

MICHAEL SCHEFCZYK &
CHRISTOPH SCHMIDT-PETRI
Editors

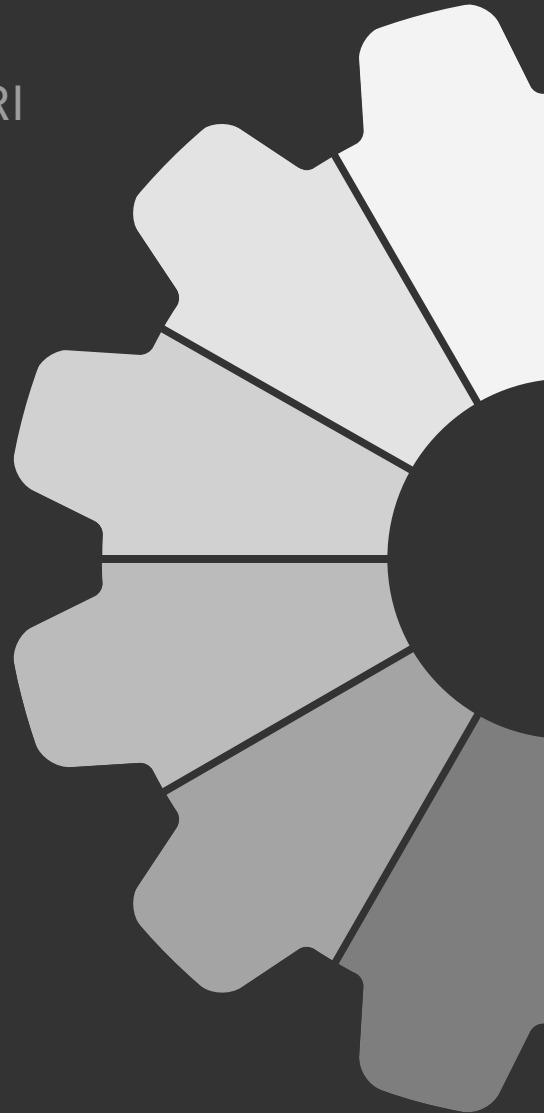
Utility,

Progress,

and Technology

Proceedings of the 15th Conference of the International
Society for Utilitarian Studies

 **KIT** Scientific
Publishing



Michael Schefczyk & Christoph Schmidt-Petri (eds.)

Utility, Progress, and Technology

Proceedings of the 15th Conference of the
International Society for Utilitarian Studies

Utility, Progress, and Technology

Proceedings of the 15th Conference of the
International Society for Utilitarian Studies

edited by
Michael Schefczyk & Christoph Schmidt-Petri

Cover design by Nico Brähler

Impressum



Karlsruher Institut für Technologie (KIT)
KIT Scientific Publishing
Straße am Forum 2
D-76131 Karlsruhe

KIT Scientific Publishing is a registered trademark
of Karlsruhe Institute of Technology.
Reprint using the book cover is not allowed.

www.ksp.kit.edu



*This document – excluding parts marked otherwise, the cover, pictures and graphs –
is licensed under a Creative Commons Attribution-Share Alike 4.0 International License
(CC BY-SA 4.0): <https://creativecommons.org/licenses/by-sa/4.0/deed.en>*



*The cover page is licensed under a Creative Commons
Attribution-No Derivatives 4.0 International License (CC BY-ND 4.0):
<https://creativecommons.org/licenses/by-nd/4.0/deed.en>*

