdoing with how persons—as individuals, not in the aggregate, and including at least some non-human animals alongside many humans—are *affected* for better or worse.

On the face of things, PBA is a compelling way of thinking about how moral law is structured. It seems plausible that an act's moral status can be tied directly to what that act does—or, at least, can be expected to do—for people. Moreover, the proposition that moral theory is not person-based—that it must be “impersonal” at least in part—creates its own set of challenges. Impersonal, or mixed, theories cannot easily make sense of straightforward questions of moral obligation.² Nor can they serve as the basis of a Dworkinian “political morality” we would like to be able to appeal to in addressing hard questions in respect of which the text of the law is indeterminate.³

If we do find PBA—including PBI—at least credible, the nonidentity problem will be of grave concern to us. That problem purports to show that at least some “bad” acts are “bad for” no one at all. In this paper, I will focus on one especially powerful type of nonidentity problem—what I will call the “can't-expect-better” problem.⁴ That problem type includes Kavka's slave child and pleasure pill cases, Parfit's depletion and risky policy cases and Sher's and Shiffrin's transgenerational compensation puzzles.⁵ Global warming is perhaps the world's biggest can't-expect-better problem, compelling Broome, for one, to jettison PBA in favor of an impersonal approach.⁶ I want to concede that, *if* these problem cases have been understood correctly—if, in particular, the wrong acts I concede they involve really do harm *no one*—then these problem cases decisively refute PBI and hence PBA. My argument here, however, is that these problem cases have not been understood correctly: we *think* they establish the “no harm done” result they purport to establish *only* because we have committed a certain fallacy in connection with our reasoning about just when what we do today harms, or makes things worse for, persons who will not exist until tomorrow.⁷ If I am correct, then this particular type of nonidentity challenge against PBI fails.

A very different type of nonidentity problem—the “can't-do-better” problem—includes, among others, Parfit's two medical programs case and cases of “wrongful disability.”⁸ The argument made in connection with the can't-do-better problem appeals to the fact that, in some cases, it is simply not biologically, or physically, possible for the particular person who has been brought into existence by the act under scrutiny to exist and *not* suffer a particular impairment. From there it is inferred that the procreative act itself does not make things worse for, or harm, that person. In the rare case, then, in which that same act harms *no one else*—*no other existing or future person*—PBI dictates that the act itself cannot be wrong. Here, the result that the procreative act does not harm the impaired child, so long as that child's life is not less than worth living, is one that I believe we cannot avoid. I argue, however, that the further result that the act is not wrong is one that we can accept. If I am correct, then this second type of nonidentity challenge against PBI fails as well.

This is not to say that the can't-do-better problem is not a difficult one. But my own view is that the can't-expect-better problem is still less tractable. In the can't-expect-better problem, the position that the act under scrutiny is permissible

is simply not plausible. Moreover, the phenomenon that Kavka describes as the “precariousness” of existence, which the can’t-expect-better problem appeals to for purposes of demonstrating that the act under scrutiny—the “bad” act—is “bad for” no one, is riveting in its own right and seems unassailably linked to the question of harm.⁹ Each of us, Kavka notes, has made it into existence *against all odds*. Eliminating any act—including the clearly “bad” act—in the causal sequence of highly particularized acts and events that ends in the conception of any one person, and substituting in the place of that act any clearly permissible act, very probably will result not in a different or better life for that one person but rather in *no life at all* for that one person. As Parfit puts it, how many of us would have existed had “motor cars . . . never been invented”?¹⁰ But the precariousness insight applies not just to earth-shaking events like the invention of motor cars or Hitler’s baser acts not having been countermanded earlier on, but also to things like whether a couple dines at one restaurant rather than another—or pauses, in one of Kavka’s cases, to take a teratogenic “pleasure pill”—just before conceiving their first child.¹¹ Any little variation in the causal sequence can easily affect, at least in some slight way, the timing and manner of conception—and the slightest variation in the timing and manner of conception would have all but assured that the particular person who exists and suffers would have been taken off-track for existence altogether. The permissible act might have meant the conception of another, “nonidentical,” better off person in place of the one. But it would, it seems, not have made things any *better for* the particular child who in fact exists and suffers as a result of the clearly “bad” act. If anything, the permissible act, by, very probably, leaving that child out of existence altogether, would, very probably, just have made things *worse* for that child.

According, then, to the can’t-expect-better problem, once we recognize the precariousness of existence, we are forced to accept that the clearly “bad” act is not “bad for” what might seem to be its clearest victim. Surely, after all, it is better to have an existence that is, if flawed, nonetheless worth having than it is never to have existed at all. One act in fact produces the former outcome for a person, while any other act, very, very probably, would have produced the latter. How, then, can that one act be “bad for” that person? If the broken arm inflicted as a matter of necessity in the context of a rescue is *not* a harm, then how can the suffering that accompanies as a matter of *near-necessity* all that is precious about life from the perspective of the one who lives constitute a *harm*?

I will argue, however, that the gap between necessity and near-necessity is much wider than many theorists have taken it to be. My argument will not involve any particularly creative account of when it is that an act is “bad for”—or *harms*—a person. I, instead, take it for granted that the comparative approach exploited by the can’t-expect-better problem is, in itself, unproblematic. According to that account, an act *harms* a person if that act makes things “worse for” that person than they needed to have been—if, that is, an alternative act would have made things “better for” that person than they in fact are. An act harms a person, in other words, if that act creates less wellbeing for that person when the agent—or group of agents—had the alternative of creating more. Thus, I harm you (under ordinary circumstances¹²) when I shoot you in the arm, not because (1) you have ended up shot in the arm and you suffer,¹³ and not because (2) you “would have been” better off had I not shot

you in the arm,¹⁴ but rather because (3) I had the alternative of not shooting you at all and that alternative would have made things better for you than you are, my having shot you in the arm.

More specifically, my argument will be that the can't-expect-better problem fails, not when it insists on a comparative approach to harm, but when it beguiles us into drawing haphazardly from a potpourri of *actual* and *expected* wellbeing levels ("values") and, on the basis of a comparison between some such pair of values, deciding that an act does not make a person any worse off than that person would have been under any alternative act. When we instead take care to select our value pairs in a more discriminating way from a more orderly array—as we naturally do and always should—we come to quite different results on both betterness and harm. We come to results that seem both intuitively plausible (surely, e.g., the child *is* harmed when its parents take the teratogenic pleasure pill prior to conceiving that child, and, just as surely, taking the pill *is* wrong) and not at all at odds with either PBI or PBA.

This is not to suggest that it is a mistake to bring expected value to bear in determining betterness. Doing so allows us to construct a moral theory that determines on a *prospective* basis what our moral obligations are. And we will consider that capacity critical if we think that morality has an action-guiding function. But betterness can be tricky, in two ways that are of particular import for purposes of evaluating the can't-expect-better problem. First, betterness between acts cannot be reliably determined by a comparison between actual and expected values. And, second, even if we do discipline ourselves to compare (just) expected value against expected value (or actual value against actual value), any calculation of expected value will come with its own hazards. In particular, once we know the future has unfolded in a certain way, it can be very hard, as a practical matter, to calculate *as though* we have no knowledge of that fact whatsoever. Yet that is exactly what we must do, if our expected value against expected value comparison is to determine in any reliable way whether one act is worse for a person than another.

The *two-envelope problem* exploits our epistemic vulnerabilities in the reverse way. We are urged to imagine that the future has *not* unfolded in a particular way when we are quite aware that it has. We are urged to think that we do *not* know a certain thing that we in fact know quite well. Again, as a practical matter, it is very hard to calculate as though we *do* know a certain thing when the very design of the case strenuously urges us to think we *don't*.

Unlike the nonidentity problem, the two-envelope problem has made barely a dent in the thinking of philosophers aiming to understand the structure of morality. But perhaps more note should have been taken of that problem, for it so nicely demonstrates just how much can go wrong in how we think about when it is that one act is better for a person than another. Illicit assumptions, whether they cover things we know but are not "supposed" to know, or things that we really are *supposed* to know but think we are not, quietly create chaos in our expected value calculations. But without those essentially confabulated calculations, we never obtain the problematic results we would so love to avoid: that clearly "bad" future-directed acts are "bad for" no one at all, and that (endlessly) switching from one envelope to the other can somehow constitute a worthwhile endeavor.

10.2 A Person-Based Approach to Procreative Choice

10.2.1 *The Choice Not to Conceive a Child*

The person-based intuition (“PBI”) provides a sensible account of many issues relating to procreative choice. According to PBI, an act (including any omission) is wrong at a world *only if* it creates less wellbeing for a person who does or will exist at that world when the agent had the alternative of creating still more wellbeing for that same person at some other world.¹⁵ I will reserve the term “harm” for the case in which that necessary condition is satisfied. Correspondingly, an act that (arguably) harms *only* those persons who count as *merely possible* relative to a world, by way of failing to bring those persons into existence to begin with, must be deemed permissible. According to PBI, then, losses incurred by the merely possible in virtue of their never having existed at all are without moral significance: they cannot make an otherwise permissible act wrong.

Important practical implications ensue. Consider, for example, the choice whether to conceive a child. That that child—that any child then conceived—would have a life worth living—a positive lifetime wellbeing level—if he or she were brought into existence does not, according to PBI, put *any* moral pressure on us to bring that new person into existence.¹⁶ The choice *not* to procreate becomes—in many instances—a clearly morally permissible alternative.

In contrast, impersonal, aggregative forms of consequentialism suggest, in surprising scenarios, that the choice to procreate is obligatory. Such views focus on whether the choice to bring the new person into existence increases total or average aggregate wellbeing or, under pluralism, increases aggregate wellbeing enough to counterbalance any values or ideals that weigh against that choice.¹⁷ The implication (often) will be that we are wrong not to bring the new person into existence—even if the choice not to bring the new person into existence makes things better for some persons who do or will exist (e.g., the woman who bears the child) and worse for none. That seems implausible. That moral law imposes on us such stringent procreation obligations—to produce a first child, a fifth child, a tenth child—seems highly implausible, however happy the child we might have had would have been.

10.2.2 *Abortion*

Not conceiving a child is one thing. Aborting a fetus *may* be quite another—depending on the timing of the abortion. Until that point in the pregnancy at which a *person* has come into being, PBI will deal with abortion just as it does non-conception. The implications of the woman’s abortion choice for the *non-person fetus* will not, according to PBI, create any grounds whatsoever for a moral objection to that choice. The abortion choice will be treated differently, however, at that point in the pregnancy at which a new person has commenced existence. In that case, PBI leaves the door open for a finding that the choice is wrong. But, as a simple necessary

condition on wrongdoing, PBI cannot do any more than leave that door open, even in the case of abortions performed in the last minutes of the full-term pregnancy when (it seems) we clearly do have a person. Yet PBI can be understood to be part of a broader view—a *person-based approach* (“PBA”) that includes certain sufficient conditions as well as additional necessary conditions for wrongdoing.¹⁸ One such sufficient condition would be the following (Paretian) principle: an act is wrong at a world when there exists an alternative to that act that creates additional wellbeing for at least some persons who do or will exist at that world without creating less wellbeing for any person who does or will exist at that world and without bringing any additional persons into existence. According to that principle and on the assumption that the late pregnancy involves a person whose life will be worth living, if there is no cost to the woman or to anyone else who does or will exist in allowing the pregnancy to continue, and if it’s not the case that (somehow) allowing the pregnancy to continue would mean that one or more additional persons would be brought into existence, then the abortion is wrong.

Of course, there may well be some diminution in the woman’s wellbeing involved in forgoing even the very late-term abortion and having the child. The very late abortion could, in other words, involve a *tradeoff*—a situation in which the agent can increase wellbeing for one person only by decreasing wellbeing for someone else. While I will not try to articulate a set of person-based tradeoff principles here, there is certainly no reason to think that PBA will not include them.¹⁹ PBA determines wrongdoing by reference to morally significant, person-based facts—and tradeoff scenarios are replete with just those sorts of facts. Thus, where the agent’s only alternatives are between (1) reducing the fetal-person’s lifetime wellbeing to the very low level it will have as a function of having had only a very abbreviated time in existence and (2) reducing the woman’s lifetime wellbeing in some more modest way, PBA can be expected to say that the abortion is wrong.

PBI and PBA thus serve to center the abortion debate on issues we intuitively take to have moral relevance to that debate, including, for example, the neurological status of the developing fetus, the emergence of consciousness and, most generally, when it is during the pregnancy that the developing organism counts as a *person*. In contrast, the aggregative approach sets aside, or at least minimizes, issues that intuitively seem of grave moral significance in favor of issues that seem peripheral at best. Thus, under both totalism and averaging, the non-person fetus has just about the same moral status as the fetal-person does. Similarly, under pluralism, the distinction between aborting the non-person fetus and aborting the fetal-person becomes far less relevant than it seems that it ought to be.

10.2.3 The Moral Significance of Merely Possible Persons

The appeal of PBA is increased when we realize that it can, and I think should, be understood not to place on a moral pedestal either (1) *actual* persons—persons who do or will exist at the *actual* world—or even (2) persons who *would exist* were the

act under scrutiny in fact performed. The fact that less wellbeing has been created for a person when the agent had the alternative of creating more can bear on the moral status of a given act, according to PBA, even if that person is neither actual nor someone who would exist were that act performed. Thus, PBA does not claim that the losses incurred by the merely possible *never* have moral significance or that the merely possible *never* matter morally. Those losses *do* matter morally, so long as they are incurred at worlds where the persons who incur them do or will exist, for purposes of determining the moral status not just of acts performed at those worlds but of acts performed at still other worlds as well. Thus, PBA recognizes that the plights of the merely possible may imply that a wrong has been done—not at the world at which those persons are merely possible but rather at a second world where those same persons do or will exist and suffer as compared to how those same persons fare at still a third world. Under PBA, the “genocidal adventures of nonactual dictators” remain a no–no, even where the victims of genocide are all “nonactual” as well.²⁰ Moreover, there is no reason to think that the fact that the merely possible and their losses matter *there* does not bear on the moral status of an act performed *here*. An act may thus be deemed perfectly permissible at a world in the case where the only way to create more wellbeing for the relatively well-off persons who do or will exist at that world without creating less wellbeing for any of those same persons is to bring into existence still additional persons who are then treated very badly relative to some third world—who are, for example, made to serve as slaves or organ donors for the rest of us.²¹

10.2.4 *The Moral Status of Future Persons*

It is also a plus for PBI that PBI is consistent with—and PBA *requires*—our denial of a view Narveson seemed to suggest decades ago and that other theorists have at least flirted with since: that, if the “children produced have a good chance at a good life, we think people should have them if they want them.”²² But we do not think that a procreative choice is *always* permissible so long as the new persons are “happy” or that it is *always* enough to give our own offspring a “good chance at a good life.” Completely independent of the nonidentity problem, we think that making “happy people” often isn’t morally neutral.²³

PBA, which I have set forth here as an inherently *maximizing* approach (that it does not aggregate does not mean it does not, on a person-to-person basis, maximize), reflects that stringent standard. Suppose the difference between a child being born with spina bifida or not is (just) the difference between the woman’s taking a vitamin just after conception or not. Clearly, the woman’s refusal to take the vitamin harms the child then born with spina bifida. That is so even where that child’s life will be unambiguously worth living if the woman does not take the vitamin.²⁴ But once we establish that the act harms an existing or future person, we immediately avoid, under PBI, the implication that no wrong has been done. If the case is also one in which the woman’s taking the vitamin does not harm any other existing or future person, including the woman herself, and does not bring still other persons

into existence who are then maltreated, PBA can be expected to do still more—that is, to imply that what the woman has done is wrong.

10.2.5 Genetic Interventions

A last point in favor of PBA is that it adjusts its results nicely in the face of technological change. Suppose that a genetic technology has been developed that corrects the Huntington's gene in the newly formed embryo. And suppose that that technology has been made available to the couple but they refuse to allow it to be used to correct "their" embryo's mutated gene. However ample their resources and unambiguously worth living the child's life will be even without the correction, the couple cannot rely on PBI to make the case that what they have done is permissible. For their choice clearly harms their child: they have created less wellbeing for that child when they could have created more. Moreover, depending on the tradeoffs—if any—that are involved, PBA will imply that the couple's refusal to allow the correction to be made is wrong.

10.2.6 Fit with Law

The particular branches of the law that most directly address issues relating to future persons—constitutional privacy law, tort law and family law—themselves seem intransigently person-based in nature. Tort law ties wrongdoing to the harming of some person or another. And constitutional privacy law allows the state to regulate what it considers to be wrongdoing only when the state can tie that wrongdoing to the harming of persons—existing or future.²⁵ Family law focuses on risks to offspring. Where would-be parents can demonstrate their conduct is devoid of any risk to any existing or future person, the provisions of family law that would normally permit agents of the state to intervene in the parent-child relationship or remove children from the custody and control of their parents have no application.

What this means is that a person-based approach has the capacity to provide guidance in dealing with hard cases involving future persons in respect of which the law itself is indeterminate. In contrast, the legal system and totalism (or averagingism or pluralism) will be like two great ships passing in the night (and one of them will be the Titanic). Moral theory will generate findings of harmless wrongs that courts will find legally irrelevant—or will consider acts to be permissible that courts will view with alarm. More generally, an impersonal, aggregative form of consequentialism will not have a sufficiently good "fit" with the law to be useful in helping us to sort through the complex array of future person cases we now face—from abortion to the "custody" of frozen embryos to supernumerary pregnancy to global warming.

We could, of course, always restructure the law—a mere human product—along impersonal, aggregative lines. But such a restructuring is one that we should at least hesitate to undertake. PBA is well-entrenched within the law, and the law itself is a product of thousands of years of intense human effort. Moreover, the aggregative

approach itself is hardly without blemish. None of this means PBA is correct, of course. But it does mean that it is worth a close look.

10.3 The Nonidentity Problem

10.3.1 *Two Types of Nonidentity Problems*

The nonidentity problem is really just a collection of different problems displaying distinct logical features. Those problems, accordingly, can be typed. When we fail to sort nonidentity problems in accordance with their types, we may well think “the nonidentity problem” shows that some “bad” acts are “bad for” no one. In contrast, when we analyze the problems in accordance with their type, we come to quite different results. We can then see that (1) the problems that really do demonstrate “no harm done” to any person—the “can’t-*do*-better” problems—are exactly those in respect of which it never becomes quite clear that a wrong has been done, and (2) the problems that involve acts that are clearly wrong—the “can’t-*expect*-better” problems—are exactly those in respect of which it never becomes clear that that same person has not been harmed.

10.3.2 *The Can’t-Do-Better Problem*

The can’t-*do*-better problem arises when the agent’s procreative choice harms no existing or future person—that is, when it creates at least as much wellbeing for each existing and future person as any alternative choice the agent might have made instead.²⁶ Suppose that a couple’s choice to bring a Huntington’s child into existence meets that condition. Suppose, also, that the couple had the alternative of bringing into existence a nonidentical, genetically healthier child in place of the Huntington’s child. The can’t-*do*-better problem successfully challenges PBI only if we agree that, on those facts, the couple’s choice is wrong. But is it really so clear to us that it is?

In answering this question, it is critical to note just how rare the bona fide “no harm done” case really is. Even if the *procreative* effect of the couple’s procreative choice does not constitute a harm to the impaired child, the *distributive* effects of that very same choice may well constitute harms—to the impaired child’s older or younger siblings or to the impaired child or to both. Distributive harms will arise when the couple scrimps on the resources—time, energy and money—they would otherwise expend on the impaired child in order to insure that that child’s older and younger siblings are protected from any ill effects of their choice. Alternatively, the couple may scrimp on expenditures for those other children in order to insure that the impaired child has the care that he or she requires. Either way, they have put themselves in a moral bind by making the procreative choice in the way that they have. They have insured that harm will befall some or all of their offspring.²⁷

But if there is harm—if the act is “bad for” someone or another—then the case is not an instance of the can’t-do-better problem. The can’t-do-better problem arises *only if* the couple’s choice harms *no one—only if*, that is, it is *maximizing* for each existing and future person. It seems to me that, in that rare case, any sentiment that we might have that the couple’s choice is wrong will itself begin to fade. If so, then our conviction that the case proves PBI false must begin to fade as well.

10.3.3 *The Can’t-Expect-Better Problem*

In contrast to the can’t-do-better problem, the can’t-expect-better problem includes many cases in which the choices under scrutiny seem incontrovertibly wrong. In any such case, rescuing PBI requires a showing that the logic that takes us to the “no harm done” result is itself mistaken.

I will focus on just one instance of the can’t-expect-better problem here—Kavka’s slave child case.²⁸ There, a couple enters into a binding, enforceable contract with a wealthy man according to which the couple will conceive and bear a child who will be transferred at birth to the wealthy man as a slave. In exchange, they will receive \$50,000, which they will use not to save the world but to buy a yacht. The couple then produces a child as a slave. Let’s call that child “p.” Despite p’s status as slave, p’s life is worth living.

Has the couple’s act of bringing a child into existence in this particular way *harmed* p—created, that is, less wellbeing for p when the couple had the alternative of creating more?²⁹ The couple, of course, had a number of alternatives—including not entering into the contract and still taking steps to produce a child, and not entering the contract and *not* taking steps to produce a child. The latter of those two alternatives obviously would not have generated any additional wellbeing for p. But what about the former?

Kavka, citing the “precariousness” of existence, argues that it, too, has little to offer p.³⁰ After all, had the couple not acted just as they did, what would the chances have been that the very same gametes that happened to combine to produce p would still have combined? What would the chances have been that the very same sperm (out of hundreds of millions!) would still have inseminated the very same egg? Practically none at all. Surely, then, from p’s own point of view, it is better for the couple to enter into the contract. Surely their choice, if anything, benefits p rather than harms p.

10.3.4 *A Closer Look at the Can’t-Expect-Better Problem*

We need to look more closely at this argument. Let “A” be the couple’s act of entering into the contract and taking steps to produce a child (their entering into the contract and, let’s suppose, having sex). That is what the couple has in fact done. Let “B” be their act of *not* entering into the contract yet still taking steps to produce a child. B clearly existed as an alternative for the couple at the critical moment just prior to their performance of A. That is so, even if it is also true that, had the couple

not performed A, they *would not* have chosen B—even if, that is, they would have instead refrained from producing a child altogether (“C”).³¹

We should agree that the *probability* that p will come into existence, given B, is very, very low.³² That is simply to recognize the phenomenon of the precariousness of existence. Moreover, by hypothesis, A confers on p an existence worth having. One act, B, generates for p an unbelievably tiny chance of ever coming into existence at all, while another act, A, generates for p an existence worth having.

These points seem clearly correct. Nonetheless, if we try to determine on the basis of just these points whether A really is at least as good for p as B is, or whether A harms p, then we allow ourselves to be rushed down the garden path. This is so, for two reasons. (1) The claim that p’s chances of existence are very small, given B, at least bears rewriting. For surely just how small p’s chances are of coming into existence, given B, depends on just how B itself is realized. (2) Moreover, betterness presumably is not established by comparing the *probability* of p’s achieving a certain wellbeing level given one choice against the *actual wellbeing level* in fact generated for p by another. After all, probability is just a number between 0 and 1. Yes, the actual wellbeing level that A generates for p may well be substantially greater than any such number. But it would be premature to conclude, on that basis alone, that A is better than B for p. We can take these two points in turn.

(1) *Choice of B as a way into existence for p.* Let’s call the highly particularized way in which the couple in fact realizes the generic A—how they realize A at, e.g., the *actual world*—“A*^{act}”. We can note, then, that there are a lot of ways in which the couple could have realized A that do not bring p into existence. It is their concrete performance of A in all its identity-influencing detail—their performance of A* in place of A^{act}, A^{act}—that brings p into existence. But those details have *nothing* to do with the fact that A* is an act of entering into the slave child contract and *everything* to do with the various spatial-temporal-mechanical characteristics of A* that put the right sperm in the right place at the right time.

That, in turn, means that the alternative act B *can* equally well put the right sperm in the right place at the right time. It will all depend on *just how* the couple goes about performing B. Clearly, among the many, many alternative ways of realizing the generic B, there exist some that *perfectly mimic A* in all respects relevant to p’s coming into existence*—that include, that is, the very same spatial-temporal-mechanical identity-influencing details we find in A*. Call some such concrete, particular way of realizing B “B*.”

We should note, as well, that B* clearly exists as an alternative for the couple at that critical time just prior to performance. Nothing in natural law or the acts of other agents prevents the couple from performing B* in place of A*. They *can* perform B* just as easily as they *can* perform A*.

Moreover, as between A* and B*, there is no basis for thinking that the couple’s performance of B* will end in p’s never coming into existence at all. If A* fits into how the future will otherwise unfold in a way that brings p into existence as a slave, then B* will fit into how the future will otherwise unfold in a way that will bring p into existence as a non-slave. If A* makes p’s coming into existence likely,

then so does B*. If A* makes p's coming into existence a tiny bit less than highly improbable, so does B*.

Can't we then, after all, say that the couple has created less wellbeing for p when they could have created more—that A* is not, after all, at least as good for p as B* is—and hence that what the couple has done in performing A* harms p?

We *might well* say that—but only if we are *actual value* consequentialists.³³ But we might not be actual value consequentialists. We might think instead that moral law is based not on *actual* value but rather on *expected* value. For we might (not implausibly) believe that the moral assessments that we make, at least in theory, can have an *action-guiding function*. An assessment of wrongdoing based on actual value is one that can be reached only in hindsight (if at all). That approach leaves the agent open to charges of wrongdoing even when the agent is conceded to have made the best possible choice given all the information within the agent's grasp at that critical moment just prior to choice. The great thing about expected value is that it is something we can calculate, at least in roughly, before we act. It takes into account that we cannot know precisely how the future will unfold and, indeed, that how the future will unfold may not be a determinate matter of fact.

If we think that expected value is critical to the issue of permissibility, then we will find in the slave child case a far more formidable challenge to PBI. We agree that A* and B* are on par in terms of bringing p into existence—that if A* does so (and it does), B* would have done so as well. Yet even then the can't-expect-better problem may seem able to argue its way to the results that A* is at least as good for p as B* is and that A* does not harm p.

(2) *Relevance of probability*. If probability is important in connection with the slave child case, it is important because it tells us something about how much *expected value* B generates for p. And hence we call the particular type of non-identity problem that relies on probability the “can't-expect-better” problem.

Let's just pause here to note an important constraint that our analysis will be subject to if we do indeed opt to jettison an actual value theory in favor of an expected value theory. For a moral theory to have the action-guiding function that the concept of expected value is meant to make possible, expected value must itself be something that we can calculate, at least roughly, *just prior to choice*. That fact, in turn, constrains—and, indeed, helps us to identify—just which probabilities can properly be understood to bear on that calculation. In particular, we must accept the following rule:

Calculation-prior-to-choice constraint: The probabilities relevant to the calculation of expected value are just those determined on the basis of the information within the agent's grasp at the critical moment just prior to performance.

This constraint sometimes does not make itself felt in any obvious way. It does not, for example, lead to any surprising perturbations in how we calculate the expected value for the generic B. Let “t0” be the moment just prior to performance. Suppose that as the future unfolds and given the couple's choice of A*, p's level of wellbeing is +100. And suppose that p's level of wellbeing at a world where B is chosen over A* and p (against all odds) exists as a nonslave is (estimating wildly)

+200. We have already conceded that the probability of *p*'s coming into existence, given *B*, is very low. And that assessment does not change when we explicitly restrict the information on which the assessment is based to just that information within the couple's grasp at *t*₀. The probability of *p*'s coming into existence, given *B*, at *t*₀ remains very low. Suppose that that probability (estimating still more wildly now) is .0000000001.

Adapting, then, the standard expected value formula for purposes of PBA, we take the summation of the actual values for *p* of each possible outcome of that act *B* multiplied by the probability that that outcome will obtain, given the performance of *B*. On the assumption that the value of never having existed at all for *p* is itself zero, we then can write:

$$EV(B, p, t_0) = .0000000001(200) + .999999999(0)$$

And we conclude—as anticipated—that the generic *B* generates precious little expected value for *p*.

We are now in a position to compare—not a mere probability, a number between 0 and 1—but a *value*—the *expected* value generated by *B* for *p*. But against what? What the can't-expect-better problem seems to ask us to do is compare that very low *expected* value generated by *B* for *p* against the very high *actual* value generated by *A** for *p*—that is, +100—and then to obtain that *A** is, after all, at least as good for *p* as *B* is.

Now, I argue, in part 10.3.5 just below, that we cannot validly reach that particular betterness result. But first let's complete the argument, to see just how the can't-expect-better problem is *supposed* to achieve its ultimate result that *A** does not harm and is not “bad for” *p*.

It would be premature to infer, from the bare fact that *A** is at least as good for *p* as *B* is, that *A** does not harm *p*. We must first determine more generally that *A** is at least as good for *p* as *any* other alternative is for *p*. And in that connection, we must obviously consider *B**, which includes all the identity-influencing spatio-temporal-mechanical details that we find in *A**. It has been conceded that *p* would have come into existence, given the substitution of *B** in for *A**. It might thus seem that, while $EV(B, p, t_0)$ is very low, $EV(B^*, p, t_0)$ should be substantially higher.

But that would be a mistake. Here, the calculation-prior-to-choice constraint does make itself clearly felt. Under that constraint, the optimistic probability of *p*'s coming into existence that we may think attaches to *B**—a number that, in some objective sense may well be correct and that we have no reason to abandon³⁴—must nonetheless be set aside. If expected value is to have the action-guiding function that is its very purpose, then the probabilities we rely on in calculating expected value are limited to those that are determined *on the basis of just that information that is within the agent's grasp at t*₀. But think about how little is within the couple's grasp at *t*₀—just how little is settled for them at *t*₀! Perhaps it is settled for them at *t*₀ that they will choose *A* rather than *B* or *C* and will together produce some child or another—call it “*p*”—as a slave.³⁵ It is not at all settled for them that the probability of *p*'s coming into existence, given *B**, is anything more than “very low.” Perhaps it is true that (in some objective sense) *p*'s coming into existence, given *B**, is greater

than “very low.” The problem is that that is nothing the couple can grasp at t_0 . The couple has no basis for correlating their performance of B^* with the coming into existence of p . For all they know at t_0 , B^* is just one of the many, many ways of realizing B —ways that include B^{**} , B^{***} etc.—that would have taken p off track for existence altogether. Calculated, then, on the basis of *just that information that is within the couple’s grasp at t_0* , the probability of p ’s coming into existence, given B^* , remains very, very low.

Things would be quite different if the couple could grasp at t_0 that the particular way they will realize A will be by performing A^* . For then they would have a basis for correlating their performance of B^* with the coming into existence of p . That is: having made the correlation between A^* and p , they could then make the correlation between B^* , which exactly mimics A^* in all of its identity-influencing spatial-temporal-mechanical details, and p . The problem is that, at t_0 , it remains highly unsettled for them, and may not then even be a determinate matter of fact, that their choice of A will be realized by A^* . For all they then know, they will realize A not by the performance of A^* but rather by the performance of A^{**} or A^{***} or etc.³⁶

We must now concede that the probability that p will come into existence, given B^* and calculated on the basis of *just the information within the couple’s grasp at t_0* , remains very low. But that means that the *expected* value B^* generates for p is very low as well. Yet the *actual* value A^* generates for p has not changed. It is still quite high (+100). Yet B^* constitutes p ’s last, best hope of coming into existence as a non-slave. And thus the argument concludes: no alternative is better for p than A^* is—and A^* thus does not harm p .

10.3.5 The Mistake in the “Can’t-Expect-Better-than-Get” Problem

The difficulty with the argument we have just described is that we cannot reliably determine betterness, and ultimately harm, on the basis of a comparison between two radically distinct sorts of value—the very low *expected* value generated by B (or B^*) for p against the relatively high level of *actual* value that has in fact been generated for p by A^* . We can certainly compare those two numbers and obtain the result that $AV(A^*, p) > EV(B^*, p, t_0)$. But it is a mistake to think that comparison has anything to do with betterness, or harm.

And we can count the ways why that is so. For one thing, a rule that determines betterness by comparing actual against expected values is inconsistent. After all, $AV(B^*, p)$ may be quite high. Suppose that it is. Yet $EV(A^*, p, t_0)$ is very low (we come back to this point in what follows; for now, it is enough to note that, for the same reason $EV(B^*, p, t_0)$ is very low, so is $EV(A^*, p, t_0)$). If we think betterness can be established by a comparison between actual and expected values, we must now say that B^* is better for p than A is. But this way of thinking has already committed us to the view that A is at least as good for p as B^* is. Given that betterness is (surely) anti-symmetric, we now face an inconsistency.

For another, it intuitively seems that the only way we can apply an actual-against-expected comparison to get the result that A^* is at least as good for p as B^* is will be to equivocate on “value.” “Value” has to mean, first, *actual* value and, second, *expected* value if the comparison is to connect in the right way to the facts of the case. But betterness for a person has to do with producing *more of a certain stuff* for that person. It is true that the *number* representing $AV(A^*, p)$ is greater than the *number* representing $EV(B^*, p, t_0)$. But it does not follow that A^* is better for p than B^* is, any more than a comparison between five sifted and four unsifted cups of flour demonstrates betterness in respect of—i.e. more—flour.

Once we have the two numbers—the actual value of A ; the expected value of B^* —in which we have a great deal of confidence—and we do—it then becomes very hard *not* to compare those two numbers and think we are validly coming to an accurate result on betterness, and ultimately harm. But we aren’t.

We can call this variation on the can’t-expect-better problem the “can’t-expect-better-than-*get*” problem. It is a fallacy—a bit of reasoning that seems compelling on its face but in fact is mistaken and should be rejected.

10.3.6 A Variation on the Can’t-Expect-Better Problem

A critic might object that I have misconstrued the slave child case. Perhaps the “no harm done” result is supposed to be derived not from the problematic actual-against-expected comparison but rather from a seemingly more legitimate expected-against-expected comparison. We can call this variation on the argument the “can’t-expect-better-than-*expect*” problem.

We have said that the probability, calculated on the basis of just the information within the couple’s grasp at t_0 , that p will come into existence, given B^* , is very low. It may seem that A —and surely A^* !—makes p ’s coming into existence significantly more likely than *that*. It may seem, then, that we can conclude that $EV(A, p, t_0) > EV(B^*, p, t_0)$ —or at least that $EV(A^*, p, t_0) > EV(B^*, p, t_0)$. Either way, we can conclude, now under the more legitimate expected-against-expected comparison, that A , or at least A^* , is better for p than B^* is and thus does not harm p .

10.3.7 The Mistake in the “Can’t-Expect-Better-than-Expect” Problem

This argument, however, fails as well. There are three questions that we should closely consider.

(1) *How much expected value does A generate for p ?* The difficulty with this argument is that the probability of p ’s coming into existence, given A and calculated on the basis of just that information that is within the couple’s grasp at t_0 , is itself very low.

Suppose, as before, that it is settled for the couple at t_0 that they will choose A rather than B or C and that, as a result of that choice, they will together produce

some child or another p whose existence as a slave will be flawed yet still worth having. That information is within their grasp at t_0 . And it is a lot of information: on the basis of that information, the couple can, for example, reasonably project at t_0 that the *actual* value A will generate for that child p will in fact be relatively high—around $+100$.

But the *actual* value A will generate for p is no longer at play. Pertinent now is the question of how much *expected* value A will generate for p . To reach the result that A generates so much *expected* value for p that A is at least as good for p as B^* is, the couple would also need to grasp that the probability of p 's coming into existence, given A , is itself significant (is greater, that is, than the very low probability the couple is in a position to assign to p 's coming into existence, given B^*). And, according to the calculation-prior-to-choice constraint, that would be something they would need to grasp prior to choice—prior, that is, to t_0 . But they can't. At least, on the highly plausible assumption that it remains unsettled for them, at t_0 , that they will happen to realize A by performing A^* rather than A^{**} , A^{***} , etc., all they can grasp at t_0 is that there are a lot of ways of performing A that will mean that p will never exist at all.

In short: all the couple can grasp at t_0 is this: some child or another will be brought into existence by their choice of A , *and* whichever child that happens to be—call it " p "—will have been brought into existence *against all odds*. But since the probability of p 's coming into existence given A , calculated on the basis of just that information within the couple's grasp at t_0 , is very low, so is $EV(A, p, t_0)$.

(2) *How much expected value does A^* generate for p ?* Do things change when we turn to A^* ? Is $EV(A^*, p, t_0) >$ than $EV(A, p, t_0)$? More to the point, is $EV(A^*, p, t_0) >$ than $EV(B^*, p, t_0)$? It is always hard, when we know just how some question of fact has ultimately been settled, to keep in mind that we must think about things as though it has not. In calculating the expected value that A^* generates for p , however, that is exactly what we must do.

We retain the (highly plausible) supposition that it remains unsettled for the couple, at t_0 , that they will happen to realize A by performing A^* rather than A^{**} , A^{***} , etc. We then find that, calculated on the basis of just that information within the couple's grasp at t_0 , the probability of p 's coming into existence, given A^* , remains very low. That is so, for exactly the same reasons that we said before that the probability of p 's coming into existence, given B^* , is very low. We concede (as we did for B^*) that the performance of A^* in place of A^{**} , A^{***} , etc. increases (in some objective sense) the probability of p 's coming into existence; we concede that it is true at t_0 that p will, or probably will, come into existence, given A^* . Under the calculation-prior-to-choice constraint, however, that probability must be set aside as irrelevant to the calculation of expected value. The relevant probability is, instead, what can be calculated on the basis of *just the information within the couple's grasp at t_0* . But at *that* time the couple has no basis for correlating their performance of A^* with the coming into existence of p . For all they know at t_0 , A^* is just one of the many, many ways of realizing A —ways that include A^{**} , A^{***} etc.—that will take p off track for existence altogether. Calculated, then, on the basis of *just the information within the couple's grasp at t_0* , the probability of p 's coming into existence, given A^* , remains very low.

(3) *Abandoning the plausible supposition.* In addressing questions (1) and (2), we have supposed that the couple does not happen to know in advance that they will realize A by performing A* rather than A**, A***, etc. Would that bit of foreknowledge have changed the analysis?

No. We still cannot reach the result that A* is better for p than B, or B*, is—or the critical “no harm done” result. Since the couple now is understood to grasp that they will realize A by performing A*, they can also grasp just how to go about performing B in a way that will perfectly mimic A* in all its critical identity-influencing, spatial-temporal-mechanical detail. Suppose, e.g., that part of A* is the fact that the couple will expend exactly six seconds actually signing the slave child contract. All of that is within their grasp since (by supposition) it is within their grasp that they will realize A by way of performing A*. They, moreover, understand that that particular sequencing will bring their child p “a step closer” to coming into existence. They are thus now in a position to identify an act B* that has exactly those identity influencing features—an act B* that involves, e.g., *feigning* to sign the slave child contract for exactly six seconds. The upshot is that—still calculating on the basis of just the information that is within the couple’s grasp at t0—whatever the probability of p’s coming into existence, given A*, will also be the probability of p’s coming into existence, given B*.

But since the actual value p will enjoy, given that B* is performed and p comes into existence, will be far greater than it is under A*, we now can obtain that $EV(B^*, p, t_0) > EV(A^*, p, t_0)$. We thus again block the inference to the result that A* is at least as good for p as B* is and that A* therefore does not harm p.³⁷

10.3.8 *Caveat*

My claim here is not that failing to maximize expected value for p in itself *harms* p. Rather, it creates a *risk*. It increases the chance of opening the door to a causal chain that will end badly for p. Where that risk eventuates—as it does in the case at hand; p has not somehow lucked out and gotten to exist as a nonslave; p’s actual wellbeing level at the end of the day remains avoidably diminished by p’s status as slave—we can say that what the couple has done *harms* p.

10.4 The Two-Envelope Problem

10.4.1 *The Argument for Switching*

In the two-envelope problem, two amounts of money are covertly placed in distinct envelopes. The envelopes are displayed to the subject, who is reliably told that one amount is twice the other. The subject is also told that he (or she; let’s suppose he) may choose one of the two envelopes to keep. He arbitrarily chooses an envelope. He is then offered the option of switching. Where we let “S” stand for the act of switching and let “m” designate the value of the contents of the selected envelope

(the “in-hand” envelope), and where we understand that the probability is .5 that the “out-of-hand” envelope contains the greater amount, we calculate as follows:

$$EV(S) = .5(2m) + .5(.5m) = 1.25m$$

Since the value of not switching, that is, holding (“H”), is m , and since $1.25m > m$, we conclude that S is better for the subject than H is. But that is an odd result. Surely the subject’s initial arbitrary selection of the one envelope does not really involve a mysterious creation-at-a-distance of new value in the out-of-hand envelope.

Things get worse. Let “ n ” designate the contents of the out-of-hand envelope. We now obtain that $EV(H) = 1.25n$. Since $1.25n > n$, we also obtain that H is better for the subject than S is. But since betterness is anti-symmetric, we face a contradiction—just as we did in the can’t-expect-better-than-*get* variation on the slave child case.

Why not deny here, as we did there, that betterness can be reliably determined on the basis of an actual-against-expected comparison? We then remain free to take either the view that betterness is to be determined by an actual-against-actual comparison or the view that betterness is to be determined by an expected-against-expected comparison. Whichever view we take, however, we (appropriately) find ourselves unable to reach any betterness result at all.³⁸ We thus avoid both the odd result and the contradiction.

But not for long. Just as we did in the context of the slave child case, we can rewrite the two-envelope problem as an expected-against-expected problem. And we know independently that if we jettison actual-against-expected comparisons we will need to be willing to take the position that such a rewriting is sometimes plausible. Consider the idea that “a bird in hand is worth two in the bush.” If we think actual-against-expected comparisons do not reliably determine betterness, we can simply rewrite this perfectly cogent aphorism as an expected-against-expected comparison. It’s just that the probability of ending up with one bird is *extremely high*—close to 1—where the agent opts to hold onto the bird in hand. The EV of that choice thus approaches its AV.

A parallel approach to the two-envelope problem generates the following: $EV(H)$ is surely m (or at least very close to m).³⁹ $EV(S)$ remains $1.25m$. We again infer that S is better than H —and by an analogous string of inferences that H is better than S .

10.4.2 *The Mistake in the Argument*

What triggers the two-envelope problem is the inference from the standard formula to the result that $EV(S) = 1.25m$. But we can infer on the basis of that same formula that $EV(S) = n$.⁴⁰ Those two facts are on their own sufficient to establish a problem—at least on the assumption that, calculated for a given time, agent and world, $EV(S)$ constitutes a unique value. For it then follows that $1.25m = n$ and hence that $n > m$ —all of which can be established prior to the point at which the subject chooses whether to switch. But it can’t. The subject has *no idea* that

the actual value of the out-of-hand envelope is greater than the actual value of the in-hand envelope. Moreover, it may not even be true that $n > m$. In fact, the chances are 50–50 that it *isn't* true. One of the two candidates for $EV(S)$ must go.

But which? We originally let “ m ” rigidly designate the value of the contents of the in-hand envelope and “ n ” the value of the contents of the out-of-hand envelope. In doing so, we (implicitly) supposed the referents of “ m ” and “ n ” to be fixed—that is, that the contents of the two envelopes cannot themselves change from moment to moment or indeed at any time over the course of the game. Various theorists have relied on that fact to challenge the inference from the standard formula to the result that $EV(S) = 1.25 m$.⁴¹ On that basis, we can rule out $1.25 m$ as a candidate for $EV(S)$. We are then left to compare n and m since those are, after all, the two expected values that we can legitimately calculate. But that comparison obviously provides no grounds to think that it is better for the subject to switch since, by supposition, we do not know what the referents of “ n ” and “ m ” are and hence can make no assessment regarding how they relate.⁴²

It is worth explicitly noting why the supposition that the referents of “ m ” and “ n ” are fixed blocks the inference to the result that $EV(S) = 1.25 m$. The point can be put somewhat roughly as follows: that the referents are fixed, together with the fact that subject has made his initial choice between the two envelopes, means that it is *not* an unsettled matter of fact what outcome will obtain in the case where the subject chooses to switch. One of the two outcomes that the problem urges us to consider possible outcomes of switching has in fact been taken off the table. The subject does not know which outcome has been taken off the table, but does know that one, or the other, has. That means that there is no future left to unfold in one possible way as opposed to another. The future has already unfolded.