Print on Demand 2021 – Gedruckt auf FSC-zertifiziertem Papier

ISBN 978-3-7315-1108-3

DOI 10.5445/KSP/1000134479

<i>Emily Lanman</i> UTILITARIANISM AND THE ENGLISH POOR LAW REFORM	111
<i>Cheng Li</i> SAMUEL ROMILLY AND JEREMY BENTHAM'S DECISIONS OF PUBLICATION	123
<i>Fayna Fuentes López</i> KILLING ANIMALS: THE BADNESS OF DEATH, VALUE AND REPLACEABILITY	131
<i>Christoph Lumer</i> FROM UTILITARIANISM TO PRIORITARIANISM	139
<i>Christoph Lumer</i> HOW TO DEFINE 'PRIORITARIANISM' AND DISTINGUISH IT FROM (MODERATE) EGALITARIANISM	153
<i>Susanne Mantel</i> THE REASONS OF OBJECTIVE CONSEQUENTIALISM AND COLLECTIVE ACTION PROBLEMS	167
<i>Vincent Emmanuel Mathon</i> A SHELL GAME THEORY – RECONNECT MANKIND WITH NATURE TO CREATE WEALTH	175
<i>Ricardo Miguel</i> AGAINST ANIMAL REPLACEABILITY: A RESTRICTION ON CONSEQUENCES	183
<i>Tim Mulgan</i> WHAT EXACTLY IS WRONG WITH HUMAN EXTINCTION?	193
<i>Ryu Okazaki</i> HEGELS BEGRIFF DER NÜTZLICHKEIT: ZUM ZUSAMMENHANG VON RELIGIONSKRITIK UND TERROR	205
<i>Filimon Peonidis</i> JAMES MILL ON OFFENCES COMMITTED BY THE PRESS	213
<i>Ingmar Persson</i> PARFIT'S REORIENTATION BETWEEN <i>REASONS AND PERSONS</i> AND <i>ON WHAT MATTERS</i>	223
<i>Ingmar Persson</i> PRIORITARIANISM AND THE MORAL NEGATIVITY BIAS	231

<i>Giuseppe Rocché</i> ABOUT THE BADNESS OF EXISTENCE AND THE PROSPECT OF EXTINCTION	239
<i>Michael W. Schmidt</i> SIDGWICK, REFLECTIVE EQUILIBRIUM AND THE TRIVIALITY CHARGE	247
<i>Shingo Segawa</i> IST DER PERSONENBEGRIFF WIRKLICH ÜBERFLÜSSIG FÜR DIE BIOMEDIZINISCHE ETHIK?	259
<i>Adam Shriver</i> IS HEDONISM A VERSION OF AXIOLOGICAL MONISM?	269
<i>Koji Tachibana</i> NEUROFEEDBACK-BASED MORAL ENHANCEMENT AND MORAL REASON	283
<i>Piero Tarantino</i> CREATING AN OBLIGATION: BENTHAM AND THE NORMATIVE QUESTION.....	293
<i>Hiroki Ueno</i> DOES ADAM SMITH’S MORAL THEORY TRULY DIVERGE FROM HUMEAN UTILITARIANISM?	305
<i>Satoshi Yamazaki</i> PIGOU’S THEORY ON WELFARE ECONOMICS IN THE NARROW AND BROADER SENSES: BASED UPON THE INDIRECT UTILITARIAN STRATEGY	315
<i>Alexandra Zinke</i> TWO WAYS TO SATISFY (AND NO WAY TO SATISFY UTILITARIANS)	325
Panel Discussion	
HARE’S UTILITARIANISM, VARNER’S ANIMALS	335
<i>Gary Varner</i> OVERVIEW OF THE BOOK.....	336
<i>Alastair Norcross</i> ON THE MORAL SIGNIFICANCE OF PERSONS, NEAR-PERSONS, AND THE MERELY SENTIENT	339
<i>Adam Shriver</i> COMMENTARY ON VARNER’S <i>PERSONHOOD, ETHICS, AND ANIMAL COGNITION</i>	343

About the Badness of Existence and the Prospect of Extinction

Giuseppe Rocché, University of Palermo, Italy

Abstract

In this paper I consider the case of people who find The Mere Addition Principle counterintuitive. Their particular intuitions may be understood as instances of The Principle of Intrinsic Disvalue of Existence (PIDA). Following this idea, Contractualism seems to be an appropriate method to solve population ethics dilemmas. Still, I show that their rejection of The Mere Addition Principle – if understood as an instance of PIDA – is not enough to avoid these dilemmas and to reach a stable equilibrium among their intuitions. In fact, if their denial of The Mere Addition Principle is grounded on PIDA, the consequences which would follow are likely to be unacceptable for many of them. In particular, either they hold that we have a duty in favor of extinction, or they cannot take PIDA seriously enough. Rejecting the Mere Addition Principle without endorsing PIDA seems the best they can do in order to reach a stable equilibrium among their intuitions.

According to the *Mere Addition Paradox* (Parfit 1984, 419-40; 2004) we cannot consistently hold both

(i) *The Mere Addition Principle*: if a number of people with positive wellbeing is added without affecting the original people's wellbeing, the resulting population (A+) is at least not worse than the original population (A) (Arrhenius 2000, 250)

(ii) *Non-Antiegalitarianism*: considered two populations of the same size, if in the first (B) there is both a higher average wellbeing and more equality than in the second (A+), the first is better than the second (Ng 1989, 238)

(iii) *The Principle of Transitivity and Substitution* (Temkin 1987, 143-44)

and

(iv) *The Denial of the Repugnant Conclusion* – a population (A), of at least ten billion people, in which all its members have a high level of wellbeing is better than a much bigger population (Z) in which all its members have lives that are barely worth living.