We can put the point more precisely.⁴³ The original articulation of the problem takes it for granted that the facts of the case support the following application of the standard formula for expected value:

$$EV(S) = .5(2 m) + .5(.5 m) = 1.25 m$$

Clearly, however, by its own terms this application generates the result that $EV(S) = 1.25 m$ only if *both* the probability is .5 that switching will yield $2 m$ *and* the probability is .5 that switching will yield $.5 m$. But those are propositions the subject does not know. Rather, what the subject actually knows is the following conjunction: *both* the probability is 0 that switching will yield $2 m$ or the probability is 1 that switching will yield $2 m$; *and* the probability is 0 that switching will yield $.5 m$ or the probability is 1 that switching will yield $.5 m$. The subject knows, in other words, that *whichever* envelope he has initially selected—whether the lower or the higher-valued envelope—the preceding conjunction will be true. But that knowledge is itself a function of the subject’s understanding that the value of the contents of the out-of-hand envelope is a settled matter of fact—that, in other words, n is fixed and will not vary over the course of the game. But that the subject knows that much—knows, that is, that the probability is 0 or 1 that switching will yield $2 m$ *and* 0 or 1 that switching will yield $.5 m$ —*precludes* the subject’s *also* knowing that the

probability is .5 that switching will yield 2 m and the probability is .5 that switching will yield .5 m. And without that knowledge, the conditions on the particular application of the standard formula set forth above will remain unsatisfied.⁴⁴

The upshot is that the two envelope game misuses the standard formula. It has taken it for granted that the conditions in the application set forth above have been met—in particular, that the probability is .5 that switching will yield 2 m and .5 that switching will yield .5 m—when they aren't. It is not that the standard formula *cannot* be correctly used in the context of the two envelope problem. We can, for example, use the formula to infer that the expected value for switching approaches n. But we cannot infer anything from the standard formula when its conditions aren't met. We thus are never entitled to reach the result that $EV(S) = 1.25 m$.

Why did we think, even for a moment, that we know that the probability is .5 that switching will yield 2 m and the probability is .5 that switching will yield .5 m? Because, *coming into the game*, the probability is .5 that the envelope the subject will select is the higher-valued envelope *and* .5 that the envelope the subject will select is the lower-valued envelope. So it's a game worth playing—even if the price of admission happens itself to be whatever value the lower-valued envelope happens to contain. The problem then tries to beguile us into thinking that we know that those same statistics apply even after the initial selection of an envelope takes place. But they don't. That initial selection changes everything, when it is paired with the fact that the value of the contents of the out-of-hand envelope does not change.

Critical to this way of eliminating the basis for the claim that $EV(S) = 1.25 m$ is the very natural supposition that “m” and “n” are fixed for the course of the game—that is, that the values of the contents of the two envelopes cannot change as the game progresses. A look at an alternative, less natural (one might even say preposterous) supposition clarifies, I think, just how inapt the standard formula is—inapt in the sense that its conditions are not satisfied—when we make the more natural supposition instead.⁴⁵ What (we said) blocks the application of the standard formula in the original problem was the supposition that the contents of the two envelopes are fixed throughout the course of the game—that “m” and “n” rigidly designate particular quantities. If we reject that supposition, and put in its place the less natural supposition that the contents of the out-of-hand envelope *can* change, even from moment to moment, as the game progresses, then our analysis changes dramatically.

Suppose, then, that the value of the contents of the out-of-hand envelope *remains unsettled* even *after* the subject has made the initial selection between envelopes. Suppose, in particular, that the contents of the out-of-hand envelope shift in a way that makes the following claim—false in the context of the original problem—true: that, independently of whether the envelope initially selected (the “in-hand” envelope) contained the greater- or lesser-valued contents as of that moment immediately prior to selection, the probability that the value of the contents of the out-of-hand envelope will be twice the value of the contents of the in-hand envelope is .5 and the probability that the value of the contents of the out-of-hand will be half the value of the contents of the in-hand envelope is also .5. (We might also have supposed that the contents of the out-of-hand envelope shift in a way that makes it true that the

probability of the value of the contents of the out-of-hand envelope will be twice the value of the contents of the in-hand envelope is, say, .8; but we don't.) If that is the supposition we make—if that is the way the game works—the standard formula applies quite nicely. The referent of “m” continues to be fixed (we are not supposing that the value of the contents of the in-hand envelope will vary); and we can then calculate that $EV(S) = 1.25m$; with $EV(H)$ approaching m , we then obtain the result that it is better for the subject to switch.

But—on this new and unnatural supposition—it *is* better for the subject to switch! As McGrew et al. put it, in that case the standard calculation is “correct, but not at all paradoxical.”⁴⁶ Now, as between our two candidates for $EV(S)$, it is n we must set aside. For we have now abandoned the supposition that the value of the contents of the out-of-hand envelope remains fixed over the course of the game; there is now no basis on which to think that, as the game progresses, what we call “n” will not fluctuate. But that means that the standard formula will not generate the result that $EV(H) = 1.25n$. For that, we would need a fixed “n.”

Whether we stay with the original, natural supposition, or jettison it in favor of the new and unnatural supposition, we find ourselves arguing to inconsistency *only if* we fail to see the inconsistency between the two suppositions and make the mistake of making inferences on the basis of both in constructing the problem.⁴⁷ We run into trouble, in other words, only when, having made the more natural supposition, and introduced the terms “m” and “n” in accordance with that supposition, we then bring to bear the less natural supposition and calculate as though the value of the contents of the out-of-hand envelope may shift over the course of the game. When we keep our inconsistent suppositions apart from one another, as we naturally do and always should, we avoid the inconsistency.

Still another scenario can also help us see just how inapt the particular application of the standard formula—the one that generates the result that $EV(S) = 1.25m$ —in fact is in the context of the original problem. This scenario involves, not an unnatural supposition, but simply a distinct supposition that we would quite naturally make had we happened to be playing a quite distinct game. We can call it the “four envelope game.”⁴⁸ In this new game, four amounts of money are covertly placed in distinct envelopes. But just two envelopes are displayed to the subject—the “middle” two. The subject is reliably told that one amount is twice the other and that he may choose one of the two envelopes to keep. He arbitrarily chooses an envelope. A third envelope is then added to the game. The envelope that is added to the game is the envelope that contains the largest amount of money *if* the subject has selected the envelope that, between the two middle envelopes, contains the larger amount; and the envelope that is added to the game is the envelope that contains the smallest amount of money *if* the subject has selected the envelope that, between the two middle envelopes, contains the smaller amount. Where we let “m” designate the value of the contents of the in-hand envelope (and understand the referent of “m” to be fixed for the course of the game), the subject is reliably told that one of the two out-of-hand envelopes contains $2m$ and one contains $.5m$ and is then given the option of switching the in-hand envelope for either one of the (now) two out-of-hand envelopes. What, then, is the $EV(S)$? As the subject understands quite

well, independent of how the value of the contents of the in-hand envelop relates to the value of the single envelope that it was paired with when the subject made his initial selection, the probability is now .5 that the subject, if he switches, will switch to the envelope containing 2 m, and .5 that the subject, if he switches, will switch to the envelope containing .5 m. But that means that the condition established by original problem's application of the standard formula is now met—and we can thus calculate that $EV(S) = 1.25$ m. So it is better to switch than to hold.

But now it *is* better to switch. Moreover, we avoid the result that it is also better to hold than to switch. That is so, since, at the moment just before choice, the value of “n” is not fixed—“n” simply abbreviates “what the subject gets if he switches.” There is thus no “one thing” that “n” can be thought to stand for—and hence no basis on which to argue that it is also better for the subject to hold rather than to switch.

10.5 Conclusion

Both the can't-expect-better problem and the two-envelope problem have us determine betterness by reference to a comparison between values that we have been beguiled into haphazardly selecting from a potpourri of actual and expected values that can be—under different suppositions about what is settled and what is not—attached to the acts under scrutiny. The truth is that calculating betterness, and harm, is a delicate matter. When we confuse what is and what is not settled for the agent at a given time, our expected value calculations become unreliable. The can't-expect-better problem imports into its scenarios critical claims that the future will unfold in a particular way when a proper calculation of expected value will set those same claims to the side. The two envelope problem reverses things. There, we are urged to think that the future has *not* unfolded in any particular way when in fact it already has. And a proper calculation of expected value will take that fact into account. It will yield not that it is better to switch, but rather that the two expected values that we can legitimately calculate—m and n—provide us with no basis for determining whether it is better to switch or not.

When we take care to select our values in a discriminating way from an orderly array—as we naturally do and always should—we come to betterness results that are not disconcerting at all. Interestingly, our results then coincide with the “naive” impressions we might have had about the problem cases when we first examined them (long ago): the slave child has been harmed, and it doesn't matter a whit whether the subject switches envelopes or not. I think in the end our best theories and best thinking lead us to exactly the same results.

Acknowledgements I owe many people thanks for their extremely helpful comments on earlier versions of this paper, including, in particular, David Wasserman, Nils Holtug, Fred Feldman, Peter Vallentyne, John Sisko, Pierre LeMorvan, Wlodek Rabinowicz, Larry Temkin and Alan McMichael. I presented an earlier version of one part of this paper at a DeCamp Seminar (University Center for Human Values, Princeton University, Oct. 2006). I owe Jeff McMahan, Peter Singer, Elizabeth Harman and other participants in that seminar thanks for their extremely helpful

comments. I presented another version at the University of California at Riverside (Feb. 2008), and owe members of the philosophy department there a large debt for their useful and stimulating discussion.

Notes

1. Parfit (1987), p. 363 and generally pp. 351–79.
2. Traditional, impersonal, aggregative forms of consequentialism, such as totalism and averaging, thus face riveting population problems, including the repugnant conclusion, the mere addition paradox, the infinite population problem and extreme inequality problems. See Parfit (1987), pp. 381–90 and 419–41; Vallentyne and Kagan (1997), pp. 5–26; and Roberts (2002), pp. 322–23. Temkin’s remains the best general introduction to this set of problems and to a view, which I will here call “pluralism,” that is intended to address the population problems while avoiding the nonidentity problem. See Temkin (1993). Pluralism can be viewed as a form of consequentialism that emphasizes a plurality of values or ideals, including the maximization of aggregate wellbeing as well as individual human flourishing and autonomy, equality and improving the lots of the least well off. See Temkin (1993), pp. 221–27. Pluralism seems plausible on its face. However, while the articulation of totalism and averaging is well underway, the articulation of pluralism is more challenging. We must identify the relevant values and provide an account of how those values are to be balanced against each another. For that reason, pluralism is difficult to test and remains hard to apply in any practical setting—for example, by judges working their way through hard cases with respect to which the text of the law is indeterminate or by women thinking through the ethics of early abortion or even contraception.
3. See e.g. Dworkin (1986).
4. Hanser, as well, distinguishes among types of nonidentity problems. See Hanser (2009).
5. See Kavka (1981), pp. 98–101; and Parfit (1987), pp. 361–66 and 371–74. See also Smolkin (1999), pp. 195–96; and Sher (2005), pp. 185–200 (discussing the nonidentity problem in the context of transgenerational compensation, the African slave trade and aboriginal land appropriation cases). See also Shiffrin (2009).
6. See generally Broome (1992). I borrow from Jamieson here, who describes global warming as the “world’s biggest collective action problem.” Jamieson (2008).
7. The “fallacy” I elsewhere describe in this paper is presented in more detail in connection with Parfit’s depletion example. See Roberts (2007). One main aim of this present paper is to provide a clearer and more well-grounded account of just why an accurate expected value calculation will not in fact yield the “no harm done” results that are usually attributed to the can’t-expect-better problem. The case I focus on here is Kavka’s slave child case, which I also discussed in Roberts (2003b). See generally Roberts (1998), Chapter 3.
8. Parfit (1987), pp. 366–71. For a brief discussion of wrongful disability, see note 27 below.
9. Kavka (1981), p. 93.
10. Parfit (1987), p. 361.
11. Kavka (1981), p. 98.
12. Extraordinary circumstances would include those in “The Negotiator,” in which the Kevin Spacey character shoots the Samuel L. Jackson character in the shoulder, thereby saving the latter from certain death at the hands of the true villains. “Harm” is open to distinct conceptions. I believe, however, that there is an ordinary, comparative sense of “harm” in which the shooting in this case does *not* constitute harm. In this sense, the fact of impairment (or, e.g., serious bodily injury) alone is not sufficient to establish harm. Similarly, I will say that (1) the meticulously performed open-heart surgery that is necessary to save a person’s life does not impose a harm at all, notwithstanding the suffering, disability and substantial period of recuperation, in the case where the life saved is worth having, whereas (2) exactly that same

meticulously performed open-heart surgery does impose a harm, in the case where an aspirin alone would have done the patient just as much good as the surgery.

13. See note 12 above.
14. I thus will not appeal to a “counterfactual,” or “but for,” account of harm. The notion that an act harms a person only if “but for” that act that person would have been better off has been clearly refuted. Suppose I shoot you in the arm, and that (I was so angry that) had I not shot you in the arm I would have shot you in the heart. I still *harm* you when I shoot you in the arm. A better account is to say that an act performed by an agent (or group of agents) harms a person if and only if that agent (or group) has in fact created less wellbeing for that person through the performance of that act when they could have (by performing an alternative act) created more. For that account of harm to be plausible, we need to recognize that the fact that an act harms a person does not, on its own, mean that the act is wrong.
15. The term “agent” must be understood to include both agents acting individually as well as groups of agents acting collectively (though not necessarily collaboratively or in concert). We will otherwise miss important instances of harm. See Roberts (2007).
16. We do face here an asymmetry, but one that seems untroubling in view of the underlying person-based principles. But see Persson (2009) and McMahan (2009).
17. For a description of pluralism, see note 2 above.
18. A handful of person-based principles are stated in more detail in Roberts (2003a, b, 2002).
19. See note 18 above.
20. This is Caspar Hare’s example. Hare (2007), pp. 498–511.
21. Hare concedes that PBA should not be interpreted as a form of what he calls *strong* moral actualism. But, contrary to Hare, there is also no reason to think that PBA should be interpreted as a form of what he calls *weak* moral actualism, a view that would assess as impermissible the act that fails to bring the additional persons into existence in the circumstances described here. But that reading of PBA seems unnecessary. PBA should, instead, be understood to take into account the highly person-affecting fact that the additional person will suffer—in an entirely avoidable way—if brought into existence as a slave or an organ donor. If failing to increase wellbeing for an already relatively well-off person *p* is necessary to avoid bringing another person *q* into existence who will then be treated very badly, PBA, to remain credible, should be understood to imply that failing to increase wellbeing for *p* is permissible. Tradeoffs, including trans-world tradeoffs, are going to be an important part of any plausible form of PBA.
22. Narveson (1976), p. 73. See also Narveson (1967), p. 65. See too Heyd (2009).
23. Narveson (1976), p. 73. See also Narveson (1978), pp. 55–56 (adopting impersonal principle in response to the nonidentity problem).
24. Her refusal to take the vitamin harms the child, as well, where the woman accurately claims that, if she had been required to take the vitamin, she would never have had the child at all. See note 14 above (on the counterfactual account of harm).
25. Philip Peters agrees that tort law embraces a person-affecting approach but argues that constitutional privacy law is best understood to include both person-affecting and impersonal values. See Peters (2009). In taking the position that constitutional privacy law is best understood as person-affecting in nature, I am adopting a view that John Robertson has described in substantial detail. See Robertson (2004) and Robertson (1994), pp. 22–42, 75–76 and 168–71.
26. Another example of the can’t-do-better problem is Parfit’s “two medical programmes” example. Parfit (1987), pp. 366–71.
27. Sometimes, of course, the impaired child exists as a result not of parental choice but rather of health care provider negligence. This happens, when the provider fails to diagnose or inform couples of their elevated risk of producing a genetically or congenitally impaired child. For the same reasons, however, that the couple can harm their own offspring by choosing to produce the impaired child in place of the healthier child, so can the provider harm the impaired child or that child’s older or younger siblings or both. We must, in other words, take into account not just the *procreative* effect of the provider’s negligence on the impaired child, but also the many *distributive* effects of that negligence. On those grounds, I argue that the impaired child—in conjunction with any siblings—may have a valid cause of action

- against the provider even in the case where the existence itself is clearly worth having. See Roberts (2008). Using the terminology of Buchanan et al., we perhaps most accurately call this claim “wrongful disability” rather than “wrongful life.” See Buchanan et al. (2002), pp. 222–57.
28. I have discussed the version of the nonidentity problem elsewhere as well. See note 7 above.
 29. I assume here that no interesting reading of the nonidentity problem will rely on the narrower, counterfactual (or “but for”) account of harm. See note 14 above.
 30. Kavka (1982), p. 93.
 31. Thus, the fact that p would have never existed at all had the couple not chosen A is not relevant to the question of whether the couple, in choosing A, harms p.
 32. See Kavka (1982), p. 100 n.15 (“It is enormously improbable that the couple . . . could succeed in producing the same child . . . even if they had tried. For it is unlikely they could arrange conditions of conception similar enough to ‘what would have been’ to insure that the very same sperm would have fertilized the same egg.”).
 33. An actual value consequentialist (whether he or she adopts a total, average or person-based approach) can thus take the following position: B* is better for p than A* since it produces, for p, more actual value than A* does; A* thus harms p; and, finally, that harm (given the tradeoffs that are involved; given, e.g., that p’s being born into slavery does not somehow save the world but merely enables the couple to buy a “yacht”) is itself a wrong. That account assumes that, if B* were performed, then p would exist and be better off than p in fact is given A*. On that basis, B* is said to produce more “actual” value for p than A* does. Since B* is just like A* in all its identity-influencing, spatial-temporal-mechanical features, that assumption has strong support.
 34. We should concede that the probability of p’s coming into existence, given A* (or B*), may well be substantial—or at least greater than the very low probabilities we must restrict ourselves to in calculating expected value—in some objective sense. That increase, in effect, is a function of the fact that as the future unfolds in one way rather than another, p’s chances of coming into existence will (ordinarily) increase.
 35. Of course, the couple cannot, at t0, “know who p is”—and someone might think that that means that there is no possible act the couple could perform at t0 that could harm, or wrong, p. But it is implausible that a condition on harming a person—or on acting in a way that is morally impermissible in respect of a person—is that we know “who that person is”—or that we have any *de re* attitudes in respect of that person at all. Suppose, e.g., a man shoots into a bustling crowd of shoppers at a mall. He may not know in advance who he will end up shooting, but he plausibly knows that he will end up shooting someone or another. Call that person “p.” Whoever p happens in the end to be, it will be true that the man has harmed p, in virtue of the fact that he could have created more wellbeing for p and has instead created less.
 36. I put this point differently in Roberts (2007). There was some suggestion there that we needed to adopt a concept according to which the relevant probability (for example, the probability of p’s coming into existence, given B*) would itself change over time. The better view, which I have described here, is that what may change over time is simply the particular basis on which the probability is calculated. What changes over time, in other words, is the information that is within the grasp of the couple. The upshot is that the probability we calculate on the basis of what the couple grasps at t0 may be very low even if it is in point of fact true that the probability of p’s coming into existence, given B* and the fact that B* mimics A* in all its identity-influencing details, is very high.
 37. In addition to actual value consequentialism and the “action-guiding,” or expected value, consequentialism, still a third account of betterness, and harm, makes use of what we can describe as “objective” probabilities—probabilities, that is, that the agent just prior to choice may well have no way of grasping and thus cannot rely on for action-guiding purposes. On this view, we have no need for the assumption that, if B* were performed, then p would exist. What is relevant, rather, is that A* and B* have the same identity-influencing, spatial-temporal-mechanical features, such that, whatever the agents think, A* and B* are equally likely to bring p into existence, and that that is so even if (given that much else about the future also

remains unsettled, at the moment just prior to choice, in addition to whether A^* or B^* is to be performed) p 's coming into existence in point of fact remains highly improbable. This third view still leaves us unable to reach the result that A^* is at least as good for p as B^* is or that A^* does not harm p . Whatever the probability (relative to a particular time and world) that p will come into existence, given A^* , that is the probability (relative to that time and world) that p will come into existence, given B^* . Calculating value, then, on the basis of such "objective" probabilities, we again never reach the result that A^* is at least as good for p as B^* is. We obtain, instead, just the reverse result—that B^* is better for p than A^* is, and that (given that the risks eventuate; that p is born a slave) that A^* harms p .

38. If, as above, $1.25m$ represents the *expected* value of switching, m obviously must represent the *actual* value of not switching. (Otherwise, $EV(S)$ cannot be $1.25m$ since that amount is calculated by reference to the outcomes associated with switching. If m is not the actual value of not switching, those outcomes will be something other than $1/2$ and $2m$, and $EV(S)$ something other than $1.25m$.) It is thus an expected-against-actual comparison that has generated the betterness result in this case.

The relevance of this value— $1.25n$ —has been widely noted, often for the purpose of underlining that, once we determine that it is better for the subject, having chosen, to switch, we then determine that it is better for the subject, having switched, now to *switch back*, and so on ad infinitum. Gjelsvik, however, explicitly derives the contradiction. See Gjelsvik (2002). Chase, as well, relies on the symmetrical status of the two-envelopes in his description of what he considers a non-probabilistic version of the two-envelope problem to derive a contradiction. He writes: "Since $n > n/2$, it follows . . . that the amount you will gain, if you gain on the trade, is greater than the amount you will lose, if you lose on the trade. But an exactly parallel argument, which begins by dubbing the amount of money in the *other* envelope $\$n$, leads to the contrary conclusion that the amount you will gain, if you gain on the trade, is less than the amount you will lose, if you lose on the trade." Chase (2002), p. 158.

39. Suggestions that the comparison on which the two-envelope problem is based is expected-against-expected are scattered throughout the discussion of the two-envelope problem. However, if we do rewrite the problem in this way, we must understand "m" to designate both an expected and an actual value. For if m is not construed as an actual value in the initial construction of the problem, then $EV(S)$ would have to be something other than $1.25m$.
40. It's just that, as in the bird-in-hand case, the probability that switching will yield the value of n is very high. For all practical purposes, it is 1.
41. See especially McGrew et al. (1997), p. 29 (under the assumption that the amount in the selected envelope remains fixed, but that the "total amount involved in the game" is not, then the standard calculation is "correct, but not at all paradoxical"; under the assumption that the total amount is fixed, the "amount in the selected envelope cannot be taken as fixed. If the (fixed) total amount is, say, $3x$, then the selected envelope contains x if it contains the smaller amount, but it contains $2x$ if it contains the larger amount . . . And this means that the [standard] calculation, which assumes that the selected envelope contains *the same fixed amount* whether it is the higher or the lower envelope, is illegitimate"). See, also, Cook (2002), pp. 47 and 49.

Alternative resolutions of the two-envelope problem are grounded in mathematical considerations—relating, e.g., to the calculation of expected utility in the case where there exists an upper limit to the value of the contents of the two envelopes and to the question of whether cases in which there exist no upper limit are possible. See, e.g., Clark and Shackel (2003), pp. 691–98; Meacham and Weisberg (2003), pp. 685–87; Clark and Shackel (2000), pp. 415–28; Arntzenius and McCarthy (1997), pp. 42–45; Scott and Scott (1997), pp. 37–38; and Broome (1995), pp. 6–10. For purposes here, however, we may set these discussions aside.

42. Having ruled out $1.25m$ as a candidate for $EV(S)$, we might alternatively adopt the following widely-respected account of why switching is not better than holding: let "z" designate the amount in the lower-valued envelope. Then, there are two possible outcomes, $n = z$ and $n = 2z$, each having a probability of .5. $EV(S)$ is then just $1.5z$. The parallel calculation yields

- the identical result for EV(H). We find no contradiction and no basis to switch. Gjelsvik, e.g., sets out this line of reasoning, which seems plausible. Gjelsvik (2002), p. 354. See also Schwitzgebel (2008). Of course, a full resolution of the problem also seems to require an understanding of why the standard expected value calculation does not also yield that $EV(S) = 1.25$ m.
43. A number of theorists have suggested related grounds for setting aside the results of the standard formula in the two envelope context. See note 41 above. I, however, owe my own appreciation of why the standard formula is inapplicable—and, indeed, the entire contents of the paragraph in the text that contains this note—to a constructive dilemma suggested (though not endorsed) by Ed Gettier.
 44. See note 43 above.
 45. That multiple suppositions are (arguably) consistent with the original description of the problem has been observed by various theorists. See Chase (2002), pp. 159–80; McGrew et al. (1997), pp. 28–30; Cook (2002), p. 47; and Markosian (2005). There are still other suppositions to consider as well—e.g., that the amounts in both envelopes can change as the game progresses. Working with the problem under each of those alternative suppositions—one at a time, since only one can be true in a given case—we also avoid contradiction.
 46. McGrew et al. (1997), p. 29. Markosian makes a similar point. See note 48 below.
 47. “In sum, the ‘paradox’ arises simply from a conflation of assumptions (a) and (b). Where assumption (a) is appropriate, the above calculation is legitimate, but its result is straightforwardly, and quite unparadoxically, correct. Where assumption (b) is appropriate, on the other hand, the above calculation is illegitimate.” McGrew et al. (1997), p. 30.
 48. Markosian has previously usefully contrasted the set-up we see in the original two envelope case with alternate games (e.g., what he calls “doubles or halves”). Seeing clearly just how appropriate the standard expected value calculation is in the latter case helps make vivid just why the calculation is inappropriate in the former case. See Markosian (2005).

References

- Arntzenius, F. and D. McCarthy. 1997. The two envelope paradox and infinite expectations. *Analysis* 57(1): 42–50.
- Broome, J. 1992. *Counting the costs of global warming*. Cambridge: White Horse.
- Broome, J. 1995. The two-envelope paradox. *Analysis* 55(1): 6–11.
- Buchanan, A., D. W. Brock, N. Daniels and D. Wikler. 2000. *From chance to choice: Genetics and justice*. Cambridge: Cambridge University Press.
- Chase, J. 2002. The non-probabilistic two envelope paradox. *Analysis* 62(2): 157–60.
- Clark, M. and N. Shackel. 2000. The two-envelope paradox. *Mind* 109: 415–42.
- Clark, M. and N. Shackel. 2003. Decision theory, symmetry and causal structure: Reply to Meacham and Weisberg. *Mind* 112(448): 691–700.
- Cook, M. 2002. Getting clear on the two-envelope problem. *Southwest Philosophy Review* 18(1): 45–51.
- Dworkin, R. 1986. *Law's empire*. Cambridge, Mass.: Oxford University Press.
- Gjelsvik, O. 2002. Paradox lost, but in which envelope? *Croatian Journal of Philosophy* II(6): 353–62.
- Hanser, M. 2009. Harming and procreating. In this collection.
- Hare, C. 2007. Voices from another world: Must we respect the interests of people who do not, and will never, exist? *Ethics* 117: 498–523.
- Heyd, D. 2009. The intractability of the nonidentity problem. In this collection.
- Kavka, G. 1981. The paradox of future individuals. *Philosophy & Public Affairs* 11: 93–112.
- Markosian, N. 2005. A simple solution to the two-envelope problem. Unpublished manuscript.

- McGrew, T. J., D. Shier and H. S. Silverstein. 1997. The two-envelope paradox resolved. *Analysis* 57(1): 28–33.
- McMahan, J. 2009. Asymmetries in the morality of causing people to exist. In this collection.
- Meacham, C. J. G., and J. Weisberg. 2003. Clark and Schakel on the two-envelope paradox. *Mind* 112: 685–89.
- Narveson, J. 1967. Utilitarianism and new generations. *Mind* 76(301): 62–72.
- Narveson, J. 1976. Moral problems of population. In *Ethics and population*, ed. M. D. Bayles, 59–80. Cambridge, Mass.: Schenkman Publishing Co.
- Narveson, J. 1978. Future people and us. In *Obligations to future generations*, eds. R. I. Sikora and B. Barry, 38–60. Philadelphia: Temple University Press.
- Parfit, D. 1987. *Reasons and persons*. Oxford: Clarendon Press.
- Persson, I. 2009. Rights and the asymmetry between creating good and bad lives. In this collection.
- Peters, P. 2009. Implications of the nonidentity problem for state regulation of reproductive liberty. In this collection.
- Roberts, M. A. 1998. *Child versus childmaker: Future persons and present duties in ethics and the law*. Lanham, Maryland: Rowman & Littlefield.
- Roberts, M. A. 2002. A new way of doing the best that we can: Person-based consequentialism and the equality problem. *Ethics* 112: 315–50.
- Roberts, M. A. 2003a. Can it ever have been better never to have existed at all? Person-based consequentialism and a new repugnant conclusion. *Journal of Applied Philosophy* 20(2): 159–85.
- Roberts, M. A. 2003b. Is the person-affecting intuition paradoxical? *Theory and Decision* 55(1): 1–44.
- Roberts, M. A. 2006. Supernumerary pregnancy, the problem of collective harm and the nonidentity problem. *Journal of Law, Medicine & Ethics* 34(4): 776–92.
- Roberts, M. A. 2007. The nonidentity fallacy: Harm, probability and another look at Parfit's depletion example. *Utilitas* 9 (2007) 267–311.
- Roberts, M. A. 2008. What is the wrong of wrongful disability? From chance to choice to harm to persons. *Law and Philosophy* (pub. online May 17, 2008; forthcoming in print).
- Robertson, J. A. 1994. *Children of choice: Freedom and the new reproductive technologies*. Princeton: Princeton University Press.
- Robertson, J. A. 2004. Procreative liberty and harm to offspring in assisted reproduction. *American Journal of Law and Medicine* 30(7): 14–15.
- Schwitzgebel, E. 2008. The two envelope paradox and using variables within the expectation formula. *Sorites* 20: 135–40.
- Scott, A. D. and M. Scott. 1997. What's in the two envelope paradox? *Analysis* 57(1): 34–41.
- Sher, G. 2005. Transgenerational compensation. *Philosophy & Public Affairs* 33: 185–200.
- Shiffrin, S. 2009. Reparations for U.S. slavery and justice over time. In this collection.
- Smolkin, D. 1999. Toward a rights-based solution to the non-identity problem. *Journal of Social Philosophy* 30: 195–96.
- Temkin, L. 1993. *Inequality*. New York: Oxford University Press.
- Vallentyne, P. and S. Kagan. 1997. Infinite value and finitely additive value theory. *Journal of Philosophy* 44: 5–26.

Part V

**Is the Morality of Parental Reproductive
Choice Special? Can Intentions
and Attitudes Make an Act that Harms
No One Wrong?**

Chapter 11

Reproduction, Partiality, and the Non-identity Problem

Hallvard Lillehammer

Abstract This paper argues that there are reasonable grounds for scepticism about the idea of a uniquely integrated account of the ethics of human reproduction on either partialist or impartialist terms.

Keywords Partiality · Impartiality · Human reproduction · Non-identity problem.

11.1 The Liberal View of Reproduction

Much work in contemporary bioethics defends a broadly liberal view of human reproduction. I shall take this view to comprise (but not to be exhausted by) the following four claims.¹ First, it is permissible both to reproduce and not to reproduce, either by traditional means or by means of assisted reproductive techniques such as IVF and genetic screening. Second, it is permissible either to reproduce or to adopt or otherwise foster an existing child to which one is not biologically related. Third, it is permissible either to bring into existence a child with the greatest chance of a life of maximum human flourishing or to bring into existence a child with a life worth living but with less than the greatest chance of a life of maximum human flourishing. Fourth, it is impermissible to bring into existence a child whose life is either certain or likely to fall below some baseline of a human life minimally worth living.

There is much controversy about which moral theory makes best sense of the liberal view. A number of theories currently espoused in the literature claim to cohere with the liberal view.² The measure of such coherence is arguably twofold. First, coherence requires that the theory in question is *extensionally adequate*. In other words, its practical implications must cohere with individual judgements of permissibility or impermissibility entailed by the liberal view. In this paper, I shall make the generous assumption that all theories currently on offer can meet this

H. Lillehammer (✉)

Faculty of Philosophy, University of Cambridge, Cambridge, UK
e-mail: Hallvard.Lillehammer@kings.cam.ac.uk

constraint. Second, coherence requires that the theory is *intensionally adequate*. To be intensionally adequate a moral theory must give a reflectively coherent explanation of its practical implications for individual judgements of permissibility and impermissibility. In this paper, I argue that an important class of moral theories may struggle to meet this second constraint on coherence. This is a class of theories that take an impartial perspective on beneficence as uniquely fundamental to the ethics of human reproduction. I shall take one recent and sophisticated formulation of consequentialism as a paradigm representative of theories in this class.³ This choice is partly for ease of exposition, but also partly in response to the central place of consequentialism in contemporary moral philosophy. However, my argument will arguably also apply to other moral theories that regard some impartial perspective as uniquely fundamental. I do not wish to argue that all impartialist moral theories should be rejected in favour of a uniquely partialist moral theory. My hypothesis is that partial and impartial perspectives on ethical evaluation are both irreducibly fundamental. I therefore reject the idea of a uniquely integrated moral theory on purely partialist or impartialist terms. One consequence of this conclusion for the ethics of human reproduction is that different and incompatible perspectives may reasonably be adopted in different reproductive scenarios. In particular, I suggest that some values that reasonably govern public policy in matters of human reproduction may differ from values that reasonably govern individual reproductive choice on a smaller scale. If so, the integrated theoretical approach favoured by many philosophers writing on the ethics of human reproduction is actually misleading and potentially misguided.

11.2 Reproduction and Non-identity

Reproductive choices are identity-affecting. A choice is identity affecting if it determines which among some set of possible items will actually come to exist.⁴ Reproductive choices are identity-affecting in a particularly problematic way. Many reproductive decisions involve choices between possible human lives where, depending on the choice made, different human beings will come to exist. In such cases, there is no individual person whose existence remains constant across the different choice options and about whom we can say that he or she would be better off in one scenario than in another. This claim rests on the assumption that human identity depends on genetic origin and constitution. It follows from this assumption that some of the actual spatiotemporal properties of a person's origins are essential to her identity. This assumption is rarely challenged in the literature in spite of the fact that pre-theoretical beliefs about genetics, origin, and identity are widely confused and unreliable. The genetic information embodied in the cells of one's body could be replicated and embodied in the cells of a different body. Identical twins are (near enough) genetic clones. The properties of an adult human being, including her personality and sense of identity, are not determined by her genetic constitution but are a causal product of a complex interplay of genetics, pre-natal development, and post-natal environment. Nevertheless, many people remain attached to some

kind of origin essentialism about personal identity. Perhaps this is due to difficulties of individuation arising in cases where the creation of more than one person is in question. In such cases, intuitions about identity may lose their robustness. Suppose, for example, that the parents of an only child could have had two children instead of one, each child looking pretty much identical to their actual child. Which, if any, of these possible children would be identical to their actual child? Some form of actualist essentialism about origin might be thought to offer the least painful way out of this, and related, difficulties of individuation.

Identity-affecting choices include the choices of prospective parents to reproduce at a given time, with a certain partner, and in a given way. More controversially, they include the decisions of prospective parents and health professionals to select for children with given traits by means of genetic technology, either in view of the health of the resulting child or in view of the health of some existing child for whom the resulting child can act as a donor of bone-marrow or the like. Finally, identity-affecting choices include the choices of public institutions and governments to implement policies that will affect the identity and living conditions of generations to come.

The cluster of ethical difficulties that arise from identity-affecting choices underlies what has come to be known as the ‘non-identity problem.’⁵ This problem has attracted increasing amounts of attention in recent years, both in moral philosophy and elsewhere. One central aim of moral theory as applied to the non-identity problem has been to produce an integrated account of identity-affecting choices, both in the context of human reproduction and elsewhere. On an integrated account, the non-identity problem would receive the same treatment across different domains, including the identity-affecting choices of prospective parents and groups of kin on the one hand, and governments and other public institutions on the other. The advantage promised by this strategy is obvious. A systematic and theoretically unified model of moral explanation, applicable to all scenarios where identity-affecting choices could potentially arise would simplify the intellectual challenge faced by individuals or groups when making difficult choices in which life and death are at stake. One of the most promising versions of this strategy in contemporary moral theory is that offered by sophisticated forms of consequentialism. *Qua* consequentialist, such theories construe right reproductive choice as some function of the good, considered impartially. *Qua* sophisticated, such theories construe right reproductive choice as determined by its consistency with norms or principles the near universal acceptance of which would be impartially beneficial at a given place and within a given time-frame. Such forms of sophisticated consequentialism would arguably rule out the choice of radically sub-optimal reproductive options without thereby requiring prospective parents to always reproduce for the greatest impartial benefit. They would therefore seem capable of respecting many of the partial commitments that characterise normal reproductive projects, such as the desire to have one’s own child with whom one wants, when one wants, and in the way one wants (subject to generally acceptable standards of reproductive service either licensed, provided, or otherwise accommodated by the state). To this extent, a sophisticated consequentialist approach coheres with the liberal view of reproduction.

11.3 Reproduction and Partiality

Many decisions relevant to human reproduction are normally subject to impartial constraints. Government decisions on reproductive health care policy are subject to ethical criticism if they are systematically biased in favour of one group of citizens over another. It is also natural to ask that governments consider the consequences of current reproductive policy on future generations. Other reproductive decisions relevant to human reproduction are not normally subject to comparably impartial constraints. Individual couples are not normally subject to ethical criticism for not considering the interests of all citizens equally when deciding whether to have children, with whom to have children, how many children to have, and so on. It follows that the features of options regarded as ethically most important will normally differ between general social policy on the one hand, and individual family planning on the other. This difference in ethical focus is emphasized in a recent paper by David Wasserman, who writes that “we are inclined to see the role-specific duties of parents or caretakers as quintessentially personal. After all, when philosophers attempt to justify a partiality that defies the alleged imperative to maximize aggregate welfare, they typically adduce the duties of parents to their children; the contrasting paradigm is the cold, inflexible bureaucrat.”⁶ This claim should not be taken to imply that public and familial contexts of ethical choice have no features in common. In modern liberal societies with centrally provided health-care the options available for individual reproductive choice are legally constrained by impartial considerations of potential risks and benefits. Thus, there are widely accepted restrictions on who can become a sperm donor for the purposes of IVF, for example. Nevertheless, in many modern liberal societies reproductive decisions on a familial scale are granted a significant degree of autonomy from impartial social concerns. A sophisticated consequentialist theory will be consistent with this degree of reproductive autonomy to the extent that such autonomy has a general impartial rationale. Yet why should we think that the ethical credentials of individual reproductive autonomy are hostage to its explanation on impartial terms? What, if anything, speaks in favour of assigning unique ethical priority to impartial considerations in matters of reproductive decision-making?

One way to think of reproductive choices is to think of them as the realisation of reproductive options, where options are objects of possible desire, only some of which will ever be actualised or made real. In choosing between options, ethically serious persons will consider their ethically relevant features and realise the options that appropriately instantiate them. One central question for moral theory is therefore what makes something an ethically relevant feature of options. One set of features normally endorsed as ethically relevant is the set of features that make options good in some way. Thus, the option of having children can be good insofar as its realisation will be productive of happiness. Yet options can be good both partially and impartially. The creation of a given child may not promote happiness impartially even if it is a source of great happiness for the child or its family (or vice versa). So why think the impartial goodness of an option is uniquely fundamental?

One way to think of impartiality is as an attitude of neutrality between objects in a given domain. On this understanding, partiality consists of a preference directed toward a proper subset of objects within that domain. Thus, a parent can be impartial with respect to the attention he gives to his children, or he can give some of his children special treatment. It is widely agreed that good parents are, in some sense, impartial between their own children (even if it is less often agreed what such impartiality requires in practice). It is also widely agreed that good parents need not be impartial towards all children. A parent would normally be expected to give special treatment to his own children, except in special circumstances, such as when overseeing an organised activity like a football-game. Different partial and impartial concerns coexist across a wide range of activities in the lives of ethically serious parents.

The distinction between partiality and impartiality is easiest to grasp where the options considered all involve actually identifiable individuals, such as existing children. In cases where the existence and identity of merely possible individuals are at stake the distinction between partiality and impartiality is harder to pin down. The difficulty derives in part from the idea of being partial or impartial with respect to an indefinite number of possible objects, not all of which are realisable together, and only some of which will actually exist while others will not. A human reproductive choice is but one example of a situation exemplifying each of these properties. Thus, it is natural to expect ethically serious persons to be partial toward those possible objects (in this case, people) that either do, or actually will, come to exist. Yet this partiality does not preclude the expectation that ethically serious persons should be impartial when deciding whom, among merely possible people, to cause to exist. A partial bias towards the actual is compatible with an impartial attitude towards the (as yet) merely possible.

Ethical impartiality consists in being disposed to realise those among possible options that would appropriately instantiate ethically relevant features. Suppose, for example, that human wellbeing were the one and only ethically relevant feature. If so, ethical impartiality might be thought to consist in realising those among possible options that would maximize human wellbeing. Ethical partiality, on the other hand, might be thought to consist in realising those among possible options that maximally instantiate some subset of human well-being, such as the wellbeing of the agent himself, his friends and loved ones, or some social group towards which the agent has a special attachment.

Some recent discussions of the ethics of human reproduction have focused on the distinction between person involving and non-person involving concerns in order to clarify the ethics of identity-affecting choices.⁷ Person involving concerns can be thought of as concerns for benefits and harms to identifiable individuals. Non-person involving concerns can be thought of as concerns for good and bad states of affairs involving individuals, regardless of whether any individuals can be antecedently identified as involved in those states of affairs. Non-person involving concerns are obviously central to the ethics of reproductive choice, given the absence from such choices of antecedently identifiable individuals to play the role of primary beneficiary or victim. Of course, reproductive choices always involve some identifiable

individuals, including, most obviously, the prospective parents. Person involving concerns are therefore also always present in reproductive choices.

The centrality of non-person involving concerns in reproductive choice is neutral with respect to the relative priority of partiality and impartiality. A choice can be person involving and impartial, as when someone acts to benefit living people anywhere. A choice can also be person involving and partial, as when someone acts in favour of living members of their family. Likewise, a choice can be non-person involving and impartial, as when someone acts to leave as much as possible for future generations. Finally, a choice can be non-person involving and partial, as when someone acts to leave as much as possible for future members of his family. It follows that the potential conflict between partial and impartial concerns cannot be reduced to a conflict between person involving and non-person involving concerns. This fact is significant for the ethics of human reproduction. Thus, it might be tempting to think that the ethics of reproductive choice must be fundamentally impartial because reproductive choices are identity affecting and the values at work in identity affecting choices are fundamentally non-person involving. This would be a sound argument if all non-person involving values were fundamentally impartial. But they are not. So this line of thought is mistaken. What is arguably not mistaken is to think that, at least in a wide range of cases, the affinity between partial concerns and person involving considerations is closer than the affinity between partial concerns and non-person involving considerations. Thus, it might make more sense for prospective parents to adopt an impartial perspective when thinking about possible future children than it does for actual parents to adopt an impartial perspective when dealing with their actual children. Even so, it does not follow that all non-person involving values are fundamentally impartial. This is shown, for example, by the high value placed by many parents on the option of having a child that that is “their own” in more than the obvious and trivial sense that applies to all parents (I am grateful to David Wasserman for pressing me on this point).

The conflict between partial and impartial values in human reproduction derives partly from the ethical significance of what is sometimes referred to as “special ties.” As we have seen, such ties cut across the distinction between person involving and non-person involving concerns, and include ties to individuals and groups, as well as to other personal or communal projects. Most moral theories give some place to special ties in an ethically good life. Nevertheless, ethical impartiality presents a challenge to the place of special ties. For on impartialist terms, the existence of special ties is hostage to the fortune of their coherent integration into a theoretical framework in which their place is at best contingent.

Conflicts between partial and impartial values are avoidable if one set of values is derivable from the other. Unfortunately, there is no prospect of such a derivation either way a priori. Genuine conflicts between partial and impartial values are obviously possible as a matter of logic. The prospect of some kind of a posteriori derivation might seem better, at least if suitably restricted in terms of time, space, and relevant personnel. As previously noted, the project embodied in sophisticated forms of consequentialism provides a paradigm instance of the strategy of deriving ethical partiality on ethically impartial terms.⁸ The project of justifying altruism on

self-interested terms might be thought to provide an instance of an attempt going in the other direction.⁹ The downside of such attempts is their frequent failure to capture the explanatory ambitions of the set of values reduced. Both partial and impartial values appear on the face of it to be ethically relevant in themselves. It is therefore *prima facie* unclear how any reductive project of this kind can be intensionally adequate. In the next two sections, I shall examine a series of arguments that favour the uniquely fundamental ethical status of either partiality or impartiality. I shall conclude that none of these arguments is successful on its own terms.

11.4 The Case Against Impartiality

Let us suppose that some sophisticated consequentialist account of reproductive choice is extensionally adequate. Such an account would account for the extensional correctness of partial reproductive preferences on impartial terms. Why should anyone think that more is required in order to fully explain the partial values embodied in reproductive choice?

One argument against the impartialist project is based on the principle of “no sacrifice without compensation.”¹⁰ Thus, in the case of some intrapersonal sacrifices it can be argued that sacrificing present satisfaction for future satisfaction is justified because the agent doing the sacrificing will be compensated for the sacrifice in the future. Not so for many interpersonal sacrifices, including individual sacrifices in aid of impartial good. In the interpersonal case there will be no compensation in the absence of contingent social arrangements. To this extent, interpersonal sacrifices are ethically problematic and stand in need of special justification.

The “no sacrifice without compensation” principle does not issue in blanket support for the liberal view of reproduction. It only supports this view where making a different reproductive choice would result in significant cost to the prospective parents. In cases where prospective parents could make impartially more beneficial choices at little or no cost to themselves, their reproductive autonomy is not protected by the “no sacrifice without compensation” principle. Nevertheless, the principle does offer some protection for the reproductive autonomy of prospective parents whose reproductive interests do conflict with impartial beneficence. This principle could therefore be invoked in defence of the reproductive projects of a number of minority groups whose procreative freedom would arguably be sub-optimal from a more impartial perspective.

“No sacrifice without compensation” is a substantial and non-obvious ethical principle. Its soundness is therefore in need of explanation. The coherence of sophisticated forms of consequentialism shows that this explanation could take an impartial form. If so, the mere soundness of the principle does not support a uniquely partialist theory. Furthermore, if the explanation takes a partialist form, defenders of ethical impartiality will complain that it begs the question. Thus, in one of its more plausible formulations, the principle is explained by appealing to the idea that persons are separate entities with a unique, finite, and irreplaceable first personal

perspective on the world. Yet from an impartial perspective, no person is either more or less unique, finite, or irreplaceable than any other. It is therefore unclear how the separateness of persons can be invoked in defence of ethical partiality without begging the question on the partialist's behalf.

A more plausible argument against the impartialist project derives from the role of partial considerations in pre-theoretical ethical thought. Many choices involve an impartially disproportionate concern for special ties. Consider a parent who sees a group of children threatened by fire. The parent rushes to save his child. It might be natural to think that a sufficient explanation of why the parent saved one child rather than another is that it was *his* child: 'It was my child' would be taken by many to constitute a sufficient reason to both explain and justify the parent's action. The answer: "It was my child; I care more about my child, and in this situation considerations of impartial beneficence show that to prevent people from saving those they care about most would be counterproductive" would make the parent vulnerable to the charge of having "one thought too many."¹¹ Similar considerations apply to reproductive choices. Consider a couple who learn that any child of theirs will require a higher than average economic investment in order to attain a statistically normal level of material flourishing. The couple decide to have a child anyway. They cite their desire to have a child of their own as their explanation for not selecting an alternative way of investing in the next generation. Their desire for a child of their own would for many constitute both a sufficient explanation and justification of their partially driven reproductive choice. The answer: "It will be our own child; we have a strong desire to have our own child, and in this situation considerations of impartial beneficence show that preventing couples like us from having their own child would be counterproductive in the long run" would make the parents vulnerable to the charge that they have "one thought too many." Just as parents care directly and partially about the lives of their children, prospective parents care directly and partially about their projects of reproduction. These direct and partial concerns for familial projects have deep roots in pre-theoretical ethical thought. People who do not exhibit a direct and partial concern for familial projects are often regarded with suspicion (unless they deliberately isolate themselves from familial affairs in the way members of some religious orders do, for example). Any account that fails to make sense of this direct and partial element of pre-theoretical ethical thought is a non-starter.

In spite of its intuitive appeal, the argument from explanation is not decisive. At best the argument creates a presumption in favour of the fundamental status of ethical partiality. One obvious response to the argument is to distinguish between impartial beneficence as a criterion of correctness on the one hand, and impartial beneficence as an element of a deliberative procedure on the other.¹² It does not follow from the fact that there is an impartialist account of why it is better for potential parents to have partial reproductive projects that those individuals should be thinking in terms of this account in the course of explaining their reproductive choices. Impartial concerns could favour a set of deliberative norms that primarily appeal to partial values endorsed in pre-theoretical ethical thought. If so, there is no sound basis for the charge that ethical impartiality requires ethically serious

people to have one thought too many. While this response gives rise to notorious issues about transparency and the potential for divided moral selves, its coherence in principle is enough to stop any a priori inference from the pre-theoretical centrality of partial considerations to the uniquely fundamental status of ethical partiality.

In any case, it would be mistaken to claim that pre-theoretical ethical thought is exclusively partial. A parent who showed exclusive concern for his own children would normally not be considered as admirable as someone who, while giving priority to his own children, would also be concerned to help as many other children as the situation allows. Furthermore, a parent who after saving his own child walked serenely away without attempting to help any other children would arguably be subject to more censure than someone who would save their own child first on the condition that doing so is impartially justified. Similar concerns apply to reproductive choices where there is not as yet any identifiable other to play the role of primary beneficiary or victim. This absence of an identifiable other might tempt some to conclude that only impartial considerations are relevant to such choices. I showed in the previous section that this temptation should be resisted. Yet any couple pursuing a reproductive project regardless of impartial cost, and thus in complete disregard of its consequences for actual or potential others, would arguably be subject to social censure in a similar way to parents who show no concern for children other than their own. Insofar as it suggests otherwise, the argument from explanation fails to do justice to the pre-theoretical role of impartial values. Thus, while the argument from explanation might provide a presumptive case in favour of the fundamental status of partiality, it does not undermine the idea that impartial values are equally fundamental.

A third argument against the impartialist project derives from the indispensable role of partial considerations in all practical thought.¹³ Suppose you know that Hallvard is infertile. Unless you know that you are Hallvard, you cannot use this information to benefit your family planning. Self-identification is necessary for all rational agency. Yet self-identification depends on the presence of strongly perspectival elements in thought. The fundamental perspective of the human agent is that of an individual with personal projects, acting as *I*, from *here, now*. It follows that no coherent form of self-understanding can be based on purely non-perspectival thought alone. To insist that ethical thought improves in proportion to the relative absence of perspectival elements is also implausible. Yet the perspectival elements embodied in ethical thought inevitably introduce a significant element of partiality into such thought. The place of partial considerations in ethical thought about human reproduction is therefore not a matter for philosophical argument to decide. It is both fundamental and indispensable.

This is a weak argument. The indispensability of perspectival elements in all ethical thought does not entail the soundness of a fundamentally partial ethical outlook. A fundamentally impartial ethical outlook can admit the unique significance of perspectival elements in ethical thought insofar as these elements indispensably specify the origin of agency from which impartial beneficence is to be promoted. Even if human agency is essentially perspectival in conception, it does not follow that the

acting individual cannot adopt a more or less impartial perspective when thinking about what to do. A prudentialist egoism is not an unavoidable feature of human agency. Any plausible account of human agency will be consistent with the fact that individuals, including prospective parents, have a range of available options regarding the partiality of their perspectively conceived projects. We should therefore be suspicious of any attempt to establish the character of an ethical outlook on the basis of claims about the structure of practical thought alone. The indispensability of perspectival elements in practical thought might be genuinely revealing of deep truths in semantics and the philosophy of mind. It is not thereby equally revealing of deep truths in moral philosophy.

Finally, someone might wish to defend the claim that partial values are in general more fundamental than impartial values on the grounds that impartial value exists only as a function of integrating partial values. How can partial values be ethically less fundamental than impartial values if the existence of the latter is a function of the existence of the former? It might be thought that this asymmetry gets the relationship of ethical dependence the wrong way round. Unfortunately, this argument is unconvincing. For as it stands, this argument runs together two distinct forms of priority, namely ethical priority and metaphysical priority. It may well be true that partial value is metaphysically prior to impartial value. It does not follow without further argument that partial value is ethically prior to impartial value.

11.5 The Case Against Partiality

There are a number of notable responses to the claims of ethical partiality in the current literature on human reproduction. One such response, namely that some impartialist accounts are extensionally adequate, has already been discussed and considered insufficient. A more direct response in favour of ethical impartiality derives from one of the more attractive ideas embodied in the utilitarian tradition.¹⁴ Ethical impartiality is partly a function of the unbiased integration of partial evaluative perspectives. The fact that some partial evaluative perspective does not win out in the calculation of overall goodness does not show that it has not been fairly taken account of in the process of constructing an integrated ethical perspective of beneficence. To insist that some partial perspective be given extra weight in competition with other partial perspectives is to defend an unreasonable system of selective double counting.

This argument begs the question. True, it cannot be an objection to an ethics of impartiality that some particular partial perspective fails to win out in a calculation of impartial betterness. In any conflict between perspectives, some perspective is bound to lose out on pains of inconsistency. Yet it cannot simply be assumed that the candidacy of a partial perspective to win out in ethical deliberation depends on its winning out in a process of impartial integration. Defenders of partiality deny that impartial integration is ethically exhaustive. The partialist does not need to claim either that impartial integration is impossible or that some partial values

should benefit from double counting. The reasonable partialist denies that integration amounts to reduction. He will claim that the ethical significance of partial value can survive the process of impartial integration, if only at the cost of residual conflict. The potential for such conflict is arguably reflected in the dilemmas experienced by ethically serious persons who face ethical dilemmas where individual or familial concerns conflict with the impartial values of a wider social group. The reasonable partialist claims that in some such dilemmas the ethical conflict goes “all the way down.”

A third impartialist response appeals to the sub-optimality of choices that fail to maximize impartial value.¹⁵ It is natural to think it wasteful not to maximize a value, at least when it would be possible to do so at little or no extra cost. Suppose it would be better if all people were maximally strong physically. It might then seem wasteful to bring less than maximally strong children into existence so long as children who are stronger could be caused to exist instead at little or no extra cost. To this extent, the actions of prospective parents deciding to produce less than maximally strong children could reasonably be considered ethically worse and/or wrong.

Even if it is better impartially if all children created are maximally strong, there could be partial perspectives from which it is not better that all children are created as maximally strong. Consider a currently flourishing group of less than maximally strong people who would fail to flourish in a society dominated by maximally strong people. Members of this group might prefer to live in a community where less than maximally strong people can flourish. They might prefer this to the option of being physically stronger themselves. From their perspective, it would be wasteful (or worse) to create children who are maximally strong. The intuitive idea that prospective parents should aim to maximize value is therefore ambiguous. The ethics of impartiality requires more than that prospective parents maximize value. It requires that prospective parents maximize value impartially. This claim only follows from the thought that it is wasteful to not maximize value if this thought is given an impartialist reading. The unique correctness of this reading remains to be established. Now arguably all rational agency involves the maximization of value in some sense. Perhaps not maximizing value so understood is paradigmatically irrational. Even so, maximizing value so understood is only contingently related to the maximization of impartial value. It therefore does not follow that it is irrational to fail to maximize impartial value.

A fourth impartialist response takes account of the ambiguity exposed in the previous paragraph by appealing to a distinction between rightness on the one hand, and goodness or betterness on the other.¹⁶ Ethical goodness, on this view, is impartial. Questions of ethical goodness are settled by what is better impartially considered. Answers to such questions constitute what is sometimes called “the ethics of beneficence.”¹⁷ This claim is consistent with questions of rightness not being a simple function of goodness. Questions of rights, intention, agency, and the like may function as genuine constraints or permissions with respect to promotion of the good. If so, it does not follow from the fact that some action would be better that this action is right. Suppose it would be better if no children were born with less than maximally developed muscle strength. It does not follow that it would be wrong to

have children with less than maximally developed muscle strength, provided there is a permission protecting the autonomy of prospective parents to decide whether to select for offspring with maximum muscle strength. No ethically serious person would claim that such permissions are absolute. Prospective parents should not be granted a permission to create children with severe muscle deficiency merely in order to watch them suffer. Furthermore, while it could be permissible to select for children with less than maximum muscle strength above some basic threshold, it would still be worse to do so.

Some distinction between rightness and goodness is deeply rooted in pre-theoretical ethical thought. A failure to observe it may have led some liberal-minded philosophers astray about the ethical status of impartially sub-optimal reproductive decisions. This distinction does not, however, support a fundamentally impartialist moral theory. First, some permissions and constraints on promotion of the good could have a uniquely partialist grounding. If so, the “ethics of rightness” would be fundamentally partial. Second, the notions of goodness and betterness are themselves open to partialist interpretation. Conflicts between partiality and impartiality sometimes arise because what is better from one evaluative perspective is not better from another. What is better for an individual is often not best from the more impartial perspective of his or her group, or from what Sidgwick somewhat opaquely referred to as “the point of view of the universe.”¹⁸ What is better for an individual couple, a family, an ethnic minority, or a nation is often not best from a perspective of greater impartiality. If ethical priority belongs to some maximally impartial perspective this must be established by means of an argument within the ethics of beneficence.

A fourth impartialist response claims that the idea of irreducible partiality within the ethics of beneficence is inconsistent. The argument from inconsistency is sometimes given to undermine various forms of relativism about value, whether the relativity in question obtains with respect to persons, time slices of persons, or individual projects.¹⁹ Partial value is a form of relative value, obtaining as it does relative to some less than fully integrated impartial perspective. Thus, the partial reproductive values of prospective parents may generate inconsistencies both with respect to the values of other currently existing people and with respect to the values of future, and therefore currently only possible, people. To accept that some values are irreducibly partial is to accept that some values cannot be consistently integrated into a theory of moral betterness. For some philosophers, this is sufficient reason to reject partial values in favour of an integrated account of impartial betterness without partial residue.

Some forms of evaluative relativity are incoherent. Thus, there are radical forms of time-relativity entailing that an agent’s actions are beyond rational criticism even if they are instrumentally self-defeating. This degree of value relativity is arguably irrational. Yet not all forms of value relativity have these irrational implications. Thus, the reproductive autonomy embodied by the liberal view of reproduction does not entail instrumental self-defeat. Consider a couple with less than average economic resources deciding to have their own child. Their reproductive project may conflict with values embodied by the partial perspective of possible or existing

children they might help if they did not have their own child. It might also conflict with an impartial perspective of universal beneficence. Yet it does not conflict with itself. Indeed, there is nothing internally inconsistent about either this or a plethora of other partially motivated familial reproductive projects. The fact that some forms of partial value are incoherent is therefore not an argument against all forms of evaluative partiality. The fact that all forms of partial value involve conflicts with other forms of (partial or impartial) value shows only that some partial values must lose out in a process of impartial integration. As I have already shown, this does not demonstrate that impartial integration has unique ethical priority.

A fifth impartialist response claims that fundamental partiality is unethically arbitrary. The idea that some arbitrary preferences are ethically problematic is also deeply rooted in pre-theoretical ethical thought. Widely accepted claims about justice, fairness, and equality are sometimes motivated by the idea that their contraries result from making arbitrary distinctions between people who are all worthy of comparable ethical concern. This idea has obvious implications for the ethics of human reproduction. It is widely agreed in liberal societies that individuals are entitled to some degree of reproductive autonomy. Yet the value of personal autonomy applies to all citizens equally. It is therefore wrong to prioritize the autonomy of one individual over another without explanation. Any such explanation must appeal to ethically relevant features of that individual's circumstances to distinguish them from those of others. In the absence of such an explanation, the priority granted to this individual is morally wrong.

Ethically reasonable preferences are based on ethically relevant features of options. Perfect ties aside, it is only ethically reasonable to treat options differently if there is some ethically relevant difference between them. Arguably, all features of options that make them better than others are ethically relevant. Thus, it is ethically relevant that some reproductive option will increase human wellbeing. The precise location of the realisation of some reproductive option on a spatial or temporal continuum, on the other hand, is not ethically relevant, at least not intrinsically (it might become ethically relevant because of its extrinsic properties). In general, it might be thought that the more strongly perspectival the feature of an option, the less ethically relevant it is. Thus, the purely indexical truths that something will happen either *here* or *there*, *now* or *then*, to *us* or *them* are arguably less ethically relevant than whether that something will involve happiness or pain to human beings, how much happiness or pain to human beings it will involve, and so on. It might therefore be thought that only features describable without the use of strongly perspectival terms such as the paradigm indexicals can reasonably explain a preference for one option over another. If so, many forms of partiality are ethically unreasonable as they stand. Such forms of partiality are ethically reasonable, if at all, only because they have a fundamentally impartial rationale. This claim can be forcefully applied to the issue of human reproduction. There are no non-perspectival features of most individual reproductive projects that distinguish them from any other actual or possible reproductive projects as candidates for ethical priority. Thus, if reproductive partiality is ethically reasonable, this must be because it has an impartialist explanation, perhaps along the lines offered by sophisticated consequentialism.

One response to the arbitrariness argument is to claim that it begs the question.²⁰ Consider the case of egoism. The egoist is accused of making an arbitrary distinction between his own interests and the interests of others as objects of ethical priority. There is no non-perspectival way for the egoist to distinguish his own projects from those of others as objects of particular concern. To this extent, the egoist's preferences are arbitrary. Yet in what sense are they unreasonably so? It is not as if the egoist denies the same privilege to anyone else. At a higher-order of reflection the egoist makes no unreasonable exception for himself, as a separate individual with projects and plans of his own. The egoist can agree that everyone else is in the same position and is therefore reasonably permitted to prioritize their own interests over those of others. Of course, it cannot be assumed that the interests of different individuals can be impartially integrated without remainder. Yet the ethical priority of impartial integration is supposed to be the conclusion of the arbitrariness argument, not a premise of it. Analogous considerations apply to groups of individuals relative to which a set of partial values can be defined. Such groups include couples, families, and larger social groups with a sense of their own, or common, good. In each case, there is no conclusive case for ethical impartiality based on arbitrariness alone. In any case, the arbitrariness argument is arguably too strong. Two complaints are relevant here. First, to the extent that a human perspective is a partial perspective, the arbitrariness argument entails that prioritizing human benefit is an ethically unreasonable prejudice.²¹ This claim is likely to remain controversial not only in the general human case, but also in cases where ethical priorities are made on a large scale within the human domain, such as the priorities made by national governments in favour of their citizens. Given the ethical significance of considerations involving "special ties" and the like, the pluralism that follows from the apparent irreducibility of partial values to impartial values could reasonably be regarded as applying not only at the level of individuals, but also at the level of groups. Second, we have already seen that no form of ethical self-understanding is possible in the absence of some perspectival characterisation of an individual's options. Consider the case of modality. On some accounts of the semantics of natural language, "actually" is an indexical term, fixing the reference of thoughts and sentences to the actual as opposed to merely possible worlds.²² Given this account of modals, any ethical outlook denying the explanatory role of perspectival specifications of favoured options would entail that it is ethically unreasonable to care more about actual as opposed to merely possible people. This is not the potentially reasonable claim that we should be equally concerned with the interests of those currently non-existing people who, depending on what we do now, will come to exist in the future. On a natural interpretation, these are all actual people. The claim is rather that it is unreasonable to care more about people who either do exist or will exist in the future than about people who could possibly, but who will never actually, come to exist (the inhabitants of the merely possible worlds discussed by the modal logicians). This is not a claim that most defenders of ethical impartiality would want to be committed to merely in virtue of claiming that fundamental partiality is ethically unreasonable. Now perhaps "actually" is not best interpreted as an indexical term.

Yet some terms will come out as indexical on any plausible account of their use. At least some of these terms are essential to self-identification and rational agency. The basic question is therefore which indexical characterisations of options are ethically relevant, and how. This question remains a source of ongoing controversy.

11.6 A Pessimistic Conclusion?

Given the partial aspects of any perspective and the vagaries of talk about (some merely possible) ethically relevant beings of the future, one might reasonably question both the determinacy and coherence of the idea of a maximally impartial ethical perspective. In this paper, I have bracketed this question by talking of “more”, “less”, and “maximally” impartial perspectives. When interpreted in this way, the arguments considered so far suggest that the partial and impartial values at work in human reproduction are sometimes in irreducible conflict. This is in some ways a pessimistic conclusion. What follows from this conclusion for the ethics of human reproduction? One thing that does not follow is that prospective parents are ethically unconstrained to pursue their reproductive projects partially. Impartial beneficence is ethically relevant wherever it applies, including cases of identity-affecting reproductive choice. Nor does it follow that wherever partial and impartial values conflict it is always reasonable to prioritize a partial perspective. Thus, it may be reasonable for public policy makers to prioritize an impartial perspective that takes account of the interests of people in general across several generations over the partial perspectives of a small minority practicing bias in favour of themselves and their own offspring. Likewise, it may be reasonable for individual couples to prioritize their own reproductive potential over impartial beneficence, at least where doing so is consistent with avoidance of serious impartial harm or the creation of children with little or no prospect of a decent human life. That doing so may be impartially sub-optimal does not show that it is thereby unreasonable. This case for ethical pluralism suggests that it is a mistake to claim that all ethical thought should make the same fundamental prioritizations between evaluative perspectives in the course of weighing up different reproductive options. Thus, whereas moral thinking on a large scale about public policy and future generations might reasonably prioritize a perspective of impartial beneficence, moral thinking on a small-scale about familial reproductive projects might reasonably prioritize a comparatively partial perspective. Such partial perspectives are arguably as ethically fundamental as any perspective of impartial beneficence.

What does follow from the pessimistic conclusion is the possibility that some conflicts between partial and impartial values are impermeable to reasonable resolution on ethical terms. Such conflicts, it might be said, are instances of the well-known phenomenon of value incommensurability.²³ While potentially misleading, there is something to be said in favour of this claim. On the one hand, we are not dealing with conflicts between values of radically different kinds (music versus skiing, for example). In the present case all the values in question are reproductive

values. Nor are we dealing with a single scale on which we are unable to determinately fit the values concerned (Mozart versus Beethoven as good composers, for example). We are dealing with at least two, irreducibly distinct, scales. Nor is it simply that of two values we cannot say that either one is greater than the other or that they are equally great. Each value is both lesser and greater than the other, but from different perspectives. Nevertheless, the conflict between partial and impartial value does arguably share one feature of paradigmatic cases of value incommensurability. Even if there is a significant range of questions on which no determinate answer exists as to which set of values wins out, it is reasonable to expect there to be a non-empty set of questions outside this range on which either some partial or impartial perspective wins out, all things considered. Thus, there could be some range below which impartial sub-optimality is ethically monstrous. No ethically serious parents would want to expose either their potential offspring or anyone else to a life of unbearable suffering. (This claim is neutral with respect to the question whether or not the wrongness of exposing potential offspring to a life of unbearable suffering can be accounted for in person involving terms.²⁴) Likewise, there could be some range above which impartial sub-optimality is ethically innocent. Thus, it is arguably unreasonable to demand that agents always maximize impartial value, no matter the cost to themselves.

The existence of irreducible ethical conflict presents a troubling challenge for the integrationist project of constructing a unified moral theory. A moral theory with nothing to say about choices between conflicting evaluative perspectives is about as good as a moral theory with nothing to say. As we have seen, however, matters are not that desperate. The existence of value incommensurability is consistent with the existence of reasonably determinate answers in a significant range of cases. It does therefore not entail that any reproductive choice is as reasonable as any other.

It is natural to expect that most reasonable reproductive choices will be based on a combination of partial and impartial values. Moreover, in some fortunate cases both partial and impartial values will favour the same conclusions. In such cases, it is reasonable to think that the combined force of partial and impartial values will be greater than the force of either partial or impartial values on their own. While nothing I have argued in this paper is inconsistent with this claim, the arguments of the previous two sections inject a heavy dose of pessimism with respect to the idea that the combined force of partial and impartial values is one that can be calculated as a matter of algorithm on a single and uniquely privileged evaluative scale. Indeed, if the arguments given above are cogent, the very idea of such a scale is inherently problematic. I have argued above that it is at least as difficult to compare partial and impartial values in cases where they conflict. Nevertheless, if some prospective parents have no time for impartial values at all, it could still be reasonable for impartial public institutions to coerce them, provided their reproductive choices are such as to create a significant risk of great suffering to actual or potential others. Although this claim obviously constitutes a threat to the unconstrained reproductive autonomy of prospective parents, it coheres with the liberal view of reproduction as set out at the beginning of this paper.

Notes

1. Glover (2006).
2. Buchanan et al. (2000), McMahan (2001), Broome (2004), Mulgan (2006).
3. Mulgan (2006).
4. Lillehammer (2005).
5. Parfit (1984).
6. Wasserman (2005), p. 142.
7. McMahan (2001, 2005).
8. Mulgan (2006).
9. Gauthier (1984).
10. Brink (1989).
11. Williams (1973, 1982).
12. Railton (1984).
13. Williams (2006).
14. Broome (2004).
15. Broome (2004).
16. Parfit (1982), Broome (2004).
17. Parfit (1982), Woodward (1986).
18. Sidgwick (1907), p. 382.
19. Broome (2004).
20. Sidgwick (1907).
21. Williams (2006).
22. Lewis (1984).
23. Williams (1982), Broome (2004).
24. Roberts (1998).

References

- Brink, D. O. 1989. *Moral realism and the foundations of ethics*. Cambridge: Cambridge University Press.
- Broome, J. 2004. *Weighing lives*. Oxford: Oxford University Press.
- Buchanan, A., Brock, D. W., Daniels, N., Wikler, D. 2000. *From chance to choice*. Cambridge: Cambridge University Press.
- Gauthier, D. 1984. *Morals by agreement*. Oxford: Oxford University Press.
- Glover, J. 2006. *Choosing children*. Oxford: Oxford University Press.
- Lewis, D. 1984. *On the plurality of worlds*. Oxford: Blackwell.
- Lillehammer, H. 2005. Benefit, disability, and the non-identity problem. In *Philosophical reflections on medical ethics*, ed. N. Athanassoulis, 24–68. New York: Palgrave Macmillan.
- McMahan, J. 2001. Wrongful life: Paradoxes in the morality of causing people to exist. In *Bioethics*, ed. J. Harris, 445–75. Oxford: Oxford University Press.
- McMahan, J. 2005. Causing disabled people to exist and causing people to be disabled. *Ethics* 116: 77–99.
- Mulgan, T. 2006. *Future people*. Oxford: Oxford University Press.
- Parfit, D. 1982. Future generations: Further problems. *Philosophy & Public Affairs* 11: 113–72.
- Parfit, D. 1984. *Reasons and persons*. Oxford: Oxford University Press.
- Railton, P. 1984. Alienation, consequentialism, and the demands of morality. *Philosophy & Public Affairs* 13: 134–71.
- Roberts, M. 1998. *Child versus childmaker: Future persons and present duties in ethics and the law*. Lanham: Rowman & Littlefield.
- Sidgwick, H. 1907. *The methods of ethics. Seventh Edition*. London: Macmillan.

- Wasserman, D. 2005. The nonidentity problem, disability, and the role morality of prospective parents. *Ethics* 116: 132–52.
- Williams, B. 1973. A critique of utilitarianism. In *Utilitarianism: For and against*, ed. J. J. C. Smart and B. Williams, 77–150. Cambridge: Cambridge University Press.
- Williams, B. 1982. *Moral luck*. Cambridge: Cambridge University Press.
- Williams, B. 2006. The human prejudice. In *Philosophy as a humanistic discipline*, ed. A. W. Moore, 135–52. Princeton: Princeton University Press.
- Woodward, J. 1986. The non-identity problem. *Ethics* 96: 804–31.

Chapter 12

Two Varieties of “Better-For” Judgements

Peter Herissone-Kelly

Abstract This paper argues against Julian Savulescu’s principle of procreative beneficence. It maintains that prospective parents have no obligation at all to choose the child, out of a range of possible children, who is likely to lead the best life. This is because a standpoint that the author labels “the internal perspective” is a perfectly appropriate one for parents to adopt when thinking about their own future children. It is only policy makers who are obliged to take up an opposing standpoint—“the external perspective”—and to be motivated by the sorts of “better for” judgements that that perspective delivers.

Keywords Principle of procreative beneficence · Savulescu · Parenthood · Prospective parents.

12.1 The Principle of Procreative Beneficence

In his much-discussed paper “Procreative Beneficence: Why We Should Select the Best Children,”¹ Julian Savulescu gives admirably clear expression to, and argues forcefully for, a principle that commands the assent of many bioethicists. He states that principle as follows:

Couples (or single reproducers) should select the child, of the possible children they could have, who is expected to have the best life, or at least as good a life as the others, based on the relevant, available information.²

Savulescu labels this the Principle of Procreative Beneficence. For ease of reference, I will in this chapter refer to it as PPB.

Savulescu happily admits that PPB asserts only a *prima facie* obligation, or one that is capable of being overridden by other moral reasons. He also maintains that, despite their being *prima facie* morally obliged to select a child who is likely to

P. Herissone-Kelly (✉)

Philosophy Section, International School for Communities, Rights, and Inclusion, University of Central Lancashire, UK; University of Bolton, UK
e-mail: pnherrissone-kelly@uclan.ac.uk

have the best life, prospective parents ought not to be compelled to do so, either legally or in any other way. If a prospective parent neglects to choose the best child she can, then she, other things being equal, makes a moral mistake. But it is one that she ought to be allowed to make. Others—medical staff for instance—are justified in trying rationally to persuade her not to make it, but they have no moral licence to coerce her. The demands of a principle of reproductive autonomy entail that parents must be given room to err in their decisions, so long as their errors will not harm their children. And while it is typically wrong to flout the requirements of PPB, Savulescu thinks that in doing so, a parent does not harm her child.

It is my aim in this chapter to argue, against Savulescu and those who agree with him, that prospective parents, precisely by dint of their filling that role, are in fact not bound by PPB. That is, I will urge that no prospective parent has even a *prima facie* obligation to select the child, from the range of her available possible offspring, who is likely to lead the best life, or at least as good a life as the others. Ultimately, I want to claim that this is because the best life would not be better *for* any particular child. But that does not mean that I can be numbered amongst those who straightforwardly oppose Derek Parfit's No-Difference View.³ On the contrary, I think that there *can* be moral reasons to bring about better lives that are not better for any particular person or set of persons. It may even be the case that there *usually* are such reasons. Crucially, however, I think that those reasons do not apply to prospective parents in respect of their potential offspring. There is something about someone's being a prospective parent that, we might say, exempts her from any requirement to create the best state of affairs she can through her reproductive decisions.

My argument for this conclusion will involve the drawing of a distinction between two perspectives that we can take up when considering the lives of possible future people. These I call "the internal perspective" and "the external perspective." From the latter perspective, we may see ourselves to have reasons to bring about states of affairs that are, as it were, "better *period*," rather than better for any particular person or set of persons (though I will later suggest that a judgement that some state of affairs is better *period* can be characterised as a disguised "better-for" judgement of a distinctive kind). From the former perspective, by contrast, no such reasons are visible. And, I will endeavour to show, the former standpoint is one that it is morally appropriate and fitting for a prospective parent to adopt and to take seriously.

Before I present that argument, however, I want briefly to make mention of another type of objection to PPB that finds expression in the literature. It is an objection of considerable importance, and one with which I have a good deal of sympathy. Its main thrust is that it is just impossible, at least in the vast majority of cases, to determine which of a range of possible lives is the best. Clearly, if we cannot determine that, then we have no prospect of selecting the possible child who is likely to lead the best life, just because we will not know which possible child that is. And on the overwhelmingly plausible assumption that we

cannot be obliged to do what it is impossible for us to do, we cannot be bound by PPB.

Why might it be thought impossible to achieve a defensible judgement concerning which of a range of possible lives will be likely to be best? It is certainly not the case that we are completely unable to tell, for instance through pre-implantation genetic diagnosis, what the person who would develop from a particular embryo will be like. We are by no means wholly in the dark here, and it is to be expected that further developments in genetic technology will only bring increasing light. We can, for example, reliably identify in an embryo certain genetic traits that would lead to disability in the person who, under the right conditions, would develop from it. The point of the objection, rather, is that even when we have this information, we are unable (for a substantial range of cases) to tell whether an embryo that possesses those traits is likely to develop into a person with a worse life than a person who would develop from an embryo free of those traits. This is in part because, as Michael Parker has argued in a recent exchange with Savulescu, any realistic candidate for a notion of the good life will inevitably display a not inconsiderable degree of complexity.⁴ Just knowing, say, that a potential person A would be able-bodied, while another potential person B would not, gives us no reason to suppose that A’s life is likely to be better than B’s. To think otherwise is to adhere to an unacceptably simple and coarse-grained conception of what it is that makes a life go well.

Parker’s thought seems to allow that PPB would be a perfectly good ethical principle in a possible world in which our epistemic circumstances were very different. That is, prospective parents would, or at least could, be obliged to bring about the best lives, if and only if they were to possess perfect knowledge of what constitutes the good life (which of course would entail either that there were some definite universal truth of the matter here, or that it were possible to discern what a future person’s own conception of the good life would be), *and* were unfailingly able to tell when one possible life would be better than another (which would entail that pre-implantation genetic diagnosis or some alternative procedure were able to deliver much more detailed information than it in fact can, or perhaps than it is in the actual world likely ever to do). But as things stand such knowledge is unattainable. Consequently, prospective parents cannot be bound by PPB.

As I have hinted, I am strongly inclined to accept that the mere fact that A would be able-bodied and B disabled provides us with insufficient grounds for judging that A’s life would be likely to be better than B’s. And the same would be true if the claim were that A would be more intelligent, or more athletic, or more attractive than B. However, I want to make, so to speak, a more destructive criticism of PPB than this, or one that strikes more directly at its foundations. I want to maintain that even if our epistemic situation were monumentally altered, and we were able invariably and accurately to determine when one possible future life would be better than another, prospective parents would still have no obligation to choose a child with the best life. That being the case, I will sidestep questions about the feasibility of ranking possible lives on a scale from best to worst, and assume, for the sake of argument and to some degree against my intuitions, that this is something we can do.

12.2 PPB and the Non-Identity Problem

Some people, upon encountering Parfit's non-identity problem, may understandably be tempted to dismiss PPB in its light. That is, they might come to believe, against Parfit, that the possibility of producing an outcome *O* that is better without being better *for* anybody—an outcome that is, as I put it earlier, better *period*—can provide us with no reason to seek to bring about *O*. So, for example, suppose that we have before us two embryos, A and B, and that we have to decide which to transfer to a womb. Let us stipulate that it is abundantly clear that A, if it were transferred and were successfully to implant, would develop into the person with the better life (let us, then, call A the better-life embryo and B the worse-life embryo). Since the better life here would not be better *for* either the person to develop from A or the person to develop from B, we might feel that there can be no obligation to bring it about. To be more exact, we may hold one of two views. The stronger of these views would maintain that the very notion of a betterness that is not a betterness *for* anybody is simply incoherent. The weaker view would be that, although betterness *period* is a perfectly coherent notion, it is unable to ground any moral obligations of the sort expressed by PPB.

I think that both the stronger and the weaker views are mistaken. That is, I want to claim that betterness *period* can be shown to be entirely coherent, and that it can in addition ground moral obligations. But before I go on to argue for this claim, I think it will be helpful to spell out in a little more detail just why selecting the better-life embryo will not be better for anybody, and selecting the worse-life embryo will not be worse for anybody, and why these facts might tempt us to reject PPB.

Where A is a better-life embryo and B a worse-life embryo, it may seem sensible to ask *for whom* the better life would be better, and *for whom* the worse life would be worse. We may feel inclined, at least before we have thought about it in any depth, to say that the better life would be better for the person—call him Fred—who would develop from A. However, this seems to imply something that is straightforwardly false, namely, that Fred could live either of the two lives available, but that happily he would as a matter of fact get to live the better. But that is not the case. If B were transferred, the worse life would be lived, but it would not be lived by Fred. So how would the better life be better *for him*?

Correlatively, the worse life would not be worse for the person—call him Barney—who would develop from B. To say that it would be worse for him is to say that he somehow has the possibility of living either of the two lives available, but that as a matter of fact he would be saddled with the worse. And again, that misrepresents the situation. Had A been transferred, the better of the two lives would have been lived, but it would not have been lived by Barney. So it would not be worse for Barney if B were to be transferred.

Of course, there is a sense in which it would be wrong to conclude from this that the better life here is necessarily not better for *anyone*. It may be better for a parent to have the child who would develop from the better-life embryo (though one can also envisage circumstances in which it might be better for the parent to have the child who would develop from the worse life embryo). This, however, is not

the sense of “better for” in which I am interested here; I am concerned only to deal with the question of whether a better-life embryo, if transferred, will develop into someone whose life is somehow better *for him or her*. And it seems that the answer is plainly “No.”

Now, the worry here is that principles like PPB rely on the idea that one life can be better *period* than another. And, the thought goes, a life that is better *period*, since it is not better for either the selected or the deselected child, displays at best a nebulous, free-floating betterness, which is so abstract, so unrelated to any concrete person’s interests as to be morally inert, even if it is not of dubious coherence. Suppose, for example, that in the mid-1960s a very sophisticated form of pre-implantation genetic diagnosis had been available. Had that been the case, it seems uncontroversially true that my parents could have selected an embryo that would have developed into somebody with a life superior in many ways to mine. Assuming for a moment that these are ingredients of a good life, that person might have been more athletic than me, or more organised, or of an all-round sunnier disposition. But since that life would have been better neither for me, nor for my athletic, efficient, jovial counterpart, what does it matter? What is it to me that someone could undoubtedly have lived a better life than mine, especially since my life is at the very least perfectly acceptable? And how could anyone be obliged to bring about a betterness that is not better *for* anyone, even assuming the coherence of such a notion?

That is the position that the existence of the non-identity problem might tempt us to adopt. It holds that PPB and principles relevantly like it must be illegitimate, since they rely on an at worst incoherent and at best morally inert conception of best or better lives, a notion of lives that are not best or better for anyone, but just best or better *period*. My own position is that if the non-identity problem tempts us in this way, then we ought to resist the temptation. In so far as I believe this, I agree with Parfit. I think, that is, that the notion of lives that are better *period* is entirely coherent, and that the consideration that one life or set of lives would be better than another life or set of lives can provide us with a moral reason to select the former. But note that I hold only that it *can* provide us with such a reason; importantly, I do not maintain that it invariably *does*.

12.3 Internal and External Perspective “Better-For” Judgements

The existence of the non-identity problem ought not to lead us to deny the coherence or moral force of better-*period* judgements, at least in part because those judgements are, at bottom, disguised better-for judgements of a distinctive kind. Admittedly, to say that Fred’s life would be better than Barney’s is not to say that it would be better either for Fred or for Barney if embryo A were to be transferred. Nonetheless, it *is* to say that Fred’s life would be better for Fred than Barney’s life would be for Barney. To make this last claim is, plausibly, just to assert that Fred’s interests would be better met by Fred’s life than Barney’s interests would be met by Barney’s life. This

is the sort of claim that could be true, whereas I take it that the former could not. It is also the sort of claim that we could easily envisage giving us a moral reason to select embryo A over embryo B.

It has to be conceded, however, that the latter claim represents a very different variety of better-for judgement to the former. The way in which it differs is best explained, I think, by characterising the former as an internal perspective better-for judgement, and the latter as an external perspective better-for judgement. The internal perspective/external perspective distinction can be explained as follows.

In thinking about the life of a possible person A, we adopt the internal perspective when we (i) “imaginatively inhabit” that life, imagining what it would be like to live it, and (ii) make the sort of “better” and “worse” judgements that we would make about A’s life if we were A (hence the *internal* perspective). All the “better for” or “worse for” judgements made from the internal perspective will be judgements about what will be better for or worse for A. When we adopt the internal perspective, we remain unmoved by the talk of better and worse lives that features in principles like PPB, as such talk belongs only to the external perspective.

In thinking about the life of a possible person A from the external perspective, by contrast, we (i) imaginatively inhabit that life, attempting to gain a sense of what it would be like to live A’s life; (ii) imaginatively inhabit the life of another possible person B, attempting to gain a sense of what it would be like to live B’s life; and (iii) *draw back from the perspective of both A and B* in order to make a judgement about which life is better. This, as we have seen, involves judging whether A’s interests would be better met by A’s life than B’s interests would be met by B’s life. The perspective from which the judgement is made here is “external” because it stands outside either life. It is, so to speak, transpersonal.

The objection to PPB that we might frame upon understanding the non-identity problem fails, because it assumes that the only sort of better-for judgements there can be are internal perspective better-for judgements. But that is not the case. There are perfectly coherent external perspective better-for judgements as well, and there is no a priori reason to suppose that their being true cannot give rise to moral obligations. Of course, in itself the bare fact that these two different varieties of better-for judgements are possible tells us nothing about which we should take seriously in trying to determine what we are morally obliged to do. If I take seriously the external perspective judgement that A is the better life embryo, and so that A’s life would be better for A than B’s life would for B, then it may well seem to be the case that I am morally obliged to transfer A in preference to B. But if, on the other hand, I take seriously only internal perspective better-for judgements, I will seem to have no obligation to choose to transfer either A over B, or B over A.

The notion of “taking seriously” is important here, and is being used as something of a term of art. When I take up the internal perspective on the sort of life that Barney would lead, I will not be entirely blind to the truth of the external perspective judgement that his life would be worse for him than Fred’s life would for Fred. Similarly, when I adopt the external perspective in imaginatively inhabiting and comparing the lives that Fred and Barney would be likely to lead, I will be no less aware that the fact of Fred’s life being the better of the two would be of no

consequence to Barney. In other words, I will fully understand that it would not be better for Barney (or indeed for Fred) if embryo A were transferred. In thinking about the lives of possible future people, both perspectives are always accessible to me, and I am able simultaneously to appreciate the truth of internal and external perspective better-for judgements. But if a better-for judgement is to help me to decide what I ought to do, I will need first to know which variety I ought to take seriously. That is, I will need to know which should take primacy in guiding my decision. I need to know, just because different considerations show up as reasons from each perspective. That A is the better-life embryo is, from the external perspective, a reason to select it in preference to B. But from the internal perspective, it is no such reason, since A’s being selected will produce a life that is better neither for Fred nor for Barney.

The question of what we are obliged to do when presented with the opportunity to bring one (or one set) out of a range of possible persons into existence, then, boils down to the question of which perspective ought to be taken seriously when selecting. Savulescu clearly opts for the external, as PPB is rooted in that perspective. I think, however, that there is no one, invariant answer to this question. That is because, crucially, the perspective that it is appropriate to take seriously will fluctuate with the identity of the decision maker, and with the context in which the decision is being made.

For example, it seems to me that it is most probably *prima facie* obligatory for political decision makers, who choose social policies that determine both who will come into existence in the future, and the quality of those future persons’ lives, to take seriously the external perspective, and so to act on principles rooted in that perspective. Such people should allow external perspective better-for judgements to influence their decisions. To illustrate what I mean here, there is no better example than that provided by Parfit in Chapter 16 of *Reasons and Persons*.

Parfit’s example runs as follows.⁵ Suppose that a choice has to be made between the adoption of a policy of conservation of natural resources on the one hand, and a policy of depletion of those resources on the other. Adoption of the depletion policy, Parfit stipulates, would improve the quality of our lives for the next few centuries, but after that time would lead to a significantly lower (though not unbearable) quality of life for all subsequent generations. Adoption of the conservation policy, on the other hand, would result in our quality of life remaining exactly where it is, and would keep it constant for all successive generations.

Importantly for the point that the example aims to demonstrate, Parfit thinks it safe to assume that adoption of the depletion policy in preference to the conservation policy will alter future events to such an extent that people will meet different partners, conceive at different times, and so on, than they would have had the choice gone the other way. The upshot of this is that, in a couple of centuries’ time, a wholly different set of people will exist than would have existed had we opted for conservation.

Now, it should be clear that, if the depletion policy is adopted, there will not be anyone, once the general quality of life starts to deteriorate, who will live a worse life than they would have lived had the conservation policy been adopted. That is

because, had the conservation policy been adopted, those people would not have existed to lead a better life. So, it will not be worse *for* anybody (any set of people) if the depletion policy is adopted. And, similarly, it would not be better *for* anybody if the conservation policy were adopted, because there would be no set of people who would be made better off than they otherwise would have been. (Of course, we might want to say that those born after adoption of the conservation policy *would* be better off, because they would exist. But, it should be noted that, if we accept this point, the same could be said of those born as an indirect result of the depletion policy's adoption). Despite the fact that there is no set of people *for whom* adoption of the conservation policy would be better, it still seems, as Parfit notes, intuitively clear that the conservation policy is the one that ought to be chosen. That is, the decision makers should opt for that policy, the implementation of which will result in the existence of a set of people whose lives are best able to meet their interests.

In his defence of PPB, Savulescu uses an example of the same general type as Parfit's.⁶ But the judgements taken seriously in that example are external perspective judgements. And I want to argue that, although the external perspective may well be the proper one for political decision-makers to take seriously, it is appropriate (that is, it is at least permissible) for prospective parents to take seriously the internal perspective when thinking about the lives of their potential children.