Just what a solution to the Mere Addition Paradox consists in is a contentious issue. Speaking of our intuitions about A+ and A, Parfit said that

To avoid the paradox we must believe, without considering the rest of the argument, that A+ is worse than A. [...] To the extent that we find this hard to believe, we still face a paradox (Parfit 1984, 428; 2004, 16)

These words may be taken to mean that we look for a *psychological* solution: solving the puzzle is finding a principle that – given our present attitudes – is *intuitive* (a substantive methodological choice as far as it rules out the use of *bullet biting* – a strategy that Parfit himself endorsed, see Parfit 2016, 120). In this framework, given a clash among the aforementioned four judgments, we settle the issue by showing that, despite the appearances, we do not find one of them really intuitive. In this paper, I propose to wonder whether we feel psychologically compelled by The Mere Addition Principle. Is The Mere Addition Principle really engraved in our psychological outlook? Parfit exposed some natural properties of a Mere Addition – namely that additional people’s lives are worth living and that they do not lessen other people’s wellbeing. Now we should ask whether we happen to care about these natural properties or not.

Some authors take on this challenge arguing that we attach *intrinsic* value to the existence of additional people. Hints of this axiological attitude are – for example – that we think it would have been bad if the happiest share of humanity of the past had never been born (Rachels 1998, 103); that we think would have been a terrible loss if some land on earth had never been populated (Ord 2014, 51); that we would regret being alone in the universe (Tännsjö 2004, 231-32; Rachels 1998, 103). Without discussing these arguments at length, I point out that in some cases our alleged recognition of the intrinsic value of existence is affected by what we may call – in the absence of anything better – *aesthetic* features of lives (Sumner 1996, 21-23). Insofar as we aim at a comprehensive ranking of possible populations, aesthetic features should be taken into consideration, but if we are interested just in *welfarist* axiology, we should neglect them. My proposal is to ask whether we attach value to the creation of people with worthwhile but totally anonymous – devoid of aesthetic value – lives.

Imagine a technological or biological machine capable to create additional people with positive wellbeing without affecting anyone. The machine is currently turned off, but we can easily turn it on so that it will start creating new anonymous lives. These will be long lives with many pleasures and some peaks of sheer bliss, interrupted by some pains, pain will increase as the end gets closer. What to do in this case? I conjecture that I would not turn the machine on. Without holding that this would be the reaction of many people, I assume

that there is a moral tribe¹ which would share my same reaction. The question is whether the members of this tribe have a set of intuitions which avoid population paradoxes.

The refusal to activate the machine could be explained on the basis of different principles, here I shall understand it as an instance of the idea that existence is bad – call it *The Principle of Intrinsic Disvalue of Additions* (PIDA). According to PIDA mere additions are all things considered bad, whereas additions are *prima facie* bad but their badness can be counterbalanced by other goods, above all the *instrumental* value of additions for existing people. Now a problem arises, i.e. that the Repugnant Conclusion is implied even though we recognize that additions of people with a positive wellbeing are intrinsically bad. For we can easily imagine additions, dubbed *Benign Additions* (Huemer 2008), which raise the wellbeing of the original population but lower the average wellbeing of the total population.

Contractualism – choice under a veil of ignorance – offers an answer to this predicament. Imagine (Case 1) you have to choose in self-interested terms whether to be member of a world in which people have a welfare level of 100 or of a world in which those people have a welfare level of 105 and many more people have been added whose lives are barely worth living (level 5). It may be argued (Tännsjö 2004, 211-12) that a self-interested decision-maker in different-number choices has to consider the risk not to exist at all and, therefore, could have reasons to choose the bigger population with a lower average wellbeing. Still, if PIDA is our starting point, parties rather than being averse to the risk to not exist, would be averse to the risk to exist. This idea does not imply the very radical conclusion that (Case 2) given two populations the first with a very high wellbeing, the second with a very low wellbeing but slightly smaller than the first, parties would choose the second world. Parties consider existence as a risk in itself, but not every existence is equally risky and they may be more averse to a low risk of existing and having a life which is barely worth living, than to a high risk of existing and having a life which is well worth living. Then in Case 1 parties would choose the less populated world with higher average wellbeing, whereas in Case 2 they would go for the more populated world.