12.4 Prospective Parenthood and the Internal Perspective

Imagine that a prospective parent is trying to decide which of, say, two embryos should be selected for transferral. In order to do this, she must "imaginatively inhabit" both possible future lives. Now, there is here, I think, a point that has been astonishingly absent from the literature. That is, when the prospective parent does this, she is imaginatively inhabiting the lives of her potential offspring. It is, therefore, appropriate that she should relate to those potential children's lives in a way that is proper for a parent. In saying this, I am in agreement with Simo Vehmas' claim that attempting to produce children at all "puts the potential parents *morally* in the position of parenthood."⁷ In selecting at all, our decision-maker is selecting as a prospective *parent*, and *what* she is selecting will, if all goes well, develop into her *child*.

Now, the prospective parent's relating to her possible children's lives in a way that is proper for a parent will, I suggest, involve her in an extremely close identification with the possible subject of each life she imagines. That identification will be so close that it will include, and count as significant, the sorts of assessment that each offspring would be likely to make of his or her own life. The nature of this identification is, I think, well captured in Aristotle's claim that parents properly relate to their children as "other selves."⁸ This is not to say simply that they recognise their children to be selves, just as they themselves are. Presumably, parents recognise this to be true of all those creatures who happen to be selves. Rather, it is to say that they relate to their children in some sense as if they *are* them, or as if their children's interests are their interests.⁹ In short, the fitting standpoint

for a prospective parent to focus upon and count as significant will be one that includes just those features from which the external perspective prescind: it will be the internal perspective. The better-for judgements accessible from this perspective will be concerned only with what would be better for the person to develop from each embryo. The transpersonal better-for judgements of the external perspective will not be afforded any significance.

Two objections are possible at this point. The first, which can I think be dealt with reasonably swiftly, can be stated as follows. It could be argued that it is simply never legitimate to give any weight to the internal perspective when thinking of merely potential lives, as that perspective involves the imaginative inhabitation of a life of someone who does not exist, and so who has no interests. But my response to this is that the making of transpersonal, external perspective judgements about better and worse lives—which is something that we must engage in if we are to be guided by PPB-like principles—also involves the imaginative inhabitation of the lives of people who (as yet) do not exist or have interests. Secondly, and much more importantly, it may be thought that all I have done so far in this section is to describe what a prospective parent will as a matter of fact do, or be likely to do, in thinking about the lives of her potential children. The thought here may be that I have only said to what perspective it would be natural, or understandable for a prospective parent to give primacy in her deliberations. And that prospective parents may naturally, as a matter of fact gravitate towards an internal perspective that blinds them to reasons that can be seen to hold from the external perspective is no proof at all that that is what they *should* do.

My response to this latter objection is crucial to my entire argument. My claim is emphatically not that a prospective parent simply *will* take the internal perspective seriously, or that she is *likely* to, or that it is *understandable* that she might. Rather, my point is that it is appropriate and fitting for her to do so. It is quite proper for a prospective parent *qua* prospective parent to take the internal perspective seriously. The role of parent is not merely biological; indeed, in some cases it is not biological at all. It also has a significant normative element to it. There is a correct, morally admirable way for a parent to relate to her child, a way of relating that it may well be a moral fault for a parent to lack, and this is a way of relating that is, as it were, built into the very concept of the good parent. A good parent is, *inter alia*, one who is able to see things from her child’s perspective, who makes the child’s interests her own, who gives weight to the child’s view of his own life. She is one who relates to her child as to another self, in the Aristotelian sense.

Of course, at the time that the prospective parent is in a position to decide which embryo ought to be transferred, there is no child, and the prospective parent is not yet a parent (at least not to either of the children who might develop from the available embryos). Even so, to make a decision at all, the prospective parent must, as we have seen, imaginatively inhabit the lives of each possible future person. And in this very process of imaginative inhabitation and decision making, she is not, and cannot be obliged to be, a detached, neutral person. It is true that in making her selection she is a merely *prospective* parent. But she is also a prospective *parent*, and this is a role that she fills essentially—she would not be making that decision at all were

it not for the fact that she will be the parent of whatever child results from it. It is not only understandable, but perfectly proper, perfectly fitting, that she should take seriously the internal perspective in thinking about the lives of her possible future children. And from that perspective, there is just no reason to select a better life embryo in preference to a worse life embryo.

There is a point that should be noted here. I have urged that it is proper, or fitting, or appropriate for a prospective parent to take internal perspective better-for judgements seriously. And I have claimed that this is so on account of the normative character of parenthood, both prospective and current. Now, to hold that it is appropriate for a prospective parent to give primacy to the internal perspective is not to endorse the strong claim that she is *obliged not* to give primacy to the external. Instead, it is (or at least need be) only to endorse the weaker claim that she is *not obliged* to do so. It is therefore only to maintain that she is not morally bound by PPB, and not that she is morally bound to ignore it.

Now, Savulescu might argue against the position I have set out here in the following way.¹⁰ A world in which I select a child without asthma, for example, will be a better world than one in which I select a child with asthma. This, surely, is a reason for selecting a non-asthmatic child, especially as there is no corresponding reason at all for selecting an asthmatic child. The problem with this objection, as I see it, is that it already assumes the legitimacy of privileging the external perspective over the internal, since it is only from the former perspective that I can be judged to have a reason to bring about a better world that is not better for any of its inhabitants. My point is that, before I can determine how I ought to behave with regard to possible future people—and so whether I have a responsibility to bring about the better of two worlds—I need to answer the prior question of which perspective I ought to take seriously. And I have suggested that a prospective parent, *qua* prospective parent, is not obliged to give primacy to the external perspective. Bringing about the best world she can through her reproductive actions is not an obligation that attaches to a prospective parent. A parent may have some sort of duty to make the world as good a place as she can for any child that results from her reproductive actions, but that is an entirely different matter.

12.5 Principled Procreation

I argued in Section 12.3 that there can be an obligation to act in such a way that better lives are brought about, even when those lives are not better *for* anyone. Just such an obligation attaches to the decision makers in Parfit's example, who are faced with a choice between a policy of conservation and a policy of depletion. The chief reason for believing this, it seems, is just that it is intuitively obvious that the conservation policy ought to be adopted, even though there is nobody whose life will be improved by its adoption. As Parfit says:

The great lowering of the quality of life must provide *some* moral reason not to choose Depletion. This is believed by most of those who consider cases of this kind.¹¹

I have denied, however, that prospective parents have any obligations of this kind: they are not obliged to produce the child, from amongst their possible children, who will be likely to lead the best life. There is a problem for my claim, however, given that I acknowledge the intuitive force of Parfit’s conservation/depletion example. That is, Parfit deploys additional examples, which are on the face of it equally intuitively compelling, in order to show that we can be obliged to bring about better lives that are not better for anyone. But these examples concern the reproductive decisions of prospective parents.

For instance, Parfit imagines a woman who is told that she has a temporary medical condition, such that if she were to conceive within the next three months, she would have a child with a disability, though with a life that would be worth living. If she conceives after the three months, she will have an able-bodied child with a better life. Parfit takes it that this prospective parent ought to wait before conceiving.¹² He also considers the case of a 14-year old girl who plans to have a child. Because this would make her an exceedingly young mother, it is plausible to suppose that her child would have a bad start in life. If instead she were to wait and to conceive in several years’ time, the resulting child—who would of course not be identical with the child she could have now—would be likely to enjoy a considerably better quality of life. Again, Parfit maintains that the girl ought to wait.¹³

Faced with these examples, it might seem that I have only three options. First, I can deny the intuitive force of those of Parfit’s examples that feature prospective parents, while retaining my belief that the conservation/depletion example accords with intuition. Second, I can deny that either type of example demonstrates what Parfit takes it to. I find neither of these two alternatives especially appealing, though the first is rather more attractive than the second. My third option, of course, would be to abandon my central claim that prospective parents are not bound by PPB. Fortunately, I think there is no need for me to take this drastic step. In order to explain why not, I first need to prepare the ground a little, by introducing a principle that I take to be defensible in a way that PPB is not.

I have argued that it is wholly appropriate for a prospective parent to give primacy to the internal perspective in thinking about the lives of her possible future children. And it might seem that, from the internal perspective, it is a matter of complete indifference whether a better-life embryo or a worse-life embryo is selected. Indeed, I do think that the *mere* fact that A is a better-life embryo does not oblige a prospective parent to select it in preference to a worse-life embryo B. However, that is conspicuously not to say that there cannot be other reasons for rejecting B in favour of A.

Clearly, the fact that the internal perspective is an appropriate standpoint for prospective parents to take seriously does not preclude the possibility of other standpoints and principles that are rooted in the normative character of prospective parenthood. For instance, it could plausibly be maintained that a good parent is obliged to hold a principle of the form “I will not allow any child of mine to have a quality of life below *L*,” where *L* is a level of acceptable outlook. This “principle of acceptable outlook” (let us call it PAO) will typically find expression in attitudes such as “I do not want any child of mine to suffer unacceptably,” attitudes that we

would deem it proper for parents and prospective parents to hold, and improper for them to lack.

The most straightforward application of PAO involves cases in which the life of a possible child would not be worth living. Such a life would clearly fall below any plausible candidate for *L*. However, PAO could also apply to other cases: cases in which the possible child's life would be worth living, but would contain what the prospective parent considers a greater load of suffering than she is prepared to allow a child of hers to endure. It may at first sight seem curious to maintain both that a child's life could be worth living, *and* that it could fall below a reasonable candidate for *L*, and so should not be brought into being. However, I take it that the following point from Jonathan Glover is a pertinent one:

Should not potential parents be under some moral pressure, at least, to consider whether it is right to bring into the world a child whose life is, by a small margin, just worth living?¹⁴

We might restate Glover's point in this way: would not someone accurately describable as a good parent set *L* at quite some distance above the point at which a life just becomes worth living? After all, if in raising my existing children I were to supply them with lives that were just worth living I would be setting myself up for legitimate censure. So, while PAO is not itself an internal perspective principle (though nor, as I shall argue presently, is it rooted in the external perspective), it can function as a constraint on which internal perspective judgements may legitimately be taken seriously. For example, from the internal perspective it might seem entirely unproblematic for me to bring into the world a child whose quality of life would be just above the level at which it would cease to be worth living. But if I am a good prospective parent I will, in addition to apprehending my possible future children's lives from the internal perspective, subscribe to PAO. And that principle will be likely to tell me that I really ought not to produce children with such a low, though bearable, quality of life.

Of course, precisely where *L* ought to be set is not a question that will admit of any very definite answer. It can only be for each parent to decide, though of course her decision will not in principle be immune from the criticism of others. And this, I think, is capable of accounting for the fact that we might judge (a) that both prospective parents in Parfit's examples ought to delay conception, while at the same time (b) no parent is obliged to select the best possible child she can. That is, there are any number of ways in which a parent might come to set *L*. She may do it with reference to the average quality of life in her community. She may do it by considering her own life history (if for instance she has a disability that she has not found to have a significantly negative impact upon her life). Or she may do it by considering the sorts of child she is able to have. In other words, she may, for example, set *L* higher than she otherwise would, if she knows that there is the prospect of her having a child with a very high quality of life.

This last claim may provoke the suspicion that PAO is capable of collapsing into PPB. Does not the setting of *L* here involve an appeal to an external perspective judgement? I do not think that it does. In this case, the decision to set *L* at a particular

point is a choice that is informed by the parent’s knowledge of the sort of child she could have. But it is not, as an external judgement is, informed by her comparing one possible child’s life with another, and reaching a conclusion about which will be better. It is simply the setting of a standard of quality of life, such that should a possible child’s life seem likely to fall below that standard, then that child will not be selected. In so far as this is the case, it does not differ from the setting of L by reference to an average quality of life, or the parent’s own life history, or what have you.

What is more, once L has been set, any judgements made in its light concerning the lives of possible future children are not external perspective judgements either. They do not involve the comparison of the life of one possible child with that of another. Rather, they involve the comparison of the life of any possible child with a level of acceptable outlook.

Suppose once again that a prospective parent is faced with the choice between implanting a better life embryo A or a worse life embryo B. And suppose too that B would be likely to develop into someone whose life would fall foul of PAO—whose quality of life would be below L . In such a case, if guided by PAO the prospective parent will reject the worse life embryo B in favour of the better life embryo A. But note that her reason for this choice is not that A is the better life embryo, but that B’s quality of life would be unacceptably low. So, even though the result would be the same as it would have been had she acted on PPB, she would in fact have acted on PAO.

Why does it matter which principle she acted upon if the results are the same? The point here is that the results will not be the same *in all cases*. For instance, where the expected quality of life of both embryos exceeds L , PAO will not oblige the prospective parent to select the better life embryo. So certain choices that under PPB would count as morally forbidden come out as morally permissible under PAO.

On the other hand, certain choices that PPB allows will be forbidden by PAO. Suppose that, for whatever reason, it is only possible for a couple or single reproducer to produce one embryo, and that that embryo will develop into a person with a quality of life below L . Whereas PPB would allow such an embryo to be implanted, PAO, pretty obviously, would not. This may, I think, make PAO a little more intuitively acceptable than PPB, though admittedly Savulescu might respond by saying that PPB has in such cases to be supplemented or constrained by a threshold principle such as PAO.

The important point to grasp here, and one that shows PAO to be strikingly different to PPB, is this. If a prospective parent has before her an array of embryos from which to choose, all of which would be likely to develop into people whose lives would either reach or exceed L , then other things being equal (if, for instance, she could cope equally well with being a parent to any of the potential children) it is a matter of complete moral indifference which she selects. This is the case even if the array displays an enormous range of prospective qualities of life, from perfectly acceptable to surpassingly excellent. In this situation, the toss of a coin might be a wholly fitting way of selecting an embryo for transfer.

12.6 Summary

I have argued that prospective parents, simply in virtue of the fact that they *are* prospective parents, are not bound by PPB. This is because PPB is rooted in what I have called the external perspective, while a proper and fitting perspective for prospective parents to take seriously is the internal. From this latter perspective, the mere fact that A is a better-life embryo supplies no reason to select it in preference to a worse-life embryo B.

That does not mean, however, that a prospective parent's choice about which of a range of possible children to select should be entirely unprincipled. Instead of being guided by a principle of procreative beneficence, a prospective parent should make her choice on the basis of a principle of acceptable outlook. Such a principle would, significantly, make it a matter of indifference whether a better or worse life embryo were selected whenever both embryos are likely to develop into people with a quality of life that will not fall below *L*.¹⁵

Notes

1. Savulescu (2001), p. 415.
2. Savulescu (2001), p. 415.
3. Parfit (1984), pp. 367ff.
4. Parker (2007). See also the response to Parker in Savulescu (2007).
5. Parfit (1984), pp. 361–64.
6. Savulescu (2001), pp. 417–18.
7. Vehmas (2001), pp. 433–40.
8. Aristotle (1984), 1161^b29.
9. An illuminating account of the Aristotelian notion of other selves can be found in Stern-Gillet (1995).
10. Savulescu put this point to me in person at the Joint BIG/ESHRE International Conference on the Ethics of Stem Cell Research and Moral Responsibility in ART, held in Ghent, Belgium in November 2007.
11. Parfit (1984), p. 363.
12. Parfit (1976), pp. 373–74.
13. Parfit (1984), pp. 357–61.
14. Glover (2006), pp. 54–55.
15. This chapter expands upon themes, arguments, and examples to be found in Herissone-Kelly (2006).

References

- Aristotle. 1984. *Nicomachean ethics*. In *The complete works of Aristotle, volume two*, ed. J. Barnes. Princeton, NJ: Princeton University Press.
- Glover, J. 2006. *Choosing children: The ethical dimensions of genetic intervention*. Oxford: Clarendon Press.
- Herissone-Kelly, P. 2006. Procreative beneficence and the prospective parent. *Journal of Medical Ethics* 32(3): 166–69.
- Parfit, D. 1976. Rights, interests and possible people. In *Moral problems in medicine*, ed. S. Gorovitz, 373–74. Englewood Cliffs, NJ: Prentice-Hall.

- Parfit, D. 1984. *Reasons and persons*. Oxford: Oxford University Press.
- Parker, M. 2007. The best possible child. *Journal of Medical Ethics* 33(5): 279–83.
- Savulescu, J. 2001. Procreative beneficence: Why we should select the best children. *Bioethics* 15(5/6): 413–26.
- Savulescu, J. 2007. In defence of procreative beneficence. *Journal of Medical Ethics* 33(5): 284–88.
- Stern-Gillet, S. 1995. *Aristotle’s philosophy of friendship*. Albany, NY: State University of New York Press.
- Vehmas, S. 2001. Assent and selective abortion: A response to Rhodes and Häyry. *Cambridge Quarterly of Healthcare Ethics* 10(4): 433–40.

Chapter 13

Harms to Future People and Procreative Intentions

David T. Wasserman

Abstract This paper argues that agents' intentions can play a critical role in the moral appraisal of policies that result in the creation of shorter or harder lives than the lives that would have resulted from other policies. When agents, whether parents or policymakers, intend to create people whom they believe can only lead shorter or harder lives than others they might have created, the people they create will lack a complaint against them—a complaint they might have had if the agents lacked those intentions.

Keywords Intentions · Partiality · Policy · Justification · Harm · Nonidentity problem.

13.1 Introduction

We can be criticized for two kinds of actions with respect to future people. We can cause harms to future generations, like floods, famines, and nuclear catastrophes. And we can create people who will have shorter or harder lives than those of other people we might have created.

My interest in this paper is in exploring the relationship between these two kinds of actions in a class of situations where they converge: where future people are harmed or impaired by the very actions or policies that are necessary for their worthwhile existence. Derek Parfit offers such “identity-affecting” cases to suggest that the underlying moral offense in both kinds of action is the same: the creation of a world which is, from an impersonal perspective, worse than it might have been—with less overall well-being, happiness, or utility.¹ The future people harmed by policies that are necessary for their worthwhile lives are not wronged, the argument goes; rather, they are the incidental beneficiaries of policies that can be only be criticized for making the world as a whole worse off. But if the moral offense is

D.T. Wasserman (✉)
University of Maryland, University Park, MD, USA
e-mail: dwasserm@umd.edu

impersonal in this class of cases, it may also be impersonal in standard “identity-preserving” cases, where harm is inflicted on people who could have existed without suffering it.

Several philosophers have sought to deny Parfit’s first claim, that the harm in identity-affecting cases is impersonal, so as to forestall any attempt to generalize the impersonal account to standard cases of harming. They argue that actions and policies that harm future people may violate their rights, or wrongfully harm them, even if they are necessary for their worthwhile lives.²

In this paper, I will argue that the wrong in policies that cause “necessary” harms to future people—harms they could not have existed without suffering—cannot be analyzed as *either* a standard rights violation or a lack of beneficence. The wrong in such policies is unlike a standard rights violation, since it is found only in the absence of a certain kind of intention—an intention to create an individual or kind of individual whom one knows can only live a truncated or disadvantaged life. But for much the same reason, the wrong cannot be seen as a simple failure to act in a beneficent manner, since there is no wrong, or only a distinct and lesser one, in policies that express such procreative intentions, even policies expected to yield consequences impersonally worse than the alternatives.

13.2 Must We Create the Best? Partiality Toward Future Imperfect People

Several approaches to the non-identity problem have retained or acquired currency in the 30 or so years since its introduction. Most seek to explain or accommodate the strong intuitions elicited by a handful of influential hypotheticals introduced by Parfit—e.g., *The Risky Policy* and *A 14-Year Old Girl*.³ These approaches differ in whether they seek to justify (1) the judgment that the agents act wrongly and (2) the judgment that they wrong those who owe their worthwhile lives to the very policies or actions that harm them. Some attempt to do both, by explaining the wrong in these cases in person-affecting terms; in terms of a harm or wrong done to the individuals who owe their existence to the agent’s conduct. These approaches identify a violation of rights or breach of duty in the intentional, knowing, or negligent creation of future people expected to face certain harms or fall below some threshold of well-being.⁴

Other approaches seek to defend only the first judgment—that the agents act wrongly—while denying that they wrong those who owe their worthwhile lives to the very policies and actions that harm them. These latter approaches share Parfit’s view that non-person-affecting principles are needed to explain such cases, but they typically seek to limit recourse to such principles to choices where the number of future people will be the same in each alternative.⁵

Still other approaches distinguish two kinds of identity-affecting cases. In “macro” cases like *The Risky Policy*, current policy decisions have far-reaching effects on the identity of the people who will exist in future generations and how

well they will live. In “micro” cases like *A 14-Year Old Girl*, a parent makes a reproductive choice that results in her having a child with an impairment instead of a child without one.⁶ One approach, developed by Melinda Roberts, claims that in standard macro cases, the non-identity problem does not arise because the wrong can be explained in person-affecting terms. The affected individuals *are* made worse off, since they *could* have existed without the harm: they had the same vanishingly small ex-ante odds of coming into being whether or not the harm-causing policy was selected.⁷ On this approach, the non-identity problem is raised largely by micro cases, in which the individual could not, at least as a practical matter, have come into existence without the harm. In such cases, the agent does not harm the individual she creates, and the judgment that she does is either mistaken or misplaced: there may be harm to third parties, but not to the individual created.

The approach I favor, and will develop here, offers a more qualified judgment in micro cases. It can be understood in terms of a distinction drawn by Hallvard Lillehammer in this volume: between “person-involving” concerns—about “benefits or harms to identifiable individuals,” and partial concerns—about benefits or harms to “a subset of humanity, for example, the agent’s family, to which the agent has a special attachment.” As Lillehammer argues:

[T]he potential conflict between partial and impartial concerns cannot be reduced to a conflict between person involving and non-person involving concerns. This fact is significant for the ethics of human reproduction. Thus, it might be tempting to think that the ethics of reproductive choice must be fundamentally impartial because reproductive choices are identity-affecting and the values at work in identity affecting choices are fundamentally non-person involving. This would be a sound argument if all non-person involving values were fundamentally impartial. But they are not.⁸

I will argue that in macro as well as micro cases, the agent can choose to create some people rather than others out of partial concerns. I will also contend—and here I may part company with Lillehammer and many others—that if an agent acting from such partial concerns creates people with shorter or more disadvantaged lives than others she might have created, she will not wrong those people, as she would have if their creation had resulted from purely selfish or other ulterior motives.

These claims are easier to make out in micro cases, where prospective parents may have strong preferences about the kind of child they want to bring into existence and into their families. I have argued elsewhere that if parents act on such a preference, they will not wrong the child they create, even if they believe that a child of that kind can only live a truncated or disadvantaged life.⁹ Some, but by no means all, parents will have such reasons in creating children whom they know or expect will lead especially difficult lives, e.g. parents who choose to have a child with whom they will have a genetic tie, when any such child is likely or certain to have, by virtue of that tie, genetically-based impairments.

It might, however, seem difficult to attribute such reasons to policymakers, whose decisions affect the population of distant generations. Why would they have any reason for preferring the unidentifiable individuals who would exist hundreds of years later under one policy over the different individuals who would exist under another? I will suggest that such reasons can be found in the ordinary partiality

that policymakers display towards members of their own political community or culture. Such partiality might well give them reasons to populate a future world with more of their own kind, even if doing so meant that the population of the future world would have shorter or less advantaged lives than if they had chosen impartially.

Such partiality may be even harder to defend than the partiality that individuals have to future people with their own genes, but I will argue that it plays a similar role in the moral appraisal of identity-affecting decisions. It denies a complaint to the objects of that partiality—people who will live shorter or less advantaged lives than the people who would have existed in their place under an impartial policy. The policymakers in such cases are still subject to criticism, because they make the world an impersonally worse place by acting on a suspect partiality. But this criticism is weaker than that to which the decision makers in Parfit's Risky Policy case are subject, since it is not accompanied by even the semblance of a complaint from the affected individuals. In light of the policymakers' intentions, those individuals appear to be the beneficiaries rather than the victims of an impersonally inferior policy.

Without a complaint from the affected individuals, the impersonal objection to the chosen policy has much less force. But this suggests that much of the outrage provoked by the Risky Policy reflects a judgment that the individuals harmed by that policy have also been wronged, and thus that the objection to the Risky Policy cannot be completely explained in impersonal terms. I will argue that this judgment is warranted. The Risky policymakers, having no intention to create individuals who can only live shorter and less advantaged lives, display a disrespectful indifference to the hardships faced by those individuals. In contrast, policymakers who act partially in creating future individuals expected to face similar hardships show no such disrespect to those individuals. I will argue that such differences in intention are relevant to our moral judgment of policies, as well as parental decisions, that cause harm to people who would not have existed without them.¹⁰

I will begin with a case in which policymakers make the impartial decision to delay the extinction of humanity by adding a generation with truncated lives. I will argue that a standard rights-based account cannot account for our approval of the decision to create an additional generation which will be painlessly annihilated after 35 years. An impersonal account has an easy explanation of our favorable judgment in this different-number case. But, as well as challenging a rights-based account of the Risky Policy, this case will set the stage for several others that challenge impersonal accounts.

An impersonal account will have trouble explaining our judgments in same-number variations of this case, in which the policymakers prefer the truncated lives of future people with whom they have a strong affinity to the longer lives of other people with whom they lack such affinity. People who owe their worthwhile lives to the agent's choice to create them instead of different people, or no people at all, have no grounds to complain about the harm they unavoidably suffer as a result of their creation. The agents may be criticized on other grounds, e.g., for chauvinism,

but those are not grounds on which the affected individuals will have any special standing to complain. These cases should suggest the need to take account of the agents' intentions in the moral appraisal of identity-affecting harm. Standard rights-based approaches cannot account for the lack of a complaint; impersonal approaches cannot account for the less severe criticism to which the policymakers are subject in the absence of such a complaint.

13.3 Justifying Harm in Different-Number Cases: The Generation-Adding Policy

Some of the difficulties in explaining the moral offense committed by the Risky policymakers in person-affecting terms are apparent in considering a variation I will call the Generation Adding Case (GA).

The apparent world-wide decline in sperm count turns out to be real. The only way to prolong fertility is to generate low level radiation, which will yield nuclear waste and eventually cause a nuclear explosion that will painlessly kill everyone then alive. We must choose between two policies. Under the first, where we do not intervene, humanity will last only through the 299th generation, because that generation will be born sterile. Under the second, we keep that generation fertile with radiation, but the members of the resulting 300th generation will, 35 years after their birth, be painlessly annihilated in the explosion caused by radioactive waste.

I think that most people would find it at least permissible to adopt the GA policy, leading to the painless annihilation of the 300th generation. We would find it acceptable to create a generation that would have worthwhile but truncated lives, if the only alternative was that no one would be born after the 299th generation. To accommodate misgivings about this policy, we can push back the time of the explosion, to 60 or 70 years, or reduce its impact, so that it causes permanent injury but not death. At some point, that policy will be acceptable to virtually everyone, although it will still cause harm to the last generation that it would ordinarily be wrong to inflict on other people without their consent. How can such a policy be acceptable, if the Risky Policy is not?

Some might respond that GA is described too schematically to elicit reliable intuitions. Generations are not literally cohorts, marching in chronological step, so in any actual world a nuclear explosion would kill people of all ages, many of whom had never had a chance at a minimally decent life. We could alter the hypothetical so that the waste build- would reach a "critical mass" at a certain date, causing a mutation that painlessly killed every person born thereafter in his 35th year. We could also stipulate that this date would be known, or knowable, so that the last generation could plan its lives accordingly. The reason I will stick with the nuclear disaster is that it clearly causes non-comparative harms,¹¹ thereby making it easier to claim a rights-violation despite the overall goodness of the truncated lives.¹² If no violation can be found in these circumstances, it seems, a fortiori, that none could be found in cases where the policymakers cause the last generation a mere deficiency rather than a discrete injury.

13.4 Justifying Harm in Same-Number Cases: The Chauvinist Policy

GA poses no obvious difficulties for proponents of an impersonal account of the Risky Policy. Unlike that policy, the GA policy involves a different-number choice, one that results in an impersonally better outcome than the alternative: instead of humanity dying out, another generation of people lead worthwhile but truncated lives. In contrast, the Risky Policy yields the same number of people as its alternative, but leaves the last generation with truncated lives. Our approval of the GA policy might thus appear to reinforce the claim that our real objection to the Risky Policy is its creation of an impersonally worse state of affairs.

In contrast, the GA policy appears to challenge accounts that locate the wrong done by the Risky policymakers in their violation of a right against the harm they cause to the affected individuals—their midlife annihilation.¹³ The “victims” in GA suffer the same harm and same benefit as the victims of the Risky Policy—a worthwhile life that is abruptly truncated—but lack a similar complaint.

The defender of a rights account might seize on the difference in overall outcome between GA and the Risky Policy to claim that any rights violation in GA is justified by the impersonal benefits the policy confers.¹⁴ But the increase in aggregate well-being resulting from the GA policy is clearly not sufficient to preclude a rights-violation. The same impersonal benefits claimed to justify or override the rights violation in GA would not justify the rights-violations where the same harm of midlife annihilation was inflicted on anyone *except* the beneficiaries of the policy. Thus, most of us would not regard it as justifiable to annihilate the next generation in mid-life in order to add a 300th generation; we would see the victims as seriously aggrieved.¹⁵

More important for my purposes, a net benefit is not *necessary* to avoid a rights-violation. There are policies that cause the same harm to the same number of individuals as the Risky Policy, with the same loss in aggregate well-being, that do not even appear to violate the rights of the affected individuals or give them grounds for complaint. Such policies will, for this reason, pose a challenge to impersonal accounts of the Risky Policy.

Consider a same-number variation of the GA policy in which an impersonal justification is clearly lacking. Call this the Chauvinist Policy (CP):

American policymakers can irradiate either the U.S. or Russia to preserve the fertility of its 299th generation. Irradiating either nation will yield about same number of people in the 300th generation. But irradiating the U.S. will cause a nuclear disaster in that generation, while irradiating Russia will not. Members of the last Russian generation would live normal life spans, although they could not reproduce. The policymakers, wanting the last generation be American, chose to irradiate the U.S. instead of Russia.

The policymakers might reasonably be criticized for their excessive partiality; for their chauvinism in valuing a generation of truncated American lives more than a generation of full Russian lives.¹⁶ But while the policymakers might be criticized from an impersonal perspective as well, for failing to choose the better available

to them, their failure to do so would not even appear to violate the rights of the last American generation. The members of that generation might, if they knew of their fate and the policy responsible for it, also criticize the policymakers for their chauvinism or lack of beneficence. But they would lack any special standing to complain, and the harm they suffered would be no more of a rights violation than the harm suffered by the last generation in GA, where the policymakers displayed no partiality.

This contrast suggests that if the Risky Policy violates the rights of the last generation, it cannot be merely because it lacks an impersonal justification—CP lacks that justification as well, yet it does not appear to violate the rights of those it kills.¹⁷ In both cases, the “victims” get the same net benefit from the policy that causes them harm—worthwhile but truncated lives. Yet the fact that the last American generation owes that benefit to the policymakers’ partiality appears to deny them a complaint against their benefactors, a complaint they might have if they were only the incidental beneficiaries of a policy adopted for other reasons.

If there is a rights violation in the Risky Policy, then, it seems that it can only be because the Risky, unlike the Chauvinist, policymakers, did not intend to create the people killed or injured in the nuclear accident. To explain why the rights of the last generation’s members are violated by the GA but not the Risky Policy, we would have to posit rights not against specific kinds of injury, but against having those kinds of injury inflicted in the absence of certain intentions. And although I think the wrong to the last generation is best understood in terms of the policymakers’ intentions, I am not sure it is helpful to understand that wrong as violation of a right to be acted on or against only in the presence or absence of certain intentions.

13.5 An Impersonal Explanation of the Chauvinist Policy

As I noted, CP also poses a challenge for impersonal accounts of the nonidentity problem. This challenge may not be obvious, since our disapproval of the policymakers appears to be readily explained in impersonal terms: they make the world worse off than it might have been by bringing into existence a last generation of truncated rather than full lives. The difficulty lies in explaining why we judge the Chauvinist policymakers less harshly than the Risky policymakers, even though they bring about consequences that do not differ from an impersonal, impartial perspective.

Some of those who defend an impersonal account of the Risky Policy, like Phillip Peters, urge us to reject the commonsense moral judgment that those harmed by the Risky policy have a personal complaint against the policymakers, and to see the wrong as one to the *class* of people who are worse off than they could have been—here, the potential members of the last generation.¹⁸ CP, however, inflicts the same harm on the same class as the Risky Policy without even appearing to give its members a complaint. (Perhaps earlier generations of *Russians* would have a complaint about the devaluation of Russian lives implicit in the policymakers’ choice.) This suggests that the impersonal objection to the Risky Policy does not

fully explain the disapproval it faces. Our sense that the people killed by that policy are wronged as well as harmed, and that they have a complaint that others lack, cannot be explained by the policy's consequences alone.

The defender of an impersonal account might argue that the difference in moral appraisal between the Risky and Chauvinist policies can be explained impersonally, in terms of a second baseline of comparison for the latter: the Chauvinist policymakers also had the option of not irradiating anyone, which would have resulted in the extinction of humanity at the close of the 299th generation. Irradiating the United States yielded an outcome that was impersonally superior to this, since the last generation of Americans has lives worth living. The Chauvinist policymakers did not do as well as they could have, but they also did better than they might have. In contrast, the Risky policymakers simply chose the worse of two options. It is this difference, the defender of an impersonal account might argue, that accounts for the difference in their moral appraisal.

The first problem with this response is that it may be a costly recourse for someone who seeks to explain our condemnation of the Risky Policy in impersonal terms while denying a general duty to produce the best consequences. One intuitively appealing strategy for narrowing the scope of the duty violated by the Risky policymakers involves restricting the comparison to alternatives involving the creation of the same number additional people. Thus, Jonathan Glover offers a principle, akin to Parfit's Q, that does not require us "to create extra happy people" but holds that "when we are going to add to the population, where the choice arises we must always prefer to add a happier than a less happy person."¹⁹ This principle is contravened as fully by the Chauvinist as the Risky policymakers; to condemn the latter more severely than the former, it appears that we must appeal to an impersonal principle of broader scope, one which takes account of the different-number alternative available to the former but not the latter, of letting humanity end in the 299th generation.

Moreover, the difference in the disapproval faced by the two policies is qualitative as well as quantitative. The Risky policymakers do not merely seem less beneficent or more selfish than the Chauvinist policymakers; they appear to wrong specific individuals in a way that the Chauvinist policymakers do not.

This difference is highlighted by a variation on CP in which the alternatives yield the same net utilities, but the people harmed appear to have a complaint against the policymakers. In the Ultra-Chauvinist Policy (UCP), the policymakers' motivation is no more selfish and the benefits and harms resulting from the policy are the same, but those benefits and harms are differently distributed:

The policymakers can bring about a 300th generation of Americans only by irradiating Russia, foreseeably causing the 299th generation of Russians to die in mid-life from a nuclear disaster. The 300th generation of Americans, which will have roughly the same number of people as the 299th generation of Russians, will live a normal life span. If the policymakers used a lower level of radiation, they could escape a nuclear disaster, but then the 300th generation would be Russian; if they did nothing, the world's population would die out peacefully in the 299th generation. The members of the 299th Russian generation owe their existence to the higher level of radiation, since without it, different Russians would have come into existence in that generation.

It is unclear how the proponent of an impersonal account would distinguish this case from the original CP. An impersonal inventory of the harms and benefits resulting from the two policies does not reveal a net difference.²⁰ Compared to the alternative of doing nothing, which would allow the extinction of humanity at the end of the 299th generation, both CP and UCP cause one harm, the nuclear disaster, and add the same number of total “life years”—UCP adds more life-years than CP in the 300th generation, but takes away the same number in the 299th generation. Compared to a non-Chauvinist alternative that adds a 300th generation with full lives, both policies involve one harm, the nuclear disaster, that befalls the same number of people and causes the same loss of life years—in the 300th generation in CP; in the 299th generation in UCP.

The critical difference between the two policies cannot be that CP only inflicts harm on those who benefit.²¹ The 299th-generation Russians annihilated in UCP benefit in the same way as the 300th-generation Americans annihilated in CP—they receive worthwhile but truncated lives. What distinguishes the policies is the fact that the policymakers *intended* to confer this benefit on the 300th generation of Americans in CP; but confer it on the 299th-generation Russians as an unintended side effect in UCP. These Russians, like the last generation destroyed by the Risky Policy, appear to have a substantial complaint against the policymakers, a complaint that the 300th-generation Americans lack. The intention of the policymakers in CP to confer a greater benefit on those who will suffer the lesser harm appears to give them a moral immunity that the UCP policymakers lack, for the harm caused to future people by the actions that bring about their worthwhile existence.

13.6 Limits on the Role of Procreative Intentions in Justifying Harm

The defender of an impersonal account might challenge the significance of intention for avoiding rights-violations with a different variation on the GA policy. Consider the Selfish Generation-Adding Policy (SGA):

The policymakers are as selfish and shortsighted as the Risky policymakers, equally unwilling to conserve energy in order to prevent a buildup of radioactive waste. The expected outcome of their policy, however, is the same as in GA—the accumulation of nuclear waste will extend fertility before it causes a disaster, so that its net effect will be to add a disaster-truncated generation of people with worthwhile lives.

Clearly, we would not praise the SGA policymakers, although the foreseen effects of their policy were beneficent. But we might not condemn them either, at least as harshly as the Risky policymakers, nor regard the last generation as having a complaint against them. The defender of an impersonal approach could argue that since these policymakers are as selfish as the Risky policymakers, the only basis for judging them less harshly is the more beneficial effect of their action.

Admittedly, the SGA policymakers, unlike the Risky Policymakers, do not manifest a willingness to cause harm by bringing about a comparatively worse outcome.

We may judge them less harshly because we (charitably) assume that they were unwilling to cause certain bad outcomes, although they were not motivated to produce certain good ones.²² But if they would not have caused a nuclear disaster without a net impersonal benefit, then they escape blame only because they accept an impersonal constraint on their conduct.²³

If impersonal factors play this even this indirect role in denying a complaint to the 300th generation in SGA, they appear to limit the significance of procreative intentions for the moral appraisal of harm to future generations. The Risky and SGA policymakers both lack such intentions in causing lethal harm, but the victims have a complaint only against the former. The beneficial consequences of the SGA policy appear to block the complaint despite the policymakers' lack of appropriate intentions.

And yet not all impersonally beneficial outcomes would do so—the 299th Russians irradiated in UCP would have had no less of a complaint if the only alternative policy would have been less beneficent, resulting in the quiet extinction of humanity at the end of the 299th generation. Although the policy that annihilates them is impersonally superior to that alternative, since it adds a 300th generation of Americans at the price of truncating the 299th generation of Russians, the Russians could claim to be sacrificed for the good of others, even though their worthwhile existence also results from the policy.

It seems that the beneficence of the policy only forestalls a complaint when the benefits *could* have been intended for those harmed. The SGA policy could have been adopted with such intentions, as the policy in GA was, but it is doubtful that the UCP policy could have been: what possible reason could the policymakers have had for preferring the specific 299th-generation Russians killed by that policy to the 299th-generation Russians who would have lived full lives under the alternative? This contrast suggests that those who are harmed by an impersonally beneficial policy will, if they lead worthwhile lives, have no grounds for complaint when they *could* have been its intended beneficiaries; when the different people who would have been created under alternative policies were distinguishable from them in ways that *could* have had motivational significance for the policymakers.

It might seem, then, that the impersonal benefit secured by SGA makes no independent contribution to the immunity enjoyed by the SGA policymakers, but merely provides, in the generation it adds, a plausible object for procreative intentions. This, however, would be too strong a conclusion. If a policy is *not* impersonally beneficial, the mere fact that it could have been intended to benefit those victims would not deny them a complaint. Consider CP, which is impersonally worse than its alternative. Although my intuitions on this are not strong, I think that the 300th generation of Americans created under that policy would have grounds for complaint about their mid-life annihilation if the policymakers had not intended to create them, e.g., if they had produced the radiation solely for their own comfort. In sum, these cases suggest that a policy that is neither impersonally beneficent, nor intended to benefit those it unavoidably harms, does not confer immunity on the policymakers for the harm.

At most, then, impersonal considerations *may* block a complaint from those who are harmed by policies that could have been, but were not, intended for their benefit. But those considerations cannot explain the substantial differences we find

in the evaluation of policies with impersonally similar outcomes. In particular, they cannot explain why the American “victims” of CP do not have a complaint, while the Russian victims of UCP do. An impersonal account simply lacks the resources to explain the significance of procreative intentions in the moral appraisal of identity-affecting policies.

13.7 An Alternative? Asymmetrical Responsibility for Harms and Benefits

The difficulties already facing the notion of procreative intentions might suggest that the variable moral assessment of the array of policymakers I have discussed has a simpler explanation. Early in the debate on the non-identity problem, Matthew Hanser argued that we condemn the Risky policymakers because they are responsible for the harm but not the benefit to the last generation.²⁴ Hanser suggested that the conditions for assigning responsibility for harm and benefit are different; without attempting to specify those conditions, he found it plausible to conclude that the Risky policymakers satisfy the former conditions but not the latter. Their knowledge that their policy will cause the nuclear disaster suffices to make them responsible for the harm they cause, while their indifference to and ignorance of the specifics of, the policy’s identity-affects denies them responsibility for the worthwhile existence of the nuclear people. They are responsible for the nuclear disaster, which they foresee, but not for the worthwhile existence of the last generation, since they are indifferent to the existential consequences of their policy. Hanser contrasted policymaking with ordinary reproduction, claiming that a parent who knowingly has a child with a harmful genetic condition is responsible for the benefit of that child’s existence but not the harm²⁵—a claim I will return to later in the paper.

Hanser’s account has obvious affinities with, and a deep influence on, my account of procreative intentions. But I do not think the piecemeal attribution of responsibility for harm and benefit can explain the range of cases I have presented. Although I think that Hanser is correct in claiming distinct responsibility conditions for harm and benefit in general, I will argue that the agents in the cases I have considered are responsible for *both* the intended benefit and the expected harm. In those cases where they intend to bring into existence people whom (they believe) can only exist with those harms, they can invoke the benefit to justify the harm. And I will argue that this is equally true for policymakers and prospective parents.

Hanser’s account was schematic; he did not spell out responsibility conditions, but merely argued for an asymmetry between the responsibility conditions for harm and benefit. I will begin, then, by considering the responsibility conditions for benefits in more detail.

Two features of the relationship of the Risky policymakers to its putative victims might explain why we do not hold the former responsible for the worthwhile existence of the latter: first, the policymakers’ lack of intention or even foresight about the beneficial effects of their actions, and second, the very oblique, insubstantial causal role their actions play in bringing about those effects.

Joel Feinberg famously compared the Risky policymakers to a cabdriver whose negligent driving causes his passenger to miss a flight that later crashes.²⁶ We regard the driver as responsible for making the passenger late but not saving his life. The obvious difference, however, between the taxi driver and the Risky policymakers is that the latter could have foreseen, even if they did not intend, the benefit to the nuclear people: they could well have known that the policy they chose would alter the identity of the members of future generations and thus bring about the worthwhile lives of the people it harmed.

Hanser might respond that the creation of the *specific* people who come into being in the 300th generation is not foreseeable, and that for this reason the Risky Policymakers are *not* responsible for their creation. But this would also seem to deny credit to the GA and CP policymakers (unless one can intend a result one cannot foresee), since neither has any knowledge about the specific people their actions will cause to exist: the GA policymakers merely know that additional people will live worthwhile lives if they act in one way, none if they do not; the CP policymakers merely know that additional Americans will lead worthwhile lives if they act in one way, while additional Russians would lead worthwhile lives if they acted otherwise.

A more tenable distinction between the policymakers might be made on the basis of their causal contribution to the benefits enjoyed by the last generation. The causal role of the GA and CP policymakers, although mediated by the reproductive choices of 299 intervening generations, appears to be more direct and substantial than that of the Risky policymakers: the actions of the former have a direct, if delayed, effect on the fertility of the 299th generation. But however attenuated it may be, the causal role of the Risky policymakers in altering the identities of the people in the last generation would suffice to make them responsible for benefits they *intended* to bring about, e.g., if they sought to prevent the creation of a 300th generation Hitler by actions that altered the mating patterns in the next generation. Unless the conditions for causation vary with the agent's intentions, it is hard to deny the Risky policymakers responsibility for the benefit on the ground that their causal contribution is not adequate.

But while an intention to confer the benefit of a worthwhile existence may be necessary to give the policymakers responsibility for that benefit,²⁷ it is not sufficient to allow them to invoke that benefit to justify serious harm. The piecemeal attribution of responsibility cannot explain, for example, why the offset of harm by benefit is permitted in CP but not in UCP.²⁸ Presumably, part of the reason the policymakers would not be blamed for the harm to the Americans in CP is that it is outweighed by the benefit of their worthwhile existence, for which the policymakers are also responsible. Although the UCP policymakers are not responsible for benefiting the Russians they harm, they are responsible for the benefit of a full life they confer on the 300th generation Americans—a an even greater benefit than they confer on the last generation of Americans in CP. The question is why they cannot “transfer” their credit from the intended to the unintended benefit, or offset the blame they get for the harm by the credit they get for the benefit, even though the harm and benefit are to different individuals. Excluding credit when the victims and beneficiaries are different people requires something more than the separate attribution of responsibility

for harm and benefit; it requires an account of the moral significance of the intended or foreseen *relationship* between the harms and benefits.²⁹

Hanser might claim that all that was needed was a restriction on permitted offsets to harms and benefits received by the same person. An account that permitted A's harm to be offset by B's benefit would fail to respect the moral boundaries between people. But a limitation on set-offs cannot explain why we do not always give off-setting credit for the benefits of existence when the policymakers are responsible for harms and benefits to the same people; it cannot account for cases where future people appear to have a complaint about the harm caused by policies that were intended in some sense to bring about their worthwhile existence.

Such cases can arise in both same- and different-number contexts. In the Risky policy, which is a same-number case, we can increase the responsibility of the policymakers for the distinct existence of the nuclear people without significantly weakening the complaint those people have against them. Take a case where the policymakers seek, on the advice of counsel, to assure a totally different population in the 300th generation, by choosing, among equally risky policies, the one most certain to ensure no overlap with the population that would have existed under the alternative Safe policy. They may even be willing to endure some small discomfort to ensure there is no overlap. We would still condemn the policymakers, and still see the last generation as having a complaint against them. Yet Hanser could hardly deny that they were responsible for the separate existence and worthwhile lives of the people whom their policy destroyed. Their responsibility for the benefit, however, would not confer even partial immunity for the harm.³⁰

Second, such immunity is lacking even in some different-number cases, where the policymakers intend to create *additional* people. Consider a variation on GA in which the policymakers, though willing to create an additional generation, are unwilling to incur even a slight sacrifice of their own comfort to give its members full lives. The Very Selfish Generation Adding (VSGA) policymakers could ensure the fertility of the 299th generation with a much lower level of radiation that would not risk nuclear disaster in the 300th generation but which would require a significant economic sacrifice to produce. Instead, they use a high level of radiation, with its attendant risks, that can be produced without sacrifice. As in the Risky Policy, the choice of the higher dose has far-reaching effects, resulting in completely different people in the last generation than there would have been with a lower dose of radiation.

Like the original GA policymakers, the VSGA policymakers appear to be responsible for the benefit as well as the harm to the last generation, because their only reason for producing radiation in the first place is to extend fertility. But we would still condemn them almost as harshly as the Risky policymakers, whom Hanser holds responsible for the harm but not the benefit.³¹ We would condemn them even if, like the liability-avoiding policymakers described above, they had selected among equally risky policies the one most certain to alter the make-up of the last generation. We might condemn them less harshly than the Risky policymakers because they did more good than they might have, by creating even a truncated 300th generation, and because they did intend that benefit. But their responsibility for the existence

of the people harmed would not give them the moral immunity enjoyed by the GA policymakers.³² By incurring even a slight sacrifice to create additional people, they would appear to benefit those people in a morally relevant sense. But that benefit does not offset the harm in the way that it does for the GA policymakers. Hanser's account lacks the resources to explain this difference.

13.8 The Limitations of Asymmetrical Responsibility

Hanser, as noted, reaches a different conclusion about responsibility in micro than in macro cases. When harm to an individual, such a genetically-based impairment, arises from what he calls "intrinsically reproductive acts," he suggests that the agent will be responsible for the benefit—the individual's worthwhile existence—but not that harm. Unlike the Risky policymakers, who cause the harm with no intent to benefit, the parent intends (in continuing if not in initiating the pregnancy) to confer the benefit and merely "transmits" the harm.

I agree that the moral evaluation of policies and procreative acts will often be different, but I do not think the piecemeal assignment of responsibility is any more adequate for micro than macro cases. The limitations of Hanser's account can be seen by contrasting two sets of parents. Both can avoid having a child with a mild cognitive impairment simply by waiting a year to conceive. For the first set of parents, that delay will cause a significant loss of income or other goods. Call these the burdened parents. The second set of parents *wants* to have a mildly impaired child because their closeness with their own mildly impaired parents have given them a strong affinity with, and partiality towards, people with similar impairments. Call these the partial parents. In light of their close relationship with their own parents, they are confident that they would be devoted and skillful caregivers. Neither generation, however, would find it a sacrifice to forego such having a cognitively impaired child; everyone would be happy if the child were not impaired.

The first case involves a substantial burden but no reason (except avoiding that burden) for creating an impaired child; the second involves a reason but no substantial burden.³³ While many people would disapprove of both sets of parents for not waiting to conceive, I suspect that most people would criticize the burdened parents more strongly than the partial ones. But even those who criticized the partial parents as strongly would have trouble attributing a complaint to the child. A desire to create and raise a particular kind of child appears to forestall or mitigate a complaint more than a substantial burden.

This difference is especially striking because the second set of parents *intends* for the child to have an impairment, while the first set merely foresees the impairment as a consequence of conceiving within a year.³⁴ Although the two sets of parents seem equally responsible for the benefit of the child's existence, the partial parents seem, if anything, more responsible for the harm. Hanser's account would have trouble explaining this difference, since the comparative blameworthiness of the two sets of parents correlates so poorly with their comparative responsibility for the benefits and harms.

Moreover, the desire for a retarded child seems mitigating in a different way than the burden of avoiding one. The burdened parents are engaged in a kind of balancing, while the partial parents are not—the question is whether their reason for desiring a child with the impairment is an appropriate one, not whether the burden they suffer in failing to satisfy that desire is substantial enough. (I have stipulated that it is insubstantial.) Finally, the child of the burdened parents has more of a basis for resentment, since he can see his impairment as the unintended consequence of his parents' insensitivity to the limitations and frustrations they expected him to experience. The child of the partial parents cannot see his impairment in this way; it is not a consequence of his parent's disregard but (part of) the reason he was brought into existence. That awareness may be bittersweet for a child if he is frustrated by his impairment, but any resentment should be tempered by the recognition that he is alive because his parents wanted a child "like him" to enjoy the goods of existence.

13.9 Selectivity, Partiality, and Procreative Intentions

In denying that the decisions of the Chauvinist policymakers and the partial parents give rise to a complaint, I have suggested a particular way of looking at their procreative acts: as picking from an array of potential creatures those who will be brought into existence. The agent may know that some of the potential people will, if made actual, suffer various harms or limitations, but it seems inappropriate to blame him for harming them, or for causing them to exist in a harmful state, by the very acts that select them for existence—even if those acts cause or contribute to the harm, and the agent knows that they do.³⁵

There are, of course, some selections it would be wrong to make, e.g. of people who will live with no pleasure and unremitting pain. And there are, of course, many bad reasons for selection, like those based on Chauvinist, racist, or sexist preferences. Indeed, I have argued elsewhere that almost any trait-based selectivity represents a departure from the parental ideal of unconditional welcome.³⁶ But the use of a bad selection criterion may not give the individual selected a complaint with respect to the harm he unavoidably suffers.³⁷

Clearly, not all agents who knowingly cause people to come into existence can be seen as choosing them from an array of potential people. Most ordinary parents cannot, since they exercise no such selectivity. The agents who adopt the Risky policy cannot be seen as making such a choice either, because they have no intentions with respect to the "existential consequences" of their actions. In contrast, the adoption of the GA policy does involve a similar choice—of people who will lead truncated lives over no people. The requisite intent can be found in both same- and different-number cases.³⁸

But the mere intention to create additional people is not enough. Although the VSGA policymakers clearly have procreative intentions, they cannot be seen as choosing people expected to suffer certain harms from an array of potential people. Apart from the lesser burden on them, they have no reason to prefer the future people

who would be harmed over future people who would not be, so they cannot be seen as choosing the former in selecting the higher level of radiation.

Nor can such a choice be based on the mere intention to create different people than those who would have existed under a policy that had less harmful consequences. In acting to ensure that the people who come into existence in the 300th generation under the Risky Policy are not the same people who would have come into existence under a safe policy, the Risky policymakers do not select one group of people from an array of potential people. The problem is not that they select by exclusion, as they would in adopting a policy aimed at creating “anyone but Russians.” It is rather that their selection criterion picks out people only with reference to the outcome of an alternative policy, and not with reference to any independently-identifiable characteristic of the people chosen among. While “non-Russian” can be a criterion for selection, “different from any person who would have come into existence under the alternative policy” cannot be.

13.10 Conclusion

The immunity conferred by procreative intentions may still seem mysterious. We can hardly assimilate partiality toward some possible future children to partiality toward one’s actual children. A parent does not need to give reasons for preferring her children to those of other parents. But the partiality that she is morally permitted or even required to show is grounded in a relationship she had already established. It is not clear how she can invoke such partiality to justify the decision to establish such a relationship in the first place.

This difficulty in justifying partiality towards a kind of future child does not arise from the difference in moral status between potential and actual children—if potential children can even be candidates for moral status (as in-vitro embryos can be). Rather, it arises from the different relationships between chooser and chosen. This is suggested by comparing prospective biological parents with prospective adoptive parents. Although the latter may choose among actual children, they also have not yet developed a relationship with those they choose among. For this reason, trait-based partiality toward an existing adoptive child is no easier to justify than such partiality toward a future child. Adoption agencies may have a good moral reason to permit such partiality and allow parents to select children of their own race, or of a particular gender—to save more children from abandonment, hunger, or physical harm. But it is not clear that the adoptive parents have any stronger moral basis for exercising such partiality than prospective biological parents. *If* adoptive parents are morally permitted to exercise such partiality, it is for the same reasons that biological parents may—that in choosing the individual with whom they will establish this uniquely intimate relationship, they are unconstrained by impersonal considerations—in their case, by considerations of how much good they could do by adopting one child rather than another.³⁹

The similarity in the latitude enjoyed by prospective biological and adoptive parents in selecting among children suggests that the causal role of the former in causing “necessary” harm to the children they select has little moral significance.

What does seem to matter is whether the child is “wanted” in part for some attribute that it can only have in a harmful state, or even wanted for that harmful state itself. While the desire to create a person with such a trait or in such a condition may sometimes reflect a morally deformed conception of the good, those who owe their worthwhile existence to that desire will have no special standing to complain.

Notes

1. Parfit (1984), pp. 351–79.
2. See, for example, Woodward (1986), Steinbock’s and McClamrock (1994) and Tooley (1998).
3. Parfit (1984), *op. cit.*
4. See references in note 2.
5. See, for example, Brock (1995) and McMahan (1998).
6. There are several related distinctions between macro and micro cases as described: between the agents—policymakers vs. individual parents; between the acts—policy decisions and their implementation vs. “inherently reproductive acts” (Hanser, 1990); and between the intentions with which the acts are done—whether or not the agent intends to affect the existence, number, or identity of future children. I will distinguish macro and micro cases by differences between agents and actions, but not by intentions; I will claim that policymakers can act, in choosing long-term policies, with intentions that are similar in morally relevant respects to the intentions parents sometimes act with in having children.
7. Roberts (2007).
8. Lillehammer (2009), in this collection. Roberts’ position, as outlined above, is that macro cases can be “person-affecting,” in making people worse off than they would have otherwise been, even if they are not “person involving” in Lillehammer’s sense, since the affected individuals cannot be identified at the time of the decision or action.
9. Wasserman (2005).
10. Wasserman (2005).
11. Shiffrin (1999) and Harman (2004).
12. GA itself appears to be motivated in part by non-person affecting considerations—it is not for the sake of specific human beings, or even kinds of human beings, that the policymakers prolong humanity, but from a combination of impersonal beneficence, historic preservation, and species-partiality. Still, the 300th generation people can be seen as the policymakers’ intended beneficiaries if part of their reason for adding a generation was to allow its members, whoever they might be, to enjoy the goods of existence.
13. This would include Woodward’s (1986) account, which involves a right against that harm, but not Steinbock’s and McClamrock’s (1994), which is based on a threshold of well-being that the members of the last generation may well exceed even if they are annihilated in mid-life.
14. The defender might argue that although the GA as well as the Risky policy violated the rights of the last generation, that violation was overridden in the former by a weighty justification lacking in the latter. The first thing to notice is that a justification in these terms for GA would not be a strictly person-affecting one, since the benefit to the people harmed is the same in both cases: their worthwhile existence. The harm in the former case is offset by the impersonal benefit of creating an additional generation, a benefit the Risky Policy does not confer.

But if GA was distinguished from the Risky Policy only by the availability of an impersonal justification for the harm caused the last generation, the explanation for the wrongness of the Risky Policy would not be person-affecting “all the way down”: while that policy would be wrong because there was an unjustified rights violation, there would be such an violation only because the harm to the last generation was not outweighed by the impersonal benefits flowing from the policy. This might be acceptable to some proponents of a person-affecting account, as

a way of integrating (or personalizing) the impersonal differences adduced by Parfit: while the rights-violation is personal to the victim, it can be overridden if, on an impersonal assessment, its benefits outweigh its harms.

15. Although it might be less egregious to annihilate the next generation in order to add a 300th generation than to do so in order to substitute one set of people in the 300th generation for another, particularly if the two sets were now indistinguishable. The midlife annihilation of the next generation to produce a 300th, of course, would also pose an issue for an impersonal account, but only the familiar one of utility trumping rights.

Perhaps a net benefit, assessed impersonally, is sufficient to preclude a rights violation in cases where the victims are also the beneficiaries. But this would yield a peculiar hybrid account. It could not justify the harm either in terms of the capacity of a sufficiently great impersonal benefit to override important rights—the next-generation case rules that out—nor in terms of the victims' hypothetical consent to the harm—such consent would be forthcoming from the victims of the Risky Policy as well. Moreover, as I go on to argue in the text, even in cases where the victims are also the beneficiaries, a net impersonal benefit may not be *necessary* to avoid a rights violation.

16. We could modify the example to make their chauvinism more or less reasonable by making it more or less “naked,” e.g., by varying the expected cultural continuity and genetic affinity between present-day and 300th generation America. Even if we believed that people had a prerogative not only to favor their present countrymen but to perpetuate their culture and even their genes, we might still balk at respecting their strong preference for future people who had nothing in common with them genetically or biologically, just the same national affiliation.
17. The limited relevance of the impersonal justification of an action or policy to the occurrence of a rights violation is also apparent from cases involving existing people. Thus, if a surgeon can only save an accident victim by a transplant that will cause a fatal immune reaction in a year, he would have an impersonal justification for doing if no one else needed the organ, even if the victim injured himself by drunk driving. But the transplant might not be justified if it denied the organ to another person blamelessly injured in the accident, who could be saved by the transplant without any reaction. While many people would favor a lottery between the two victims, most would condemn the surgeon for choosing the first simply on whim, or because the transplant procedure was slightly less time-consuming. His choice would result in a severely truncated life for one patient rather than a full life for another, arguably more deserving one. In impersonal terms, the fatal immune reaction suffered by the patient he chose lacks a justification.

But the surgeon would not appear to violate the rights of the organ recipient any more than he would if no one else needed the organ. The fact that the patient's own life was the one saved would preclude us from finding that his rights were violated by the fatal immune reaction the transplant caused. This would be so even if the patient himself was critical of the surgeon after the operation for choosing him when he could have saved someone more deserving with a much better expected outcome.

18. Peters (2004).
19. Glover (1992).
20. We could also stipulate that all knowledge concerning the policy choice would be lost long before the 299th generation, so that the nuclear disaster would appear equally accidental to its victims in both cases.
21. In any case, we can stipulate that the radiation causes sufficient genetic changes over the Russian generations to make the existence of any given individual in the 299th generation more likely, even *ex ante*, with the radiation than without it. The increased *ex ante* odds of coming into existence make it more plausible to claim for CP than for the Risky Policy that the policymakers do as well as possible for the affected individuals (see Roberts, 2007).
22. What if the SGA policymakers had expected to bring about the same outcome as the Risky policy but unexpectedly brought about the same outcome as the GA policy? I think we would judge them less harshly than the Risky policymakers, but only because actual consequences, and thus moral luck, play a significant role in moral appraisal.

23. The opponent might insist the constraint need not be an impersonal one. It might be that the SGA policymakers would be unwilling to harm anyone who would not thereby receive a greater benefit, whether or not the net effect was beneficial. But if we impute this constraint to them, why not to the Risky policymakers as well? What if the Risky policymakers would have balked at causing a nuclear disaster if they believed that it would befall people who would otherwise have led full lives? I can see no satisfactory way of distinguishing the SGA policymakers from the Risky policymakers on the basis of their intentions, actual or counterfactual, except in terms of an impersonal constraint that is violated by the Risky but not the SGA policymakers.
24. Hanser (1990).
25. Hanser (1990).
26. Feinberg (1986).
27. However, I argued in Section 13.6. that the intention to confer the benefit may not be necessary, merely the possibility of having such an intention.
28. There is a threshold difficulty in even applying Hanser's account to the Chauvinist Policies—he assumes the policymakers “could not possibly make a Same Number Choice [one affecting the identity but not the number of future people] in order to ensure one group of people's coming into existence rather than another.” Hanser (1990), pp. 61–2. This is because “prior to the people's actual conception there can be no foreknowledge of who will come into existence if one policy is adopted and who will come into existence if another is adopted.” The Chauvinist Policies, however, involve just such a choice, about what group of people will come into existence. All that is required for the policymakers' choice is that they have foreknowledge of the *type* of people who will come into existence under different policies.
29. The question of whether responsibility for benefits and harms to future people can be assessed piecemeal may be related, in a way I cannot explore here, to the question of whether there is a difference in the reason-giving force of bads and goods in the lives of future people. According to McMahan, the commonsense view is that the expected bads of a future life have both a “cancelling” function – to balance out the expected goods in a future life in assessing its overall value for the person – and a “reason-giving” function – to provide the agent with a reason to not bring that person into being if the balance is negative. McMahan (2009). In contrast, the expected goods of a future life have only a cancelling function—to justify (or failing to justify) the agent in exposing that person to the unavoidable hardships of his life.
30. The persistence of the victims' complaint in the face of the policymakers' intention to alter the identity of the 300th generation also raises doubts about Feinberg's suggestion that the nuclear people resent the Risky policymakers for their manifest willingness to inflict harm on members of their class. That resentment persists in this variation, despite the fact that the Risky policymakers only manifested a willingness to harm those whom they benefit overall.
31. The selfish policymakers could be compared to the woman in Parfit's (1984) case who conceives a child with a genetic impairment now because she is too impatient to wait a month to conceive an unimpaired one. (We could increase the similarity between the cases by having the act of conception itself contribute to the impairment, so that the mother would clearly be responsible for the impairment as well as for conceiving an impaired child.) She, too, is responsible for the existence of the person she helps create, as well as for the harm he will suffer. If we do not judge her as harshly as the VGSA policymakers, it may be because she confers the benefit of existence so much more directly, or because she chooses to carry the fetus to term, or because her impatience expresses a commendable longing to confer that benefit, a longing that she expresses “not wisely but too well.”
32. Hanser, of course, would not be compelled to reject this conclusion. In fact, he expresses doubts about a principle that would grant such immunity to the VSGA policymakers, to wit: “If a choice benefits someone (in the morally relevant sense) more than it harms him (in the morally relevant sense), the choice is not objectionable owing to the harm it does them.” But he does not explain why this principle is unsatisfactory.

33. The burdened parents are like the VGSA policymakers, but with a more direct procreative role and a greater sacrifice required to avoid harm; the partial parents are like the Chauvinist policymakers, with a more direct procreative role and a better reason for creating a person they know can only exist in a harmful state.
34. The Chauvinist policymakers would fall in between, since they see the (more clearly) harmful condition into which they place the last generation of Americans as necessary for its existence.
35. Whether we can even be said to harm someone by the “very act” by which we intentionally bring about his worthwhile life seems to depend on how we individuate acts. It is difficult to see the act of (deliberately) creating a person as harming him, even if it unavoidably causes him harm, since the relevant counterfactual comparison is between the person’s whole life and his non-existence. The question appears to be whether a single complex set of physical movements can be an act of harming as well as an act of creation when it is logically but not physically possible to have the life without the harm. Hanser (this volume) raises distinct objections to the claim that parents can be said to harm their children by bringing about their harmful states merely by genetic transmission.
36. Asch and Wasserman (2005).
37. Thus, to modify a case of Parfit’s (1984) imagine that Ruth will pass on her fatal disease only to sons, but that she so much wants a son that she employs preimplantation screening to assure that she gets one. That son has no grounds for complaint against Ruth if he has a worthwhile life, even if we would condemn Ruth for such an extreme sex preference. In contrast, the son in Parfit’s original case, where the mother does not act on a sex-preference, may have a complaint, because his mother’s indifference to the sex of her child makes the harm he suffers seem gratuitous, despite the fact that he could not have existed without it. Even if we think that this complaint would be a weak one, the contrast between the two cases highlights the immunizing role of procreative intentions.
38. We can classify grounds for choice into descriptive or historical. A choice based solely on expected attributes that are desired by the parent or the procreative agent is descriptive, and it raises the now-familiar spectre of a genetic supermarket. A choice based on the donor of the gamete source is descriptive to the extent it seeks the donor’s traits, but is historical to the extent that it seeks to express or extend a relationship with the donor, e.g. a woman who chooses to implant an embryo with a genetic impairment conceived with her deceased husband rather than an embryo without the impairment but conceived with a stranger, because she would rather have an impaired child “with” her husband than an unimpaired child from a stranger. The choices made by the Chauvinist policymakers and the couple with cognitively impaired parents are hybrids as well: they are historical to the extent they concern the relationship between the potential person and other people; they are descriptive to the extent that they concern the expected attributes of the potential person.
39. But by the same token, if trait-based selectivity is morally problematic for adoptive parents, it is not because they owe equal consideration to the children they choose among. If, as I would be inclined to argue, they should flip a coin among those children, it is not because they owe them a right to equal consideration or equal chances, but because they should not condition membership in the intimate association of their family on any of the characteristic they could screen for.

References

- Asch, A. and D. Wasserman. 2005. Where is the sin in synecdoche? In *Quality of life and human difference: Genetic testing, health care, and disability*, eds. D. Wasserman, J. Bickenbach, and R. Wachbroit. New York: Cambridge University Press.
- Brock, D. 1995. The non-identity problem and genetic harms: The case of wrongful handicaps. *Bioethics* 9: 269.
- Feinberg, J. 1986. Wrongful life and the counterfactual element in harming. *Social Policy and Philosophy* 4: 145–78.

- Glover, J. 1992. Future people, disability, and screening. In *Justice between age groups and generations*, eds. P. Laslett and J. Fishkin. New Haven: Yale University Press 127–43.
- Hanser, M. 1990. Harming future people. *Philosophy and Public Affairs* 19: 47–70.
- Harman, E. 2004. Can we harm and benefit in creating? *Philosophical Perspectives* 18: 89–113.
- Lillehammer, H. 2009. Reproduction, partiality, and the non-identity problem. In this collection.
- McMahan, J. 1998. Wrongful life: Paradoxes in the morality of causing people to exist. In *Rational commitment and social justice: Essays for Gregory Kavka*, eds. J. Coleman and C. Morris. Cambridge: Cambridge University Press.
- McMahan, J. 2009. Asymmetries in the morality of causing people to exist. In this collection.
- Parfit, D. 1984. *Reasons and persons*. Oxford: Clarendon Press.
- Peters, Jr., P. 2004. *How safe is safe enough? Obligations to the children of reproductive technology*. New York: Oxford University Press.
- Roberts, M. 2007. The non-identity fallacy: Harm, probability and another look at Parfit's depletion example. *Utilitas* 19: 267–311.
- Shiffrin, S. 1999. Wrongful life, procreative responsibility, and the significance of harm. *Legal Theory* 5(2): 117–148.
- Steinbock, B., and R. McClamrock. 1994. When is birth unfair to the child? *Hastings Center Report* 24: 15–21.
- Tooley, M. 1998. Value, obligation, and the asymmetry question. *Bioethics* 12: 111–124.
- Wasserman, D. 2005. The nonidentity problem, disability, and the role morality of prospective parents. *Ethics* 116: 132–52.
- Woodward, J. 1986. The non-identity problem. *Ethics* 96: 804–31.

Part VI
Is the Person Affecting Approach
Objectionable Independent of the
Nonidentity Problem?

Chapter 14

Can the Person Affecting Restriction Solve the Problems in Population Ethics?

Gustaf Arrhenius

Abstract The person-affecting restriction, in its slogan form, states that an outcome can be better than another only if it is better for someone. It has a strong intuitive appeal and several theorists have suggested that it avoids certain counterintuitive implications in population ethics. At the same time, the restriction has highly counterintuitive implications and yields non-transitive orderings in some nonidentity cases. Many theorists have taken this criticism to be decisive. Recently, however, there have been some reformulations of the restriction, suggesting that the restriction survives this “old” criticism. This paper investigates the viability of those reformulations, which are versions of “Comparativism,” and argues that most of them either have counterintuitive implications or are extensionally equivalent with impersonal theories, but that “Soft Comparativism” seems to have an advantage over impersonal theories.

Keywords Person-affecting restriction · Population ethics · The repugnant conclusion · Broome · Parfit · Temkin · Comparativism · Non-transitive orderings.

14.1 Introduction

The Person Affecting Restriction, in its slogan form, states that an outcome can only be better than another if it is better for someone.¹ The restriction has a strong intuitive appeal and several theorists have suggested that the counterintuitive implications in population ethics of so-called “impersonal” welfarist theories, such as classical utilitarianism’s implication of Derek Parfit’s well-known Repugnant Conclusion, could be avoided by adopting it.²

However, the Person Affecting Restriction has been subjected to some severe criticism by, among others, Derek Parfit, John Broome, Larry Temkin, and me.³ This criticism turns around the implications of the restriction in non-identity cases,

G. Arrhenius (✉)

Department of Philosophy, Stockholm University, Swedish Collegium for Advanced Study, Sweden

e-mail: Gustaf.Arrhenius@philosophy.su.se

that is, cases in which the identity of the people who will exist in the future is at stake. It has been shown that, in a number of such cases, the restriction has highly counterintuitive implications and yields non-transitive orderings.

Many have taken this criticism to be decisive and thus accept that we have to abandon the restriction. Recently, however, there have been some interesting reformulations and defences of the restriction, including one of my own making, suggesting that the restriction survives the “old” criticism. I shall investigate the viability of these proposals.

These new theories are all versions of a view I call Comparativism so I shall start by introducing this view. I shall then discuss different versions of the restriction and how these are related to Comparativism. The rest of the paper discusses the implications of different versions of Comparativism in population ethics. I shall show that most versions of this view either have counterintuitive implications or are extensionally equivalent with impersonal theories. The exception is a version of what I call Soft Comparativism. Albeit not without its own problems, it seems to have an advantage over impersonal theories.

14.2 Comparativism

According to Comparativism, we should draw a distinction between *uniquely* and *non-uniquely realisable* people. The former people exist in only one out of two compared outcomes, whereas the latter exist in both of the compared outcomes.⁴ The idea is that we should take the well-being of non-uniquely realisable people into account in a different way as compared to the well-being of uniquely realisable people.

Consider the following condition for same sized population:

Neutrality: If there is a one-to-one correspondence from outcome A to outcome B such that every person in A has the same welfare as their counterpart in B, then A and B are equally good.

Standard welfarist axiologies, such as the axiological part of Total Utilitarianism, count everyone’s welfare equally and thus satisfy Neutrality. Comparativism, however, counts people’s welfare differently depending on whether they are uniquely or non-uniquely realisable and thus violates Neutrality.

A strict comparativist only counts the welfare of non-uniquely realisable people and completely disregards the welfare of uniquely realisable people. A soft comparativist, on the other hand, counts the welfare of everybody but gives more importance to the welfare of non-uniquely realisable people.

According to a third view, among people with positive welfare, we should only count the welfare of non-uniquely realisable people, but among people with negative welfare, we should count the welfare of everybody. In other words, this view respects Neutrality in regard to populations with negative welfare. The reason behind this move is to incorporate an idea called *Asymmetry*: we have no moral

reasons to create people with positive welfare, other things being equal, but we have reasons not to create people with negative welfare, other things being equal.

In many cases, the motivation behind drawing one or the other of the above distinctions is the Person Affecting Restriction. The idea is that the restriction is supposed to entail one or another of the above distinctions. How this entailment is supposed to work is by no means clear and depends, of course, on how one understands the Person Affecting Restriction, to which we shall now turn.

14.3 The Person Affecting Restriction

The Person Affecting Restriction has been formulated in different ways by different authors. I shall take as my point of departure the slogan above, originally formulated by Temkin. It is quite vague, however, and open to several interpretations, and I think it can be worthwhile to discuss some of them, since it might be that the intuitive support for the Person Affecting Restrictions derives from these weaker and less controversial principles.

The restriction could be understood as an idea about which kind of objects moral goodness supervenes on, for example, that goodness is essentially related to the interests of human beings. All moral claims would thus necessarily involve a reference to humans: outcome A is better than outcome B since people have higher welfare in the former as compared to the latter outcome, or since in the former but not in the latter outcome people's rights are fulfilled, or in the former but not in the latter people have equal opportunities, and so forth.

Examples of putative moral claims which are ruled out by this restriction would thus be: outcome A is better than outcome B since the scenery is beautiful in the former but ugly in the latter outcome, or since the ecosystem is in balance in the former but not in the latter outcome, and so forth. Roughly, this interpretation of the Person Affecting Restriction, which we could call the Human Good Restriction, claims that two outcomes can only differ in value if they differ in regard to some aspect of human goods.⁵

This restriction is pretty reasonable and I think that much of the appeal of the Person Affecting Restriction derives from the Human Good Restriction.⁶ It is clearly insufficient, however, to yield any kind of distinction between the value of people that exist in more than one outcome and uniquely realisable people.

One can give a stronger interpretation of the Person Affecting Restriction than the one given above. One can stress an individualist aspect of value: all moral goods are *personal goods* which, roughly, are goods attached to individuals. These goods are features of people's lives that make them go better or worse. One version of this view is a form of Welfarism which claims that an outcome can only be better than another outcome if there is someone in the first outcome with a higher welfare level than someone in the other outcome (given the same number of people in the compared outcomes).

Consider the following two outcomes: in A, Krister and Erik are equally happy. In B, they are both happier than in A but Krister is happier than Erik. An egalitarian

might argue that B is worse, or at least in one respect worse, than A, since although both Erik and Krister are better off in A than in B, B involves inequality whereas there is perfect equality in A. One might say that B is worse in regard to one aspect of human goods, namely its distribution. “Worse for whom?” some theorists ask rhetorically. Perhaps they endorse a reading of the Person Affecting Restriction, which we could call the Personal Good Restriction, to the effect that an outcome cannot be worse than another, if it isn’t worse in regard to personal goods.⁷

The egalitarian concern above is grounded in a good that isn’t attached to one individual but to a relation between two individuals: What is bad about outcome B is that one person is worse off than another person. Consequently, this concern is ruled out by the Personal Good Restriction. Since B is not worse than A in respect to personal goods, B cannot be worse than A. In other words, if we find this restriction plausible, then we have another reason for rejecting Welfarist Egalitarianism.⁸

The Personal Good Restriction doesn’t imply, however, any value distinctions between uniquely and non-uniquely realisable people. It is compatible with such distinctions: one might decide, perhaps on purely intuitive grounds, that only personal goods belonging to non-uniquely realisable people count. It is, however, equally compatible with principles which don’t distinguish between uniquely and non-uniquely realisable people. Total Utilitarianism, for example, entails the Personal Good Restriction.

The next step to take is to stress the individualist aspect of value even more by claiming that morality is essentially *person comparative*: if an outcome is better (worse) than another, then it is better (worse) for at least one person. We shall formulate this view with a little bit more content.

The Person Affecting Restriction

- (a) If outcome A is better (worse) than B, then A is better (worse) than B for at least one individual.⁹
- (b) If outcome A is better (worse) than B for someone but worse (better) for no one, then A is better (worse) than B.¹⁰

This is the principle that I shall henceforth refer to as the Person Affecting Restriction (or “the restriction” for short). Strictly speaking, clause (a) above expresses what I take to be the restriction proper whereas clause (b) is a dominance condition. For expositional simplicity, I’ve added clause (b) to the statement of the restriction. It is quite an intuitively compelling condition in the present context and both the proponents and the opponents of the person affecting view in the literature seem to endorse such a dominance condition (I shall discuss a possible objection to it below).

It is not that the above statement of the restriction by itself captures what all authors have had in mind. As we shall see, however, most authors’ views can be captured by the above restriction in combination with some further specification of what it means for an outcome to be better or worse for an individual.

In cases involving only the same people in the compared outcomes, this view is roughly extensionally equivalent to the Personal Good Restriction. In cases

involving people whose existence is contingent on our choices, however, it becomes ambiguous. An outcome A is better than B for Peter if Peter has, for example, higher welfare in A as compared to B (we are assuming that if a person has higher welfare in one outcome as compared to another, then the former outcome is better for that person, other things being equal). But what if Peter exists in outcome A but not in outcome B? Is outcome A then better than outcome B *for Peter*? This is the crux of the matter. Depending on the answer to this question, different versions of Comparativism result.

14.4 Strict Comparativism

One possible answer to the question whether existence can be better or worse for a person is to claim that non-existence is neither better nor worse than existence for a person: non-existence and existence are either equally as good or incomparable in value for a person. This answer in combination with the restriction yields *Strict Comparativism*: we should completely disregard the welfare of *uniquely realisable* people, that is, people who only exist in one out of two compared outcomes.

This seems to be John Broome's take on the restriction. "Suppose [an alternative X] contains a certain number of people, and [an alternative Y] contains all the same people and some more as well . . . Then [the person-affecting view] is that [X] is at least as good as [Y] if and only if it is at least as good for the people who exist in both."¹¹

At times, this also appears to be David Heyd's view. He argues that the welfare of future possible people has "no direct moral significance and cannot be decided in ethical terms."¹² Furthermore, he holds that "the very comparison of the welfare of two possible children is based on the fallacious notion of an abstract, impersonal quantity of happiness in the world which should be maximized" and that "[e]xcluding the welfare and interest of future merely possible person . . . is a necessary consequence of a coherent person-regarding theory of value." He suggests that we can solve the problems in population axiology "by simply rejecting the logical legitimacy of comparisons between the welfare of a possible population A and a possible population B (when they consist of *different* people)."¹³

Strict Comparativism has such a counterintuitive conclusion that it is hard to believe that anyone seriously has endorsed it. Consider the *Bliss or Hell Case*:

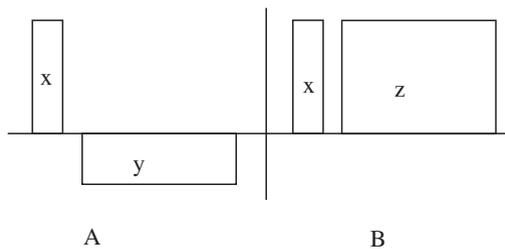


Fig. 14.1

Assume that we can either see to it that all the people in the future have hellish lives (the y-people in outcome A) or that they have excellent lives (the z-people in outcome B). Assume further that these two possible futures consist of different but the same number of people and that these two outcomes are equally good for us, the present x-people. Most of us, I presume, would consider outcome B clearly superior to outcome A and hold that we ought to realise B rather than A.

Since the y- and z-people are uniquely realisable people, outcome A and B are incomparable or equal in value for these people. Consequently, outcome A is neither better nor worse for the y- and z-people as compared to B. Moreover, the two outcomes are equally good for the x-people. Hence, according to the first clause of the Person Affecting Restriction, A cannot be better than B since it is not better for at least one individual (and vice versa). Consequently, Strict Comparativism ranks these outcomes as either equally good or as incomparable in value. But that is clearly the wrong answer to the Bliss or Hell Case.

14.5 Asymmetrical Comparativism

Strict Comparativism is a distraction. A perhaps more intuitively appealing answer to the question whether existence can be better or worse for a person is to claim that existence with positive welfare is equally as good as non-existence for a person, or incommensurable in value for a person, but a life with negative welfare is worse for a person than non-existence.

Melinda Roberts seems to have something like this in mind when she writes that “the choice to *bring* extra people into existence is . . . a morally neutral choice, so long as . . . the only people who are affected by that choice are the extra people themselves and those extra people themselves have *lives worth living*”¹⁴ whereas “the bringing into existence of a person whose life is likely to be anguished is one that can indeed be coherently said to harm the victim, with the precise measure of harm consisting of the difference between the zero level of wellbeing implied by nonexistence and the negative level likely to be suffered.”¹⁵

Asymmetrical Comparativism yields the right answer in the Bliss or Hell Case but it will fall foul for a version of it:

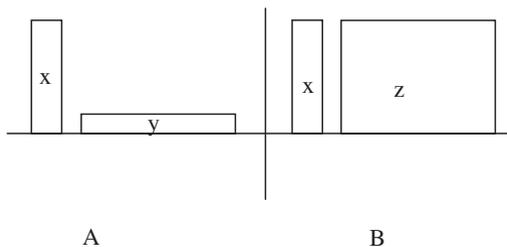


Fig. 14.2

In outcome A, the future y-people have lives barely worth living whereas the future z-people in B enjoy the same very high welfare as the x-people in B.¹⁶ Again, since the y- and z-people are uniquely realisable people, outcome A is neither better nor worse for the y- and z-people as compared to B. As before, the two outcomes are equally good for the x-people. Hence, according to the first clause of the Person Affecting Restriction, A cannot be better than B since it is not better for at least one individual (and vice versa). Consequently, although A involves inequality and is much worse than B in regards to people’s welfare, Asymmetrical Comparativism ranks these outcomes as either equally good or as incomparable in value. Again, we have reached the wrong answer.

In addition, it is not clear that it is possible to uphold this kind of asymmetry. The standard argument for why existence isn’t better for a person than non-existence is well expressed by John Broome:

[I]t cannot ever be *true* that it is better for a person that she lives than that she should never have lived at all. If it were better for a person that she lives than that she should never have lived at all, then if she had never lived at all, that would have been worse for her than if she had lived. But if she had never lived at all, there would have been no her for it to be worse for, so it could not have been worse for her.¹⁷

This argument works, however, equally well against the idea that existence could be worse for someone than non-existence. If it were worse for a person that she exists than that she should never have existed, then it would have been better for her if she had never existed. If she had never existed, then there would have been no her for it to be better for, so it could not have been better for her. Thus, it cannot be true that it could be worse for a person to exist than not to exist. In other words, it doesn’t look possible to uphold an asymmetry here by an appeal to Broome’s argument.¹⁸

14.6 Inconsistency

Let’s consider a last problem for Strict and Asymmetrical Comparativism which also will be of interest when we consider Soft Comparativism below. A common charge against the Person Affecting Restriction is that it yields non-transitive orderings. This is indeed true for Strict and Asymmetrical Comparativism, since taken as transitive population axiologies, they are inconsistent. Consider the following case:

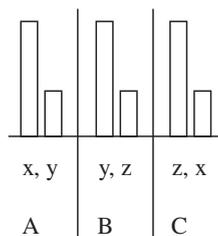


Fig. 14.3

The x- and y-people exist in outcome A, the y- and z-people exist in B, and the z- and x-people exist in C. Assume that all of these people have positive welfare, but that the y-people are better off in B as compared to A, the z-people are better off in C as compared to B, and the x-people are better off in A as compared to C.¹⁹

Since the x-people don't exist in B, B is neither worse nor better than A for them. Similarly, since the z-people don't exist in A, A is neither worse nor better than B for them. However, B is better than A for the y-people. Consequently, B is better than A according to the second clause of the Person Affecting Restriction. The same reasoning yields that C is better than B, and A is better than C. But if B is better than A, and C is better than B, then transitivity yields that C is better than A. Consequently, A is both better and worse than C, which cannot be true.

The earlier arguments against Strict and Asymmetrical Comparativism relied only on clause (a) in the Person Affecting Restriction. The above argument, however, involved the dominance condition in clause (b). Thus, a strict or asymmetrical comparativist might here reply that the above inconsistency only follows because we have attributed too strong a dominance condition to her. Rather, she might claim, a strict comparativist believes that if outcome A is better than B for some people, and *at least as good* for everybody else, then A is better than B, but if A and B are incomparable in value for someone, then A and B are also incomparable in value.

This is an untenable answer for several reasons. Firstly, this theory yields strongly counterintuitive conclusions that also are at odds with the person affecting intuition. Consider the following case. In outcome D, there are a vast number of people, the x-people, suffering horribly. In E, all the x-people have excellent lives and there is an extra person, also with an excellent life. With the revised dominance clause above, Strict and Asymmetrical Comparativism implies that D and E are incomparable in value since it is incomparable in value for the extra person in E, and although all the non-uniquely realisable x-people are much better off in E as compared to D. This is hardly believable.

In response to this objection, the strict comparativist might reply by further revising the dominance clause and claim that if A and B are incomparable in value for some sufficient number of people, then A and B are also incomparable. But this position leads to more problems than it explains, it seems to me.

Firstly, even if there were some great number of extra people with excellent lives in outcome E, it is still very counterintuitive, in my mind, that D and E are incomparable, and at odds with the person affecting intuition.

Secondly, this kind of comparativist owes us an explanation how it can be that the incomparability of outcomes is a function of the number of people for which the outcomes are incomparable in value. To the best of my knowledge, there is no such credible account.

Finally, this theory is a strange hybrid which I think falls between the chairs of Strict and Soft Comparativism. The motivating idea behind Strict Comparativism is that since the relative goodness of outcomes are determined only by how they are better or worse for individuals, and since outcomes cannot be better or worse for uniquely realisable people, we can ignore them when ranking outcomes. According

to the hybrid theory, however, we cannot ignore uniquely realisable people, since sometimes their mere existence can make outcomes incomparable. On the other hand, and in contrast with the soft comparativist, we can ignore the *welfare* of uniquely realisable people since it is irrelevant for the goodness of outcomes.

14.7 Personal Good Restriction Regained

The last answer to the question whether existence can be better or worse for a person that I shall consider is the claim that existence with positive welfare is better for a person than non-existence and that existence with negative welfare is worse for a person than non-existence.

As my discussion above indicates, I worry that such statements really don't make sense when the "better for" relation is understood in ordinary welfarist terms, that is, in terms of what is better or worse for the individual concerned rather than in terms of what we ought to do or what makes the world better or worse.

There is, however, a way of analyzing the "better for" relation that makes sense of this idea although it might not capture all aspects of the ordinary notion of "better for."²⁰ One could explicate "better for" in terms of what a benevolent impartial observer or guardian angel would choose for a person when she is only considering the wellbeing of the person under consideration (as all guardian angels do). According to this view, a state X is better for a person than state Y if and only if this is what a guardian angel would choose for her. If the guardian angel has a choice between bringing someone into existence with negative welfare and not bringing her into existence at all, she would of course choose the latter. Moreover, it doesn't seem unreasonable to say that if the guardian angel had the choice between bringing someone into existence with positive welfare and not bringing her into existence, she would choose the former. On this view, this is all that it means to say that non-existence is better or worse for a person.²¹

Consequently, given these understandings of the "better for" relation, one can claim that it is worse for a person to exist with negative welfare than not to exist at all, and that it is better for a person to exist with positive welfare than not to exist at all, without implying any absurdities.

In combination with the restriction, this answer to the question whether existence can be better or worse for a person doesn't have any of the disagreeable implications of the versions discussed above. It doesn't have any force, however, and doesn't imply any value distinctions between uniquely and non-uniquely realisable people. Actually, it is just a restatement of the Personal Good Restriction discussed earlier.²² Hence, it is compatible with, for example, Total Utilitarianism and other axiologies that respect Neutrality.

On a more positive note, the reasoning above shows how one can avoid the pejorative rhetoric of an "impersonal ethics," which is often applied to Total Utilitarianism and other views that respect Neutrality, since it is possible to couch one's principles in person affecting terms but still count everybody's welfare equally.

14.8 A Normative Version of the Restriction

A possible rejoinder to the non-transitivity problem for Comparativism presented in Fig. 14.3 is to argue that the restriction should be couched in normative rather than axiological terms, supplemented with the plausible claim that there is no analogue to the transitivity of “better than” for those normative concepts.²³

Here’s how we shall formulate a normative version of the restriction:

The Normative Person Affecting Restriction:

- (a) If it is impermissible to choose outcome A in a certain situation, then there is another outcome B in the situation such that A is worse than B for at least one individual.
- (b) If outcome A is better than B for someone but worse for no one, then in any situation involving a choice between A and B, it is wrong to choose B.