How much parties – who have been constructed on the basis of PIDA – want to avoid the risk of existing? Different answers to this question shape different conceptions of Contractualism in population ethics. A possible conception of Contractualism is the use of what we may call “*The Same-Number Restriction*” (SNR). According to SNR, parties who are about to choose between different-number scenarios – alternative populations of different sizes – ignore this feature of their choice, so that they think to be choosing between same-number scenarios. In Case 1 parties are facing the choice between two worlds one of them much

¹ I borrow this expression from Greene 2013.

more populated than the other. Anyway, they ignore this fact and know just that in the first world everyone has a very high level of wellbeing, whereas in the second a tiny share of people is even better-off than in the first world but the vast majority has barely worth-living lives.

SNR gives horrible results when negative well-being is concerned. This can be easily proved through Parfit's *Hell 1* and *Hell 2* thought experiment (1984, 392): a world in which a handful of people is suffering hellish torments for fifty years would be worse than a world in which billions of people are suffering the very same torments for fifty years minus a day. In fact, under SNR parties would ignore the fact that in the second world many more miserable people exist. Still, these problems may be avoided by introducing some exceptions into SNR. I do not discuss this point here.

Other cases pose problems which are harder to be solved by means of exceptions. Imagine that – Case 3 – you have to choose between a world in which everyone has a very high welfare level – say 100 – and a world in which the welfare level of the people of the first world has been raised – to, say, 110 – and some people with a welfare level even higher – say, 111 – have been added. We can call these cases *Fair Benign Additions* (FBA), additions in which original people's well-being is raised, so it is the average well-being, and – to exclude cases like Case 1 – additional people have a positive welfare level higher than original people's welfare level after the addition. For many people, intuitively, FBA are never bad, moreover thinking that they are never bad does not entail The Repugnant Conclusion – insofar as the result of FBA is an increase of average well-being. For many people their intuitions about FBA are unproblematic because they do not yield counterintuitive results – as Mere Additions do – when they are put together with their other intuitions.

Still, if we think that existence is somehow bad – see PIDA –, we may also think that some FBA are bad. Imagine – in Case 3 – that the number of additional people with a well-being level of 111 is huge. If – following SNR – we think that the addition should be performed, our endorsement of PIDA is somehow shaky. Actually, it seems one of the weakest endorsement possible. We would recognize that it would be good adding every number of people in order to slightly improve the well-being of an already existing person – by, say, giving him an additional lollipop –, when the additional people are better-off than him. In other words, SNR mirrors the psychological outlook of those who are averse to the risk of existing but are *lexically* more averse to the risk of existing and living a life which is worse than the life they

could have lived². If you endorse PIDA but not this lexical priority, then you should drop SNR.

Hence, if you think that you would refrain from adding billions of people – whose lives would be full of pleasures but also with some pains, especially at their end – when by doing so the well-being of an existent person would be slightly improved – an additional lollipop –, then you feel the need to take the badness of existence – PIDA – more seriously than how SNR implies. Now a major risk is that, once we have found a principle which takes more seriously the idea of the badness of existence, we are bound to accept that we have the duty to stop procreating and to cooperate to realize the extinction of mankind. Some people think we have this duty. This duty is for them intuitive (Benatar 2006, 207), even though they pretend to prove it by means of considerations other than intuitiveness – namely, *bullet biting* and *evolutionary debunking arguments* (Benatar 2006, 202-7). Other people – who endorse PIDA – lack the intuition that we have a duty to bring about the extinction of mankind. A relevant question is whether these people’s intuitions are unstable like those of people who accept The Mere Addition Principle but want to avoid the Repugnant Conclusion.

Perhaps a solution could be provided by a plausible description of what would happen if we head towards extinction. The process realistically would determine that most of last people’s lives would have negative well-being: the prospect of being the last people in the universe fills many of us with anguish. According to a *moderate* lexical account, the badness of FBA can overwhelm the disvalue of a decrease in existing people’s well-being when their well-being remains anyway high and positive, but the badness of FBA cannot overwhelm the disvalue of the production of tormented lives – as we postulate last people’s lives would be.

Even this proposal may be unsatisfying for those who have the intuition that existence is intrinsically bad. Imagine a case – Case 4 – in which just a handful of people exists. If they do not reproduce, then they will have tormented lives. If they have children, their lives will be worth living but they will create billions of people whose lives will be barely worth living. In a case like this, taking PIDA seriously seems to imply that the relevant gain in existing people’s wellbeing cannot make up for the huge disvalue of the addition of billions of people with lives barely worth-living.