As with the axiological version of the restriction, clause (a) above expresses what I take to be the normative restriction proper whereas clause (b) is a weak dominance condition.

We can now combine this restriction with different answers to the question whether existence can be better or worse for a person to get different normative versions of Comparativism. Assume now that we combine it with the claim that non-existence and existence are equal or incomparable in value for a person to get a normative version of Strict Comparativism. As is easily seen, this view implies that none of the outcomes in the Bliss or Hell Case are impermissible because of clause (a) in the Normative Person Affecting Restriction. What will be the implication of this theory in regard to the case in Fig. 14.3?

Again, since the x-people don’t exist in B, B is neither worse nor better than A for them. Similarly, since the z-people don’t exist in A, A is neither worse nor better than B for them. However, B is better than A for the y-people. Consequently, A is an impermissible choice according to the second clause of the Normative Person Affecting Restriction. The same reasoning applied to outcome B and C yields that these outcomes are also impermissible choices. Hence, Normative Strict Comparativism implies that all the available alternatives are wrong and we are facing a moral dilemma: whatever act we perform we are going to act wrongly.²⁴

I’m sceptical about such moral dilemmas in general, but be that as it may, I think it is unlikely that the case described in Fig. 14.3 can credibly be described as a moral dilemma. It would be odd to claim that, if faced with the situation in Fig. 14.3, we would do something wrong if we choose, say, outcome A. The claim is odder still if we assume that everyone in Fig. 14.3 enjoys very high welfare in all outcomes and that the difference between the best and worst off is small (notice that this assumption wouldn’t change the structure of the case).

Moreover, we would get a moral dilemma even if the y-people are only slightly worse off in A, whereas the z- and x-people are much worse off in B and C, respectively. In such a case, it seems clear that it is at least permissible to choose A since

it involves much less harm and much more well-being than B and C. Yet, according to Strict Normative Comparativism, all the alternatives are wrong.

Lastly, consider a version of the case in Fig. 14.3 where there is another outcome F consisting only of people with very low positive welfare, the *w*-people. Again, Normative Strict Comparativism implies that alternatives A, B and C are wrong. In addition, the first clause of the restriction implies that it cannot be impermissible to choose outcome F, since there is no alternative outcome in which the *w*-people are worse off. Hence, assuming that there are no deontic gaps, Normative Strict Comparativism implies that alternative F is obligatory.

The above discussion points to a general problem for those who hope to evade non-transitive orderings by switching to a normative framework. As I have showed elsewhere, non-transitive value orderings most often translate into moral dilemmas on the normative level.²⁵

Melinda Roberts seems to apply this strategy when she tries to show that in cases where the axiological version of the restriction yields non-transitive orderings (which she calls the “inconsistency argument” against the Persons Affecting Restriction), a normative interpretation of the restriction will yield reasonable results.²⁶ She shows that this can be done in some versions of the Mere Addition Paradox. To that end, she formulates a theory called *Personalism* (“ $W(p, X)$ ” is the numerical representation of person *p*’s welfare in outcome X):

Personalism:

- P1: An alternative X is impermissible, or wrong, for agents if and only if some existing or future person is or will be wronged by agents at X.
- P2: A person *p* is not wronged at X if, for each Y that is an alternative to X, $W(p, X) \geq W(p, Y)$.
- P3: A person *p* is not and will not be wronged at X if *p* does not and will not exist at X.
- P4: A person *p* is wronged at X if
- (i) there is some Y such that $W(p, Y) > W(p, X)$,
 - (ii) there is no *q* who exists or will exist at Y such that $W(q, X) > W(q, Y)$
and
 - (iii) for each *q* who exists or will exist at Y, *q* exists at some time at X.²⁷

Roberts adds an implicit fifth principle to these four, namely that “[i]f no one is or will be wronged at a given alternative, then we know that the alternative itself represents a permissible choice.”²⁸

What implications does Personalism have for the inconsistency argument described in Fig. 14.3? Consider the *y*-people in A. Clause (i) and (ii) in P4 are satisfied since the *y*-people are better off in B and there is no person in B who has higher welfare in A. However, clause (iii) is not satisfied since the *z*-people only exist in B. Consequently, because of clause (iii), Personalism is silent about this case.

This result makes Roberts' solution to the inconsistency argument seem a bit unsatisfactory. A satisfactory solution to the inconsistency argument should show, I think, that in all cases where the axiological restriction yields non-transitive orderings, the normative version yields a somewhat reasonable evaluation. Moreover, the case described in Fig. 14.3 is the classical argument for the non-transitivity of the Person Affecting Restriction.

Notice also that Personalism is silent in cases where it seems clear what we should and shouldn't choose from a person affecting viewpoint. Assume that, in another situation, we have a choice between only outcome D and E:

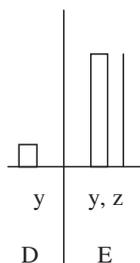


Fig. 14.4

Outcome D consists of the y-people with very low positive welfare whereas outcome E consists of the y-people and one z-person, all with very high welfare. Hence, the y-people are much better off in E as compared to D and, intuitively, the y-people would be very much wronged if we were to choose outcome D. Moreover, there is an extra person enjoying very high positive welfare in E. Clearly, the right result here is that D is impermissible and E is permissible. Personalism, however, neither implies that the y-people are wronged nor that D is impermissible. Again, clause (iii) is not satisfied, this time because the z-person exists in only one of the compared outcomes.

One might think that Personalism would at least imply that E is permissible. As it is stated, however, it doesn't. To know that E is permissible, we must know that no one is wronged in E. It follows from P2 that the y-people are not wronged in E. But it doesn't follow from P2 or any of the other principles that the z-people are not wronged in E. It doesn't follow from P2 since it is not true for any z-person z_i that $W(z_i, E) \geq W(z_i, D)$ because the z-people doesn't have a welfare level in D as they don't exist there.

In Roberts' discussion of other examples, it appears that she does assume something to the effect that one is not wronged by being brought into existence with positive welfare as long as the alternative is non-existence or existence with lower welfare.²⁹ One way to achieve this result is to claim, a bit nonsensically in my mind, that one has zero welfare in the outcomes in which one doesn't exist.³⁰ In another paper, Roberts seems to claim exactly this: "Nora does not have any properties at all at any alternative at which she does not exist, and, second, that, where Nora has no properties at all, all the properties she does have—the empty set—add up to a zero level of wellbeing."³¹ With this assumption, P2 will yield the desired result and

E will be permissible. Still, however, it doesn't follow that the y-people are wronged in D and this outcome is not classified as impermissible.

Since Roberts has only stated a sufficient condition for when people are wronged, she could respond by adding some condition to the effect that the y-people are wronged in D in Fig. 14.4 and in A in Fig. 14.3. Such a condition would also imply that the z-people are wronged in B in Fig. 14.3, and that the x-people are wronged in C (we would get the same result if we dropped clause (iii) in P4). It follows from P1 that all the available alternatives are wrong and we are again facing a moral dilemma.

Such a version of Personalism would also imply moral dilemmas in all the cases above in which Normative Strict Comparativism implied such dilemmas. Moreover, it would imply that outcome F is obligatory in the last case that we discussed in connection with Normative Strict Comparativism.

A better response is to add some condition to the effect that the y-people are not wronged in A in Fig. 14.3, and likewise for the z-people in B and the x-people in C. We would then get the result that outcome A, B, and C, are all permissible, which seems intuitively correct.³²

A rather intuitive way of achieving this result, I think, is to claim that the y-people are harmed in A although this is not sufficient to settle the matter whether the y-people are wronged or not and whether A is permissible or not, since other people might be harmed (e.g., the z-people in B) in the alternative outcomes. To settle which alternatives are permissible, we would then have to weigh together how much people are harmed in the different outcomes. Moreover, since it seems reasonable to claim that the y-people are as much harmed in A as the z-people in B and the x-people in C, such a view could yield that all the outcomes have the same deontic status and that they are all permissible. This view is similar to what I call Soft Comparativism and the Person Affecting Principle of Comparative Harm, which we shall discuss in the next section.

14.9 Soft Comparativism

According to the soft comparativist we should count the welfare of everybody but give more weight to the welfare of non-uniquely realisable people in some manner. Here's how I think we can develop a kind of person affecting version of this theory.³³ The central idea is that a person is harmed if she is worse off than she could have been, and it makes an outcome worse if people are harmed, or it might make it impermissible to choose the outcome. Another way to put this idea is to say that such people have a legitimate complaint or grievance and this makes the outcome worse or might make it impermissible. In addition to the wellbeing of everybody, this proposal takes the badness of legitimate complaints, or comparative harms, into account.

Consider the following case from Parfit:

Two Medical Programmes: If a pregnant woman has medical condition J, which a simple treatment could cure, this will cause the child she is carrying to have a certain handicap. If a

woman has condition K when she conceives a child, this child would get the same handicap. Condition K cannot be treated but disappears after two months. There are two medical programmes: Pregnancy Testing for J (PTJ) and Preconception testing for K (PCTK). In PTJ, women would be tested during pregnancy and those found to have condition J would be treated. It is predicted that if we implement PTJ, 1000 children who would otherwise have been handicapped will be born without the handicap. In PCTK, women would be tested when they intend to become pregnant, and those found to have K would be advised to postpone conception for at least two months. It is predicted that if we implement PCTK, 1000 children will be born without the handicap rather than 1000 (different) handicapped children. We only have funds for one of the medical programmes. Which one should we choose?³⁴

Since both programs would reduce the number of handicapped children by 1000, many would, like Parfit, consider these programs equally good. Likewise, this is the implication of Total Utilitarianism and other theories that respect Neutrality.

If we choose to implement PCTK, however, there will be 1000 children with a handicap that they wouldn't have had if we had chosen to implement PTJ instead—these people are non-uniquely realisable and will exist irrespective of our choice. These people can therefore be said to have been harmed and thus have a legitimate complaint. If we choose to implement PTJ, there will be 1000 children with a mild handicap but since these children owe their existence to our choice—they are all uniquely realisable people—they cannot be said to have been harmed or made worse off and thus don't have a legitimate complaint.

Consequently, although the effect on people's welfare is the same for both programs, PCTK is worse in one respect since it will cause people to be worse off than they could have been and thus there will be people who are harmed or can legitimately complain. Soft Comparativism captures this intuition and ranks PTJ as better than PCTK.

As is easily seen from our discussion above, Soft Comparativism doesn't follow from the Person Affecting Restriction as we have understood it here (it only follows that either only non-uniquely realisable people, or only non-uniquely realisable people and people with negative lifetime welfare count, or everybody counts and counts equally). I suggest that the interpretation of the person affecting idea that figures here is not a necessary condition for an outcome being better or worse but rather a condition for when people are harmed:

The Person Affecting Principle of Comparative Harm: If a person would be worse off in an outcome A as compared to an alternative outcome B, then she would be comparatively harmed if A rather than B came about.³⁵

So the person affecting version of Soft Comparativism is then the view that the value of an outcome, or the deontic status of the act yielding the outcome, is determined by both the value of the welfare in the outcome and the value of the comparative harm in the outcome. Henceforth, I shall just refer to this view as "Soft Comparativism."

As with the Person Affecting Restriction, we will get different versions of the comparative harm principle depending on which answer we give to the question whether existence can be better or worse for a person. If we claim that non-existence is neither better nor worse than existence for a person, then we get the result that

people are not harmed by being brought into existence. Alternatively, we could again introduce an asymmetry and claim that existence with negative lifetime welfare is worse than non-existence for a person whereas existence with positive welfare is not better than non-existence for a person.

This is an interesting idea but there are problems ahead. Consider again Fig. 14.3. All the outcomes in the figure are equally good in respect to the amounts of people's well-being. However, since the *y*-people are worse off in A as compared to B, the *y*-people would be harmed if A came about. In this respect, A is worse than B. Consequently, all things considered, A is worse than B. The same reasoning yields that B is worse than C, and C is worse than A. But if A is worse than B, and B is worse than C, then transitivity yields that A is worse than C. Again, it follows that A is both better and worse than C, which cannot be true.

There is, however, another way of explicating Soft Comparativism which doesn't imply non-transitive orderings. When determining the value of an outcome we should consider both people's well-being and whether they are harmed in the sense of being worse off than they could have been. The value of an outcome is determined by the value of the well-being in the outcome reduced by a factor that reflects whether people are harmed.³⁶

Here's an example of how this could be done. Assume that we represent well-being on a numerical scale and that the total wellbeing of the best-off people in Fig. 14.3 is 10 units and the total wellbeing of the worst-off people is 5 units. The value of outcome A would then be 15 minus some factor *h* that represent the fact that the *y*-people are worse off than they could have been. Intuitively, this factor should correspond to how much worse off the *y*-people are in A as compared to B.³⁷ Similarly, the value of outcome B and C would be 15 minus *h*. Consequently, on this view all the outcomes in Fig. 14.3 are ranked as equally good which seems to be the intuitively correct answer.

Since this version of Soft Comparativism assigns a specific numerical value to all outcomes, it cannot imply non-transitive orderings. Hence, we here have a person affecting theory that refutes the inconsistency argument.

Moreover, in regard to the two medical programmes, Soft Comparativism would pick PTJ since the two programs are equally good in regard to people's welfare but PCTK is worse in one respect since it will cause people to be worse off than they otherwise could have been.

Lastly, in regard to the Bliss or Hell case in Fig. 14.1, irrespective of whether we consider the *y*-people harmed or not in outcome A, Soft Comparativism will rank B as better than A since it also takes into account the well-being of everybody, which is clearly superior in outcome B.

14.10 Dominated Outcomes

Although Soft Comparativism neatly captures some people's intuitions regarding the two medical programs and avoids the threat of non-transitivity, it also has implications which one might consider counterintuitive. Consider the following three outcomes:

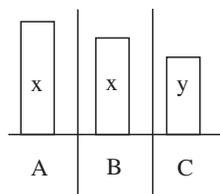


Fig. 14.5

Each of the three outcomes contains the same number of people. Everyone in A is better off than everyone in B, and everyone in B is better off than everyone in C. Again, the x -people would be harmed if outcome B came about since they would be worse off than they otherwise could have been, that is, if A had been the case instead of B. Let h represent the total value of the harm done to the x -people in B. Let d represent the total difference in well-being between the x -people in B and the y -people in C. The difference in value between outcome B and C will then be d minus h .³⁸ Consequently, if h is greater than d , then Comparativism will rank C as better than B although everyone in the latter outcome is better off than everyone in the former. This seems a bit counterintuitive.

Yet, this result might not perturb the comparativist. She might say that since she believes in the negative value of comparative harm, she's willing to trade off some welfare to avoid such comparative harm. She might point to other non-welfarist axiologies such as desert theories where the value of an outcome is determined both by the receipt of welfare and the fit between receipt and desert. On such theories, there might arise some cases where we have to forego some welfare to achieve a better fit between receipt and desert.³⁹ Likewise for other pluralist axiologies.

One might think that this shows that Soft Comparativism really isn't a person affecting theory as we have understood it here. Notice again that everyone in C is worse off as compared to everyone in B. According to the Person Affecting Restriction, if outcome C is better than B, then C should be better than B for at least one individual, which it isn't. So, rather than being a person affecting theory, Soft Comparativism is a pluralist axiology that sometimes will trade off people's well-being against the value of avoiding comparative harms.

As we already noted above, however, Soft Comparativism doesn't follow from the Person Affecting Restriction. Again, Soft Comparativism is based on an interpretation of the person affecting view as an idea regarding when people are harmed in different outcomes.

There are, however, further problems with this view if we turn to its normative implications. Assume some weak form of consequentialism, or consequentialist part of a deontological theory, to the effect that we ought to choose the best outcome in the cases currently under discussion.⁴⁰ Assume further that A is a very unlikely outcome. If we try to achieve A, we are most likely to fail and end up with C. If we aim at B or C, we will succeed. Still, Soft Comparativism tells us that it would be wrong to choose B, just because there is an unlikely outcome A in the choice set.⁴¹

One might think that this problem can be fixed by letting the harm-factor depend on the probability of the better alternative. Instead of letting h represent the total

value of the harm done to the x -people if we were to choose outcome B, it should be represented by ph , where p is the probability that A will be the case given that we choose A, and h is the harm done to the x -people were we to choose outcome B when we could have chosen A with certainty of success. The difference in value between outcome B and C will now be d minus ph so still, if ph is greater than d , which is clearly possible, then Comparativism will rank C as better than B although everyone is better off in B.

A better solution might be to let the harm-factor depend on people's expected welfare given a certain action which with certain probabilities brings about certain outcomes. On this view, people are harmed if their expected welfare is lower than it could have been given a different choice of action. Assume that if we choose an action a_A aiming at bringing about A, then the probability that A will be the case is 0.10 and the probability that C will be the case is 0.90, whereas if we choose action a_B , then B will be the case with certainty. Assume further that the total well-being in A is ten and in B eight. Then the x -people's expected welfare if we choose a_A is $0.10 \times 10 = 1$, whereas it is 1×8 were we to choose action a_B . Consequently, on this formulation of Soft Comparativism, the x -people are not harmed if we choose a_B since their expected welfare is higher if we choose that action rather than a_A . Hence, by switching to expected welfare and defining harm in terms of expected welfare, the problem of unlikely outcomes disappears.

On the other hand, if we went for a_A and A actually came about, then we would still have harmed the x -people and done the wrong thing since their expected welfare (at the time of the choice) was lower than it would have been had we chosen a_B , although the x -people are better off since A rather than B actually came about. This might strike some as implausible. But a possible rejoinder is to claim that it was wrong to choose a_A since that action exposed the x -people to a risk of getting nothing.⁴²

This problem, however, is not peculiar to Soft Comparativism but analogous to the old dispute among consequentialists regarding whether one should adopt a formulation of consequentialism in terms of the actual or probable outcomes of actions, so I shall not discuss it further here.⁴³ It is noteworthy, however, that Soft Comparativism seems more compatible with a probabilistic rather than an actualistic formulation of consequentialism.

Let me end this section with two other objections to Soft Comparativism.⁴⁴ In Fig. 14.3, this theory correctly ranked all the outcomes as equally good. One might object, however, that we cannot know this without knowing that exactly these three outcomes are the only ones available in the situation since, according to Comparativism, the value of an outcome depends on the set of possible outcomes in the situation. Suppose, for example, that there was another outcome D with only the x -people at level 15. This would not only yield that D was the best outcome in the situation but also change the ranking of A, B, and C, since the x -people in C will be more harmed than the y -people in A and the z -people in B. Hence, C will be ranked as worse than A and B.

The first objection is that it is absurd that one and the same outcome can both be worse than and equally as good as another outcome. This seems to be the case

here since, when D is not present in the set of outcomes, C is ranked as equally as good as A and B whereas, when D is present, C is ranked as worse than A and B. Hence, it looks like the same outcome, C, is both worse than and equally as good as A and B.

This would surely be absurd but the obvious rejoinder is to deny that these outcomes are the same outcomes. We can just partly individuate outcomes by the situation to which they belong. Hence, if we add another outcome to the situation described in Fig. 14.3, then we have a new situations with, say, alternatives A', B', C' and D and it is B' which is better than C' which doesn't contradict that outcome B and C in the original situation are equally good.

The second objection is that in practice we could never be epistemically justified in limiting the number of possible outcomes as we have done in the examples above. Hence, since the Soft Comparativist ranking depends on the possible outcomes in the situation, we cannot be justified in believing in the ranking.

It is true that this makes Soft Comparativism a bit special as an axiology since most axiologies, such as the axiological component of classical utilitarianism, yield context-insensitive rankings of outcomes. However, this problem appears for these theories on the normative level, since which outcome is the best one, and thus which one we ought to choose, depends on which other outcomes are available in the situation. Hence, this alleged particular problem with Soft Comparativism reduces to the old problem of whether consequentialist theories ought to be and can be action-guiding and is thus no special problem for Soft Comparativism. The same standard responses come in handy here. For example, we could make a sharp distinction between criterion theories and decision methods and claim that Soft Comparativism is a criterion theory that has no claim to be used as a decision method other than indirectly in the choice of which decision methods that we should use in everyday life.⁴⁵

14.11 Future Populations and Trade Offs

So far, so good. The main problem with Soft Comparativism is, however, that it will not be of much help in regard to other problems in population axiology. A difficulty shared by all versions of Soft Comparativism is that all the problems afflicting neutral theories will reappear in the specification of the method for aggregating people's welfare. Summing implies Parfit's Repugnant Conclusion; averaging implies that it can be better to create miserable rather than happy people, and so forth.⁴⁶

For instance, assume, as we did above, that the value of an outcome is determined by the value of the total well-being in the outcome reduced by a factor that reflects whether people are harmed. In all cases involving only uniquely realisable people, Soft Comparativism determines the ranking by the total sum of people's welfare since such cases don't involve any harm. Consequently, like Total Utilitarianism, it will imply the Repugnant Conclusion in respect to future populations where there is no overlap of individuals in the compared populations.

This might be an acceptable implication for some comparativists, however, since what matters to them is that people are not worse off than they otherwise could have been. Nevertheless, Soft Comparativism will imply similar conclusions even in cases that involve overlaps and thus involve great losses for non-uniquely realisable people (e.g., present people) and a lot of comparative harm.

Here's a numerical illustration of this point, involving a micro-version of the Repugnant Conclusion:

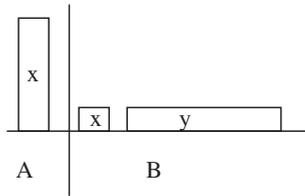


Fig. 14.6

Assume that we have a choice between outcome A with ten persons, the x -people, and outcome B with the x -people and an additional two hundred persons, the y -people. In outcome A, the x -people have very high lifetime welfare. Assume that this high welfare corresponds to a ten units of welfare. Consequently, the value of outcome A is $10 \times 10 = 100$.

In outcome B, the x - and y -people have very low positive lifetime welfare. Assume that this very low welfare corresponds to one unit of welfare. Since the x -people have much lower welfare in B, they are harmed in B. Assume that the harm factor for each x -person corresponds to her difference in welfare between outcome A and B. Thus, the value of the harm to the x -people in B is $10 \times (-9) = -90$ whereas the value of their welfare is $10 \times 1 = 10$. Taken all together, the value of the x -people in outcome B is $10 - 90 = -80$.

However, since there are also two hundred y -persons in outcome B, the total value of outcome B, according to Soft Comparativism, is $200 - 80 = 120$ which is greater than the value of A. Hence, Soft Comparativism here opts for the outcome with very low positive welfare, outcome B, although it involves a great loss in welfare and a lot of comparative harm for the x -people, and thus implies the Repugnant Conclusion. In this respect, Soft Comparativism doesn't constitute much of an advance towards a satisfactory population axiology as compared to standard welfarist axiologies such as Total Utilitarianism.

One might think that this result depends on the weight given to comparative harm. However, as long as the harm factor is represented by a finite number (that is, as long as we don't give lexical priority to comparative harm), Soft Comparativism will have the above implication. Here's a general demonstration:

Assume that h is a positive finite number that represents the weight given to the comparative harm of an individual due to the fact that she is worse off than she could have been. Let A consist of n non-uniquely realisable persons with very high welfare u_1 . Let B consist of a mixed population of n uniquely and m non-uniquely realisable people with very low positive welfare u_2 . The value of A is thus nu_1 and

the value of B is $n(u_2 - h) + mu_2$. Now, for any value of h , there is an m such that $nu_1 < n(u_2 - h) + mu_2$, that is, a value of m that makes B better than A, namely $m > n(u_1 - u_2 + h)/u_2$.

If we give lexical priority to avoiding comparative harm, then the Repugnant Conclusion will be avoided. Given such an extreme negative weight on comparative harms, however, Soft Comparativism will yield counterintuitive results in a version of the case described in Fig. 14.2, and will face an extreme version of the case described in Fig. 14.5.

Recall that in Fig. 14.2 the y-people in A have lives barely worth living and the z-people in B enjoy the same very high welfare as the x-people in A and B. Assume further that one of the x-persons has slightly lower welfare in B as compared to A. If one gives lexical priority to avoiding comparative harms, then A is better than B according to Soft Comparativism, since no one is harmed in A but one of the x-people is harmed in B.⁴⁷

Likewise, consider a version of the case depicted in Fig. 14.5 in which the y-people in C have lives barely worth living and only one of the x-people has slightly lower welfare in outcome B as compared to outcome A. Again, if we give lexical priority to avoiding comparative harms, C is better than B.

Soft Comparativism might be caught in a dilemma here. Only if it gives lexical negative weight to comparative harm will it avoid the Repugnant Conclusion in cases that involve great losses for non-uniquely realisable people and thus a lot of comparative harm. However, with such a negative weight on comparative harms, it will imply clearly unacceptable versions of the cases described in Figs. 14.2 and 14.5. It seems hard to find an acceptable way out of this dilemma.

14.11 Summary

The Person Affecting Restriction has a strong intuitive appeal. It is often contrasted with Total Utilitarianism and other “impersonal” theories and it has been hoped that the restriction would help us avoid the disagreeable implications of impersonal theories in population ethics. I’ve argued that there are some versions of the idea that are quite acceptable. On the other hand, these versions are compatible with Total Utilitarianism and other neutral theories that count everybody’s welfare and count it equally. Hence, these theories cannot help us solve the problems in population ethics but can help us to escape the pejorative rhetoric of an “impersonal ethics” since they can be used to formulate our principles in person affecting terms but still count everybody’s welfare equally.

The main part of the paper considered Comparativism, the view that we should take the well-being of non-uniquely realisable people into account in a different way as compared to the well-being of uniquely realisable people. As I showed, different versions of Comparativism can be derived from one version of the Person Affecting Restriction, depending on how we answer the question whether existence can be better or worse for a person.

With one exception, all of these theories have very counterintuitive implications. The exception was a combination of the Person Affecting Restriction with the claim that existence with positive welfare is better for a person than non-existence and that existence with negative welfare is worse for a person than non-existence. Again, we obtained a theory which is compatible with Total Utilitarianism and other neutral theories.

Lastly, I considered Soft Comparativism according to which we should count the welfare of everybody but give more weight to the welfare of non-uniquely realisable people in some manner. Soft Comparativism cannot be derived from the Person Affecting Restriction but, as I suggested, it can be partly based on an interpretation of the person affecting idea as a condition for when people are harmed, namely the Principle of Comparative Harm. According to that principle, if a person would be worse off in an outcome A as compared to an alternative outcome B, then she would be comparatively harmed if A rather than B came about.

As we showed, Soft Comparativism has a number of attractive properties. It avoids the inconsistency charge that has been levelled against the Person Affecting Restriction. Moreover, it yields the right answer in a number of test cases, including cases where Total Utilitarianism seems to give the wrong answers, such as the Two Medical Programmes. We also showed that a number of possible arguments against Soft Comparativism could be rebutted. Nevertheless, as we showed in the last section, Soft Comparativism runs into problems with the Repugnant Conclusion and cases that involves tradeoffs between the welfare of non-uniquely and uniquely realisable people.

Acknowledgements I would like to thank Julian Savulescu, Torbjörn Tännsjö, Dan Wikler, Andrew Williams, and especially John Broome, Krister Bykvist, Björn Eriksson, and Melinda Roberts, for fruitful discussions and comments. An earlier version of this paper was presented at the Högre Seminariet, Department of Philosophy, Stockholm University, October 2007. I would like to thank the participants in that seminar for their stimulating criticism. Thanks also to CERSES, CNRS, for being such a generous host during some of the time when this paper were written. Financial support from the Bank of Sweden Tercentenary Foundation, STINT, and the Swedish Collegium for Advanced Study is gratefully acknowledged. Some of the ideas discussed in this paper also appear in Arrhenius (2000a, 2003b, 2008).

Notes

1. Temkin (1993a, b) claims that this restriction, which he dubs “the Slogan,” is presupposed in many arguments in moral philosophy, political theory, and welfare economics. The term “Person Affecting Restriction,” introduced by Glover (1977), p. 66 (but see also Narveson (1967)), might be misleading since many theorists would, sensibly I think, lessen the restriction to also include other sentient beings. Cf. Holtug (1996). Below, I shall only discuss applications of the restriction to human populations. Consequently, whenever I claim that a certain interpretation of the restriction is reasonable, this claim only holds for human populations.
2. See Parfit (1984), p. 388. For an overview of these counterintuitive implications, see Arrhenius et al. (2006c) and Arrhenius (2000a).
3. See Arrhenius (2000a, 2000b, 2006b), Broome (1992, 2004), Parfit (1984), Temkin (1993a,b).
4. I’ve taken the term “uniquely realisable person” from Bykvist (1998) albeit his usage is slightly different from mine. There are a number of related but different views that appear in the

literature (albeit seldom made explicit) that often are conflated with Comparativism and the Person Affecting Restriction. Comparativism should be distinguished from *Presentism* which draws a distinction between presently existing people and non-existing people; *Necessitarianism* which distinguishes between people who exist or will exist irrespective of how we act and people whose existence is contingent on our choices; and *Actualism* which differentiates people who have existed, exist or are going to exist in the actual world, on the one hand, and people who haven't, don't, and won't exist, on the other. These distinctions don't amount to the same thing but there are relations among them. A presently existing person is also a necessary and actual person but not the other way around since necessary and actual people may be located in the past and the future. A necessary person is also an actual person but a future actual person may be contingent on our choice. A uniquely realisable person is also a contingent person, but a contingent person is not necessarily uniquely realisable in respect to all pairs of outcomes in a choice situation since she can exist, for instance, in two out of three outcomes. I discuss these distinctions at length in Arrhenius (2000a, 2006b).

5. Perhaps it is this restriction that is at stake in Moore's criticism of Sidgwick at the turn of the century. It can be seen as a denial of Moore's idea in *Principia Ethica* that an unpopulated beautiful world is intrinsically better than an unpopulated ugly world, and a reaffirmation of Sidgwick's view that all moral goods must be of "Human Existence." See Moore (1903), Section 50, and Sidgwick (1907), Bk. I, Chapter IX, Section 4.
6. Cf., however, Note 1 above.
7. I have taken the term "personal good" from Broome (1991), Chapter 8. The Personal Good Restriction is not, however, equivalent to his principle of personal good.
8. See Arrhenius (2009) for a discussion of Welfarist Egalitarianism and population ethics. Broome (1991), pp. 180–81, suggests a way of understanding the goodness of equality that turns it into a personal good.
9. An interesting question is whether the restriction should be supplemented with a necessary condition for outcomes being equally good. I would suggest the following condition: If outcome A is equally as good as B, then A and B are equally as good for at least one individual, or A is better (worse) and B is worse (better) for at least one individual. I'm grateful to Melinda Roberts for spotting an error in an earlier formulation of this principle.
10. I'm here making the trivial assumption that if A is worse (better) than B for no one, then B is better (worse) than A for no one. Cf. Note 22.
11. Broome (1992), p. 124. Broome rejects the Person Affecting Restriction.
12. Heyd (1988), p. 157. See also Heyd (1992), pp. 124–25. Cf. Narveson (1967), p. 67 (emphasis in original): "If you ask, 'whose happiness has been increased as a result of his being born?', the answer is that nobody's has. —Remember that the question we must ask about *him* is not whether he is happy but whether he is happier as a result of being born. And if put this way, we see that again we have a piece of nonsense on our hands if we suppose the answer is either 'yes' or 'no.' For if it is, then with whom, or with what, are we comparing his new state of bliss? Is the child, perhaps, happier than he used to be before he was born? Or happier than his alter ego? Obviously, there can be no sensible answer here."
13. Heyd (1988), pp. 159–61 (emphasis in original). Heyd says that his view is "grounded in an 'anthropocentric' conception of value according to which value is necessarily related to human interests, welfare, expectations, desires and wishes—that is to say to human volitions" (p. 164). How this "volitional concept of value" is supposed to generate the conclusions quoted in the text is not clear to me and, as I pointed out earlier, the success of such deductions seems unlikely. Heyd also argues against Asymmetry by claiming that it "is inconsistent with a person-affecting theory as it presupposes the comparability of non-existence with life of a certain quality."
14. Roberts (2003a), p. 1 (emphasis in original except for the last one). Notice that Roberts is here only talking about mere additions of extra people.
15. Roberts (2003b), p. 179. I don't think, however, that Roberts is an Asymmetrical Strict Comparativist but rather, as I shall explain below, a Soft Comparativist.

16. This case is basically a version of the Depletion case that Parfit (1984), pp. 361–62, uses as an argument against the Person Affecting Restriction.
17. Broome (1999), Chapter 10, p. 168 (emphasis in original). See also Parfit (1984), pp. 395 and 489, and Heyd (1988).
18. For an effort to defuse Broome's argument, see Roberts (2003b). I shall say a bit more about Roberts' argument below.
19. A similar example is used by Temkin (1987), pp. 168–69, to illustrate the intransitivity of the Person Affecting Restriction.
20. For some sceptical arguments to this effect, see Bykvist (2007).
21. I first suggested this in Arrhenius (2003b). I think that Wlodek Rabinowicz suggested something like this in personal communication already back in 2000 and he has put his take on it in print in Rabinowicz (forthcoming). For a recent incisive critique of the idea, see Bykvist (2007). Another possibility has been suggested by Holtug (1996). Again, the reason why some theorists have been inclined to deny that it can be better or worse for a person to exist than not to exist is that they believe that this position implies that it can be better or worse for a "person" not to exist than to exist and they find this implication nonsensical. One can deny this implication by revising the logic of "better for" and hold that it can be better or worse for a person to exist than not to exist, but it cannot be better or worse for a "person" not to exist than to exist. It is clear that a state X is better than a state Y if and only if state Y is worse than state X (I would say that this is a conceptual truth). Holtug denies that this logic also holds for "better for," that is, a state X is better for a person than another state Y if and only if state Y is worse for the person than state X , the reason being that "better for" is only applicable when a person to which the "for" in "better for" refers to exists. I consider Holtug's radical revision of the logic of "better for" an all too high price to pay, especially in the light of my suggestion above which yields the same evaluations without any such revision.
22. With one qualification: it involves a weak dominance condition which we didn't include in the description of the Personal Good Restriction.
23. I discuss this latter claim at length in Arrhenius (2005).
24. Following Vallentyne (1988), we could call a dilemma of the above mentioned type a "prohibition dilemma." There are also "obligation dilemmas," that is, situations where more than one action is obligatory. Since the normative version of Asymmetrical Comparativism is extensionally equivalent with Normative Strict Comparativism in cases that only involve lives with positive welfare, it also implies that Fig. 14.3 depicts a moral dilemma.
25. Arrhenius (2005).
26. Roberts (1998, 2003a).
27. Roberts (2003a), pp. 10–13. See also Roberts (1998), pp. 62–65.
28. Roberts (2003a), p. 10. See also Roberts (1998), p. 64. The word "wronged" has two meanings in English which are relevant in the present context (see e.g., Oxford English Dictionary). One meaning is more or less synonymous to "harmed" or "affected harmfully," another to "unjustly treated." I take it that Roberts is using the term in the second meaning, or something close to that.
29. For example, Roberts (2003a), p. 11, where she discusses whether r is wronged in C .
30. It seems trivially true that a person doesn't have any welfare level at all in a state in which she doesn't exist (well-being presupposes, so to say, being). A better solution, in my mind, is to revise P2 along the following lines:
 A person p is not wronged at X if
 - (i) for each Y that is an alternative to X , $W(p, X) \geq W(p, Y)$; or
 - (ii) $W(p, X) \geq 0$ and for each Y that is an alternative to X and in which p exists, $W(p, X) \geq W(p, Y)$.
31. Roberts (2003b), p. 169. Moreover, Roberts (1998), p. 64, writes that "I am thus supposing that it is at least possible that s has more well-being in a world in which s does not exist than s actually has. Suppose s 's existence in X is unavoidably *less* than one worth living . . . and

- that s has, in any world in which s does not exist, a zero level of well-being. Under these conditions, s 's level of well-being at zero is actually *greater* than s 's well-being in X . . . Thus, [P2] avoids the implication that the person whose life is less than one worth living has *not* been wronged" (emphasis in original). On the other hand, she also claims that "[t]here is no need, ever, on my account of either wrongful life or the non-identity problem to assign a value, even a *value of zero*, to nonexistence for an individual who never exists in the world that is subject to appraisal." (Roberts (1998, pp. 174–75) (emphasis in original).
32. In personal communication, Roberts endorses the same view.
 33. I haven't seen the person affecting version of Soft Comparativism explicitly stated in the literature but it seems to me that ideas somewhat along these lines are put forward in Hope (2003), Meyer (2003), and Roberts (1998, 2003a, 2007), and by Andrew Williams in private communication on the lawn of Worcester College summer 2003. At least I've been inspired by these sources to develop the above theory. Julian Savulescu has also proposed such a view in private communication. In Savulescu et al. (2006), p. 163, it is, however, stated that "[a]ccording to a person affecting view of harm, a person is harmed by an act if she is made worse off than she would otherwise have been if that act had not been performed." This sounds more like the counterfactual analysis of harm which is plagued by a number of well-known problems. For a discussion of these in connection with population ethics, see Roberts (2007).
 34. Parfit (1984), p. 367. I've changed the wording of the example.
 35. Notice that the notion of "comparative harm" above is a technical notion that doesn't completely map onto our everyday use of "harm." For example, if you will enjoy an excellent life in both outcome A and B but you are slightly less happy in B, then you are comparatively harmed if B came about, but many would hesitate to say that you are harmed in the ordinary language sense of "harm" (there are many other examples). I could have used some other term to capture the idea that it makes an outcome worse if people are worse off than they otherwise could have been, but I think the technical notion of "comparative harm" is sufficiently related to the ordinary notion of harm to justify its name. An analogy would be the difference between the notion of the "consequence of an action" in ordinary language and in the formulation of consequentialism (usually the whole possible world that would be the case if the action were performed). For brevity, I will in the following sometimes use the term "harm" and its cognates although I always have in mind "comparative harm" in the above sense.
 36. Alternatively, we could represent the value of an outcome with an ordered pair (w, h) in which w represents the value of the well-being in the outcome and h represents the (negative) value of the harm in the outcome. Such a representation would leave open the possibility that comparative harm has lexical priority over total welfare in the ranking of outcomes. As I shall discuss below, such a view is not very attractive.
 37. Below, I'll assume that the harm factor is determined relative the optimal outcome for the people in question. There are of course other alternatives but I shall not pursue this question further here.
 38. Assume that the total well-being in C is m and in B thus $m+d$. The difference in value between B and C is then $(m + d - h) - m = d - h$.
 39. See Feldman (1997) and Arrhenius (2003a, 2006a, 2007).
 40. The kind of consequentialism I have in mind is what we could call *Ceteris Paribus Act-Consequentialism*: other things being equal, an action is right (obligatory) if and only if its outcome is at least as good as (better than) that of every alternative. An action is wrong if and only if it is not right. In other words, if a choice situation doesn't involve actions that are right or wrong by virtue of a certain deontic constraint or option, then the normative status of the actions are determined by the value of their respective outcomes. Most deontologists accept this form of consequentialism. For a discussion, see Arrhenius (2005).
 41. Strictly speaking, Soft Comparativism in conjunction with consequentialism has this implication. For the sake of brevity, I'll omit this qualification below. Moreover, I'm here assuming that probable outcomes are relevant for determining comparative harm.
 42. Another interesting possibility, suggested to me by an anonymous referee on another paper, is to claim that people are comparatively harmed if and only if they are worse off than they could

have been both in terms of expected and actual welfare. According to this view, if we choose a_A and A actually came about, then the x-people are not harmed since their actual welfare is maximised. A problem for such a theory, however, is that its normative prescriptions are a bit unclear. Does it direct us to choose a_A or a_B ?

43. See, for example Carlson (1995) and Feldman (2006).
44. I'm indebted to John Broome for pressing these two points.
45. See Bales (1971) for an excellent treatment of this issue. See also Brink (1986), pp. 421–27, Danielsson (1974), pp. 28–29, Kymlicka (1990), p. 29.
46. For the Repugnant Conclusion, see Parfit (1984), Chapter 17. Cf. Arrhenius (2000a), Chapter 3.
47. One could also construct examples in which the y-people have hellish lives, like the case described in Fig. 14.1. However, such counterexamples could be avoided by revising Soft Comparativism such that it counts as being comparatively harmed if you are born into a life not worth living and there is an alternative in which you're not brought into existence.

References

- Arrhenius, G. 2000a. *Future generations: A challenge for moral theory*. Uppsala: University Printers (F.D. Diss.).
- Arrhenius, G. 2000b. An impossibility theorem for welfarist axiologies. *Economics & Philosophy* 16: 247–266.
- Arrhenius, G. 2003a. Feldman's desert-adjusted utilitarianism and population ethics. *Utilitas* 15(2): 225–236.
- Arrhenius, G. 2003b. The person affecting restriction, comparativism, and the moral status of potential people. *Ethical Perspectives* 10 (3–4): 185–195.
- Arrhenius, G. 2005. The paradoxes of future generations and normative theory. In *The repugnant conclusion*, eds. J. Ryberg and T. Tännsjö, 201–18. Dordrecht: Kluwer Academic Publisher.
- Arrhenius, G. 2006a. Desert as fit: An axiomatic analysis. In *The good, the right, life and death: Essays in honor of Fred Feldman*, eds. R. Feldman, K. McDaniel, J. R. Raibley and M. J. Zimmerman. Aldershot: Ashgate Publishing.
- Arrhenius, G. 2006b. The moral status of potential people. Department of Philosophy, Stockholm University (<http://people.su.se/~guarr/>) (unpublished manuscript).
- Arrhenius, G., Ryberg, J., Tännsjö, T. 2006c. The Repugnant conclusion. *The Stanford Encyclopedia of Philosophy (Spring 2006 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/spr2006/entries/repugnant-conclusion/>>.
- Arrhenius, G. 2007. Meritarian axiologies and distributive justice. In *Hommage à Wlodek: Philosophical papers dedicated to Wlodek Rabinowicz*, eds. T. Rønnow-Rasmussen, B. Petersson, J. Josefsson and D. Egonsson, 1–43 (www.fil.lu.se/hommageawlodek).
- Arrhenius, G. 2008. Life extension versus replacement. *Journal of Applied Philosophy* 25(3): 211–227.
- Arrhenius, G. 2009. Egalitarianism and population change. In *Theories of intergenerational justice*, eds. A. Gosseries and L. Meyer. Oxford: Oxford University Press.
- Bales, R. E. 1971. Act-utilitarianism: Account of right-making characteristic or decision-making procedure. *American Philosophical Quarterly* 8(3): 257–265.
- Brink, D. 1986. Utilitarian morality and the personal point of view. *Journal of Philosophy* 83(8): 417–438.
- Broome, J. 1991. *Weighing goods: Equality, uncertainty, and time*. Oxford: Basil Blackwell.
- Broome, J. 1992. *Counting the costs of global warming*. Cambridge: White Horse.
- Broome, J. 1999. *Ethics out of economics*. Cambridge: Cambridge University Press.
- Broome, J. 2004. *Weighing Lives*. Oxford: Oxford University Press.
- Bykvist, K. 1998. *Changing preferences: A study in preferentialism*. Uppsala University (F.D. Diss.).