² Their attitude must not be confused with *leximin*. We can imagine three welfare level W1 slightly higher than W2, and W2 slightly higher than W3 and two populations P1 (W2) and P2 (W1,W3). According to the lexical principle expressed by SNR, parties could have reasons to choose P2 if the risk to have welfare level W3 is very low – because the corresponding subset of population is very small.

As a last option, perhaps we could cling to a peculiar feature of our condition. In general, we may grant that for every population with a certain welfare level, there is a number of possible people with a certain welfare level whose addition would be optimal. The more we add beyond that number – or we fail to add under that number – the worse is the outcome. In Case 4 there is a large relevant³ disproportion between the number of additional people and the number of existing people whose well-being is positively affected by the addition. But it may be argued, this is not our condition. To avoid to live tormented lives we are not required to create thousands of billions of people. We need just to add roughly as many people as we are: we need just to secure our replacement. Then, in cases like Case 4, where there is a relevant disproportion between additional people and existing people, extinction ought to be chosen; if, on the contrary, a generation can avoid a painful extinction by just “replacing” itself, this is what that population should do. Because our condition resembles this latter case, we have no duty to cause our extinction – on the contrary we should have children.

Many people would find this reasoning nothing but a cunning casuistry. It is somehow true that a generation is directly responsible only for the addition of the next generation, but as far as we can foresee that our successors will have the same bitter alternative between a painful extinction and the addition of another generation, we are indirectly responsible for the addition of that generation as well. Because we will be indirectly responsible for their successors, and for the successors of their successors – and so forth –, our condition does resemble Case 5 in which there is a stark disproportion between the number of additional people and the number of existing people whose well-being is positively affected by the addition.

Concluding, the members of the moral tribe, who share the intuition that would be bad to turn the machine on, may have discovered that their set of intuitions is as shaky as that of the supporters of The Mere Addition Principle. In fact, *if* they understand their response to the machine thought experiment in terms of the endorsement of PIDA, they are forced to accept principles like the lexical account, the moderate lexical account, or the restriction of our moral responsibility to our direct responsibility – the aforementioned “cunning casuistry”. If, on one hand, they are not glad to accept any of these principles, and, on the other hand, they are not psychologically led to revise their judgment in the machine case, they seem to be in a deadlock. My suggestion is that they should look upstream for an understanding of their response in that case different from PIDA. What about the case in which

³ I.e. adjusted considering the welfare levels involved.

the machine would be operative and we could turn it off? If they conjecture that they would leave it working, can they make any sense of this answer?

References

- [1] Arrhenius, Gustaf. 2000. "An Impossibility Theorem for Welfarist Axiologies." *Economics & Philosophy* 16: 247-66.
- [2] Benatar, David. 2006. *Better Never to Have Been*. Oxford: Clarendon Press.
- [3] Greene, Joshua. 2013. *Moral Tribes*. New York: The Penguin Press.
- [4] Huemer, Michael. 2008. "In Defence of Repugnance." *Mind* 117 (468): 899-933.
- [5] Ng, Yew-Kwang. 1989. "What Should We Do About Future Generations? Impossibility of Parfit's Theory X." *Economics & Philosophy* 5 (2): 235-53.
- [6] Ord, Toby. 2014. "Overpopulation or Underpopulation?" In *Is the Planet Full?*, edited by Ian Goldin, 46-60. Oxford: Oxford University Press.
- [7] Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Clarendon Press.
- [8] _____. 2004. "Overpopulation and the Quality of Life." In *The Repugnant Conclusion*, edited by Jasper Ryberg, and Torbjörn Tännsjö, 7-22. Dordrecht: Kluwer Academic Publishers.
- [9] _____. 2016. "Can We Avoid the Repugnant Conclusion?" *Theoria* 82 (2):110-27.
- [10] Rachels, Stuart. 1998. "Is it Good to Make Happy People?" *Bioethics* 12 (2): 93-110.
- [11] Sumner, Leonard W. 1996. *Welfare, Happiness, and Ethics*. Oxford: Clarendon Press.
- [12] Tännsjö, Torbjörn. 2004. "Why We Ought to Accept the Repugnant Conclusion." In *The Repugnant Conclusion*, edited by Jasper Ryberg, and Torbjörn Tännsjö, 219-38. Dordrecht: Kluwer Academic Publishers.
- [13] Temkin, Larry S. 1987. "Intransitivity and the Mere Addition Paradox." *Philosophy and Public Affairs* 16 (2): 138-87.