- Bykvist, K. 2007. The benefits of coming into existence. *Philosophical Studies* 135: 335–362.
- Carlson, E. 1995. *Consequentialism reconsidered*. Dordrecht: Kluwer Academic Publisher.
- Danielsson, S. 1974. *Two papers on rationality and group preferences*, Filosofiska Studier no. 21. Uppsala: Filosofiska föreningen och filosofiska institutionen.
- Feldman, F. 1997. *Utilitarianism, hedonism, and desert: Essays in moral philosophy*. Cambridge: Cambridge University Press.
- Feldman, F. 2006. Actual utility, the objection from impracticality, and the move to expected utility. *Philosophical Studies* 129: 49–79.
- Glover, J. 1977. *Causing death and saving lives*. Harmondsworth: Penguin Books.
- Heyd, D. 1988. Procreation and value: Can ethics deal with futurity problems. *Philosophia* (Israel) 18: 151–170.
- Heyd, D. 1992. *Genethics: Moral issues in the creation of people*. Berkeley: University of California Press.
- Holtug, N. 1996. In defence of the slogan. In *Preference and value: Preferentialism in ethics*, ed. W. Rabinowicz, 64–89. Lund: Lund University (Department of Philosophy).
- Hope, T. 2003. Physicians' duties and the non-identity problem, University of Oxford (unpublished manuscript).
- Kymlicka, W. 1990. *Contemporary political philosophy*. Oxford: Clarendon Press.
- Meyer, L. 2003. Intergenerational Justice. *Stanford encyclopaedia of philosophy*.
- Moore, G. E. 1966. *Principia ethica*. Cambridge: Cambridge University Press.
- Narveson, J. 1967. Utilitarianism and new generations." *Mind* 76: 62–72.
- Parfit, D. 1991. *Reasons and persons*. Oxford: Clarendon Press.
- Rabinowicz, W. Broome and the intuition of neutrality. *Philosophical Issues* (forthcoming).
- Roberts, M. A. 1998. *Child versus childmaker: Future persons and present duties in ethics and the law*. Lanham, Maryland: Rowman and Littlefield.
- Roberts, M. A. 2003a. Can it ever have been better never to have existed at all? Person-based consequentialism and a new repugnant conclusion. *Journal of Applied Philosophy* 20(2): 159–185.
- Roberts, M. A. 2003b. Is the person-affecting intuition paradoxical? *Theory and Decision* 55: 1–44.
- Roberts, M. A. 2007. The non-identity fallacy: Harm, probability and another look at Parfit's depletion example. *Utilitas* 19(3): 267–311.
- Savulescu, J., Hemsley, M., Newson, A., Foddy, B. 2006. Behavioural genetics: Why eugenic selection is preferable to enhancement. *Journal of Applied Philosophy* 23(2): 157–171.
- Sidgwick, H. 1967. *The methods of ethics* (7th ed.). London: Macmillan.
- Temkin, L. S. 1987. Intransitivity and the mere addition paradox. *Philosophy and Public Affairs* 16(2): 138–187.
- Temkin, L. S. 1993a. *Inequality*. Oxford: Oxford University Press.
- Temkin, L. S. 1993b. Harmful goods, harmless bads. In *Value, welfare and morality*, eds. R. G. Frey and C. W. Morris, 291–324. Cambridge: Cambridge University Press.
- Vallentyne, P. 1988. Gimmicky representations of moral theories. *Metaphilosophy* 19(34): 253–263.

Part VII
What are the Implications
of the Nonidentity Problem for Law
and Public Policy?

Chapter 15

Implications of the Nonidentity Problem for State Regulation of Reproductive Liberty

Philip G. Peters

Abstract Many proposed laws regulating the use of reproductive technology will not protect identifiable children from harm but will instead reduce future suffering by changing the identities of the children who are born. Because the “injuries by substitution” caused by these uses of technology are not person-affecting—they have no individual victims—the state’s interest in preventing them might be considered less powerful than its interest in preventing person-affecting harms. This paper argues, however, that the framework of U.S. constitutional democracy does not require that we think of the state’s interest in preventing impersonal injuries by substitution as less than compelling. Accordingly, the fact that such laws (including laws against incest) fail to protect offspring against person-affecting harm is not sufficient to establish that they represent unconstitutional interferences with procreative liberty. Naked “moral” objections to the exercise of a fundamental right will not give the Supreme Court ground to uphold a law, but the fact that the law is closely tied to the prevention of injuries by substitution may and should be enough to do the trick.

Keywords Fourteenth amendment · Compelling state interest · Nonidentity problem · Procreative liberty · Tort law · Injury by substitution.

Laws that infringe on fundamental rights, like the right to reproduce, are subject to “strict scrutiny” in the United States. Those laws are struck down unless they advance a “compelling state interest.” Although the protection of children from credible threats of serious harm would ordinarily be deemed a compelling interest, the identity problem associated with existence-inducing behavior complicates the analysis considerably. It will force the courts to decide whether a threatened harm must be person-affecting in order to be compelling.

Surprisingly, the Supreme Court has never articulated a clear set of criteria by which to identify a “compelling” state interest. However, the Court has made statements in *Roe v. Wade*¹ and in later cases which cast doubt on the sufficiency of moral

P.G. Peters (✉)
University of Missouri, Columbia, MO, USA
e-mail: petersp@missouri.edu

objections that are not accompanied by evidence of tangible harm to others. This limitation is also arguably implicit in the Court's decision in *Lawrence v. Texas*,² which struck down a state prohibition on consenting homosexual sex. As a result, laws that regulate reproductive conduct whose harmfulness is not person-affecting, like laws requiring the screening of sperm donors, may be vulnerable to challenge unless the Court assigns significantly greater weight to the state's consequential interest in maximizing the welfare of future children (whoever they may turn out to be) than it has given to naked moral objections. That question has yet to be answered.

15.1 Fundamental Liberties

Within the framework of American constitutional democracy, the Supreme Court has no authority to strike down laws that interfere with procreative liberty simply because it believes that the laws are unwise. In theory at least, it can only strike down laws that are prohibited by the US Constitution. Under current doctrine, most laws that interfere with individual liberty, like laws setting maximum highway speeds and statutes requiring the wearing of motorcycle helmets, are constitutionally permissible as long as the legislature has a rational basis for enacting them.³ Only rarely does the Court conclude that a law has not rational basis. Usually that occurs when the legislature's ostensible objective simply will not be advanced by the law, suggesting some unstated invidious purpose.

A select group of especially prized individual liberties, however, have been deemed so important that the state may not infringe upon them without a *compelling* justification.⁴ Many of these especially important liberties are expressly identified in the first eight amendments to the Constitution, commonly known as the Bill of Rights. Two examples are freedom of speech and freedom of religion. However, the Supreme Court has also held that there are some fundamental rights that are *not* expressly mentioned in the Constitution. These unenumerated fundamental rights include the rights to procreate, to travel, to terminate a pregnancy, to marry, and to rear children. Laws infringing upon these unenumerated fundamental rights must also serve a "compelling state interest." Relatively few laws survive this "strict scrutiny."⁵

In theory, this framework minimizes the instances in which the Court must confront the difficult task of balancing the state's interests against an individual's fundamental rights.⁶ Instead, it identifies those zones of social activity which are largely beyond the range of community regulation and insulates them from all but the most urgent state interference. Our remaining liberties are subject to limits imposed through the democratic process, where the elected legislature is largely free to use its own judgment.

Laws that interfere with a person's ability to procreate *naturally* are subject to the more demanding level of scrutiny. In *Skinner v Oklahoma*,⁷ the Court overturned an Oklahoma statute calling for the sterilization of certain repeat criminal because it unjustifiably violated the criminals' fundamental right to reproduce. However, the

Supreme Court has never been asked whether individuals have a fundamental right not only to keep and use their natural reproductive capacity, but also to use some or all of the many forms of reproductive technology that are now available. Legal scholars offer conflicting predictions about the Court's likely decision. In this paper, I will assume that the Court will eventually extend fundamental rights protections to at least the most routine forms of assisted reproduction, such as artificial insemination by donor semen and *in vitro* fertilization. As a result, any laws that materially limit access to these protected procreative activities will be struck down unless the state has a compelling interest in restricting access to them.

Despite the vulnerability of laws that seriously interfere with fundamental liberties, some laws intended to ban the use of unsafe reproductive technologies are likely to survive judicial scrutiny. When the legislature's restrictions protect identifiable future children from serious harm, the government will be deemed to have a compelling state interest. However, many proposed limitations on the use of reproductive technology would not protect identifiable children from harm. Instead, they would reduce future suffering by changing the identity of the children who are born. The practices restricted by these laws do not have individual victims. Because the "harms" caused by these practices are not person-affecting, the state's interest in preventing these harms is arguably less powerful than its interest in preventing person-affecting harms. As a result, the propriety of these laws is less certain.

15.2 Types of Harm to Future Children

Only a fraction of the risks posed by reproductive technology raise the non-identify problem. Some of the harms that can be caused by the use of risky reproductive practices are person-affecting in the traditional sense. Assume, for example, that the improper storage of an IVF embryo by an infertility clinic results in the birth of a child who has disabilities caused by the earlier carelessness. This kind of carelessness causes conventional, person-affecting harm to a class of identifiable victims. Accordingly, a state law reasonably designed to prevent clinic carelessness is likely to survive judicial challenge, despite the extra cost and inconvenience that the law imposes on infertile clinic patrons.

In a second set of cases, lawmakers may reasonably believe that a specific reproductive technique leads to the birth of children who have catastrophic disabilities. If the injuries are so serious that never existing at all would be better, then a credible argument can be made that the harm caused by the technique is person-affecting.⁸ If so, then laws limiting the use of that technique will also survive scrutiny.

I want to focus on a third set of cases. These are the cases in which a safety-motivated law is intended to protect future children from less catastrophic injuries by changing the identity of the children who are born. A law that bars the use of cloning technology by couples who are able to conceive naturally would fall into this category as would laws that require the screening of sperm or egg donors. These laws aim to reduce human suffering by changing the identity of the children who

are actually born. While this reduction in suffering improves the welfare of the next generation of children, it does not reduce the harm suffered by an identifiable person. As a result, the harm it prevents is quite different than the kinds of harm that courts have, thus far, ruled sufficient to justify restrictions on fundamental liberties.

15.3 Compelling State Interests

The Supreme Court has never articulated a clear criterion or set of criteria by which to distinguish “compelling” state interests from less powerful state interests.⁹ Indeed, the very idea that fundamental rights can be trumped when the state has a compelling justification has no roots in the constitutional text.⁷ Nonetheless, the cases provide two important clues for predicting the fate of laws regulating reproductive technology.

First, the state has a compelling interest when it acts to avoid death and, perhaps, also the infliction of serious physical and emotional harm. The abortion decisions provide the clearest authority for this conclusion. Having held in *Roe v. Wade*¹⁰ that the state has a cognizable interest in the welfare of unborn fetuses once they are viable outside of the womb, the Court has concluded that protection of these fetuses from termination is a sufficiently compelling state interest to justify the a prohibition on those abortions not needed to protect the life or health of the mother. Thus, the courts are likely to uphold a Congressional ban on existence-inducing conduct that causes persons-affecting harm, such as a ban on the implantation of more than four in vitro embryos during a single cycle or a ban on genetic-engineering techniques that cause serious birth defects.

Second, naked moral objections to the exercise of a fundamental right are not ordinarily sufficient to justify laws that seriously burden that right. Although social norms play a role in the identification of fundamental rights, the Court typically insists on a more tangible impact once a right is deemed to be fundamental.¹¹ The importance of being a fundamental right, after all, is its insulation from restrictions that the majority may want to impose. The best illustration is provided by *Roe v. Wade*.¹² In that case, the Court held that a fetus is not a person within the meaning of the Constitution and that the state’s desire to protect the potential life of a pre-viability fetus was not compelling despite the strong moral objections that lay behind the prohibition. This ruling suggests that moral objections to norm-threatening reproductive technologies, like cloning, will not justify legal prohibition if access to the treatment is determined to be a protected component of reproductive liberty.

The insufficiency of naked moral objections to abortion in *Roe* contrasts dramatically with the deference shown to community norms when the liberty interest at stake is not fundamental. Consider, for example, the Court’s treatment of this issue in the first of its two decisions dealing with prohibitions on consensual gay sex, *Bowers v. Hardwick*.¹³ After the Court concluded that private homosexual sex was not a fundamental right, it applied the rational basis test. Writing for the Court,

Justice White rejected the argument that “majority sentiments about the morality of homosexuality should be declared inadequate.”¹⁴ Instead, he stated that “[t]he law . . . is constantly based on notions of morality, and if all laws representing essentially moral choices are to be invalidated under the Due Process Clause, the courts will be very busy indeed.” As a result, the Court did not insist upon evidence that the prohibited conduct actually harmed anyone.

Seventeen years later, however, in *Lawrence v Texas*,¹⁵ the Court reversed itself on the issue of fundamental rights status for private gay sex. As a result, it struck down a Texas anti-sodomy law similar to the Georgia law which survived challenge in *Bowers v. Hardwick*.¹⁶ Thus, fundamental liberty interests appear to be governed by an approach that closely resembles J.S. Mill’s “harm principle.” Mill argued that “the only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others.”¹⁷ Like Mill, the Court is not so much distinguishing between harmless and harmful conduct as between conduct that causes physical injury or economic harm, on the one hand, and conduct which only causes offense, on the other. This distinction serves the anti-majoritarian function of fundamental rights doctrine while at the same time marking out with reasonably clarity one circumstance in which the majority must tolerate unpopular behavior. When a person exercises a fundamental right in a manner not reasonably expected to cause harm to others, then the state lacks the power to prohibit that conduct, regardless of the revulsion experienced by those who disapprove of it.

Other language used by the Court supports this interpretation. In the case reaffirming *Roe, Planned Parenthood v. Casey*,¹⁸ the Court said “Some of us as individuals find abortion offensive to our most basic principles of morality, but that cannot control our decision. Our obligation is to define the liberty of all, not to mandate our own moral code.” That language was quoted approvingly by the Court in the *Lawrence* case. In addition, the Court counseled against state attempts “to set boundaries absent injury to a person or abuse of an institution the law protects.”¹⁹ In his opinion for the Court, Justice Kennedy quoted language from the dissenting opinion of Justice Stevens’ in *Bowers* stating that “the fact that the governing majority in a State has traditionally viewed a particular practice as immoral is not a sufficient reason for upholding a law prohibiting the practice.”²⁰ In response to this language, the three dissenters (Scalia, Rehnquist, and Thomas) warned that the refusal to treat moral disapproval as a legitimate basis for restricting liberty would produce “massive disruption of the current social order.”²¹ They warned that the courts would be “very busy indeed” striking down laws against bigamy, same-sex marriage, adult incest, masturbation, adultery, public nudity, and many other offenses.²² As a result, several respected legal scholars have given the *Lawrence* decision a Millian interpretation.²³

In the summer of 2007, however, the Court cast some doubt on this interpretation. In *Gonzalez v. Carhart*,²⁴ the Court upheld a law that prohibited “partial birth abortions” at any time during a pregnancy. Although the basis for this decision is not entirely clear, Justice Kennedy seems to argue that Congress’ abhorrence of the barbaric nature of this procedure provided a sufficient reason to ban it, even

if doing so actually saved no fetal lives because abortion providers could still use another procedure. The significance of this language is muddled, however, by his hints that nascent fetal life is now entitled to more weight than it had been given in the past. In addition, Kennedy argues that the burden on reproductive choice posed by a prohibition is modest because it bans only one of two late-term abortion methods. Thus, this specific law may not have needed a compelling justification. Given these ambiguities, the Court's commitment to the harm principle in cases dealing with fundamental rights is less certain than it once seemed.

15.4 Two Different Distinctions

Even if we assume that the Court's endorsement of the harm principle has survived *Gonzalez*, that endorsement does not provide a clear barometer of the Court's likely treatment of cases that lack person-affecting harm. That is because the distinction between person-affecting and non-person-affecting actions marks a different intellectual fault-line than the distinction between actions that cause serious tangible harm to others and those that merely offend their values or tastes.

In constitutional analysis, the harm principle prohibits the enactment of laws that seriously burden the exercise of a fundamental right when the only harm caused by the offending conduct is the revulsion it produces in the eyes of those who think it immoral. Thus, it distinguishes between conduct that causes serious physical injury or economic harm from conduct which only causes offense to others. This approach provides a safe harbor for the exercise of those individual choices which we have placed, to some extent, outside the sphere of majority governance. Inside that safe harbor, pluralism and tolerance trump popular sovereignty. For these purposes, the distinction between causing offense and causing more palpable harms is both useful and convincing.

The distinction between person-affecting and non-person-affecting actions marks a very different boundary and serves a very different purpose. Thus, the Supreme Court's frequent endorsement of an approach that seems to resemble the harm principle provides a poor barometer for predicting whether the Court will also endorse the idea that cognizable harms must be person-affecting.

Rather the marking the boundary between concrete harm and mere offense, the boundary between person-affecting and non-person-affecting harms defines one of the boundaries between right-based and utilitarian philosophies. Deontological or right-based philosophers believe that morally superior decision-making results from a system that derives moral obligations from a framework of rights and obligations. This approach stands in contrast to that of utilitarians, who evaluate the merits of an act or rule by comparing its good consequences with its bad ones in order to identify choices that maximize the net common good. Rights-based philosophers see serious problems in a utilitarian approach, such as its apparent willingness to sacrifice important individual freedoms and ignore important human commitments if doing so will advance overall community welfare.

An important strain of deontological thought draws a boundary between itself and utilitarian approaches by insisting that any act claimed to be wrongful by virtue of its harmful effects have the effect of making a specific individual worse off than she otherwise would have been. The fact that an act merely reduces overall social utility is not sufficient for it to be treated as harmful if no individual is made worse off.

For these philosophers, existence-inducing acts pose a genuine dilemma. On the one hand, intuitively callous choices seem to have the potential to increase the amount of suffering in the world. On the other hand, these seemingly harmful choices sometimes cause the birth of a different child than the child who would have been born if a more virtuous path had been chosen. For the child who is born with a disability, the life that she lives is the only life she could ever have had. If her parents had chosen the safer route to reproduction, then she would not have been born at all. Instead, a different child would have lived (or no child at all). As a result, the child who was born using the dangerous reproductive technology has only been harmed in the conventional sense if her life is so catastrophic that it would have been better never to exist. Both philosophers and legal scholars have been troubled by the highly counterintuitive implications of this conclusion. Yet, many continue to insist that only person-affecting harms can be wrongful.

From a legal perspective, the unanswered question is whether courts will be similarly skeptical of state laws prohibiting conduct that does not make any specific individual worse off. In the rest of this paper, I will argue that courts should make a difference choice. Unlike laws that ban victimless crimes, laws that are reasonably calculated to protect future generations from the dangers posed by unsafe reproductive technologies will prevent unnecessary human suffering. Under proper circumstances, these laws should be upheld despite the burden that they impose on important individual liberties.

15.5 Future Children, Whoever They May Be

Several years ago, Italian authorities shut down a Florentine sperm bank that was selling the sperm of a man infected with hepatitis C and genital herpes.²⁵ Because his diseases were transmissible through his sperm, the clinic was jeopardizing the health of both the women who were inseminated with his sperm and the children conceived as a result. Few people would dispute that the Italian authorities did what they should have done. Most of us would probably agree with the decision to stop the distribution of this donor's sperm even if his sperm posed a risk only to the children so conceived. Under conventional notions of person-affecting harm,²⁶ however, this safety precaution cannot be defended.

This surprising conclusion is dictated by the fact that no individual child was made worse off by failure to screen out this donor. Better screening would not have enabled the infected children to be born without their infections. Instead, they would not have been born at all. The use of sperm from other donors would have resulted in

the birth of different children. As long as their lives were better than never existing at all, no harm was done to them by the clinic's failure to screen the infected donor out of the pool.

This conclusion offends our intuitions. Better screening prevents needless suffering. Responsible clinics already know this. They screen for both infectious conditions and a history of genetic illness. Yet, an exclusively person-affecting notion of harm cannot account for this. To remedy this oversight, we need to rethink our understanding of avoidable harm. Harms like those that necessitate donor screening are not avoidable in the conventional sense. Instead, they are avoidable only by having a different child than the child who would otherwise be born—by substituting one child for another.²⁷

Choices of this kind occur throughout the world of reproductive technology. Consider, for example, the fertility clinic which learns that its ovulation-inducing drugs have produced a dozen receptive follicles in one of its patients. As a consequence, the odds of a dangerous multiple-pregnancy are very high. Discontinuing therapy until the next cycle would protect the welfare of this patient's future children, even though it would change the identity of the resulting children. Similarly, high-risk couples who are deciding whether or not to do embryo screening are making choices about the minimization of future suffering that will, at the same time, change the identity of the child who is born. Parents choosing between two fertility treatments which pose different risks to the resulting children are making the same choice. In each instance, the decision-maker faces a choice between a riskier procedure and a safer one. Our intuitions tell us that parents and clinics alike have a *prima facie* moral duty to make reasonable efforts to give life to child who will suffer less. Thus, they have a presumptive, but rebuttable, obligation to use the safest procreative method reasonably available. Doing so will prevent injuries that are *avoidable by substitution* (of one child for another). When the participants choose the risky route (as providers might in order to maximize profits or couples might in order to maximize the odds of conception), they threaten the welfare of future children in a way that requires justification. In this important sense, the obligation to use the safest option reasonably available does "affect" persons—it obliges us to save future generations from unnecessary suffering.

The potential harmfulness of these choices lies not in the absolute magnitude of the threatened injuries (which typically fall well short of the level required to make life itself a harmful event), but in the decision to take a risky route to reproduction when a safer route is available. Although no individual child can claim to have been personally harmed, the class of children conceived as a result of these choices will suffer more than they need to have suffered. A satisfying moral theory must account for this intuition.

The English moral philosopher Derek Parfit famously illustrated the shortcomings of traditional person-affecting ethical theories by examining the behavior of a woman who is advised by her doctor not to become pregnant until she gets over a temporary illness that causes birth defects.²⁸ Although she could wait two months for the condition to pass, she ignores his advice and conceives a child who suffers the deformity. Intuitively, her choice was a harmful one. Yet, if she had waited,

another child would have been born. Parfit refers to the puzzle posed by this change in identity as the “non-identity problem.”²⁹ If the mother had waited, a different child would have been. The child born with the congenital injuries would not have existed. The analysis would be entirely different, however, if the thoughtless mother had possessed the power to prevent the birth defects by taking a medication during her pregnancy.³⁰ In that event, her failure to take the medication would unquestionably constitute harmful conduct because any child born with birth defects could have been born without them. These vastly different conclusions simply do not pass a moral gut test. Both mothers had an equal opportunity to avoid causing human misery. Both failed to do so. Parfit calls this conclusion the No Difference View.³¹

Although the No Difference View is hard to reconcile with a person-affecting ethical theory, it is fully consistent with a utilitarian approach because it will maximize the welfare of the resulting children. Parfit articulated a principle that he called “Q” to capture this insight.³²

If in either of two outcomes the same number of people would ever live, it would be bad if those who live are worse off or have a lower quality of life than those who would have lived.³³

Translated into simpler language, Q exhorts providers and would-be parents to have the child who will suffer least.

Indeed, it is remarkable that so obvious a proposition has not already become a part of the fabric of public thinking about reproductive decision-making. Although several ethicists have endorsed this idea or something close to it,³⁴ no court, legislator, ethics commission, scientific body, or administrative agency has ever explicitly relied upon this conception of harmful conduct to defend the regulation of reproductive technology. It is time for them to do so.

One barrier to wider endorsement may be the conceptual problems that are posed when the substitution that takes place is not a simple substitution of one child for another (where utility comparisons are relatively simple), but a choice between a single healthy child and several children who are less healthy. Imagine, for example, that a fertility clinic is deciding whether to proceed with the insemination of a woman who has been pharmaceutically induced to release many eggs at once. Assume further that inseminating her with sperm during this cycle will increase the risks of multiple pregnancy, premature birth, and serious medical problems. Would it be better to wait for the next cycle? Answering that question requires a choice between a world in which one healthy child is born and one in which several, less healthy children are born. The first world has a higher average utility, but the second may have a higher total utility.

Both average utility and total utility can have unappealing implications under some conditions. A preference for *total utility*, for example, seems to force us “to prefer a huge, wretched population to a smaller, happy one, as long as the quantity of the huge population allows its total utility to exceed the total utility of the smaller population.”³⁵ Parfit called this “the Repugnant Conclusion.”³⁶ In the context of reproductive choices, a total utility approach would favor multiple pregnancies over single ones, even if multiple pregnancies were strongly associated

with serious injuries. The same intuition that urges us to choose the safer form of assisted reproductive compels us to reject the Repugnant Conclusion.

We can escape the Repugnant Conclusion, while still supporting the No Difference View, by using *average utility* as the benchmark, rather than total utility.³⁷ Having the single healthy child would be the preferred choice using this approach. Regrettably, average utility also has shortcomings. For example, average utility seems to favor a population policy that will result in 10 *exquisitely* happy people over a policy resulting in ten million *very* happy people. Average utility also seems to condemn the addition of one more “happy” person to a “very happy” community if the new person’s well-being is slightly lower than the others.³⁸ Parfit called this the Mere Addition Paradox.³⁹

Parfit was ultimately unwilling to choose between these two measures of utility. He was, therefore, unable to extend his preference for having the happier child in same number cases (his “Q”) into a broader principle (which he called “X”) that would provide guidance for choices between courses of conduct that result in different numbers of lives. Nevertheless, he remained convinced that some such principle must exist.⁴⁰

The underlying difficulty with the choice between average and total utility is that each illuminates a morally relevant factor. Average utility emphasizes quality of life, while total utility emphasizes quantity.⁴¹ Because both indicia can be relevant in some circumstances, an acceptable theory of moral obligations must be able to take both into consideration.⁴² The possibility of combining the two considerations into a single formula has been most elegantly explored by Thomas Hurka.⁴³ Writing in the context of population ethics, Hurka persuasively argues that total utility declines in importance relative to average utility as populations increase. Thus, the value that an additional individual contributes to the world is not constant, but varies with the number of other humans alive. While the value of additional people would be enormous following a crisis that shrank the population dramatically, as occurred in the biblical story of Noah or as might occur following a nuclear holocaust, its value diminishes greatly when the human population reaches its current size.⁴⁴

A formula which captures this variability avoids many of the unattractive aspects that plague both the average and total utility principles when standing alone. By giving more weight to average utility when population levels are high, this approach avoids the Repugnant Conclusion except in the rare circumstances when increased numbers are genuinely more important than quality of life. As a result, it constitutes an attractive improvement over exclusive reliance on either total or average utility.

In doing this composite calculation, greater weight should be assigned to average utility than to total utility. Maintaining and improving the quality of life of people who actually exist seems more crucial than maximizing the number of persons who live, at least under current social and population conditions. As Jan Narveson notes, “we are in favor of making people happy, but neutral about making happy people . . . it seems repulsive to think that the goodness of a community is a function of its size.”⁴⁵ Similarly, John Rawls has concluded that average utility is the only version of utility obtainable using his theory of justice.⁴⁶ As a result, average utility should be weighted more heavily in our hybrid formula.

Obviously, we will rarely have the kind of quantitative data necessary to perform this calculation with real numbers. However, the theoretical exercise is important both because it explains why safer outcomes matter and because it provides a heuristic device with which to frame our own less quantitative and more visceral evaluation of risky reproductive practices. It instructs lawmakers who are considering restrictions on a particular reproductive technology to evaluate the likely impact of that technology on both the average and total utility of the resulting children, giving average utility greater weight.

This way of conceptualizing harmful conduct explains better than any alternative model why we have such a strong negative reaction to the conduct of the Italian sperm bank and of the impatient mother. It is why many legislatures have enacted laws that require the screening of sperm and egg donors. Although there are no individual victims, future children will suffer as a class if we use technologies that do not maximize their well-being. Dan Brock correctly observes that these reproductive choices are genuinely “person affecting” in the sense that they inflict unnecessary human suffering.⁴⁷ John Robertson attempts to minimize this insight by describing it as a “norm against offending persons who are troubled by gratuitous suffering.”⁴⁸ Yet, that description is hardly adequate. Rather than being a norm about *offending sensibilities*, it is a norm against inflicting *gratuitous suffering*.⁴⁹

15.6 Will the Courts Agree?

The courts have yet to confront the constitutional issue directly. While tort claims involving the identity problem have received an unfriendly reception in tort cases, this response is a poor predictor of the status that will be given to non-person-affecting harms in constitutional jurisprudence. Tort law serves a very different function than constitutional law and its compensatory role justifies a skepticism toward these claims that is not necessary in the constitutional setting.

Two tort cases have raised the issue. One involved a claim against a physician who allegedly failed to disclose the risk of fetal injury from Accutane use during pregnancy.⁵⁰ If complete disclosure would have prompted the mother to forego her use of the drug, then the failure to disclose those risks was person-affecting. Her child would have been born without birth defects. If, however, the mother would have delayed conception until her drug regimen was completed, then the failure to disclose its risks was not person-affecting. In that scenario, the injured child would not have been born and another child would have been conceived in her place at a later time. The court sidestepped this issue by assuming that full disclosure would have prompted the patient to avoid use of the drug. Under that assumption, the child who was born with Accutane-associated injuries could have been born uninjured. As a result, she was entitled to compensatory damages for the harm caused by failure to disclose.

In the other case, the court insisted that the plaintiff prove that the defendant’s negligence was person-affecting. It framed the issue this way:

Given that Mr. Gerber cannot recovery on a theory that, had Roche provided certain precautions, he would not have been born, Mr. Gerber can only argue that Shirley Gerber would not have taken Accutane in the first place if Roche's warning had been adequate. In order to prove causation under the circumstances presented in this case, Plaintiff must demonstrate that an alternative warning would have changed the physician's decision to prescribe Accutane or would have altered Shirley Gerber's decision to take the drug for her severe acne condition.⁵¹

The court's insistence upon person-affecting harms was dictated by the function of tort law. That function is to compensate individuals who have been made worse off by the tortious misconduct of others. Conduct that does not make any individual worse off than she would otherwise have been falls outside of that domain. As a result, the inhospitable treatment given to claims that involve failure to avoid injury by substituting one child for another can be attributed to the limited function of tort liability. Its very function is to provide redress for person-affecting misconduct. Consequently, the chilly reception given to cases which lack any person-affecting harm does not imply a judicial determination that the prevention of non-person-affecting injuries is a trivial or legally impermissible state objective.

To the contrary, legislatures often enact laws that are intended, at least in part, to protect future generations from harms that may not turn out to be person-affecting. Environmental protection laws fall into this category. So do recent European laws that encourage child-bearing in order to reduce the burden that future working families will bear to support the anticipated cadre of senior citizens. Population control laws, like those in China, qualify as well. These laws are often motivated, at least in part, by the assumption that today's citizens have fiduciary obligations to future generations *whoever they may be*. One cannot participate in a discussion of the federal debt, social security financing, or global warming without hearing pleas on behalf of future generations. Insofar as accomplishing this objective in an ethically satisfying manner requires formal recognition of the insights provided by utility analysis, neither legislatures nor courts are likely to suffer the qualms that beset rights-based philosophers.

Even though no specific child can claim to have been made worse off by such a failure, the *class* of children born using unnecessarily dangerous reproductive practices will suffer avoidable injuries. As a class, they are indisputably worse off. Although harm to the class does not give any specific child a claim for compensatory damages, that harm does support public health regulation. Because the goal of public health regulation is to maximize community health and welfare, it can and should take into account the unnecessary harms that risky reproductive choices can inflict on future children. The reluctance to give substantial weight to "mere offense" when laws restrict core liberties by no means compels a similar unwillingness to reject the insights of utilitarian insights about the welfare of future generations. This is one setting in which utilitarian analysis seems to provide more insights than right-based thinking. It facility for revealing the stakes of both long-term and large-scale policy questions and for comparing the implications of alternative policy choices in these settings give it a legitimate place at the legislative table.

One important clue suggests that the Supreme Court will agree. Both courts and scholars universally assume that state laws barring incestuous marriage are constitutional—even though the right to marry is fundamental. One important reason for this belief is that incest prohibitions prevent the birth defects associated with incestuous partnerships.⁵² Yet, these birth defects are not person-affecting in the conventional sense (unless the birth defect is so serious that never existing at all would be better). If the offending man and woman had respected the incest taboo, then they would have reproduced with different partners and given birth to genetically different children. Doing so would have improved the overall welfare of the children in the community, but only by substituting one set of children for another, not by permitting the children who were born in the incestuous relationship to be born without injury. The incest taboo, therefore, prevents injuries by substitution.

Nevertheless, we widely agree about the propriety of taking these safety benefits into account. The incest laws, therefore, provide a powerful illustration of the moral and legal weight to which the state interest in avoiding injuries by substitution is likely to be given. If the injuries to be avoided are common and serious, then avoiding them will (and should) constitute a compelling state interest.

Notes

1. 410 U.S. 113 (1973).
2. 539 U.S. 558 (2003).
3. *United States v. Carolene Products Co.*, 304 U.S. 144, 152 n.4 (1938) (holding that the judiciary will defer to the legislature unless there is either infringement of a fundamental right or discrimination against a “discrete and insular” minority).
4. *Zablocki v. Redhail*, 434 U.S. 374, 389-90 (1978); see also *Carey v. Population Services International*, 431 U.S. 678, 686 (1977) (“regulations imposing a burden on it may be justified only by compelling state interests, and must be narrowly drawn to express only those interests”).
5. Gunther (1972) and Chemerinsky (1997), p. 417.
6. Pildes (1994), pp. 711–15.
7. *Skinner v. Oklahoma*, 316 U.S. 535 (1942).
8. Peters (2004), pp. 45–60.
9. E.g., Aleinikoff (1987), p. 977, Gottlieb (1988), pp. 932–37, and Fallon (1993), pp. 348–50; Sheppard (1994), pp. 983–85 and Bhagwat (1997), pp. 308 and n. 34.
10. *Roe v. Wade*, 410 U.S. 113,163 (1973).
11. Sunstein (2002), pp. 990–93.
12. 410 U.S. 113 (1973).
13. 478 U.S. 186, 196 (1986), overruled by *Lawrence v. Texas*, 539 U.S. 558 (2003).
14. 478 U.S. at 196.
15. 539 U.S. 558 (2003).
16. 478 U.S. 186, 196 (1986).
17. Mill (1985), p. 80.
18. 505 U.S.833, 850 (1992).
19. 539 U.S. at 567.
20. 539 U.S.at 577.
21. 539 U.S. at 591(Scalia, J., dissenting).
22. 539 U.S. at 590, 605 (Scalia, J., dissenting).
23. E.g., Barnett (2003), Epstein (2004), and Secunda (2005).

24. 127 S.Ct. 1610 (2007).
25. St. Louis Post-Dispatch (1997); *see also Johnson v. Superior Ct.*, 80 Cal. App. 4th 1050 (2000) (trial court found that sperm bank had failed to disclose that sperm came from a man known to have a family history of an autosomal dominant polycystic kidney disease).
26. Bayles (1976), p. 297, Parfit (1976), p. 101 and Feinberg (1986), pp. 168–69.
27. Peters (2004) and Brock (1995b).
28. Parfit (1976, 1984).
29. Parfit (1984), pp. 351–56.
30. Parfit (1976), p. 103.
31. Parfit (1984), pp. 367–68.
32. Parfit (1984), p. 360.
33. Parfit (1984), p. 360.
34. Brock (1995b) p. 271 and Peters (1989).
35. Laudor (1994), p. 1679.
36. Parfit (1984), p. 388.
37. Parfit (1984), p. 401.
38. Laudor (1994), p. 1684.
39. Parfit (1984), pp. 419–20.
40. Parfit (1984), pp. 380–441.
41. Parfit (1984), p. 401.
42. Parfit (1984), p. 405. Parfit considered combining the two. He hypothesized a combination of the two factors in which quality of life would count down to a threshold amount and then be discounted entirely. Because this threshold approach was quite blunt, it produced results that are vulnerable to criticism. However, the use of a sliding scale or weighted formula should eliminate the counterintuitive implications of more blunt combinations.
43. Hurka (1983).
44. Hurka (1983), p. 497.
45. Narveson (1976), p. 73.
46. Rawls (1971), pp. 163–66. However, Rawls ultimately rejects average utility as a general moral theory in favor of his own theory of justice. Rawls (1971), pp. 167–75.
47. Brock (1995b), p. 273.
48. Robertson (1997), p. 76.
49. Brock (1995a), pp. 203–04.
50. *Hogle v. Hall*, 916 P.2d 814, 816–17 (Nev. 1996).
51. *Gerber v. Hoffmann-La Roche Inc.*, 392 F. Supp. 2d 907, 920, 921 (S.D. Tex. 2005); *see also Walker v. Mart*, 790 P.2d 735, 740 (Ariz. 1990) (“If her parents had decided to conceive, despite knowledge of probable congenital defects, the law would recognize no action on Christy’s behalf against them.”).
52. E.g., Associated Press (2002).

References

- Aleinikoff, T.A. 1987. Constitutional law in the age of balancing. *Yale Law Journal* 96: 943.
- Associated Press. 2002. Study downgrades risk of birth defects in children of couples who are cousins. Thur. April 4, 2002, *St. Louis Post-Dispatch*, p. A11 (suggesting that close cousins face an additional risk of birth defects of 1.7–2.8 percent, a smaller difference than expected).
- Barnett, R.E. 2003. Justice Kennedy’s libertarian revolution: *Lawrence v. Texas*. In *Cato Supreme Court Review 2002–2003*, ed. J. Swanson, 21. Washington, DC: Center for Constitutional Studies, Cato Institute.
- Bayles, M. 1976. Harm to the unconceived. *Philosophy & Public Affairs* 5: 292.
- Bhagwat, A. 1997. Purpose scrutiny in constitutional analysis. *California Law Review* 85: 299.

- Brock, D. 1995a. Procreative liberty. *Texas Law Review* 74: 187 (1995).
- Brock, D. 1995b. The non-identity problem and genetic harms: The case of wrongful handicaps. *Bioethics* 9: 269.
- Chemerinsky, E. 1997. *Constitutional law principles and policies*. New York: Aspen Law & Business.
- Epstein, R. A. 2004. Of same sex relationships and affirmative action: The covert libertarianism of the United States Supreme Court. *Supreme Court Economic Review* 12: 75.
- Fallon, Jr., R. 1993. Individual rights and the powers of government. *Georgia Law Review* 27: 343.
- Feinberg, J. 1986. Wrongful life and the counterfactual element in harming. *Social Philosophy & Policy* 4: 145.
- Gottlieb, S. 1988. Compelling government interest: An essential but unanalyzed term in constitutional adjudication. *Boston University Law Review* 68: 917.
- Gunther, G. 1972. Foreword: In search of evolving doctrine on a changing court: A model for a newer equal protection. *Harvard Law Review* 86: 1.
- Hurka, T. 1983. Value and population size. *Ethics* 93: 496.
- Laudor, M. 1994. In defense of wrongful life: Bringing political theory to the defense of a tort. *Fordham Law Review* 62: 1675.
- Mill, J.S., 1985. *On liberty*, ed. D. Bromwich and G. Kate. New Haven: Yale University Press.
- Narveson, J. 1976. Moral problems of population. In *Ethics & population*, ed. M. Bayles, 59. Cambridge, MA: Schenkman.
- Parfit, D. 1976. On doing the best for our children. In *Ethics & population*, ed. M. Bayles, 100–115. Cambridge, MA: Schenkman.
- Parfit, D. 1984. *Reasons and persons*. Oxford: Clarendon Press.
- Peters, Jr. P. 1989. Protecting the unconceived: Nonexistence, avoidability, and reproductive technology. *Arizona Law Review* 31: 487.
- Peters, Jr., P. 2004. *How safe is safe enough? Obligations to the children of reproductive technology*. New York: Oxford University Press.
- Pildes, R. 1994. Avoiding balancing: The role of exclusionary reasons in constitutional law. *Hastings Law Journal* 45: 711.
- Rawls, J. 1971. *A theory of justice*. Cambridge, MA: Belknap Press of Harvard University.
- Robertson, J. 1997. Wrongful life, federalism, and procreative liberty: A critique of the NBAC cloning report. *Jurimetrics* 38: 69.
- Sheppard, S. 1994. The state interest in the good citizen: Constitutional balance between the citizen and the perfectionist state. *Hastings Law Journal* 45: 969.
- Secunda, P.M. 2005. *Lawrence's* quintessential Millian moment and its impact on the doctrine of unconstitutional conditions. *Villanova Law Review* 50: 117.
- St. Louis Post-Dispatch, Diseases sperm forces fertility clinics to close, Nov. 30. 1997, p. A5.
- Sunstein, C. 2002. Is there a constitutional right to clone? *Hastings Law Journal* 53: 987.

Chapter 16

Reparations for U.S. Slavery and Justice Over Time

Seana Valentine Shiffirin

Abstract This paper extends a non-comparative account of harm to the issue of reparations for identity-affecting historical injustices, focusing on American slavery. Many object to reparations for slavery because the original victims and perpetrators of slavery are dead and because the lives and identities of the descendants of slaves are causally dependent upon the fact of slavery. This objection misses the point of reparations. What is central to reparations claims is not any idea that some particular individuals, considered in isolation from social groups, either owe or are owed. Rather, reparations arguments are centrally concerned with the relations between and histories of important social collectives and how these relations in turn affect individuals. Reparations should serve as, and be seen as, a concrete rather than merely discursive repudiation of slavery and commitment to its non-recurrence.

Keywords Reparations · Slavery · Groups · Non-identity problem · History · Justice · Time.

Many philosophers treat the reparations question as another occasion on which the non-identity problem has bite. They worry about the temporal distance between the crimes and wrongs committed and the proposed moments of reckoning with them, taking it to be a serious issue that—to put it starkly—the original victims of slavery are dead, and the original perpetrators are no longer alive either.¹ Whereas, I worry that there are no genuine philosophical problems here and that even refuting them may do little but to perpetuate interest in detours and red herrings. This is a real hazard, but nonetheless reflecting on the persistent cultural anxiety about reparations and the past may help to clarify the point of reparations and to reveal aspects of the intimate connection between our self-conception and group-identification.

The fact that the institution of U.S. slavery officially ended long ago might initially be thought to be philosophically relevant for four reasons: First, it might be thought that repair is not possible because the injured are no longer alive. Second,

S.V. Shiffirin (✉)
Law and Philosophy, UCLA, Los Angeles, CA, USA
e-mail: shiffirin@law.ucla.edu

it might be thought that repair is not possible because the perpetrators are not alive: since “we” did not participate in the crimes ourselves, we cannot sincerely make reparations for that which we did not do. Third, issues of possibility aside, some complain that it is unfair for contemporary agents—who, with respect to slavery, are non-involved innocents—to have to bear burdens for the morally criminal behavior of others who acted in the past.² Fourth, the individual identities of the descendants of slaves depends upon slavery; were it not for the institution of slavery and the social conditions it produced, people would not have met and conceived when they did and different children would have been born than in fact were. Some suggest that this poses a problem for reparations claims: even if the descendants of slaves are worse off than others and their condition is attributable to the lasting effects of slavery, no particular individuals may claim that they are worse off than they otherwise would be *because* of slavery, because they would not even exist but for slavery.³

With respect to this last concern: Although similar arguments about identity in a variety of contexts have given thinkers as esteemed and various as Derek Parfit and Frederick Hayek pause, I do not regard them as serious obstructions.⁴ As I have argued elsewhere, in the context of future generations, the so-called non-identity problem trades on an implausible, economically-influenced, caricature of what constitutes a harm or a wrong.⁵ On this view, a claim that one has been harmed or wronged is undercut by a demonstration that the overall position of the person or the overall outcome of the episode in question is beneficial; the benefits outweigh the costs and so there is no harm or no wrong. If existence isn’t a nightmare, then the conditions productive of the overall good of a human life cannot be considered an overall harm.

Such a view is, I believe, deeply mistaken. I submit that it is more plausible to think that harms and wrongs may occur when certain conditions of a sufficiently bad sort are inflicted or when boundaries of respect are transgressed; salutary outcomes of other sorts cannot nullify or negate these transgressions. An insult to one’s self-respect does not disappear if it accidentally results in a financial windfall. One may be harmed or wronged even if in some large accounting book, one is made no worse off overall. So, slavery may be the condition of a person’s existence, the life is overall a good, and still, its effects can harm and deliver the sort of disrespect that a person may legitimately complain of.

The second weakness of this position is that it is not only the descendants of slaves whose identities depend on these past crimes. It is probably fair to say of each of us—the descendants of slaves, the descendants of slave-owners, the descendants of beneficiaries of slavery, descendants of abolitionists, immigrants, and the descendants of immigrants—that we each owe our existence to conditions produced by slavery and to events stemming from slavery’s morally criminal nature. The descendants of slave-owners and other beneficiaries of slavery may not have participated in slavery directly, but inheriting the benefits of slavery as well as a debt—a moral liability to apologize and to make repair—is arguably the condition of their/our existence; so on this argument, they could not claim to be made worse off for it. If these considerations jeopardize the individualistic claims of individual descendants to repair, they seem also to jeopardize the claims of individuals to resist

reparations. One might object: we are worse off financially than we would otherwise be if we must pay reparations as compared to if we are not so obliged; but so too the descendants of slaves are worse off if reparations are not made as compared to if they are. I do not doubt that we could play this game all afternoon but I doubt we would learn much because the exchange is beside the point. What is central to any case for reparations is not the idea that some particular individual, considered in isolation from a social group but merely considered as a descendant of someone involved in the institution of slavery, either owes or is owed reparations. Rather, reparations arguments are more centrally concerned with the relations between, and histories of, important social collectives and how this affects individuals.⁶ Reparations arguments regarding slavery seek reparations from a collective social group and direct them at another collective social group (although there may be overlap in membership); they do not seek to blame or to hold particular individuals responsible.

So let us return to what I take to be the more serious, although also misbegotten, objection that the institution of slavery ended before our lives began. If the harmful effects of slavery lay entirely in the past and our society were fully egalitarian, I might agree that reparations efforts were unnecessary. But the notion that there are no contemporary victims of slavery is laughable. African Americans suffer in palpable, material ways that have strong genealogical connections to the institution of slavery, the past material deprivation it involved, and the forms of racist hatred and insidious indifference that both were its fuel and are its heavy residue.⁷ In fact, I suspect that it is part of the lingering, damaging effect of slavery that serious reparations efforts (or some reasonable facsimile) have not been undertaken. For much of the impetus of reparations is an understandable demand that it be made explicit to this and to future generations that there is a sincere, ironclad commitment to repudiating the wrong of the past, its animating rationale, its residual effects, and its repetition.⁸

This characterization makes the point of reparations to be primarily expressive and communicative. For some, this may make it seem mysterious why apologies and abolition do not suffice—why demands for resources form part of the movement. From the other side, reparations activists may think there is a risk that such characterizations gloss over the central notion of repair.

Tangible actions and not only words are central to the idea of reparations. But I believe this is because, and here I softpedal the problem, the break with slavery and its ideological underpinnings has never been decisive. At this point, to achieve the expressive aim that animates reparations claims requires greater measures. To be sure, official apologies and symbolic changes (such as rebaptizing university halls and monuments named to honor slaveholders) should not be under-rated: official acknowledgment of wrongdoing, especially when it follows heavy resistance to do so, has hard-won meaning. But in light of a shameful and long history of an ambiguous separation from slavery, declarations seem to be a cheap and insufficient means to communicate sincere regret. Making real sacrifices and engaging in significant action may be necessary to communicate genuine compunction and commitment.⁹

As to the complaint that these crimes were committed in the past by other people: this assumes that there is no ongoing complicity with these past crimes,

an assumption immediately belied by minor practices such as honoring slavery defenders and slave owners through monuments and other named structures, by much larger patterns of structural and cultural racism, and by the more general economic advantages passed on to some and not others because substantial reparations were never paid in the first place.¹⁰ But, even on this assumption, the complaint assumes a sort of misplaced individualism, that somehow our collective obligation is undermined by a demonstration that no one of us is personally responsible for what occurred and that the important events happened long ago on others' watch.

This objection must strike at least the lawyers among us as a little rich. Lawyers, through appeals to precedent and the like, are in the general business of directing that and how activity should be governed according to the directions of people long dead. If precedent and legislation undertaken long ago have some, albeit defeasible, hold on us, then why shouldn't we also bear responsibility for the criminal activity of people in the past whose actions characterized and represented the ongoing institutions in which we participate? The idea that it is relevant that these are *past* crimes may result from a confusion of reparations with recriminations—as though particular individuals or present generations are to be blamed for the actions of our ancestors. This misinterprets what I take to be the impetus for reparations: a desire to address the needs of those who suffer the vestiges of past wrongs; to repudiate in a sincere and convincing way the crimes of the past as evidence of a commitment to end their continuing influence; and to forestall their repetition. To be asked to make such a commitment does not imply recrimination or any sort of individualized blame, but rather an acknowledgment of the past of a group with which we are affiliate and identify.

For the most part, reparations claims are addressed to social institutions—such as the state, universities, and companies. This is significant because these are collective endeavors meant to endure and to build on themselves over time. Our aims in these collective enterprises are to form groups that together not only achieve individualistically-understood goals more efficiently but that generate a variety of human goods and opportunities that could not be produced alone even with more time and resources. Joint collaborative efforts allow for mutual sharing, enrichment, and other goods that essentially must involve the collaboration, interdependence, and mutual vulnerability of distinct, separate minds. The achievement conditions of these goods depend upon our adopting attitudes of collective identity and pride, by taking it for granted that we together are joint authors of the product of our contributions; by recognizing that this joint product is ours to continue to build upon (not separably the product of the individual contributors). If we took the attitude (and the actions that would follow from it) that every generation or each person starts anew (as opposed to the attitude that each generation must make its own contribution and be willing to steer us in a starkly new direction), these achievements could not occur. Universities in particular are dedicated to this view. We together, by trusting, relying upon and integrating ourselves with others' work, are building and adding to a joint pool of knowledge. The progress we make has to presuppose the work of others in the past. Through time and through collaboration with the past, we together produce

insights, knowledge, and understanding of subjects at a level that could never be done individually and that even could never been done if we attempted to account strictly for each person's contribution. Similarly, we together over generations build an infrastructure, a legal, economic and social system—a state. Feelings of pride and identification substantially contribute to and reflect what is presupposed by these forms of focused joint collaboration. But if this rough and quick sketch gestures at anything like the correct model, then it should also follow that we must take on as part of our identification and acknowledge responsibility—as a group- for our bad actions in the past as well—even if no one of us is responsible for these actions and even if we all now would forswear such actions as antithetical to our identity and commitments.¹¹ Thus, I believe it is not only possible for “us” to make reparations, but also perfectly fair for us to make sacrifices to do so.

Thus, I do not think we should be troubled by the standard objections to reparations that revolve around concerns about the connections and ruptures between past and future generations. I am more bothered by a distinct concern raised by those planks of some reparations platforms that call for economic redistribution and institutional social reform as a form of reparation. I have been assuming that the African American population suffers from economic inequality and that the suffering at least partly has its explanations in the history of slavery. But the suffering of this population, and others, deserves attention as such. It should not be a requisite for economic and social reform that the serious obstacles faced by African Americans be traced to slavery. There are other independent duties of justice to redistribute and not to leave groups behind.

It might be thought, though, that the implicit message of the reparations movement is that the reparations arguments provide an independent reason for economic redistribution, stronger than the general reasons to do economic justice that have thus far been predominately ignored. Of course, the reparations arguments may represent a political gambit—where the general obligation to share our social product has failed to motivate the country, perhaps we must turn to those reasons that appeal to the tendency to respond to blame. I hesitate about this strategy. The notion that a culpable cause is a requisite for social attention has fed a certain righteous neglect on the grounds that either the suffering part of the population is to blame or the comfortable part of the population is not. While I am loathe to indulge this impulse, I recognize that such reluctance may reflect unwarranted optimism about the future of progressive politics. Blame games may never disappear and it may be better to play them successfully than to wait indefinitely for others' epiphanies of conscience. So I do not want to challenge the pragmatics of reparations arguments although we should keep their hazards clearly in view.

From a philosophical point of view, I am more curious about whether the appeal to reparations *must* be seen as a second best move by liberals. Given that we all should engage in substantial redistribution whether the inequalities arise inadvertently from arbitrary operations of the market, geography, bad luck or from past crimes, should history matter here? Is there an independent reason for social and economic reform that appeals to the fact that these conditions have an especially heinous history?

Although the question is philosophical, it may have practical import: if reparations arguments do not merely overlap general arguments for social justice, then questions may arise about what compensation is due and to whom. Redistributing from a motive of repair may require a greater level of compensation or a sort different in kind than redistributive motives would. Of course, reparations programs need not necessarily serve different populations than more general redistributive programs. We might start from a motive of reparation but expand the pool of beneficiaries; for example, we might announce a reparations motive for establishing a tuition-free university in an area without many higher education institutions and where many under-served African-Americans live, but allow anyone to enroll. Or, if we were concerned that African-Americans have a harder time purchasing homes than white Americans because they, on average, have a harder time qualifying for low-interest loans or have smaller nest eggs for deposits, we might establish state run low-interest loan programs with lower down payment expectations; we might establish such programs from a primary motive of reparation, but make their benefits available to all buyers below a certain income level. Although programs might be structured in this way, we should, nonetheless acknowledge that in an imperfect world of limited resources and limited political will, reparations efforts might come into conflict with other redistributive aims. If so, we need an account of whether reparations arguments have independent traction and run on an independent track or whether they just point us more vividly to our general redistributive obligations.

I start from uncertainty about this question: some of us have the sense that if real economic and social equality were in place now (which it is not), there would be little reason to pursue reparations, especially if they conflicted with redistributive obligations. On the other hand, although both failures are significant, I believe the appalling wrong of slavery is of a magnitude worse than the shamefulness involved in our failure to rectify serious economic need and stratification. In the end, I think something like the following argument has weight: the wrongs of slavery are so severe, and understandably have such deep effects across generations, that we can only recover and move forward productively by facing it squarely, by acknowledging its past horror through public measures of repudiation of some large significance, even if this involves some sacrifices to furthering other compulsory ends. And, as I suggested before, I do not think that it is an accident that we do not actually confront the hypothetical problem of asking whether reparations could be required in a context of equality. I suspect we could never have this dream problem because given our past and its partial perpetuation through denial and through ongoing racism, something significant like building a public consensus around reparations and making a large gesture, symbolic and substantive, may be required to establish the sort of mutual trust and respect necessary to sustain real equality.

Acknowledgements This is a revised version of a talk for a conference at the Yale Law School entitled "Yale, New Haven, and American Slavery." The stress in the text on universities is partly a product of that occasion. I am grateful to the audience and organizers for stimulating comments and, in particular, to co-panelists Anthony Appiah, Ronald Dworkin, and Charles Fried. Thanks also to Jeffrey Helmreich, Barbara Herman, and Steve Shiffrin for constructive advice and criticism.

Notes

1. See, for example, Sher (1980, 2005), Simon (1974), and Thompson (2001).
2. Simon (1974), pp. 317–18. Simon argues that present-day Americans are in many ways “innocent bystanders” in the injustice their ancestors perpetrated against African Americans. Fletcher argues that holding someone responsible for a past injustice because of her ancestry violates the principle of equality of persons. See Fletcher (2002), p. 1548.
3. Thompson (2001), p. 117.
4. Parfit (1984), pp. 358–60.
5. Shiffrin (1999), pp. 122–23.
6. See also Taylor (1973), pp. 178–80. Taylor argues that the obligation to compensate for past injustices perpetrated against an entire group “does not fall upon any particular individual but upon the society as a whole.”
7. See e.g., Jaynes and Williams (1989).
8. See also Weyeneth (2001), pp. 15–16. According to Weyeneth, a central purpose of issuing “historical apologies” is to declare one’s accountability to the victimized group and commitment to change the relationship.
9. Charles J. Ogletree, Jr. argues that a central motivation behind of the reparations movement is to ensure that “all of American history be fully acknowledged, accounted for, and valued” and that such full acknowledgement connotes “something more than token acknowledgment of the centuries of suffering.” Ogletree (2003), pp. 282 and 318.
10. For examples of cultural honors bestowed upon past slaveholders, see Zernike (2001) and Rappleye (2006). A short discussion of alternative ways to gauge the estimated impact on current economic stratification between whites and blacks of slavery (and the failure to make sufficient reparations after the Civil War) appears in Conley (2003).
11. See also Dworkin (1986), pp. 167–75.

References

- Conley, D. 2003. The cost of slavery. *New York Times*, February 15, 2003.
- Dworkin, R. 1986. *Law's empire*. Cambridge, MA: Belknap Press.
- Fletcher, G. 2002. The Storrs lectures: Liberals and romantics at war: The problem of collective guilt. *Yale Law Journal* 111(7): 1499–1573.
- Jaynes, G.D. and R. Williams, Jr., eds. 1989. *A common destiny: Blacks and the American society*. Washington, D.C.: National Academy Press.
- Ogletree, C. Jr. 2003. Repairing the past: New efforts in the reparations debate in America. *Harvard Civil Rights-Civil Liberties Law Review* 38: 279–320.
- Parfit, D. 1984. *Reasons and persons*. Oxford: Clarendon Press.
- Rappleye, C. 2006. *Sons of providence: The Brown brothers, the slave trade, and the American revolution*. New York: Simon & Schuster.
- Sher, G. 1980. Ancient wrongs and modern rights. *Philosophy and Public Affairs* 10(1): 3–17.
- Sher, G. 2005. Transgenerational compensation. *Philosophy and Public Affairs* 33(2): 181–200.
- Shiffrin, S. 1999. Wrongful life, proactive responsibility, and the significance of harm. *Legal Theory* 5: 117–48.
- Simon, R. 1974. Preferential hiring: A reply to Judith Jarvis Thomson. *Philosophy and Public Affairs* 3(3): 312–20
- Taylor, P. 1973. Reverse discrimination and compensatory justice. *Analysis* 33(6): 177–82.
- Thompson, J. 2001. Historical injustice and reparations: Justifying claims of descendants. *Ethics* 112(1): 114–35.
- Weyeneth, R. 2001. The power of apology and the process of historical reconciliation. *The Public Historian* 23(3): 9–38.
- Zernike, K. 2001. Slave traders in Yale’s past fuel debate on restitution. *New York Times* (August 13, 2001): B1.

Name Index

A

Aleinikoff, T.A., 329 n.9
Appiah, K.A., 72, 88 n.1
Archard, D., 166, 167, 176 n.29, 176 n.30
Aristotle, 256, 262 n.8
Armstrong, D.M., 89 n.20
Arntzenius, F., 226 n.41
Arrhenius, G., xxxii–xxxiii, 289–309, 309 n.2,
309 n.3, 310 n.4, 310 n.8, 311 n.21,
311 n.23, 311 n.25, 312 n.39, 312 n.40,
313 n.46
Asch, A., 284 n.36

B

Bales, R.E., 313 n.45
Barnett, R.E., 329 n.23
Bayles, M.D., 89 n.17, 330 n.26
Benatar, D., 38, 39, 43, 44, 46 n.26, 46 n.27,
46 n.28, 46 n.29, 46 n.30, 46 n.31, 46 n.32,
46 n.33, 46 n.34, 46 n.43, 46 n.44, 46 n.45,
57, 62, 63, 67 n.8, 67 n.11, 67 n.12,
67 n.13, 67 n.14, 67 n.15, 67 n.16, 175 n.16
Bettelheim, B., 176 n.39
Bhagwat, A., 329 n.9
Brink, D.O., 247 n.10, 313 n.45
Brock, D., 169, 172, 176 n.40, 176 n.46,
281 n.5, 327, 330 n.27, 330 n.34, 330 n.47,
330 n.49
Broome, J., xxxiii, xxxv n.2, 41, 46 n.39,
89 n.18, 113 n.7, 133 n.10, 202, 223 n.6,
226 n.41, 247 n.2, 247 n.14, 247 n.15,
247 n.16, 247 n.19, 247 n.23, 289, 293,
295, 309 n.3, 310 n.7, 310 n.8, 310 n.11,
311 n.17, 311 n.18, 313 n.44
Buchanan, A., xxxvi n.4, xxxvi n.16, 21 n.10,
21 n.11, 21 n.12, 176 n.45, 225 n.27,
247 n.2
Butler, S., 93, 94, 112, 113 n.1, 113 n.2
Bykvist, K., 309 n.4, 311 n.20, 311 n.21

C

Carlson, E., 313 n.43
Carter, W.R., 90 n.35
Chase, J., 226 n.38, 227 n.45
Chemerinsky, E., 329 n.5
Clark, M., 226 n.41
Cohen, C.B., 165, 176 n.28
Collins, J., 197 n.25
Conley, D., 339 n.10
Cook, M., 226 n.41, 227 n.45

D

Danielsson, S., 313 n.45
Dasgupta, P., 133 n.12
Davidson, M., 22 n.17
Dawkins, R., 94, 99, 113 n.3
De Shalit, A., 113 n.4
DeGrazia, D., 10, 18, 21 n.13, 22 n.14, 24 n.39,
24 n.42
Dworkin, R., 202, 223 n.3, 339 n.11

E

Elliot, R., 85, 90 n.34, 90 n.36
Epstein, R.A., 329 n.23

F

Fallon, Jr., R., 329 n.9
Fehige, C., 89 n.16
Feinberg, J., 113 n.5, 113 n.6, 132 n.4, 159,
160, 161, 166, 167, 175 n.18, 175 n.21,
276, 283 n.26, 283 n.30, 330 n.26
Feldman, F., 312 n.39, 313 n.43
Fletcher, G., 339 n.2

G

Garrett, B., 90 n.27
Gauthier, D., 247 n.9
Gjelsvik, O., 226 n.38, 227 n.42

Glover, J., 88 n.4, 88 n.13, 172, 176 n.47,
247 n.1, 260, 262 n.14, 272, 282 n.19,
309 n.1

Goldberg, C., 174 n.5

Gottlieb, S., 329 n.9

Griffin, J., 133 n.17

Gunther, G., 329 n.5

H

Habermas, J., 4, 21 n.1

Hanser, M., xxvii–xxx, xxxvii n.26, 88 n.9,
148, 153 n.16, 154 n.21, 179–196, 196 n.6,
197 n.17, 197 n.20, 223 n.4, 275, 276, 277,
278, 281 n.6, 283 n.24, 283 n.25, 283 n.28,
283 n.32, 284 n.35

Hare, C., xxxvi n.23, xxxvii n.29, 22 n.24, 106,
113 n.11, 113 n.12, 113 n.13, 224 n.20,
224 n.21

Hare, R.M., 68 n.20

Harman, E., xxvii–xxx, xxxiv, xxxvi n.25,
22 n.18, 68 n.21, 137–151, 151 n.3,
151 n.4, 151 n.5, 152 n.9, 153 n.12,
153 n.16, 154 n.21, 180, 182, 183, 184,
186, 187, 188, 196 n.4, 196 n.7, 196 n.8,
197 n.11, 197 n.16, 197 n.19, 197 n.22,
197 n.26, 198 n.28, 199 n.36, 281 n.11

Harris, J., 159–161, 167, 168, 170, 175 n.13,
175 n.17, 176 n.33

Hart, H., 45 n.2

Herissone-Kelly, P., xxxi–xxxii, 249–262

Heyd, D., xxiii–xxiv, xxvi, xxvii, xxxvi n.13,
xxxvi n.19, 3–21, 21 n.2, 21 n.6, 22 n.22,
23 n.36, 67 n.9, 88 n.9, 89 n.18, 132 n.4,
133 n.7, 155, 174 n.1, 197 n.13, 224 n.22,
293, 310 n.12, 310 n.13, 311 n.17

Holtug, N., xxvi–xxvii, xxxi, xxxiii, 71–88,
88 n.4, 88 n.10, 88 n.12, 90 n.21, 90 n.23,
90 n.26, 90 n.30, 90 n.39, 90 n.40, 133 n.6,
309 n.1, 311 n.21

Hooker, B., 123, 124, 125, 127, 133 n.13,
133 n.15, 133 n.16

Hurka, T., 326, 330 n.43, 330 n.44

J

Jaynes, G.D., 339 n.7

K

Kagan, S., xxxvi n.14, 223 n.2

Kamm, F., 46 n.13, 57, 67 n.7, 175–176 n.26

Kavka, G.S., xvii, xix, xxx, xxxii, xxxvi n.6,
xxxvi n.7, xxxvi n.11, xxxvii n.27, 14,
23 n.29, 88 n.9, 202, 203, 210, 223 n.5,
223 n.7, 223 n.9, 223 n.11, 225 n.30,
225 n.32

Kripke, S., 18, 88 n.8

Kuhse, H., 90 n.38

Kumar, R., xxxii, xxxvii n.27, 11, 12, 22 n.19,
132 n.4, 133 n.19

Kymlicka, W., 88 n.2, 313 n.45

L

Laudor, M., 330 n.35, 330 n.38

Lewis, D., 90 n.20, 197 n.23, 247 n.22

Liao, M., 175 n.25

Lillehammer, H., xxxi–xxxii, 231–246,
247 n.4, 267, 281 n.8

Locke, J., 4, 34, 35, 36, 46 n.11, 46 n.12,
46 n.16

Lyons, D., 133 n.15

M

McCarthy, D., 226 n.41

McClamrock, R., 175 n.11, 281 n.2,
281 n.13

McGrew, T.J., 221, 226 n.41, 227 n.45,
227 n.46, 227 n.47

McKie, J., 89 n.17

MacLean, D., 106, 107, 108, 113 n.10,
113 n.15

McMahan, J., xxiv–xxvi, xxxvi n.4, xxxvi n.20,
10, 11, 22 n.15, 22 n.16, 24 n.41, 29, 34,
38, 39, 43, 44, 45 n.1, 46 n.7, 46 n.8,
46 n.9, 46 n.10, 46 n.25, 46 n.41, 46 n.42,
46 n.46, 49–67, 67 n.1, 67 n.3, 67 n.4,
67 n.9, 67 n.10, 68 n.17, 68 n.18, 68 n.19,
89 n.15, 90 n.35, 132 n.4, 133 n.9,
176 n.41, 196 n.2, 224 n.16, 247 n.2,
247 n.7, 281 n.5, 283 n.29

Markie, P.J., 22 n.19

Markosian, N., 227 n.45, 227 n.46, 227 n.48

Meacham, C.J.G., 226 n.41

Meyer, L., 312 n.33

Mill, J.S., 96, 97, 98, 113 n.6, 125, 130, 321,
329 n.17

Miller, D., 88 n.2

Moller Okin, S., 46 n.15

Moore, G.E., 310 n.5

Morgan, D., 175 n.10

Mulgan, T., xxvi–xxvii, xxxvi n.24, 11,
22 n.17, 115–132, 132 n.3, 133 n.5,
133 n.11, 133 n.13, 133 n.14, 133 n.15,
133 n.17, 133 n.18, 133 n.19, 133 n.21,
247 n.2, 247 n.3, 247 n.8

N

Narveson, J., xxxvi n.15, xxxvii n.29, 207,
224 n.22, 224 n.23, 309 n.1, 310 n.12, 326,
330 n.45

Noonan, H., 82, 83, 90 n.28
 Nozick, R., 46 n.16

O

Oderberg, D., 90 n.35
 Ogletree, C. Jr., 339 n.9
 Olson, E.T., 90 n.35
 Orsi, F., 133 n.19

P

Parfit, D., xiv, xvii, xviii, xix, xxii, xxvi,
 xxix, xxx, xxxv n.2, xxxvi n.5, xxxvi n.8,
 xxxvi n.9, xxxvi n.10, xxxvi n.11,
 xxxvi n.14, xxxvi n.17, xxxvi n.22, 4,
 5, 7, 8, 9, 11, 13, 14, 15, 16, 17, 21 n.7,
 21 n.8, 21 n.9, 22 n.14, 23 n.29, 23 n.38,
 30, 37, 46 n.23, 46 n.24, 60, 67 n.2, 73,
 76, 80, 88 n.3, 88 n.4, 88 n.6, 88 n.7,
 88 n.9, 88 n.11, 88 n.14, 89 n.14, 89 n.18,
 90 n.24, 90 n.25, 90 n.35, 94–98, 99,
 100, 106, 110–112, 113 n.2, 113 n.17,
 113 n.18, 113 n.19, 113 n.20, 113 n.21,
 113 n.22, 117, 132, 132 n.1, 133 n.7, 146,
 151 n.1, 153 n.14, 167, 168, 169, 176 n.35,
 176 n.40, 176 n.41, 179, 180, 181, 185,
 189, 196 n.1, 196 n.2, 202, 203, 223 n.1,
 223 n.2, 223 n.5, 223 n.7, 223 n.8, 223 n.10,
 224 n.26, 247 n.5, 247 n.16, 247 n.17,
 250, 252, 253, 255, 256, 258, 259, 260,
 262 n.3, 262 n.5, 262 n.11, 262 n.12,
 262 n.13, 265, 266, 268, 272, 281 n.1,
 281 n.3, 282 n.14, 283 n.31, 284 n.37,
 289, 301, 302, 306, 309 n.2, 309 n.3,
 311 n.16, 311 n.17, 312 n.34, 313 n.46,
 324, 325, 326, 330 n.26, 330 n.28,
 330 n.29, 330 n.30, 330 n.31, 330 n.32,
 330 n.33, 330 n.36, 330 n.37, 330 n.39,
 330 n.40, 330 n.41, 330 n.42, 334,
 339 n.4
 Parker, M., 251, 262 n.4
 Paul, L.A., 20
 Pellegrino, G., 133 n.19
 Persson, I., xxiv–xxvi, xxxvi n.18, 29–45,
 46 n.8, 46 n.20, 46 n.21, 46 n.38,
 46 n.40, 89 n.15, 90 n.37, 133 n.9,
 224 n.16
 Peters, Jr., P.G., xxxiii–xxxv, xxxvii n.30,
 172, 175 n.12, 176 n.44, 224 n.25, 271,
 282 n.18, 317–329, 329 n.8, 330 n.27,
 330 n.34
 Pildes, R., 329 n.6
 Plantinga, A., 89 n.19
 Priest, G., 113 n.7

R

Railton, P., 247 n.12
 Rappleye, C., 339 n.10
 Rawls, J., 3, 4, 6, 7, 12, 13, 22 n.22, 22 n.23,
 326, 330 n.46
 Reiman, J., xxxvi n.18, 12, 13, 14, 22 n.20,
 22 n.21, 22 n.22, 22 n.23, 106, 107,
 113 n.11
 Roberts, M.A., xiii–xxxv, xxxvii n.29, 3, 14,
 15, 23 n.27, 23 n.28, 23 n.30, 23 n.31,
 23 n.32, 23 n.33, 40, 41, 46 n.35, 46 n.36,
 46 n.37, 49, 88 n.9, 121, 132, 132 n.4,
 133 n.8, 133 n.12, 133 n.19, 158, 162, 163,
 171, 175 n.23, 176 n.37, 201–223, 223 n.2,
 223 n.7, 224 n.15, 224 n.18, 225 n.27,
 225 n.36, 247 n.24, 267, 281 n.7, 281 n.8,
 282 n.21, 294, 299, 300, 301, 309, 310 n.9,
 310 n.14, 310 n.15, 311 n.18, 311 n.26,
 311 n.27, 311 n.28, 311 n.29, 311 n.31,
 312 n.32, 312 n.33
 Robertson, J.A., xxxiv, xxxvii n.32, 158–159,
 161, 162, 164, 168, 174 n.9, 175 n.19,
 224 n.25, 327, 330 n.48

S

Sandøe, P., 90 n.40
 Savulescu, J., 249, 250, 251, 255, 256, 258,
 261, 262 n.1, 262 n.2, 262 n.4, 262 n.6,
 262 n.10, 312 n.33
 Schaffer, J., 153 n.20
 Schwitzgebel, E., 227 n.42
 Scott, A.D., 226 n.41
 Scott, M., 226 n.41
 Secunda, P.M., 329 n.23
 Sen, A., 129, 133 n.20
 Shackel, N., 226 n.41
 Sheppard, S., 329 n.9
 Sher, G., 202, 223 n.5, 339 n.1
 Shiffrin, S.V., xxvii, xxxiii–xxxv, xxxvii n.33,
 22 n.18, 148, 151 n.5, 153 n.15, 153 n.16,
 154 n.21, 180, 182, 183, 184, 185, 186,
 187, 188, 196 n.4, 196 n.9, 196–197 n.10,
 197 n.14, 197 n.19, 197 n.21, 197 n.26,
 198 n.28, 202, 223 n.5, 281 n.11, 333–338,
 339 n.5
 Shoemaker, S., 90 n.35
 Sidgwick, H., 8, 242, 247 n.18, 247 n.20,
 310 n.5
 Sikora, R., 38
 Simon, R., 339 n.1, 339 n.2
 Singer, P., 10, 11, 23 n.26, 90 n.38
 Smolkin, D., 22–23 n.25, 223 n.5
 Snowdon, P.F., 90 n.35
 Stalnaker, R.C., 89 n.19

Steinbock, B., xxvii–xxx, xxxi, xxxiv,
 155–174, 174 n.2, 175 n.11, 175 n.22,
 176 n.31, 176 n.32, 176 n.42, 281 n.2,
 281 n.13
 Steiner, H., 36, 46 n.16, 46 n.17, 46 n.18,
 46 n.19

Stern-Gillet, S., 262 n.9

Sunstein, C., 329 n.11

T

Taylor, P., 339 n.6

Temkin, L.S., 67 n.5, 88 n.9, 132 n.4, 223 n.2,
 289, 291, 309 n.1, 309 n.3, 311 n.19

Thompson, J., 21 n.4, 339 n.1, 339 n.3

Tooley, M., 31, 32, 45 n.3, 45 n.5, 90 n.38,
 281 n.2

U

Unger, P., 90 n.22

V

Vallentyne, P., xxxvi n.14, 223 n.2, 311 n.24

Vehmas, S., 256, 262 n.7

Velleman, J.D., xxxvi n.18, xxxvi n.23, 45 n.3,
 105, 106, 113 n.9

W

Waldron, J., 46 n.14

Wasserman, D.T., xiii–xxxvii, 16, 23 n.34,
 23 n.35, 171, 176 n.43, 197 n.12, 234,
 236, 247 n.6, 265–281, 281 n.9, 281 n.10,
 284 n.36

Weinberg, R., 133 n.19

Weisberg, J., 226 n.41

Wessels, U., 89 n.16

Weyeneth, R., 339 n.8

Wikler, D., 309

Williams, B., 247 n.11, 247 n.13, 247 n.21,
 247 n.23

Williams, Jr., R., 339 n.7

Wolf, C., xxvi–xxvii, 93–113, 113 n.14

Woodward, J., xxxvi n.18, 13, 22 n.25, 88 n.9,
 132 n.4, 169, 176 n.38, 197 n.11, 247 n.17,
 281 n.2, 281 n.13

Z

Zernike, K., 339 n.10

Zohar, N.J., xxvi, xxxvi n.21, 84, 85, 86,
 90 n.31, 90 n.32, 90 n.33

Subject Index

A

- Abortion, 21, 34, 65, 67, 68 n.20, 156, 173,
174 n.2, 175 n.18, 205, 206, 208,
223 n.2, 320, 321, 322
- Acceptable outlook (Principle of), xxxi,
259, 261
- Amniocentesis, xvi, xx, 173
- Asymmetry, xxiv, xxv, xxvi, 10, 11, 49, 52,
53, 54, 59, 61, 62, 64, 65, 67, 67 n.6,
68 n.21, 120, 121, 122, 153 n.15, 165,
166, 184, 224 n.16, 240, 275, 290, 295,
303, 310 n.13
- between creating good and bad lives, 29–46
- between harming and benefiting, 54, 55,
56, 57, 58, 66
- Average Utilitarianism/Average
View/Averagism, 206, 208,
223 n.2

B

- Bads, xxv, 10, 13, 15, 29–46, 49, 50, 51, 52,
53, 54, 55, 57, 59, 60, 61, 62, 63, 66,
72, 74, 76, 88 n.14, 89, 95, 97, 106,
110, 139, 141, 143, 144, 145, 146, 147,
148, 149, 150, 157, 162, 164, 166, 172,
180, 183, 185, 186, 187, 189, 192,
193, 195, 197 n.22, 198 n.32, 201,
202, 204, 209, 213, 224 n.21, 235,
259, 279, 283 n.29, 292, 301, 322,
325, 337
- Beneficence, xxxi, 13, 29, 30, 33, 37, 38, 40,
42, 44, 45, 171, 176 n.41, 232, 237,
238, 240, 241, 242, 245, 249, 262, 274,
281 n.12
- Benefits, xxv, xxvi, xxviii, xxxi, 22 n.18, 35,
55, 57, 61, 64, 65, 66, 80, 95, 123,
126, 129, 139, 140, 141, 142, 144, 147,
148, 151 n.5, 152 n.10, 153 n.15, n.17,
154 n.21, 166, 172, 181, 182, 188,

- 189, 190, 191, 210, 234, 235, 267,
270, 272, 273, 274, 275, 276, 277,
278, 281 n.14, 283 n.29, n.32, 329,
334, 338

- existence as a, 6, 37, 43, 45, 58, 62, 67, 76,
79, 88 n.14, 89, 108, 161
- compared with or balanced against
harms, 192

C

- Can't-Do-Better Problem, 202, 209, 210,
224 n.26
- Can't-Expect-Better Problem, 202, 203, 204,
209, 210, 212, 215, 222, 223 n.7
- Comparativism
- hard, 42, 296
- soft, 290, 295, 296, 301, 302, 303, 304,
305, 306, 307, 308, 309, 312 n.33,
313 n.14
- Compelling state interest, 317, 318, 319, 320,
329 n.4
- Complaint, xxxii, 19, 20, 21, 22 n.19, 23 n.35,
95, 99, 106, 107, 109, 150, 169, 244,
268, 269, 270, 271, 272, 273, 274, 275,
277, 278, 279, 283 n.30, 284 n.37, 301,
302, 335, 336
- Conception, xviii, xxxi, xxxii, 8, 11, 14, 16, 18,
20, 23 n.29, 31, 34, 40, 44, 85, 94, 96,
98, 103, 104, 106, 107, 108, 110, 111,
112, 113, 146, 156, 164, 165, 166, 170,
171, 172, 185, 197 n.17, 203, 205, 207,
223 n.12, 239, 251, 253, 260, 283 n.28,
302, 324, 333
- Consequentialism, xxxi, xxxii, xxxvi n.14, 40,
116, 118, 119, 205, 223 n.2, 225 n.37,
232, 233, 236, 237, 243, 304, 305,
312 n.35
- Constitutional Law, 327

- Creation, xiv, xxxi, 4, 5, 6, 9, 13, 14, 16, 18, 20, 21 n.5, 22 n.18, 44, 52, 60, 118, 124, 218, 233, 234, 266, 267, 268, 270, 272, 276, 284 n.35
- D**
- de dicto*, xxxvi n.23, 18, 22 n.24, 106, 107
- Definite descriptions, 93, 105, 106, 108, 109
- Depletion (case), xix, xxx, 14, 95, 96, 111, 202, 223 n.7, 255, 256, 258, 259, 311 n.16
- de re*, 18, 22 n.24, 225 n.35
- Different number case, 140, 141, 142, 150, 151 n.8, 268, 269, 277, 279
- Disability, xvii, xxxvi n.4, 9, 22–23 n.25, 38, 95, 96, 111, 139, 146, 158, 165, 170, 172, 173, 174, 179, 180, 184, 185, 186, 187, 188, 189, 190, 195, 196, 197 n.18, 198 n.29, n.31, 202, 223 n.12, 225 n.27, 251, 259, 260, 323
- Down syndrome, xv, xvi, xxix, 164
- E**
- Embryos, 84, 162, 173, 208, 252, 256, 257, 261, 262, 280, 320
- Excuse, *see* Justification
- Existence
- as a benefit, 6, 37, 43, 45, 58, 62, 67, 76, 79, 88 n.14, 89, 108, 161
 - comparisons with nonexistence interest in, 76, 77, 78, 79, 80, 83, 89 n.14
- Expected value, xxx, 23 n.30, 123, 204, 212, 213, 214, 215, 216, 219, 222, 225 n.37
- External perspective, xxxi, 250, 253, 254, 255, 256, 257, 258, 260, 261
- F**
- Family Law, 175 n.12, 208
- Fourteenth Amendment, xxxiv, 317
- Fourteen Year Old Girl/14 Year Old Girl, 9, 10, 13, 14, 16, 17, 19, 140, 167, 168, 176 n.37, 259, 266, 267
- Future
- generations, 6, 12, 17, 22 n.17, 97, 106, 107, 122, 125, 234, 236, 245, 265, 266, 274, 276, 323, 324, 328, 334, 335, 337
 - people, 5, 9, 11, 12, 13, 22 n.17, 79, 98, 110, 112, 115, 116, 117, 118, 119, 120, 121, 123, 124, 125, 126, 127, 128, 129, 130, 132, 138, 179, 182, 183, 250, 255, 258, 265–284
- G**
- Gene-therapy, xxvi, 73, 84, 85, 86, 87, 88, 90 n.29
- Genetics, 8, 9, 84, 232
- Genome, xxvi, 4, 84, 85
- Goods, xxv, 50, 52, 53, 54, 55, 57, 58, 61, 62, 64, 65, 66, 67, 163, 165, 166, 186, 278, 279, 281 n.12, 283 n.29, 291, 292, 310 n.5, 336
- Groups (social), 235, 241, 244, 335
- H**
- Harm/Harming
- comparative, xxii, xxvii, xxviii, xxx, xxxiii, 38, 185, 188, 204, 301–309, 312 n.35, 312 n.42, 313 n.47
 - compared against benefits, 22, 143, 152, 153, 192, 272, 273, 275, 277, 283
 - counterfactual, 54, 96, 159, 160, 172, 185, 224 n.24, 312 n.33
 - noncomparative/non-comparative, xxiii, xxv, xxviii, xxix, 50, 51, 52, 56, 57, 59, 61, 62, 64, 89, 185, 188, 191, 269
- Human good, 291, 292, 336
- Huntington's Chorea/Huntington's Disease, xv, xvi, 208, 209
- I**
- Ideal Code, xxvii, 123, 124, 126, 129, 130, 131
- Identity
- narrative/biographical, 17, 18, 19, 20, 24 n.39
 - numerical/metaphysical, 4, 12, 13, 18, 19, 20, 22 n.23, 71, 72, 84, 127, 130
- Impairments, 23 n.35, 99, 148, 149, 171, 202, 223 n.12, 267, 278, 279, 283 n.31, 284 n.38
- Impartiality, *see* Partiality
- Injury by substitution, 324, 329
- Interests, xx, xxviii, xxix, 6, 7, 8, 10, 11, 12, 13, 17, 19, 20, 22 n.23, 31, 33, 35, 39, 43, 44, 52, 60, 63, 72, 80, 83, 87, 88, 89 n.14, 97, 99, 106, 108, 120, 128, 158, 161, 164, 166, 172, 174 n.7, 187, 222, 225 n.29, 234, 237, 244, 253, 256, 257, 265, 290, 291, 293, 295, 310 n.9, 318, 320, 321, 338
- Internal perspective, xxxi, 250, 254, 256, 257, 258, 259, 260
- Intransitive, *see* Transitive/Transitivity
- In Vitro fertilization/IVF, 162, 173, 231, 234, 319
- J**
- Justice, 4, 5, 6, 7, 8, 12, 13, 17, 22 n.22, 96, 97, 98, 115, 125, 191, 239, 243, 321, 326, 333–339

- Justification, xxv, 17, 33, 37, 39, 54, 64, 119, 123, 131, 137, 143, 147, 154, 156, 175 n.12, 181, 183, 184, 196 n.1, 237, 238, 270, 271, 281 n.14
- L**
- Life [not] worth Living, xxxvi, 15, 49, 50, 51, 55, 64, 65, 118, 121, 122, 161, 166, 260
- M**
- Minimally-decent life, 164, 174, 269
- Moral Actualism, xxxvii, 106, 224
- N**
- No difference view, 119, 120, 121, 122, 126, 129, 130, 170, 171, 176 n.41, 325, 326
- Nonexistence, *see* Existence
- Non-person involving morality, *see* Person involving morality
- Non-transitive, *see* Transitive/Transitivity
- P**
- Parent/Parenting, xvi, xix, xxix, xxx, xxxi, 145, 156–157, 163, 164, 167, 169, 171, 194, 202, 208, 235, 238, 250, 256–261, 267, 275, 278, 280, 284 n.38
- Partiality, xxxii, 127, 231–247, 266, 267, 268, 270, 271, 278, 279, 280
- Person-affecting
- approach, 4, 5, 9, 11, 13, 14, 18, 23 n.34, 118, 119, 120, 132, 224 n.25
 - intuition, 11, 118
 - morality/ethics, 72, 75, 79, 97
 - narrow, 14, 17, 74, 75, 76, 180, 197 n.20
 - principles, 9, 10, 11, 15, 74, 75, 76, 79, 86, 97, 98, 117, 118, 126, 128, 162, 266
 - reasons, 85, 86, 87, 88, 174, 176 n.41
 - restrictions, 171, 176 n.37, 289–313
 - wide, 76, 79, 86, 88 n.11
- Person-based
- approach, 201, 205, 206, 208, 225 n.33
 - intuition, 201, 205
- Person-based consequentialism (PBC), 40–41
- Person involving morality, 235, 236, 281 n.8
- Pluralism, 205, 206, 208, 223 n.2, 244, 245, 322
- Population ethics, xxxii, xxxiii, 73, 289–313, 326
- Possible people, 11, 14, 22 n.25, 30, 38, 40, 108, 235, 242, 244, 293
- Possible worlds, 14, 73, 74, 78, 104, 112, 127, 129, 197 n.13, 244, 251, 312 n.35
- Preimplantation Genetic Diagnosis/PGD, 5, 15, 16, 156, 173
- Probability, xvi, 38, 201, 211, 212, 213, 214, 216, 217, 218, 219, 220, 225 n.36, 304, 305
- Procreation, xvii, xix, xxi, xxviii, 4, 11, 17, 30, 31, 33, 34, 36, 53, 54, 55, 58, 59, 60, 61, 62, 64, 143, 144, 145, 163, 164, 165, 167, 172, 174, 175 n.25, 179–199, 205, 258
- Procreative/Reproductive
- beneficence (Principle of), 249
 - intentions, 265–284
 - liberty, 156, 158, 317–330
 - selection/selectivity, 279
- Prospective parent, 16, 164, 165, 169, 174 n.8, 233, 236, 237, 238, 240, 241, 242, 245, 246, 250, 251, 256, 257, 258, 259, 260, 261, 262, 267, 275
- R**
- Reparations, xx, xxx, xxxiv, xxxv, 333–339
- Repugnant Conclusion, xxxiii, xxxvi n.14, 8, 15, 116, 122, 289, 306, 307, 308, 309, 325, 326
- Rights
- general, 30, 31, 32, 33, 34, 35, 36
 - negative, 30, 39
 - positive, 30
 - special, 30, 31, 33, 34, 39
 - violations, 13, 169, 266, 270, 271, 273
- Risk, xvi, 108, 156, 157, 159, 160, 168, 172, 173, 174, 208, 217, 224 n.27, 246, 277, 305, 323, 324, 327, 335
- Risky policy (case), xix, xxix, 179, 180, 181, 182, 183, 185, 193, 197 n.13, 202, 266, 268, 269, 270, 271, 273, 275, 276, 281 n.14, 282 n.22, 283 n.23
- Rule Consequentialism/Rule Utilitarianism, xxvii, 22 n.17, 115–133
- Ruth (case), 284 n.37
- S**
- Same number case, 140, 141, 142, 150, 270, 277, 326
- Same people case, 10, 11, 74, 117, 118, 119, 120, 123, 131, 277, 280, 293
- Simple Consequentialism, *see* Consequentialism
- Slave child (case), xix, xxx, 23 n.30, 201, 202, 211, 212, 215, 217, 218, 222, 223 n.7
- Slavery, xxxv, 23 n.29, 333–339
- Standing, 11, 15, 24 n.42, 53, 64, 112, 159, 186, 187, 195, 197, 219, 223 n.12, 235, 239, 244, 254, 269, 271, 326

Substitution/Replacement (of one individual by another), 82, 84, 160, 172, 173, 325, 329

T

Threshold for worthwhile life, 13, 119, 270, 284 n.37

See also Minimally-decent life

Tort law, xix, xxxiii, xxxiv, 11, 159, 175 n.12, 208, 224 n.25, 327, 328

Total Utilitarianism/Total View/Totalism, xxi, xxxvi n.14, 122, 123, 126, 129, 206, 208, 223 n.2, 290, 292, 297, 302, 306, 307, 308, 309

Transitive/Transitivity, 82, 296, 303

Two envelope problem, xxx, 201–227

Two Medical Programmes (case), 37, 98, 146, 202, 224 n.26, 301, 302, 303, 309

U

Utilitarian/Utilitarianism, 5, 7, 22 n.17, 29, 38, 115, 116, 125, 129, 130, 132, 240, 289, 290, 292, 297, 302, 306, 307, 322, 323, 325

V

Vagueness, 103, 104, 113 n.8

Value

actual, 20, 23 n.30, 204, 212, 213, 214, 215, 216, 217, 219, 225 n.33

expected, 23 n.30, 123, 204, 212, 213, 214, 215, 216, 219, 220, 222, 223, 227 n.48

instrumental, 242

intrinsic, 42, 76, 77

W

Welfare/Wellbeing/Well-being, xiv, xxi, xxii, xxvi, xxxiii, xxxvi, 13, 16, 23 n.33, 40, 51, 52, 54, 55, 57, 61, 66, 72, 73, 74, 75, 76, 77, 79, 80, 87, 89, 122, 123, 128, 129, 130, 131, 155, 156, 158, 161, 164, 165, 174, 176 n.37, 181, 189, 203, 204, 205, 207, 208, 209, 210, 211, 212, 217, 223 n.2, 224 n.14, 225 n.35, 234, 235, 243, 265, 266, 270, 281 n.13, 290, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310 n.13, 311 n.24, 312 n.36, 320, 324, 325, 326, 327, 328, 329, 348 n.1, 310 n.13, 311 n.24, 318, 320, 324, 329

Wrongful life lawsuits, 5, 9, 16, 19, 20, 23 n.34, 158, 159, 312 n.31

Wrong/Wrongdoing, xv, xxi, xxv, xxviii, 107, 170, 201, 202, 206, 208, 212, 335